nature research

Peer Review Information

Journal: Nature Ecology & Evolution
Manuscript Title: Genome sequences of 36-37,000 year-old modern humans at Buran-Kaya III in Crimea
Corresponding author name(s): Thierry Grange, Eva-Maria Geigl

Editorial Notes:

Reviewer Comments & Decisions:

Decision Letter, initial version:

12th September 2019

*Please ensure you delete the link to your author homepage in this e-mail if you wish to forward it to your co-authors.

Dear Eva-Maria,

Your Article, "The origin of the Gravettians: genomic evidence from a 36,000-year-old Eastern European" has now been seen by three reviewers. You will see from their comments copied below that while they find your work of considerable potential interest, they have raised concerns that must be addressed. In light of these comments, we cannot accept the manuscript for publication, but would be very interested in considering a revised version that addresses these concerns.

We hope you will find the reviewers' comments useful as you decide how to proceed. If you wish to submit a substantially revised manuscript, please bear in mind that we will be reluctant to approach the reviewers again in the absence of major revisions.

It appears clear from the reports that the major interest in the manuscript comes from the genomic insights into an individual of this age, rather than the Gravettian cultural attribution specifically. And given that reviewer 1 (the Gravettian archaeology expert) argues that there are reasons to be tentative of a Gravettian affiliation for the Buran Kaya III individual, it is our belief that a more nuanced discussion of the cultural attribution (along the lines suggested by reviewer 1) is to be recommended. Rather than diminishing interest in the genomic data, we feel that this would better serve it. Reviewers 2 and 3, both experienced in working with Palaeolithic genomic data, and aware of

the challenges this poses, have recommendations for better use of comparative data and statistical analyses to strengthen the current interpretations, which we feel are welcome.

If you choose to revise your manuscript taking into account all reviewer and editor comments, please highlight all changes in the manuscript text file [OPTIONAL: in Microsoft Word format].

We are committed to providing a fair and constructive peer-review process. Please do not hesitate to contact us if there are specific requests from the reviewers that you believe are technically impossible or unlikely to yield a meaningful outcome.

If revising your manuscript:

* Include a "Response to reviewers" document detailing, point-by-point, how you addressed each referee comment. If no action was taken to address a point, you must provide a compelling argument. This response will be sent back to the referees along with the revised manuscript.

* If you have not done so already we suggest that you begin to revise your manuscript so that it conforms to our Article format instructions at http://www.nature.com/natecolevol/info/final-submission. Refer also to any guidelines provided in this letter.

* Include a revised version of any required reporting checklist. It will be available to referees (and, potentially, statisticians) to aid in their evaluation if the manuscript goes back for peer review. A revised checklist is essential for re-review of the paper.

Please use the link below to submit a revised paper:

[REDACTED]

Note: This URL links to your confidential home page and associated information about manuscripts you may have submitted, or that you are reviewing for us. If you wish to forward this email to co-authors, please delete the link to your homepage.

If you wish to submit a suitably revised manuscript we would hope to receive it within 6 months. If you cannot send it within this time, please let us know. We will be happy to consider your revision so long as nothing similar has been accepted for publication at Nature Ecology & Evolution or published elsewhere. Should your manuscript be substantially delayed without notifying us in advance and your article is eventually published, the received date would be that of the revised, not the original, version.

Nature Ecology & Evolution is committed to improving transparency in authorship. As part of our efforts in this direction, we are now requesting that all authors identified as 'corresponding author' on published papers create and link their Open Researcher and Contributor Identifier (ORCID) with their account on the Manuscript Tracking System (MTS), prior to acceptance. This applies to primary research papers only. ORCID helps the scientific community achieve unambiguous attribution of all scholarly contributions. You can create and link your ORCID from the home page of the MTS by clicking on 'Modify my Springer Nature account'. For more information please visit please visit www.springernature.com/orcid.

Please do not hesitate to contact me if you have any questions or would like to discuss the required revisions further.

Thank you for the opportunity to review your work, and as ever, please don't hesitate to contact me by email if I can be of any help during the revision process.

[REDACTED]

Reviewer expertise:

Reviewer #1: Gravettian archaeology, European Russia

Reviewer #2: Palaeolithic genomes

Reviewer #3: Palaeolithic genomes, bio-informatics

Reviewers' comments:

Reviewer #1 (Remarks to the Author):

I was very interested to read this paper, which uses new results concerning the ancient genome of an individual from the site of Buran-Kaya III to draw some conclusions regarding the "origins" of the Gravettian technocomplex. Although I think that the results are certainly worth publishing, in my view it requires significant reworking before publication and here I would like to suggest a way of doing this.
First of all, I am rather surprised by the overall tone and approach of the paper, especially given recent debate on combining archaeological and genetic evidence (see e.g. Heyd 2017 - https://doi.org/10.15184/aqy.2017.21 ; Eisenmann et al. 2018 - https://doi.org/10.1038/s41598-018-31123-z ; Furholt 2018 - https://doi.org/10.1017/eaa.2017.43 ; Riede et al 2019 - https://doi.org/10.1057/s41599-019-0260-7 ; Booth 2019 - https://doi.org/10.1080/00438243.2019.1627240). The use of the term "the origin of the Gravettians" in the paper title, as well as the first sentence of the paper in particular, suggest a rather naïve approach to archaeological understandings of past populations and furthermore suggest that the authors have not properly considered the ongoing debate on aDNA and archaeology. If the paper is published in its current form it is likely to provoke many negative reactions from archaeologists who will, rightly, feel that nuanced understandings of Upper Palaeolithic technocomplexes and their relationships with past populations are being ignored in favour of a simplistic, essentialist approach. I strongly recommend that the authors rethink their title and furthermore, take care not to misrepresent present-day archaeological understandings of "archaeological cultures" etc. The first sentence of the paper is a can of worms that I could spend 5,000 words analysing, but rather than doing that I would ask the authors to carefully reconsider it in the light of the literature cited above.
Furthermore, the authors need to pay some more attention to their use of terminology: a technocomplex is not synonymous with an archaeological culture (see Clarke 1968), and it is not clear what is meant by e.g. "materially defined culture" or "archaeologically defined material industries". An archaeological culture is not supposed to represent a past "culture" in the anthropological sense but at some points it seems as if the authors are suggesting this.

This problem is made worse by the fact that the overall understanding of Gravettian archaeology presented in the paper is rather weak. For example, in SI lines 40-48 an overview of Gravettian facies and phases is provided; the most recent work cited here is from 1996 (!) and the abundant work from the last two decades on these questions is ignored. Furthermore, why only discuss the record in Central Europe here (basically Moravia and Lower Austria) when the record from Ukraine, Moldova, Russia, and Romania is clearly more relevant? There are also some basic errors that undermine the reader's confidence, e.g. lines 72-73 (main paper): Kostenki 14 is not "found in the Borshchyovo archaeological complex", the area is known as Kostenki-Borshchevo (other transliterations are available) after the two neighbouring villages.

There is a fundamental problem in this paper that needs to be addressed, although I think there is a clear way to fix it. I will try to explain this problem over the next few paragraphs. First of all, Buran-Kaya III is located in Crimea, which the authors describe repeatedly as being in Eastern Europe. However, Crimea is better treated from a geographical standpoint as being the northwesternmost extension of the Caucasus – remember that the Sea of Azov and much of the present-day Black Sea were dry land during the time in question. Ecologically, geomorphologically, etc Crimea was part of the Caucasus. Therefore, Buran-Kaya III is most easily grouped with sites in the Caucasus, *not* with sites from much further north and west on the Russian Plain, Moravia etc, which are and were very different ecological zones. And, of course, in the Caucasus we do have several EUP sites with backed lithics: the backed lithic assemblage from Buran-Kaya III should be grouped with these Caucasus assemblages, as it often has been.

On the other hand, at EUP sites in the rest of Europe we do not have backed lithics: the sites are Aurignacian. It is only later, during the MUP, that we have many sites with backed lithics, mostly dating to after around 30,000 14C BP. (The authors do cite some very old dates for European Gravettian assemblages in the SI but these are really at the very end of the timescale and in many cases question marks can be raised over these results – for me, it is only after ca. 30 ka 14C BP that we have convincing evidence for Gravettian assemblages).

The authors present the description of Buran-Kaya III as Gravettian as if it is accepted fact: it is not. I am not disputing the presence of backed lithics or the dating of the relevant layers at Buran-Kaya III, I am questioning whether we should automatically describe these assemblages as Gravettian. Hublin (2015: https://doi.org/10.1016/j.quascirev.2014.08.011) argues that we should not call the BK-III assemblages Gravettian, given that it they are so much older than the Gravettian sites of the rest of Europe, and I am inclined to agree with his opinion. We do not yet know whether there is a direct link between the Buran-Kaya III assemblage and the later assemblages in Europe, and therefore it is premature to assign them all to the same technocomplex, especially given that there is also a geographical gap between these sites (see Teyssandier and Zilhao 2018: https://doi.org/10.1007/s41982-017-0004-4 for further recent discussion of this issue).

So, what do we have? There is an EUP group of assemblages with backed lithics in the Caucasus and Crimea, and a later group of ("Gravettian") assemblages with backed lithics in Europe proper. It is absolutely valid to hypothesise a link between them but this is a link that needs to be tested: we need a better understanding of all aspects of the assemblages (lithic technology, personal ornaments, etc) that goes beyond the basic observation of the shared presence of backed lithics. And of course, one thing that can help to test this hypothesis is to look at aDNA evidence to help test whether there was a population link between the Caucasus/Crimea assemblages and the later European assemblages.

Here I would like to suggest a way forward. First of all, I plead with the authors *not* to use the term Early Gravettian as they do in this paper, since that term is already well established to describe assemblages dating between ca. 29.5 and 27 14C kya BP across Europe – using "Early Gravettian" to describe even earlier assemblages will introduce severe confusion. However, I also suggest that they do not call the Caucasus/Crimea assemblages Gravettian unless and until we can be satisfied that they

form a chronological and cultural continuum with the later European sites (and even then, there would be a case to be made for maintaining a separate term). Hoffecker has used the term "Proto-Gravettian" occasionally but to my knowledge he is the only author who has done so, and this term obviously has certain associations which to my mind are better avoided. I suggest that the authors use a neutral term like "Caucasus backed lithic assemblages" ("Caucasus BLA"?) to describe Buran-Kaya III and the other assemblages from the Caucasus. This kind of term is far more appropriate while the link between these assemblages remains a hypothesis to be tested.
As already stated, in my view the link between the backed lithic assemblages of BK-III/the Caucasus and Gravettian assemblages of Europe remains to be established, given the geographical and chronological gap between them. Therefore, the link between them should be treated as a hypothesis rather than as simple fact. This paper would be much improved by presenting this link as a hypothesis with different aspects – including the possibility that population processes underlie this link. The genetic evidence can then be presented as a way of testing the existence of various population processes, which obviously is significant for testing the overall hypothesis of a link. So, rather than presenting Buran Kaya III as the oldest Gravettian site in Europe, and then arguing that the results presented here provide evidence for "the origin of the Gravettians", why not describe the existence of these two separate groups of sites (Caucasus BLA and Gravettian) which are chronologically and geographically distinct but adjacent, and which have in common the presence of backed lithics. Then give the hypothesis of a link between the sites, and present the genetic evidence as one way of testing this hypothesis. For me, this would put the paper on far stronger ground and would also help to get rid of the more dubious essentialist conceptualisations of technocomplex = population, mentioned above. I realise that I am suggesting a fundamental rewrite of the entire paper but I also feel that this is necessary in order to get rid of the significant weak spots in the current argumentation.

A few other notes:
Sungir' – Sungir' is *not* Gravettian according to the authors' own definition of Gravettian, and it should not be presented as such. No backed blade(let)s are found at Sungir' (see Gavrilov 2017 in ERAUL 147; backed knives don't count), and the presence of backed lithics is the most basic criterion for description of a site of Gravettian. Sungir' is a late EUP site and currently best described as Streletskian. The reason that a small number of authors have previously described it as Gravettian is that the previous dating of the burials (now superseded by Marom et al's 2012 results) placed it in the MUP as typically defined; some authors therefore unfortunately decided to call it Gravettian on chronological grounds despite the fact that no backed lithics are known from the site. Of course, the fact that rich burials were also found at the site has supported this attribution, because it fits with the widespread belief that in Europe, burials began with the Gravettian. (The term "Eastern Gravettian" should also be avoided in all cases because it is ambiguous: in Russian usage it has a strict definition and describes the late Gravettian Kostenki-Avdeevo Culture, but its history of usage in English is sloppy and it does not have a clear meaning.) Accepting that Sungir' is *not* Gravettian may significantly complicate the interpretation of the overall results of the study but I will leave this to the authors to figure out.
Kostenki 12 is frequently described as "attributed to Gorodtsovian" rather than just "Gorodtsovian" (e.g. Supp Table 6) – why the extra caution concerning this attribution? To my knowledge there is no debate concerning the description of Kostenki 12/Ia as Gorodtsovian. This is an archaeological attribution based on the site's lithic assemblage and is well-established. The authors seem at some points to be almost suggesting that perhaps the site should be re-described as Gravettian but this seems premature (and there are no backed lithics at the site).
Finally, I'm not a geneticist and I'm not qualified to comment on the genetic analyses described here,

but I have to say that as an archaeologist the presentation of the results is not particularly clear compared with other papers I have read. In particular, the text in the Discussion lines 266-287 is rather difficult to make sense of and I would strongly recommend that it be revised for clarity and sentence structure. Although I can get a general sense of the results, I really don't quite understand the exact significance of the comparisons that have been made, and how robust they are, and what the remaining gaps are in our understanding.

A few comments on specific parts of the paper:
Lines 29-32: "The Gravettian technocomplex defines European Mid-Upper Palaeolithic (MUP) industries characterized by a suite of shared innovations, along with specific modalities of their production, such as stone Gravette points, backed blades and bladelets, personal ivory and shell ornaments, ochre, and antler or bone tools."
The authors don't provide any references here to support this statement, which is a pity, but it's a bit misleading in any case: the only thing that characterizes all Gravettian sites is the presence of backed lithics. I genuinely don't know what the "shared innovations" are that the authors refer to. Personal ornaments, ochre, and osseous tools are of course common to all European Upper Palaeolithic technocomplexes.

Lines 32-33: "The Gravettian became widespread throughout Europe beginning after ca. 36,000 until ca. 23,000 cal BP years ago (ca. 32-21 ka 14C BP)."
The calibration is incorrect here: 21ka 14C BP calibrates to ca. 25.5-25 ka cal BP in OxCal using IntCal13.

Lines 56-59: "The parallels between these Eastern European industries and the Gravettian appearing later in Central Europe and the Danubian Valley have led some to propose the term “Early Gravettian” to describe these industries to distinguish them from the classical Gravettian (8,16,17)."
Lines 312-314: "Given this background, the appearance of the Gravettian in Central Europe in the MUP, where it later would blossom, is likely to have resulted from this input from the east."
Forgive my directness here, but: flowers blossom, not archaeologically defined technocomplexes. Furthermore, while early Gravettian assemblages across Europe are often described as "Early Gravettian", the middle/later Gravettian is not the "classical Gravettian". This is a strangely superficial reading of the complexity of the Mid Upper Palaeolithic. I'm also curious to understand why it's in Central Europe that the authors think that "the Gravettian" reached some sort of zenith: they don't back it up and it's not supported by what I know about the Gravettian record.

Natasha Reynolds

Reviewer #2 (Remarks to the Author):

Given its age and the archaeological context, the site of Buran Kaya is highly important in the understanding if the spread of the Gravettian culture is linked to the spread of ideas or the movements/admixture of the people. The authors have taken an elaborate and admirable approach in the identification of the best extract and used most sensitive methods to generate genome-wide data from a specimen with poor endogenous DNA preservation. They are also familiar with the published literature, as well as the analytical approaches necessary to avoid the biases that stem from working with limited amounts of data.

However, there are several aspects of this manuscript that need to be clarified and explained better, or interpreted with caution.

- In the main text of the manuscript, authors indicate that the individual from Buran Kaya shows less affinity to the GoyetQ116-1 (lines 176-178). While this is correct with the point estimates of the f3-statistics, it seems that the confidence intervals of this statistics overlap with those of Vestonice16 and Kostenki14 (as it is shown in the Figure S3., with both Mbuti and Han as outgroups).
Furthermore, from the direct D-statistics presented in the Figure S6., there seems to be no significant difference in the amount of allele-sharing in the D(Vestonice16/Kostenki14/Sunghir3, GoyetQ116-1; BuranKaya, Mbuti). Therefore, is it really that the closest genetic relationship of Buran Kaya are Kostenki14, Sunghir3 and Vestonice16 to the exclusion of GoyetQ116-1? (lines 181-185)? As authors note they are unable to resolve if Buran Kaya is closer to Kostenki14 or to Sunghir3 (lines 194-198), this is also the case for GoyetQ116-1. Furthermore, the statement that Buran Kaya shares most alleles with Vestonice16 (lines 191-194) from the results of D-statistics is perhaps not fully correct, as, again, there seems to be no significant difference when doing the statistics of D(Vestonice16, Sunghir3/GoyetQ116-1; BuranKaya, Mbuti) as it is presented in the figure S6.
The statistics that would be highly informative for resolving these relationships in more detail, and that I haven't seen presented, is D(BuranKaya, X; Y, outgroup). Is there a particular reason this has not been presented/attempted?

- It is great that the authors attempted to resolve the relationship of Buran Kaya to the individuals that have less data, and are aware of the biases that stem from these types of analysis. However, even though the filtering they describe (lines 204-208) is helpful to remove the biases introduced by aDNA damage, for the individuals having small amounts of data (and especially when having two or three such individuals in the statistics) this can lead to huge statistical noise that stems from the comparisons of just a few sites across the genome! Having 50 SNPs used for the comparisons is far too few to make general conclusions. Furthermore, it is unclear if the N(SNPs) here is the number of overlapping SNPs between three/four individuals/populations or the number of actual SNPs used for the calculations - ie, since not all SNPs overlapping between individuals are informative for calculating the f3 or f4-statistics. Also, is requesting that the resulting values are within 30% of each other when using all sites and transversions only a common practice? Where does the 30% cut-ff come from? Has it been tested and in which way?
Therefore, given that Buran Kaya individual itself is restricted by having ~27,000 SNPs overlapping with the panel(s) of informative positions, it would be better to focus the analyses/comparisons of this individual to the individuals with more data, as the results will be more robust and less prone to the statistical noise.

- Related to the point above, and to one of the important points authors infer. If applied the filters authors discuss before, and ignoring the 30% overlap between the values, from the Supplementary Table 7., and Figure S8., it would seem that Buran Kaya actually shares more drift with Muierii2 than with Kostenki12 both when looking at all SNPs and transversions only. This statistics is not explained anywhere in the text. Is it because of the 30% cut-off? There are a lot of conclusions drawn about the affinity of Buran Kaya and Kostenki12, whereas the point estimate of this statistics is far smaller than the one for Muierii2 (even though the confidence intervals overlap). Following in the same way, Natufian 9, as well as Malta1, would share more drift with Buran Kaya when focusing on the transversions only than Paglicci133. Furthermore, a lot of these confidence intervals overlap as shown in the figure S8. This is just additional evidence that the statistics done on individuals with a small number of overlapping sites should be interpreted with caution.

- Also related to the points above, if the connection of Buran Kaya to Paglicci133 is inferred from the f3-statistics, then the statement in the Discussion, lines 291-293 about not detecting genome-wide allele sharing between Buran Kaya and the Natufians would not be correct as the f3 between Buran

Kaya and Natufian9 is higher than the one from Paglicci133.

General comments:
- For inferring the genetic relationships between individuals in more detail and having the statistical significance tested, it would be helpful to see if the connections and which of them are detected and significant with D(X, Y; BuranKaya, out) as well as D(BuranKaya, X; Y, out). Doing most of the comparisons with just f3 statistics and then comparing f3 point estimates to each other is less appropriate way of drawing conclusions.
- It would be more helpful to have the same order of individuals on the X-axis throughout different figures, ie fixing the order of individuals either by age or something else rather than the ones that have the highest value for a given statistics to the ones that have the lowest value. In that way it will be easier to follow and spot which things are changing when testing for different sets of questions.
- For the paper and the conclusions drawn it would be more informative to have the Figure S6. replacing the main figure 4.
- For the figures S3. and S4., there are no boxed names indicating the individuals associated with the Gravettian.
- As it has been shown in Petr et al (2019), using ancestry informative SNPs for calculating the proportion of Neandertal ancestry can inflate these values. You can try to calculate the proportion of Neandertal ancestry with the direct f4-rations by making use of the two high coverage Neandertal genomes instead. It will not change the conclusions, but it will provide you with more robust estimates.
- Line 105 - should UNG be UDG?
- Line 221 - it should be Fu et al, 2016 (not 2015)

Reviewer #3 (Remarks to the Author):

In their manuscript Bennet and co-authors report ancient genomic data from an Upper Paleolithic individual, the 36,000-year-old BuranKaya3A from the Crimea region in Eastern Europe, which represents the earliest occurrence of the Gravettian technocomplex. The key finding of the study is that the BuranKaya3A appears genetically most similar to other UP individuals from Eastern and Central Europe such as those from Sunghir and Dolni Vestonice. Based on these findings the authors argue for an eastern origin of the early Gravettian industries.

Given the comparative rarity of early modern human remains from Eurasia, novel genomic data from such an individual is naturally very exciting. The data presented here consists of capture data on the mitochondrion (MT), as well as shotgun data on the nuclear genome. The recovered DNA shows the expected damage patterns for authentic ancient DNA, and contamination rates are negligible. Unfortunately, the preservation of the specimen used for DNA extraction was poor, resulting in very low coverage on the nuclear genome. This is not unexpected for a sample of that age; however, my concern is that some of the major claims of the authors are in my view not sufficiently supported by the data as a consequence. Below I will comment in more detail on the main sections of the manuscript

Sections MT and Y chromosome; Neanderthal ancestry

The individual was determined to be a male, with MT haplogroup N1 and more limited resolution on the Y chromosome to a possible CT or C haplogroup. I agree with the authors' interpretation that a

further derived haplogroup such as those observed at Kostenki 14 or Sunghir appears unlikely. In the analysis of Neanderthal ancestry, it is a bit unclear what the unit of the 95% CI is, as the point estimate is reported as percentage but the CI as fraction. If the 95% CI represents 2.24% as I would assume then those estimates are quite uncertain, in line with the low coverage.

Section "Genomic relationships and archaeological cultures of UP Europe"

Here the authors carry out a number of contrasts using the established outgroup-f3 and D-statistics. The dataset used throughout is a merge using the ~3 million SNPs targeted by the probes reported by Fu et al. Out of those, only about 28,000 are covered in the BuranKaya3A individual, which unfortunately severely impacts the resolution of those statistics. Among the tests carried out, the only I would consider well supported are the comparisons with the high coverage samples which show a closer affinity of BuranKaya3A with the other early Europeans such as Kostenki 14 and Sunghir, compared to Ust'Ishim (e.g. D-statistics in Fig 4). The tests involving lower coverage individuals are too noisy and lack statistical support. As an example, the authors discuss identified genetic affinities between BuranKaya3A and individuals Kostenki 12 and/or Pavlov1 (p 11, lines 211-214), but the standard errors for the corresponding outgroup-f3 statistic of those individuals entirely overlap those for Tianyuan, a 40kya East Asian individual which has been shown to have ancestry related to present-day East Asians (Fig. S8).

Furthermore, the f3 statistics for different individuals would not be directly comparable if the authors use the full set of the 3 million SNPs for all tests. Some of the individuals with capture data were not captured on the full set of probes (e.g. Goyet Q116-1 only has 1240K data, see Fu et al. 2016 Table ED1). As the ascertainment differs between the different capture panels, this can cause systematic bias in the shared drift estimated by f3, so these analyses should be carried out on a subset of SNPs covered in all individuals included in the comparisons.

Section "Common West Eurasian ancestry"

In this section the authors test for the presence of "basal Eurasian" ancestry in BuranKaya3A, by using D-statistics D(Modern or Ancient East Asian, x; Ust-Ishim, Mbuti). The result for BuranKaya3A is not significantly different from 0, which suggests no "basal Eurasian" ancestry. However as in the previous section, the standard errors are high due to low coverage, so this might also just be an issue of insufficient power to detect this ancestry in BuranKaya3A. The authors also perform tests that show "The absence of the "Common West Eurasian" ancestry, as represented by Villabruna, in BuranKaya3A", which they further say "marks a key genetic distinction between the Gravettian inhabitants of Buran-Kaya III". Indeed the D-statistics D(BuranKaya3A, Han; Villabruna, Mbuti) are not significantly different from zero (Fig. S10), which is surprising given the documented closer affinity with the Kostenki 14 lineage discussed above. Again, it is difficult to conclude whether this is a real signal or just due to a lack of power because of low coverage or e.g. higher error rate causing outgroup attraction in Villabruna (discussed in Fu et al 2015).

I appreciate the challenges of DNA sequencing of these early human remains, and would like to suggest to possibly mitigate these limitations somewhat by taking full advantage of the shotgun sequencing data. In particular, I suggest the authors re-analyze their data in the context of the other shotgun-sequenced samples (e.g. Kostenki 14, Ust'Ishim, Sunghir, Loschbour) using an extended reference SNP set (e.g. those from the 1000 Genomes project) to maximize power. This approach has

the additional advantage that a much larger number of transversion SNPs would be available than those targeted on the capture panels. With the higher resolution it might also be possible to disentangle the relationship of BuranKaya3A with Kostenki 14 and Sunghir, for which the current results hint at a closer affinity to Sunghir. Here the authors could also pool the data for all four Sunghir individuals to further increase power, and those results would have direct bearing on the questions of the origins of the Gravettian.

Minor comments

Table S2, it would be good to also report average genomic coverage in addition to the total number of reads mapped

Z-scores for the outgroup f3 plots are not really necessary and can be removed from the plot, makes them easier readable

Author Rebuttal to Initial comments

Answers to the reviewer's comments

Reviewer #1 (Remarks to the Author):

I was very interested to read this paper, which uses new results concerning the ancient genome of an individual from the site of Buran-Kaya III to draw some conclusions regarding the "origins" of the Gravettian technocomplex. Although I think that the results are certainly worth publishing, in my view it requires significant reworking before publication and here I would like to suggest a way of doing this. First of all, I am rather surprised by the overall tone and approach of the paper, especially given recent debate on combining archaeological and genetic evidence (see e.g. Heyd 2017 - https://doi.org/10.15184/aqy.2017.21 ; Eisenmann et al. 2018 - https://doi.org/10.1038/s41598-018-31123-z ; Furholt 2018 - https://doi.org/10.1017/eaa.2017.43 ; Riede et al 2019 - https://doi.org/10.1057/s41599-019-0260-7 ; Booth 2019 - https://doi.org/10.1080/00438243.2019.1627240). The use of the term "the origin of the Gravettians" in the paper title, as well as the first sentence of the paper in particular, suggest a rather naïve approach to archaeological understandings of past populations and furthermore suggest that the authors have not properly considered the ongoing debate on aDNA and archaeology. If the paper is published in its current form it is likely to provoke many negative reactions from archaeologists who will, rightly, feel that nuanced understandings of Upper Palaeolithic technocomplexes and their relationships with past populations are being ignored in favour of a simplistic, essentialist approach. I strongly recommend that the authors rethink their title and furthermore, take care not to misrepresent present-day archaeological understandings of "archaeological cultures" etc. The first sentence of the paper is a can of worms that I could spend 5,000 words analysing, but rather than doing that I would ask the authors to carefully reconsider it in the light of the literature cited above. Furthermore, the authors need to pay some more attention to their use of terminology: a technocomplex is not synonymous with an archaeological culture (see Clarke 1968), and it is not clear what is meant by e.g. "materially defined culture" or "archaeologically defined material industries". An archaeological culture is not supposed to represent a past "culture" in the anthropological sense but at some points it seems as if the authors are suggesting this. This problem is made worse by the fact that the overall understanding of Gravettian archaeology presented in the paper is rather weak. For example, in SI lines 40-48 an overview of Gravettian facies and phases is provided; the most recent work cited here is from 1996 (!) and the abundant work from the last two decades on these questions is ignored. Furthermore, why only discuss the record in Central Europe here (basically Moravia and Lower Austria) when the record from Ukraine, Moldova, Russia, and Romania is clearly more relevant? There are also some basic errors that undermine the reader's confidence, e.g. lines 72-73 (main paper): Kostenki 14 is not "found in the Borshchyovo archaeological complex", the area is known as Kostenki-Borshchevo (other transliterations are available) after the two neighbouring villages. There is a fundamental problem in this paper that needs to be addressed, although I think there is a clear way to fix it. I will try to explain this problem over the next few paragraphs. First of all, Buran-Kaya III is located in Crimea, which the authors describe repeatedly as being in Eastern Europe. However, Crimea is better treated from a geographical standpoint as being the northwesternmost extension of the Caucasus – remember that the Sea of Azov and much of the present-day Black Sea were dry land during the time in question. Ecologically, geomorphologically, etc Crimea was part of the Caucasus. Therefore, Buran-Kaya III is most easily grouped with sites in the Caucasus, *not* with sites from much further north and west on the Russian Plain, Moravia etc, which are and were very different ecological zones. And, of

course, in the Caucasus we do have several EUP sites with backed lithics: the backed lithic assemblage from Buran-Kaya III should be grouped with these Caucasus assemblages, as it often has been. On the other hand, at EUP sites in the rest of Europe we do not have backed lithics: the sites are Aurignacian. It is only later, during the MUP, that we have many sites with backed lithics, mostly dating to after around 30,000 14C BP. (The authors do cite some very old dates for European Gravettian assemblages in the SI but these are really at the very end of the timescale and in many cases question marks can be raised over these results – for me, it is only after ca. 30 ka 14C BP that we have convincing evidence for Gravettian assemblages). The authors present the description of Buran-Kaya III as Gravettian as if it is accepted fact: it is not. I am not disputing the presence of backed lithics or the dating of the relevant layers at Buran-Kaya III, I am questioning whether we should automatically describe these assemblages as Gravettian. Hublin (2015: https://doi.org/10.1016/j.quascirev.2014.08.011) argues that we should not call the BK-III assemblages Gravettian, given that it they are so much older than the Gravettian sites of the rest of Europe, and I am inclined to agree with his opinion. We do not yet know whether there is a direct link between the Buran-Kaya III assemblage and the later assemblages in Europe, and therefore it is premature to assign them all to the same technocomplex, especially given that there is also a geographical gap between these sites (see Teyssandier and Zilhao 2018: https://doi.org/10.1007/s41982-017-0004-4 for further recent discussion of this issue).

So, what do we have? There is an EUP group of assemblages with backed lithics in the Caucasus and Crimea, and a later group of ("Gravettian") assemblages with backed lithics in Europe proper. It is absolutely valid to hypothesise a link between them but this is a link that needs to be tested: we need a better understanding of all aspects of the assemblages (lithic technology, personal ornaments, etc) that goes beyond the basic observation of the shared presence of backed lithics. And of course, one thing that can help to test this hypothesis is to look at aDNA evidence to help test whether there was a population link between the Caucasus/Crimea assemblages and the later European assemblages. Here I would like to suggest a way forward. First of all, I plead with the authors *not* to use the term Early Gravettian as they do in this paper, since that term is already well established to describe assemblages dating between ca. 29.5 and 27 14C kya BP across Europe – using "Early Gravettian" to describe even earlier assemblages will introduce severe confusion. However, I also suggest that they do not call the Caucasus/Crimea assemblages Gravettian unless and until we can be satisfied that they form a chronological and cultural continuum with the later European sites (and even then, there would be a case to be made for maintaining a separate term). Hoffecker has used the term "Proto-Gravettian" occasionally but to my knowledge he is the only author who has done so, and this term obviously has certain associations which to my mind are better avoided. I suggest that the authors use a neutral term like "Caucasus backed lithic assemblages" ("Caucasus BLA"?) to describe Buran-Kaya III and the other assemblages from the Caucasus. This kind of term is far more appropriate while the link between these assemblages remains a hypothesis to be tested. As already stated, in my view the link between the backed lithic assemblages of BK-III/the Caucasus and Gravettian assemblages of Europe remains to be established, given the geographical and chronological gap between them. Therefore, the link between them should be treated as a hypothesis rather than as simple fact. This paper would be much improved by presenting this link as a hypothesis with different aspects – including the possibility that population processes underlie this link. The genetic evidence can then be presented as a way of testing the existence of various population processes, which obviously is significant for testing the overall hypothesis of a link.

So, rather than presenting Buran Kaya III as the oldest Gravettian site in Europe, and then arguing that the results presented here provide evidence for "the origin of the Gravettians", why not describe the existence of these two separate groups of sites (Caucasus BLA and Gravettian) which are chronologically and geographically distinct but adjacent, and which have in common the presence of backed lithics. Then give the hypothesis of a link between the sites, and present the genetic evidence as one way of testing this hypothesis. For me, this would put the paper on far stronger ground and would also help to get rid of the more dubious essentialist conceptualisations of technocomplex = population, mentioned above. I realise that I am suggesting a fundamental rewrite of the entire paper but I also feel that this is necessary in order to get rid of the significant weak spots in the current argumentation.

A few other notes:

Sungir' – Sungir' is *not* Gravettian according to the authors' own definition of Gravettian, and it should not be presented as such. No backed blade(let)s are found at Sungir' (see Gavrilov 2017 in ERAUL 147; backed knives don't count), and the presence of backed lithics is the most basic criterion for description of a site of Gravettian. Sungir' is a late EUP site and currently best described as Streletskian. The reason that a small number of authors have previously described it as Gravettian is that the previous dating of the burials (now superseded by Marom et al's 2012 results) placed it in the MUP as typically defined; some authors therefore unfortunately decided to call it Gravettian on chronological grounds despite the fact that no backed lithics are known from the site. Of course, the fact that rich burials were also found at the site has supported this attribution, because it fits with the widespread belief that in Europe, burials began with the Gravettian. (The term "Eastern Gravettian" should also be avoided in all cases because it is ambiguous: in Russian usage it has a strict definition and describes the late Gravettian Kostenki-Avdeevo Culture, but its history of usage in English is sloppy and it does not have a clear meaning.) Accepting that Sungir' is *not* Gravettian may significantly complicate the interpretation of the overall results of the study but I will leave this to the authors to figure out.

Kostenki 12 is frequently described as "attributed to Gorodtsovian" rather than just "Gorodtsovian" (e.g. Supp Table 6) – why the extra caution concerning this attribution? To my knowledge there is no debate concerning the description of Kostenki 12/Ia as Gorodtsovian. This is an archaeological attribution based on the site's lithic assemblage and is well-established. The authors seem at some points to be almost suggesting that perhaps the site should be re-described as Gravettian but this seems premature (and there are no backed lithics at the site).

Finally, I'm not a geneticist and I'm not qualified to comment on the genetic analyses described here, but I have to say that as an archaeologist the presentation of the results is not particularly clear compared with other papers I have read. In particular, the text in the Discussion lines 266-287 is rather difficult to make sense of and I would strongly recommend that it be revised for clarity and sentence structure. Although I can get a general sense of the results, I really don't quite understand the exact significance of the comparisons that have been made, and how robust they are, and what the remaining gaps are in our understanding.

A few comments on specific parts of the paper:

Lines 29-32: "The Gravettian technocomplex defines European Mid-Upper Palaeolithic (MUP) industries characterized by a suite of shared innovations, along with specific modalities of their production, such as stone Gravette points, backed blades and bladelets, personal ivory and shell ornaments, ochre, and antler or bone tools."

The authors don't provide any references here to support this statement, which is a pity, but it's a bit misleading in any case: the only thing that characterizes all Gravettian sites is the presence of backed lithics. I genuinely don't know what the "shared innovations" are that the authors refer to. Personal ornaments, ochre, and osseous tools are of course common to all European Upper Palaeolithic technocomplexes.

Lines 32-33: "The Gravettian became widespread throughout Europe beginning after ca. 36,000 until ca. 23,000 cal BP years ago (ca. 32-21 ka 14C BP)."

The calibration is incorrect here: 21ka 14C BP calibrates to ca. 25.5-25 ka cal BP in OxCal using IntCal13.

Lines 56-59: "The parallels between these Eastern European industries and the Gravettian appearing later in Central Europe and the Danubian Valley have led some to propose the term “Early Gravettian” to describe these industries to distinguish them from the classical Gravettian (8,16,17)."

Lines 312-314: "Given this background, the appearance of the Gravettian in Central Europe in the MUP, where it later would blossom, is likely to have resulted from this input from the east."

Forgive my directness here, but: flowers blossom, not archaeologically defined technocomplexes. Furthermore, while early Gravettian assemblages across Europe are often described as "Early Gravettian", the middle/later Gravettian is not the "classical Gravettian". This is a strangely superficial reading of the complexity of the Mid Upper Palaeolithic. I'm also curious to understand why it's in Central Europe that the authors think that "the Gravettian" reached some sort of zenith: they don't back it up and it's not supported by what I know about the Gravettian record.

Natasha Reynolds

We are very grateful to Nathasha Reynolds for her careful evaluation of our manuscript and her detailed criticism, which we took very seriously. Moreover, we read many of her papers, including the most recent, which gave us a deeper understanding around the issues she found to be problematic in our previous version. In the new manuscript, we have been much more cautious in considering the archeological impact of our genetic data. Overall, we have considerably improved the genetic analyses thanks to the addition of new data, higher genomic coverage of the previously analyzed specimen, but also of a second, slightly older specimen from the same site. This renders the genetic results more robust. In the new version, we concentrated on these genetic results, the analyses of which reveal proximity of genetic ancestry used to identify source populations, movements of people, and admixture. We refrained from proposing archaeological hypotheses that go beyond the simple indication of similarities of artefacts as identified in the published record. As a result, the archaeological implications of this primarily ancient genomics study are minimized in the current version. We have specifically removed phrasings she has pointed out in her last review and attempted to respect guidelines put forth in her papers. Where there have been disagreements within the archaeological community, in particular with the interpretation of the assemblages of Buran-Kaya III at the source layers of our material, we have attempted to convey the different interpretations in a balanced way. We have been careful throughout the rewriting to address the criticisms very thoroughly with regards to the archaeological perspective, and we hope that Natasha Reynolds will be satisfied by both the change of spirit and the final outcome.

Reviewer #2 (Remarks to the Author):

Given its age and the archaeological context, the site of Buran Kaya is highly important in the understanding if the spread of the Gravettian culture is linked to the spread of ideas or the movements/admixture of the people. The authors have taken an elaborate and admirable approach in the identification of the best extract and used most sensitive methods to generate genome-wide data from a specimen with poor endogenous DNA preservation. They are also familiar with the published literature, as well as the analytical approaches necessary to avoid the biases that stem from working with limited amounts of data.

However, there are several aspects of this manuscript that need to be clarified and explained better, or interpreted with caution.

- In the main text of the manuscript, authors indicate that the individual from Buran Kaya shows less affinity to the GoyetQ116-1 (lines 176-178). While this is correct with the point estimates of the f3-statistics, it seems that the confidence intervals of this statistics overlap with those of Vestonice16 and Kostenki14 (as it is shown in the Figure S3., with both Mbuti and Han as outgroups). Furthermore, from the direct D-statistics presented in the Figure S6., there seems to be no significant difference in the amount of allele-sharing in the D(Vestonice16/Kostenki14/Sunghir3, GoyetQ116-1; BuranKaya, Mbuti). Therefore, is it really that the closest genetic relationship of Buran Kaya are Kostenki14, Sunghir3 and Vestonice16 to the exclusion of GoyetQ116-1? (lines 181-185)? As authors note they are unable to resolve if Buran Kaya is closer to Kostenki14 or to Sunghir3 (lines 194-198), this is also the case for GoyetQ116-1. Furthermore, the statement that Buran Kaya shares most alleles with Vestonice16 (lines 191-194) from the results of D-statistics is perhaps not fully correct, as, again, there seems to be no significant difference when doing the statistics of D(Vestonice16, Sunghir3/GoyetQ116-1; BuranKaya, Mbuti) as it is presented in the figure S6.

The statistics that would be highly informative for resolving these relationships in more detail, and that I haven't seen presented, is D(BuranKaya, X; Y, outgroup). Is there a particular reason this has not been presented/attempted?

- It is great that the authors attempted to resolve the relationship of Buran Kaya to the individuals that have less data, and are aware of the biases that stem from these types of analysis. However, even though the filtering they describe (lines 204-208) is helpful to remove the biases introduced by aDNA damage, for the individuals having small amounts of data (and especially when having two or three such individuals in the statistics) this can lead to huge statistical noise that stems from the comparisons of just a few sites across the genome! Having 50 SNPs used for the comparisons is far too few to make general conclusions. Furthermore, it is unclear if the N(SNPs) here is the number of overlapping SNPs between three/four individuals/populations or the number of actual SNPs used for the calculations - ie, since not all SNPs overlapping between individuals are informative for calculating the f3 or f4-statistics. Also, is requesting that the resulting values are within 30% of each other when using all sites and transversions only a common practice? Where does the 30% cut-ff come from? Has it been tested and in which way?

Therefore, given that Buran Kaya individual itself is restricted by having ~27,000 SNPs overlapping with the panel(s) of informative positions, it would be better to focus the analyses/comparisons of this individual to the individuals with more data, as the results will be more robust and less prone to the statistical noise.

- Related to the point above, and to one of the important points authors infer. If applied the filters authors discuss before, and ignoring the 30% overlap between the values, from the Supplementary Table 7., and Figure S8., it would seem that Buran Kaya actually shares more

drift with Muierii2 than with Kostenki12 both when looking at all SNPs and transversions only. This statistics is not explained anywhere in the text. Is it because of the 30% cut-off? There are a lot of conclusions drawn about the affinity of Buran Kaya and Kostenki12, whereas the point estimate of this statistics is far smaller than the one for Muierii2 (even though the confidence intervals overlap). Following in the same way, Natufian 9, as well as Malta1, would share more drift with Buran Kaya when focusing on the transversions only than Paglicci133. Furthermore, a lot of these confidence intervals overlap as shown in the figure S8. This is just additional evidence that the statistics done on individuals with a small number of overlapping sites should be interpreted with caution.
- Also related to the points above, if the connection of Buran Kaya to Paglicci133 is inferred from the f3-statistics, then the statement in the Discussion, lines 291-293 about not detecting genome-wide allele sharing between Buran Kaya and the Natufians would not be correct as the f3 between Buran Kaya and Natufian9 is higher than the one from Paglicci133.
General comments:
- For inferring the genetic relationships between individuals in more detail and having the statistical significance tested, it would be helpful to see if the connections and which of them are detected and significant with D(X, Y; BuranKaya, out) as well as D(BuranKaya, X; Y, out). Doing most of the comparisons with just f3 statistics and then comparing f3 point estimates to each other is less appropriate way of drawing conclusions.
- It would be more helpful to have the same order of individuals on the X-axis throughout different figures, ie fixing the order of individuals either by age or something else rather than the ones that have the highest value for a given statistics to the ones that have the lowest value. In that way it will be easier to follow and spot which things are changing when testing for different sets of questions.
- For the paper and the conclusions drawn it would be more informative to have the Figure S6. replacing the main figure 4.
- For the figures S3. and S4., there are no boxed names indicating the individuals associated with the Gravettian.
- As it has been shown in Petr et al (2019), using ancestry informative SNPs for calculating the proportion of Neandertal ancestry can inflate these values. You can try to calculate the proportion of Neandertal ancestry with the direct f4-rations by making use of the two high coverage Neandertal genomes instead. It will not change the conclusions, but it will provide you with more robust estimates.
- Line 105 - should UNG be UDG?
- Line 221 - it should be Fu et al, 2016 (not 2015)

We thank the reviewer for his/her helpful criticisms and insightful comments. We recognize the bulk of the reviewer's critical comments regarding the earlier version concern problems with resolution of the relationships we proposed given the low coverage we obtained for the original sample.

We have spent the years since the last version generating and analyzing more data related to the Buran-Kaya III site that has improved the robustness and scope of the original analysis and has led to a completely rewritten manuscript. As a general response to the reviewer's comments concerning the low coverage of the sample analyzed in the previous version, and the impact this low coverage had on the statistical significance of f3 and D-statistics results especially when compared to other lower covered samples, we have now more than doubled

the number of unique reads corresponding to the previously described sample and we have generated even more data (the double of the previous one) for a second, slightly older, sample from the same site. We have also changed completely the reference dataset, we have used the 1000 genome dataset to greatly extend the SNP list. After appropriate filtering of the dataset and selection of the SNPs shared with at least one of the Buran Kaya individuals, we have thus identified 740,000 SNPs, an increase by a factor of 30 compared to the previously used 27,000 SNPs set.

Furthermore, we followed the recommendation of reviewer #3 to restrict our analyses to well covered Upper Paleolithic genomes, which also resolves many of the concerns of reviewer #2 with respect to the genomic relationships we reported using lower coverage material. Accordingly, in all our comparisons, we increased the average number of SNPs involved from 900 to 82,000 and the minimum threshold for inclusion from about 100 to 8,000 SNPs (98.5% of the 464 comparisons performed involved more than 10,000 SNPs), thus increasing by almost two orders of magnitude the numbers of SNPs supporting our current interpretations. The resulting increase in power allowed us to perform more elaborate population genomics analyses relying on state-of-the-art methods, and the time that had passed since our last manuscript oversaw the publication of several key genomes and related findings that allowed us to place our samples into a richer prehistoric context. Due to in part to these recommended improvements, we came up with a significantly improved story that addresses many of the reviewers comments regarding possible over-interpretation of statistical results based on low covered material with limited SNP numbers. We hope that the reviewer will be satisfied with the improved dataset, the new analyses, results and interpretation.

Reviewer #3 (Remarks to the Author):

In their manuscript Bennet and co-authors report ancient genomic data from an Upper Paleolithic individual, the 36,000-year-old BuranKaya3A from the Crimea region in Eastern Europe, which represents the earliest occurrence of the Gravettian technocomplex. The key finding of the study is that the BuranKaya3A appears genetically most similar to other UP individuals from Eastern and Central Europe such as those from Sunghir and Dolni Vestonice. Based on these findings the authors argue for an eastern origin of the early Gravettian industries.

Given the comparative rarity of early modern human remains from Eurasia, novel genomic data from such an individual is naturally very exciting. The data presented here consists of capture data on the mitochondrion (MT), as well as shotgun data on the nuclear genome. The recovered DNA shows the expected damage patterns for authentic ancient DNA, and contamination rates are negligible. Unfortunately, the preservation of the specimen used for DNA extraction was poor, resulting in very low coverage on the nuclear genome. This is not unexpected for a sample of that age; however, my concern is that some of the major claims of the authors are in my view not sufficiently supported by the data as a consequence. Below I will comment in more detail on the main sections of the manuscript

Sections MT and Y chromosome; Neanderthal ancestry

The individual was determined to be a male, with MT haplogroup N1 and more limited resolution on the Y chromosome to a possible CT or C haplogroup. I agree with the authors' interpretation that a further derived haplogroup such as those observed at Kostenki 14 or Sunghir appears unlikely. In the analysis of Neanderthal ancestry, it is a bit unclear what the unit of the 95% CI is, as the point estimate is reported as percentage but the CI as fraction. If

the 95% CI represents 2.24% as I would assume then those estimates are quite uncertain, in line with the low coverage.

Section "Genomic relationships and archaeological cultures of UP Europe"

Here the authors carry out a number of contrasts using the established outgroup-f3 and D-statistics. The dataset used throughout is a merge using the ~3 million SNPs targeted by the probes reported by Fu et al. Out of those, only about 28,000 are covered in the BuranKaya3A individual, which unfortunately severely impacts the resolution of those statistics. Among the tests carried out, the only I would consider well supported are the comparisons with the high coverage samples which show a closer affinity of BuranKaya3A with the other early Europeans such as Kostenki 14 and Sunghir, compared to Ust'Ishim (e.g. D-statistics in Fig 4). The tests involving lower coverage individuals are too noisy and lack statistical support. As an example, the authors discuss identified genetic affinities between BuranKaya3A and individuals Kostenki 12 and/or Pavlov1 (p 11, lines 211-214), but the standard errors for the corresponding outgroup-f3 statistic of those individuals entirely overlap those for Tianyuan, a 40kya East Asian individual which has been shown to have ancestry related to present-day East Asians (Fig. S8).

Furthermore, the f3 statistics for different individuals would not be directly comparable if the authors use the full set of the 3 million SNPs for all tests. Some of the individuals with capture data were not captured on the full set of probes (e.g. Goyet Q116-1 only has 1240K data, see Fu et al. 2016 Table ED1). As the ascertainment differs between the different capture panels, this can cause systematic bias in the shared drift estimated by f3, so these analyses should be carried out on a subset of SNPs covered in all individuals included in the comparisons.

Section "Common West Eurasian ancestry"

In this section the authors test for the presence of "basal Eurasian" ancestry in BuranKaya3A, by using D-statistics D(Modern or Ancient East Asian, x; Ust-Ishim, Mbuti). The result for BuranKaya3A is not significantly different from 0, which suggests no "basal Eurasian" ancestry. However as in the previous section, the standard errors are high due to low coverage, so this might also just be an issue of insufficient power to detect this ancestry in BuranKaya3A. The authors also perform tests that show "The absence of the "Common West Eurasian" ancestry, as represented by Villabruna, in BuranKaya3A", which they further say "marks a key genetic distinction between the Gravettian inhabitants of Buran-Kaya III". Indeed the D-statistics D(BuranKaya3A, Han; Villabruna, Mbuti) are not significantly different from zero (Fig. S10), which is surprising given the documented closer affinity with the Kostenki 14 lineage discussed above. Again, it is difficult to conclude whether this is a real signal or just due to a lack of power because of low coverage or e.g. higher error rate causing outgroup attraction in Villabruna (discussed in Fu et al 2015).

I appreciate the challenges of DNA sequencing of these early human remains, and would like to suggest to possibly mitigate these limitations somewhat by taking full advantage of the shotgun sequencing data. In particular, I suggest the authors re-analyze their data in the context of the other shotgun-sequenced samples (e.g. Kostenki 14, Ust'Ishim, Sunghir, Loschbour) using an extended reference SNP set (e.g. those from the 1000 Genomes project) to maximize power. This approach has the additional advantage that a much larger number of transversion SNPs would be available than those targeted on the capture panels. With the higher resolution it might also be possible to disentangle the relationship of BuranKaya3A with Kostenki 14 and Sunghir, for which the current results hint at a closer affinity to Sunghir. Here the authors could also pool the data for all four Sunghir individuals to further increase power, and those results would have direct bearing on the questions of the origins of the Gravettian.

Minor comments
Table S2, it would be good to also report average genomic coverage in addition to the total number of reads mapped Z-scores for the outgroup f3 plots are not really necessary and can be removed from the plot, makes them easier readable

We would like to thank the reviewer for their comments and criticisms that were both insightful and helpful to redesign our study. We present a completely revised, new manuscript based on the new analyses we have done over the last years during which we generated more data for the samples from Buran Kaya III. This greatly improved the robustness and the scope of our analyses, as requested by this reviewer whose criticisms we have taken very seriously.

We especially took into account their suggestions on how to achieve more robust data: We have worked first on improving the raw data. In particular, we have more than doubled the number of unique reads corresponding to the previously described sample and we have generated twice as many data for a second, slightly older, sample from the same site. We have also changed completely the reference dataset: 1) as suggested, we have used the 1,000 genome dataset to extend the SNP list; 2) after appropriate filtering of the dataset and selection of the SNPs shared with at least one of the Buran Kaya individuals, we have identified 740,000 SNPs, an increase by a factor of 30 compared to the previously used 27,000 SNPs set. Furthermore, as recommended, we also restricted our analyses to well covered Upper Paleolithic genomes and have performed the calling using our new SNP list. Accordingly, in all our comparisons, we increased the average number of SNPs involved from 900 to 82,000 and the minimum threshold for inclusion from about 100 to 8,000 SNPs (98.5% of the 464 comparisons performed involved more than 10,000 SNPs), thus increasing by almost two orders of magnitude the numbers of SNPs supporting our interpretation. We have additionally followed the reviewer's advice to restrict f3 analyses to shared SNPs to avoid capture-related biases. The resulting increased power achieved allowed us to perform more elaborate population genomics analyses relying on state-of-the-art methods. In this way, we came up with a significantly improved story. With regards to our responses to the reviewer's specific critical comments concerning low-covered samples and resulting high-standard errors given in the previous version, the higher resolution coming from incorporating the reviewer's suggestions above for improving both SNP and sample selection, along with the major increase in raw sequence data, has removed these lower-quality samples of concern from the study. We hope that the reviewer will be satisfied with the improved dataset, the new analyses and the obtained results, and convinced by our interpretation of the results.

Decision Letter, first revision:

30th January 2023

*Please ensure you delete the link to your author homepage in this e-mail if you wish to forward it to your co-authors.

Dear Eva-Maria,

Now sending the formal decision to you!

Your manuscript entitled "Early Upper Paleolithic genomes of Crimea show migration and admixture dynamics of the first modern European ancestries in relation to climatic crisis" has now been seen by the same three reviewers, whose comments are attached. The reviewers have raised a number of concerns which will need to be addressed before we can offer publication in Nature Ecology & Evolution. We will therefore need to see your responses to the criticisms raised and to some editorial concerns, along with a revised manuscript, before we can reach a final decision regarding publication.

We therefore invite you to revise your manuscript taking into account all reviewer and editor comments. Please highlight all changes in the manuscript text file.

We are committed to providing a fair and constructive peer-review process. Do not hesitate to contact us if there are specific requests from the reviewers that you believe are technically impossible or unlikely to yield a meaningful outcome.

When revising your manuscript:

* Include a "Response to reviewers" document detailing, point-by-point, how you addressed each reviewer comment. If no action was taken to address a point, you must provide a compelling argument. This response will be sent back to the reviewers along with the revised manuscript.

* If you have not done so already please begin to revise your manuscript so that it conforms to our Article format instructions at http://www.nature.com/natecolevol/info/final-submission. Refer also to any guidelines provided in this letter.

* Include a revised version of any required reporting checklist. It will be available to referees (and, potentially, statisticians) to aid in their evaluation if the manuscript goes back for peer review. A revised checklist is essential for re-review of the paper.

Please use the link below to submit your revised manuscript and related files:

[REDACTED]

Note: This URL links to your confidential home page and associated information about manuscripts you may have submitted, or that you are reviewing for us. If you wish to forward this email to co-authors, please delete the link to your homepage.

We hope to receive your revised manuscript within four to eight weeks. If you cannot send it within this time, please let us know. We will be happy to consider your revision so long as nothing similar has been accepted for publication at Nature Ecology & Evolution or published elsewhere.

Nature Ecology & Evolution is committed to improving transparency in authorship. As part of our efforts in this direction, we are now requesting that all authors identified as 'corresponding author' on published papers create and link their Open Researcher and Contributor Identifier (ORCID) with their account on the Manuscript Tracking System (MTS), prior to acceptance. ORCID helps the scientific community achieve unambiguous attribution of all scholarly contributions. You can create and link your ORCID from the home page of the MTS by clicking on 'Modify my Springer Nature account'. For more information please visit please visit www.springernature.com/orcid.

Please do not hesitate to contact me if you have any questions or would like to discuss these revisions further.

We look forward to seeing the revised manuscript and thank you for the opportunity to review your work.

[REDACTED]

Reviewers' comments:

Reviewer #1 (Remarks to the Author):

It was very interesting to read the new version of this manuscript and I think it presents a far stronger overall argument than the original version I read. The results of the genetic analyses are very interesting and the discussion was possible for me to follow as an archaeologist with an interest in aDNA studies but no real training in this area. A few weaknesses remain in the archaeological aspects of the discussion that I think need addressing – see notes below – but this shouldn't represent too much more work. I am glad to see some interpretations and speculations about what the genetic results might mean for our understandings of human populations and migrations in the past. As someone who works in this region, it's a fascinating area and I'm intrigued to see some of your remarks line up with some ideas about population movements (from east of Eastern Europe into Europe at the beginning of the EUP) that I and colleagues have floated in the past – see e.g. Dinnis et al. 2019 (https://doi.org/10.1016/j.jhevol.2018.11.012), p. 35, second paragraph of the Conclusions section [forgive me mentioning my own work, but I think it's relevant here]. Overall, I think this is a very interesting paper that represents a real contribution to our understanding of population processes during the Upper Palaeolithic in Europe and I would be pleased to see it published after a little more work.

Lines 29-30: "Middle-Upper Paleolithic" -> this should be "Mid Upper Paleolithic" (and check rest of manuscript)

Lines 46-49: There are more recent dates available for the CI tephra – see Giaccio et al 2017, Sci Rep, doi.org/10.1038/srep45940 – the newer Ar/Ar dates are slightly older than the ones given here

Lines 246-248: I couldn't quite parse this sentence

This version of the paper emphasises the role of the CI eruption more than the original manuscript but there are some pretty obvious gaps in the literature that is cited in this discussion. I think there are a few more papers that are worth referring to here, because there is quite a rich history of study of the

impact of the CI tephra on human populations (although much of the earlier work on this was to do with the disappearance of Neanderthals, and is now quite out of date since our chronologies of the MP-UP transition have improved). I would have a look at: d'Errico and Banks 2015 (https://doi.org/10.1016/j.quascirev.2014.05.014), Fedele et al 2007 (https://hal-insu.archives-ouvertes.fr/insu-00260073), Marti et al 2016 (https://doi.org/10.1038/srep21220), maybe Dragosavac et al 2021 (https://reff.f.bg.ac.rs/handle/123456789/3687).

Lines 355-359: "The population change in Europe during this period is also highlighted by the introduction of EUP lithic innovations and cultural practices. Backed blades and bladelet industries begin to appear in Europe, different from the Levallois blade production techniques characteristic of the IUP, accompanied by an increased diversity in the raw materials incorporated into personal adornments and portable art."
Sorry but this section needs some work. First of all it wasn't entirely clear to me exactly what you mean by "this period". Furthermore, backed lithics are found in MUP European (i.e. Gravettian) assemblages, NOT EUP ones (leaving aside Chatelperronian assemblages). (The numerous references in Hoffecker 2009 to backed lithics in EUP assemblages are erroneous, I would not cite this paper here). EUP lithic assemblages – i.e. Aurignacian assemblages – are different from IUP assemblages, and MUP assemblages are different again (in all cases these differences are both technological and typological and are well characterised). I wouldn't mention anything to do with Levallois blade production techniques in IUP assemblages, it's a bit of a can of worms and not necessary to your argument here.

Line 307: "a backed-bladed, microlaminar lithic industry" - this sounds a bit strange, I suggest something like "a lithic industry containing backed bladelets"

Line 363: "the backed-blade and micro-laminar lithics" – again, I suggest something like "the backed lithic assemblages"

line 373 – "microblade" – I would stick to "bladelet"

Reviewer #2 (Remarks to the Author):

The authors have undertaken quite a substantial amount of additional work compared to the previous version of the manuscript – both by increasing the amount of useable data from the previous Buran Kaya individual and also by adding genomic data from an additional, slightly older, individual from the same archaeological site. These individuals are an important addition in understanding the genetic ancestry of human dispersals in Early Upper Paleolithic of west Eurasia. The manuscript is well written and the analyses are presented in a clear way.

Authors find that both Buran Kaya individuals share more genetic affinity with other individuals from Europe younger than 38,000 years. There could possibly be a genetic link between Oase1 ancestry and the Buran Kaya individual BuKA3A, as demonstrated by some of the f4-statistics reported in the Tables S3 and S4, as well as Figure 3A and 3B. However, even though there is an increased affinity of BuKa3A and Oase1, unfortunately none of the statistics of f4(Mbuti, BuKa3A; Oase1, UP human) are significant, except the ones including Bacho Kiro F6 620 and possibly Bacho Kiro CC7 335. The same is the case with f4(Mbuti, Oase1; BuKa3A, UP human) - would the authors not expect this particular statistics to be significant if there is indeed a portion of Oase1 ancestry in BuKa3A? The only UP individual where the statistics is approaching to be significant is Ust'Ishim, but for BuKa3C?

I would be cautious with the interpretations drawn from the qpGraph, as it was recently demonstrated that the number of possible and equally well fitting admixture graphs increases substantially when

trying to include as many individuals and admixture edges as authors attempt in the Figure 2A.

Nevertheless, what was particularly striking in this manuscript is the connection of the Buran Kaya individuals with the individuals from the Caucasus that lived some 23,000-26,000 years later and which were previously found to have large contribution of the 'Basal Eurasian' genetic ancestry. This would make the Buran Kaya individuals currently the oldest individuals to date that were identified carrying this genetic component, and that in itself is an important observation and could potentially explain the affinities between Oase1 and BuKa3A.

There are some minor comments:

What do authors mean in lines 24-25 by the 'first migration wave of humans carrying ancestry found in present-day Europeans'? Currently unclear from the sentence structure if they refer to the overall connection to later Europeans – which would still make the Kostenki14 the earliest representative of this wave or the 'first' migration wave of human ancestry found in present-day Europeans, and which is still quite distinct from the BuKa ancestry – or if the reference is to the Basal Eurasian genetic component?

Just for practical reasons of not having to jump between different files, it might be worth placing the Supplementary Tables 2 and 7 as sheets in the xlsx file with other supplementary tables rather than in the text file of the SI.

Line 520 psesudo-haploid should be pseudo-haploid.

Reviewer #3 (Remarks to the Author):

In their resubmitted manuscript, Bennet et al substantially expand their dataset reporting new data from the previously included individual BuranKaya3A as well as a new individual BuranKaya3C of a similar age. I am very appreciative of the effort going in to increasing the amount of data from these exciting but challenging to work with human remains. The authors present mostly new analyses based on these data, which make for an overall improved manuscript. However, some of my previous concerns with regards to the analyses and interpretation of the data remain in this new version.

New dataset

The authors now report data from two individuals. Data generation, contamination estimates and analyses of uniparental markers all support authenticity of the data. While the authors present fold-coverage estimates for the MT genomes (82X and 11X for BuKa3A and BuKa3C, respectively), I could not find the same overall final fold-coverage values for the nuclear genomes. Table S6 reports per-library read numbers, and number of overlapping SNPs with the ~13 million marker reference panels indicate ~0.04X and ~0.02X coverage, but it would be useful to directly report the estimated number in the main text.

For the construction of the analysis panel, the authors extract a panel of ~740,000 SNPs from the HGDP high-coverage WGS panel, for which either one of the ancient BuranKaya individuals has coverage. This 740K panel is then used for all analyses with other published ancient genomes. In my review of the previous version, I raised the concern of possible bias in joint analyses of shotgun and targeted capture samples if the SNPs covered between them differ systematically. This remains a potential problem in the revised version, as there seems to have been no further filtering for SNPs included in the targeted capture probes in the analyses. As an example, for the f3 analyses presented,

the statistics estimated from comparison between BuranKaya and another shotgun sample will be derived from a random subset of the 740K SNPs with coverage in both, whereas in comparison with a capture sample the statistics will be heavily driven by a non-random subset of the 740K which were pre-ascertained and captured in the individual. These also might differ between capture individuals analysed here, as e.g. Vestonice 16 was captured at a set of 3.7M SNPs, whereas the Krems double burial I2483/4 was only captured using the 1240K panel. The authors should restrict analyses to the smallest subset of SNPs ascertained across all individuals, to rule out these potential issues whenever jointly analysing samples with different capture schemes.

In the methods section, the authors also describe the curation of another SNP panel (page 24, line 522), but it is unclear to me if and where this panel was used in the analyses.

UP Admixture

In this section the authors discuss differences in shared ancestry between the two BuKa individuals and other UP individuals. Some of the analyses presented here suffer from possible bias as described above, and overall signals appear quite weak, so I would be cautious in interpreting these as true differences. A direct test of clade-like relationship of the two individuals as e.g. in f4(Mbuti, other UP; BuKa3A, BuKa3C) would give a significant non-zero statistic in this case, but has not been done. If the overlap of SNPs between the two low coverage samples is too low, I suggest at least to also show a plot of pairwise f3-outgroup statistics (on smallest subset of capture SNPs as discussed above) with corresponding error bars to make it easier to discern how noisy those estimates are.

The authors also present a rather complex admixture graph fitted with both BuKa individuals in the context of other IUP and UP individuals. Unfortunately, I could not find any details on how the authors arrived at this fit (as in how many / which other models were tested), nor on the parameters used for qpGraph. As complex models like this are prone to overfitting (see e.g. recent preprint by Maier et al bioRxiv 2022.05.08.491072), I would like to see some results replicated in simpler models and/or further analyses to provide a better understanding if/how many other topologies might also be supported.

Pre-CI gene flow

In this section the authors discuss results indicating the presence of remnant IUP ancestry in BuKa. However, the evidence presented is again weak. While the f4 statistics in Fig 3A do suggest a slight increase in sharing between BuKa3A and Oase over two of the Bacho Kiro individuals, the signal is much weaker or absent when using other IUP samples as comparison. Furthermore, table S5 shows the more direct test of excess Oase ancestry using f4(Mbuti, Oase; BuKa, other UP/IUP). In case of gene flow between Oase and BuKa these statistics would be expected to be significantly >0 when comparing BuKa to other UP indiivudals, but all combinations in S5 show f4=0 which is consistent with Oase being an outgroup symmetrically related to all tested sample pairs.

Post-CI / Gravettian

The evidence for differential relationship of the Gravettian samples in Vestonice and Krems suffers from similar weaknesses as described in the previous sections. The authors base their conclusions here on the slight differences in outgroup f3 statistics (with their possible ascertainment bias issues) and indirect evidence from comparing two different f4 tests. Instead, the authors should have used the direct test for differential relatedness f4 (Mbuti, BuKa; Vestonice16, Krems), results of which can be found in table S4 and do not support gene flow between Krems and BuKa (f4 not significantly

different from 0 with Z-scores -0.25 and 0.98 respectively).

CHG affinity

The final result section discusses genetic affinities between BuKa and palaeolithic hunter-gatherers from the Caucasus. I found this result to be much more well supported and interesting than the previous sections. Both individuals show consistent increase in allele sharing with CHG in the direct test shown in Fig 3C and Table S4, especially when comparing with the geographically and temporally close Sunghir and Kostenki genomes, which additionally don't suffer from shotgun / capture biases in this configuration.

Overall, my take is that the manuscript has some exciting new data, but too much focus is on discussing weak signals based on a few marginally significant f-statistics or a possibly overfitted graph. On the other hand, there is a clear result in the comparison with CHG genomes, but no further exploration of this signal in the context of the other good quality genomes from the region (i.e. Sunghir/Kostenki). Not including a more thorough exploration of these relationships with e.g. more simple and robust admixture graph fit comparisons would in my view constitute a missed opportunity for these exciting new genomes.

**********END**********

Author Rebuttal, first revision:

Reviewer 1:

It was very interesting to read the new version of this manuscript and I think it presents a far stronger overall argument than the original version I read. The results of the genetic analyses are very interesting and the discussion was possible for me to follow as an archaeologist with an interest in aDNA studies but no real training in this area. A few weaknesses remain in the archaeological aspects of the discussion that I think need addressing – see notes below – but this shouldn't represent too much more work. I am glad to see some interpretations and speculations about what the genetic results might mean for our understandings of human populations and migrations in the past. As someone who works in this region, it's a fascinating area and I'm intrigued to see some of your remarks line up with some ideas about population movements (from east of Eastern Europe into Europe at the beginning of the EUP) that I and colleagues have floated in the past – see e.g. Dinnis et al. 2019 (https://doi.org/10.1016/j.jhevol.2018.11.012), p. 35, second paragraph of the Conclusions section [forgive me mentioning my own work, but I think it's relevant here]. Overall, I think this is a very interesting paper that represents a real contribution to our understanding of population processes during the Upper Palaeolithic in Europe and I would be pleased to see it published after a little more work.

We are glad you found this new version interesting. We greatly appreciate your comprehensive input and it was a big help to improve the manuscript and study as a whole. Since our previous version, new genomic data from Gravettian assemblages from Western Europe characterized a new genetic cluster, the Fournol cluster, which turned out to be a much closer relative to the Buran Kaya III individuals than any of the available genomes in our previous analysis. The inclusion of these samples in our current work allowed us to confirm a shared admixture signal with the earliest modern human in Europe for whom we have genomic data, Zlaty Kun, which we also found to be present in a southern Europe subset of these new MUP genomes. We further identified this component in Mesolithic hunter gatherers from the Caucasus, giving a lot of new insights into early human interactions and movements. Together, we think these new data have allowed us to present the genomic analysis of these individuals in a more comprehensive context. We hope you are satisfied with the additional improvements.

Lines 29-30: "Middle-Upper Paleolithic" -> this should be "Mid Upper Paleolithic" (and check rest of manuscript)

We have made this correction throughout the manuscript.

Lines 46-49: There are more recent dates available for the CI tephra – see Giaccio et al 2017, Sci Rep, doi.org/10.1038/srep45940 – the newer Ar/Ar dates are slightly older than the ones given here

We thank you for pointing out this more recent date. We have read the suggested paper and incorporated the more recent date and corresponding citation in the text.

Lines 246-248: I couldn't quite parse this sentence
This sentence describes the results of the f4 statistical analysis. We reformulated the sentence and hope it's now easier to understand.

Original: "Through comparisons of pre-CI groups with BuKa3A, significant allele sharing could also be seen with the 31-52,000-year-old woman from Czechia, Zlatý Kůň[4] with respect to either IUP Bacho Kiro individuals or the 45,000-year-old Ust-Ishim from western Siberia[43]"

Updated: "BuKa3A additionally shared significantly more alleles with the 31-52,000-year-old woman from Czechia, Zlatý Kůň[4] with respect to either IUP Bacho Kiro individuals or the 45,000-year-old Ust'Ishim from western Siberia[48]"

This version of the paper emphasises the role of the CI eruption more than the original manuscript but there are some pretty obvious gaps in the literature that is cited in this discussion. I think there are a few more papers that are worth referring to here, because there is quite a rich history of study of the impact of the CI tephra on human populations (although much of the earlier work on this was to do with the disappearance of Neanderthals, and is now quite out of date since our chronologies of the MP-UP transition have improved). I would have a look at: d'Errico and Banks 2015 (https://doi.org/10.1016/j.quascirev.2014.05.014), Fedele et al 2007 (https://hal-insu.archives-ouvertes.fr/insu-00260073), Marti et al 2016 (https://doi.org/10.1038/srep21220), maybe Dragosavac et al 2021 (https://reff.f.bg.ac.rs/handle/123456789/3687).

Thank you for the literature suggestions regarding the CI eruption and its impacts on humans and the environment. We have expanded our reading on this topic, including the suggested material, and now include many of your suggested papers, among others.

Lines 355-359: "The population change in Europe during this period is also highlighted by the introduction of EUP lithic innovations and cultural practices. Backed blades and bladelet industries begin to appear in Europe, different from the Levallois blade production techniques characteristic of the IUP, accompanied by an increased diversity in the raw materials incorporated into personal adornments and portable art."
Sorry but this section needs some work. First of all it wasn't entirely clear to me exactly what you mean by "this period". Furthermore, backed lithics are found in MUP European (i.e. Gravettian) assemblages, NOT EUP ones (leaving aside Chatelperronian assemblages).

(The numerous references in Hoffecker 2009 to backed lithics in EUP assemblages are erroneous, I would not cite this paper here).

EUP lithic assemblages – i.e. Aurignacian assemblages – are different from IUP assemblages, and MUP assemblages are different again (in all cases these differences are both technological and typological and are well characterised). I wouldn't mention anything to do with Levallois blade production techniques in IUP assemblages, it's a bit of a can of worms and not necessary to your argument here.

We struggled to rework this sentence in a way to generalize without oversimplifying to the point of misrepresenting the breadth of the literature concerning these topics. It is a complex area, and we feel in the end that the message we meant to convey here would require much more space to do so with accuracy and is not essential to the main points of the paper. We have decided to go another way and have removed this sentence rather than to risk oversimplification that could lead to erroneous understanding. The Hoffecker 2009 reference has been removed with the above sentence. Thank you for bringing this to our attention.

Line 307: "a backed-bladed, microlaminar lithic industry" - this sounds a bit strange, I suggest something like "a lithic industry containing backed bladelets"

We have made the recommended change according to your recommendation and the text now says "a lithic industry containing backed bladelets".

Line 363: "the backed-blade and micro-laminar lithics" – again, I suggest something like "the backed lithic assemblages"

We have changed this to "the backed lithic assemblages".

line 373 – "microblade" – I would stick to "bladelet"

We have removed the word "microblade" from the manuscript and replaced it with "bladelets".

Reviewer 2:

The authors have undertaken quite a substantial amount of additional work compared to the previous version of the manuscript – both by increasing the amount of useable data from the previous Buran Kaya individual and also by adding genomic data from an additional, slightly older, individual from the same archaeological site. These individuals are an important addition in understanding the genetic ancestry of human dispersals in Early Upper Paleolithic of west Eurasia. The manuscript is well written and the analyses are presented in a clear way.

Authors find that both Buran Kaya individuals share more genetic affinity with other individuals from Europe younger than 38,000 years. There could possibly be a genetic link between Oase1 ancestry and the Buran Kaya individual BuKA3A, as demonstrated by some of the f4-statistics reported in the Tables S3 and S4, as well as Figure 3A and 3B. However, even though there is an increased affinity of BuKa3A and Oase1, unfortunately none of the statistics of f4(Mbuti, BuKa3A; Oase1, UP human) are significant, except the ones including Bacho Kiro F6 620 and possibly Bacho Kiro CC7 335. The same is the case with f4(Mbuti, Oase1; BuKa3A, UP human) - would the authors not expect this particular statistics to be significant if there is indeed a portion of Oase1 ancestry in BuKa3A? The only UP individual where the statistics is approaching to be significant is Ust'Ishim, but for BuKa3C?

We realized that in the previous version, regrettably, not all the f4 formulations were included in the supplementary data. Specifically, apart from Oase1 and CHG tests, only the formulation f4(Mbuti, Buka3; x, y) was included (even though the corresponding f4(Mbuti, x; Buka3; y) was calculated and some of these results were used in our old figure 2B). We have rectified this oversight and now include full data for both f4(Mbuti, Buka3; x, y) and f4(Mbuti, x; Buka3; y) for all samples in the supplementary table S5.

Due to the high degree of allele sharing of BuKa with other post-CIgenomes relative to pre-CI genomes, as well as additional instances of pre-CI admixture we identify, the signal of admixture is more apparent when comparing BuKa3A to two pre-CI individuals. If BuKa3A can act as an outgroup relative to these two, the resulting statistic should be close to zero. Significant deviations from zero are indicative of allele sharing between BuKa3 and a pre-CI individual. As you mentioned, we do see more significant deviations from zero comparing BuKa3A with the two individuals from Bacho Kiro, whereby BuKa3A does share more alleles with Oase1 than with Bacho Kiro (Z-scores -2.9 and -3.5 for the neandertal-masked Oase1, which we note, interestingly, are more significant than the corresponding values for the non-masked Oase1).

More broadly though, for this revised version, we now have access to several newly published MUP western European genomes (from Posth, et al., "Palaeogenomics of Upper Palaeolithic to Neolithic European hunter-gatherers", Nature 2023) and have repeated all the f-statistics and have gained many new insights into the admixture between post-CI and pre-CI genomes. A deeper analysis of these individuals compared with those of Buran Kaya III show a much more robust support for Zlaty Kun than Oase1, and these results are more reliable due in part to the much greater coverage for the shotgun Zlaty Kun data. Regarding the admixture signals, as before, we identify admixture signals with several pre-CI ancestries in both samples (we don't presume these came specifically from the source sample populations, but may represent underlying admixed pre-CI ancestry that may have been present in Europe with various relationships to these groups, which we now discuss). In the current revised version, we now provide ancestry estimates restricted to BuKa3A and BuKa3C using simplified qpGraphs, separate for each sample, which address pre-CI admixture directly, and find the best-fitting models include admixture Zlaty-Kun-related ancestry. Use of the program ADMIXTOOLS 2 was applied in this revision to ensure the qpGraph model selected was optimal. As in the previous version, BuKa3A has a

larger IUP-related component than BuKa3C. In the current version, we show models with Zlaty Kun contribution have stronger statistical support than identical models without (Figure 3B and Figure S9). With these models, we estimate the Zlaty-Kun-like contribution to be low, 6% for BuKa3C, and higher (17%) for Buka3A. Thus we would not expect to see large f4 values in our comparisons, in particular with BuKa3C.

Importantly, despite BuKa3 clustering with MUP samples, Zlaty Kun does give several significant and near-significant results of allele sharing between BuKa3 and Zlaty compared not just with IUP-contemporary BK_F6-620 and UstIshim, but with other MUP samples as well. i.e:

Mbuti	ZlatyKun	BuKa3A	Sunghir3	-0.00138	0.000419	-3.294
Mbuti	ZlatyKun	BuKa3C	Sunghir3	-0.000833	0.000365	-2.281
Mbuti	ZlatyKun	BuKa3A	Kostenki14	-0.001382	0.000637	-2.169
Mbuti	ZlatyKun	BuKa3C	Kostenki14	-0.000903	0.000468	-1.928
Mbuti	ZlatyKun	BuKa3A	Ust_Ishim	-0.001559	0.000466	-3.348
Mbuti	ZlatyKun	BuKa3C	Ust_Ishim	-0.000883	0.000411	-2.146
Mbuti	ZlatyKun	BuKa3A	BG_IUP_BK_F6-620	-0.001676	0.000777	-2.155
Mbuti	ZlatyKun	BuKa3A	BG_IUP_BK_F6-620	-0.001895	0.000627	-3.023

(complete results available in the new Table S5)

In the current revision, we have decided to lead with these more significant results, and now feature Zlaty Kun rather than Oase1 in Figure 3. We find this presentation to be more convincing, rather than only showcasing the lower contribution and less significant values of more poorly-covered Oase1-like ancestry.

In addition, reassuring this is not an artifact with our BuKa3 data, we also observe and report similar signals of Zlaty Kun admixture persisting in several of the new UP Europe genomic data from the recently released Posth et al. publication. We now incorporate the relevant samples into our study, and observe these Zlaty Kun-related admixture signals are also elevated in several Gravettian-associated samples dating between 31 and 26ky located in the southern part of western Europe. We now include these observations in the text and have put forth the possible interpretation of low-level pre-CI ancestry persisting among certain groups in Europe until the LGM.

I would be cautious with the interpretations drawn from the qpGraph, as it was recently demonstrated that the number of possible and equally well fitting admixture graphs increases substantially when trying to include as many individuals and admixture edges as authors attempt in the Figure 2A.

We agree about the danger of relying too heavily on a fitting qpGraph model, especially an overly complex model with multiple edges. In response to this comment, we have decided to avoid risking saying less by trying to say too much, we have now replaced the large multi-population model shown in the previous Figure 2A with two more simplified models using only shotgun data as the new Figure 3B, wherein we separate each BuKa3 individual and focus on modelling its specific ancestry (as mentioned above). We additionally include less well-fitting alternate models of these with no admixture, along with their Z-scores, in the supplements as Figure S9.

Nevertheless, what was particularly striking in this manuscript is the connection of the Buran Kaya individuals with the individuals from the Caucasus that lived some 23,000-26,000 years later and which were previously found to have large contribution of the 'Basal Eurasian' genetic ancestry. This would make the Buran Kaya individuals currently the oldest individuals to date that were identified carrying this genetic component, and that in itself is an important observation and could potentially explain the affinities between Oase1 and BuKa3A.

We were also struck with the link between our BuKa3 samples and CHG, a link we found was also shared with the Krems1_1-2 sample as well in the initial version. The current version explores this link more deeply and expands the amount of pre-LGM individuals sharing this ancestry. We were surprised to find this CHG link mirrors the Zlaty Kun-like ancestry we find admixed in several pre-LGM populations.

Basal Eurasian ancestry is not yet well-defined on its own. The Basal Eurasian tests used so far look for evidence of a deeper ancestry that branched before the ancestry shared between Tianyuan and Ust-Ishim. f4(Test,Tianyuan; Ust'Ishim, Chimp) will give lower values for more deeply branching ancestry (test will look more African), and the values can be related either to the amount of deeper ancestry or the depth of the branch. Satsurblia and Kotias (CHG), Natufians, Neolithic Iranians all give low values on this, which has been interpreted as Basal Eurasian content (note that Mbuti, Yoruba and Torafalt also give very low values in this test). This test has been applied to the 26k Dzudzuana samples and they also have low values, which Lazaridis et al. Bioarchives paper from 2018 interprets as the first appearance of Basal Eurasian ancestry in Europe. Lazarides, et al. propose a ~50% Basal Eurasian ancestry for the Dzudzuana samples. This paper has not yet been peer-reviewed or published in a journal and we do not have access to the data to the Dzudzuana data to compare.

This Basal Eurasian ancestry test has not been applied to older samples with possible deep admixture. One potential problem using this test with samples containing European IUP ancestry is that they are likely to have shared alleles with Tianyuan/Asians as has been shown in the Bacho Kiro genomic paper (Hajdinjak et al., 2021), which violates the requirement of Tianyuan being an independent reference. Additionally, deeply-branching pre-40k ancestry, such as the Zlaty Kun which we report in our new revised paper, may appear as being more African and can confound the results. Thus we do not attempt to test or address "Basal Eurasian" ancestry in this study. Until we have better direct and clear characterization of Basal Eurasian ancestry, we have no evidence that the Basal Eurasian component of the CHG ancestry is not in fact referencing the deeply branching Zlaty Kun component. We expect there is sure to be more information to come out regarding the Dzudzuana individuals and Basal Eurasian ancestry, and we await these results.

As for the affinities between BuKa and Oase1 being linked to the increased Neanderthal content in Oase1, we tested for this directly by masking this content using the coordinates defined in the supplements to the Oase1 paper (Fu, et al., 2015). The masked Oase1 actually increases the f4 values, as we state in the text and show in Table S7, so we have evidence that the affinities between BuKa3 and Oase1 are seen despite, rather than being due to, the increased Neanderthal ancestry of Oase1.
We explicitly state this in the paper on page 11 lines 241-246:

"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)."

There are some minor comments:

What do authors mean in lines 24-25 by the 'first migration wave of humans carrying ancestry found in present-day Europeans'? Currently unclear from the sentence structure if they refer to the overall connection to later Europeans – which would still make the Kostenki14 the earliest representative of this wave or the 'first' migration wave of human ancestry found in present-day Europeans, and which is still quite distinct from the BuKa ancestry – or if the reference is to the Basal Eurasian genetic component?

The sentence in question is the first sentence of the abstract of the old version, "We describe the genomes of two c. 37,000-year-old individuals from the site of Buran-Kaya III in Crimea as belonging to the first migration wave of humans carrying ancestry found in present-day Europeans."

The current version has modified this as follows (page 2, lines 23-25):
"We describe the genomes of two c. 36,000 and 37,000-year-old individuals from the site of Buran-Kaya III in Crimea as belonging to the earliest migration wave of humans carrying ancestry found in present-day Europeans. "

Due to word constraints in the abstract, we define this "wave" more fully in the first paragraph of the introduction, Page 3, lines 49-55:

" 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) and Sunghir (~34,000 cal BP) in the eastern European Plain and in two ~35,000-year-old individuals, GoyetQ116-1 from Belgium and BK1653 from the MUP deposits of Bacho Kiro, Bulgaria, both of whom also contain trace ancestry linked to the previous IUP inhabitants of Bacho Kiro"

Thus, the description that later defines the 'wave' specifies that this is the wave including Kostenki14, and that Kostenki14 is currently its oldest member.

Following your comment, we have explored other formulations, but we have come back satisfied that the present wording is concise and accurate, and have decided to keep it in the current version, though we remain open to other suggestions. The BuKa3 individuals are part of the same wave of ancestry that began appearing in Europe at some point after 40,000 years ago. We do not say they were the 'first', just that they 'belonged' to it, so we don't anticipate any confusion with the older Kostenki14. We also don't anticipate this would be confused with Basal Eurasian ancestry since we do not introduce this enigmatic ancestry anywhere in the text.

Just for practical reasons of not having to jump between different files, it might be worth placing the Supplementary Tables 2 and 7 as sheets in the xlsx file with other supplementary tables rather than in the text file of the SI.

This is a good idea. We now have moved the references to genomes and mitochondrial sequences used, Supplementary Tables 2 and 11, to the supplemental table excel file.

Line 520 psesudo-haploid should be pseudo-haploid.

Thank you for catching this typo. It's fixed.

Reviewer 3:

In their resubmitted manuscript, Bennet et al substantially expand their dataset reporting new data from the previously included individual BuranKaya3A as well as a new individual BuranKaya3C of a similar age. I am very appreciative of the effort going in to increasing the amount of data from these exciting but challenging to work with human remains. The authors present mostly new analyses based on these data, which make for an overall improved manuscript. However, some of my previous concerns with regards to the analyses and interpretation of the data remain in this new version.

New dataset

The authors now report data from two individuals. Data generation, contamination estimates and analyses of uniparental markers all support authenticity of the data. While the authors present fold-coverage estimates for the MT genomes (82X and 11X for BuKa3A and BuKa3C, respectively), I could not find the same overall final fold-coverage values for the nuclear genomes. Table S6 reports per-library read numbers, and number of overlapping SNPs with the ~13 million marker reference panels indicate ~0.04X and ~0.02X coverage, but it would be useful to directly report the estimated number in the main text.

Thank you for this suggestion. We have added in the supplementary table S10 the information about the fraction of the genome covered that is almost identical to the fraction of SNPs covered.

For the construction of the analysis panel, the authors extract a panel of ~740,000 SNPs from the HGDP high-coverage WGS panel, for which either one of the ancient BuranKaya individuals has coverage. This 740K panel is then used for all analyses with other published ancient genomes. In my review of the previous version, I raised the concern of possible bias in joint analyses of shotgun and targeted capture samples if the SNPs covered between them differ systematically. This remains a potential problem in the revised version, as there seems to have been no further filtering for SNPs included in the targeted capture probes in the analyses. As an example, for the f3 analyses presented, the statistics estimated from comparison between BuranKaya and another shotgun sample will be derived from a random subset of the 740K SNPs with coverage in both, whereas in

comparison with a capture sample the statistics will be heavily driven by a non-random subset of the 740K which were pre-ascertained and captured in the individual. These also might differ between capture individuals analysed here, as e.g. Vestonice 16 was captured at a set of 3.7M SNPs, whereas the Krems double burial I2483/4 was only captured using the 1240K panel. The authors should restrict analyses to the smallest subset of SNPs ascertained across all individuals, to rule out these potential issues whenever jointly analysing samples with different capture schemes.

We have taken this comment seriously and discussed various ways to best ensure our results cannot be attributed to bias comparing between shotgun and captured datasets. As suggested, we looked into limiting our analysis only to the smallest subset of captured SNPs, but although we have worked over the past years to significantly increase the coverage achieved with our Buran Kaya sample, work that has led to the discovery of an additional sample with slightly better preservation (BuKa3C), restricting our analysis to the subset of overlapping SNPs from captured samples, especially when compared with poorly covered captured samples, resulted in large error bars that reduced the power of our analyses. Instead we have applied the following controls to attempt to both limit and quantify the impact of bias when co-analyzing shotgun and capture generated data has on our results.

We have firstly attempted to characterize the existing bias when comparing captured and shotgun data by using the work done investigating such bias published in "Three assays for in-solution enrichment of ancient human DNA at more than a million SNPs", Rohland, et al., 2022. This publication identifies which SNP positions in the 1240k panel show bias when comparing with shotgun data. In separate tests in our lab, removal of these SNPs (leaving 469k of 1240k SNP panel) is sufficient to improve congruity between shotgun and capture results. We removed these SNPs from our dataset and reran our f-statistics analyses and compared the results. We found that removing bias-prone SNPs changed our f3 values by between -2.0-5.7% (an average 1.8% reduction of the f3 value), with the order of the samples giving the highest values being unchanged. For f4 results, removing captured SNPs prone to bias moved Z-scores either above or below a significance threshold of 3 in 73 out of 1260 calculations. Of these, 43 of the 45 instances that involved pre-CI genomes increased allele sharing between BuKa3 and pre-CI ancestry, and 22 out of 28 the cases involving only post-CI genomes increased allele sharing between BuKa3 and Fournol or Vestonice cluster individuals (4 of these do not involve Gravettian-associated individuals). In summary, for the 5.8% of relationships where removal of bias-prone SNPs had a significant impact, 95.6% increased the number of alleles shared between BuKa3 and pr-CI individuals and 78.6% increased allele sharing between BuKa3 and MUP Gravettian-associated individuals when comparing shotgun and captured datasets. We conclude that wherever capture bias may have played a role in our results when comparing captured and shotgun datasets together, it appears overall only to underestimate the amount of shared alleles between BuKa3 and pre-CI populations and the affinities with Gravettian-associated groups that we report. After going over these results, we decided, since the impact of bias was minimal and not contrary to our overall findings, to maximize the power of the increased SNP count and use the full panel, keeping a consistent inventory of possible SNPs for all samples.

In the current revised version of the manuscript, we now include the results of this bias test for all values calculated as additional supplemental material in supplemental tables S3 and S5.

We recognize that a subset of samples were captured with larger SNP panels that contain additional SNPs whose bias-potential has not been characterized in the Rohland et al. 2022 publication, but given the size of the differences we observe for the tests done with the 1240k panel, we hope that you may be convinced that any remaining bias present will not consequentially change our interpretations based on tests combining both captured and shotgun data.

Secondly, to reduce the role of noise in our interpretations, in this new revised manuscript we have removed our discussion of specific relationship trends backed only by results of a single test.

Thirdly, we have limited our qpGraph analysis, which statistically favors admixture with a Zlaty-Kun-like IUP-contemporary ancestry, to only considering other shotgun data. We also stress that some of our key f4 statistics showing gene flow between pre- and post-CI individuals support this whether using formulations of both shotgun-only data or mixed shotgun-capture, e.g:

Shotgun only:

Mbuti	ZlatyKun	BuKa3A Ust_Ishim	-0.001559	0.000466	-3.348
Mbuti	ZlatyKun	BuKa3C Ust_Ishim	-0.000883	0.000411	-2.146

Shotgun-Capture:

Mbuti	ZlatyKun	BuKa3A	BG_IUP_BK_F6-620	-0.001676	0.000777	-2.155
Mbuti	ZlatyKun	BuKa3C	BG_IUP_BK_F6-620	-0.001895	0.000627	-3.023

We hope that the increased testing and stringency we've applied increases confidence in our results for these specific cases.

In the methods section, the authors also describe the curation of another SNP panel (page 24, line 522), but it is unclear to me if and where this panel was used in the analyses.

We apologize for the being unclear. We now clarify that this panel was only used for the f3 calculations with present-day genomes presented in Fig. S5, and we have changed the text to make this clear
Page 25, lines 571-576:
"For *f3*-statistical comparisons with present-day genomes shown in Fig. S5, 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 v.4.1.968[70] which led to ~1.7 million biallelic autosomal positions after exclusion of heterozygous positions coming only from BuKa3."

UP Admixture

In this section the authors discuss differences in shared ancestry between the two BuKa individuals and other UP individuals. Some of the analyses presented here suffer from possible bias as described above, and overall signals appear quite weak, so I would be cautious in interpreting these as true differences. A direct test of clade-like relationship of the two individuals as e.g. in f4(Mbuti, other UP; BuKa3A, BuKa3C) would give a significant non-zero statistic in this case, but has not been done. If the overlap of SNPs between the two low coverage samples is too low, I suggest at least to also show a plot of pairwise f3-outgroup statistics (on smallest subset of capture SNPs as discussed above) with corresponding error bars to make it easier to discern how noisy those estimates are.

We thank you for your suggestion of testing directly different allele sharing as evidenced by f3 in an f4 framework. For f3(Mbuti; BuKa3A, BuKa3C), the low overlap between the two BuKa samples unfortunately results in overly large error bars, however, we have now also performed the f4 test requested, which gives satisfactory results. We find both BuranKaya3 individuals show a statistically clade-like relationship with each other in regard to all samples tested, demonstrating that they may be treated as a population in some analyses.
For example, the highest and lowest scores in this test:

Mbuti	BE_Aur_GoyetQ116-1	BuKa3ABuKa3C	-0.007047	0.006651	-1.059
Mbuti	BG_UP_BK-1653	BuKa3ABuKa3C	0.003926	0.003422	1.147

In light of the additional concerns mentioned in this comment, in the current version we have removed the detailed descriptions of less statistically supported sample-specific relationship differences, including all those supported by the large multi-sample qpGraph, which is now also removed.

As mentioned above, a requirement of using only the smallest subset of capture SNPs would greatly reduce the resolution and limit the amount of useful information we can show. However, we do agree that a pairwise f3-outgroup plot with error bars will be useful to easily judge the confidence of any given result. Based on this suggestion, in the current revision we now include a biplot of the f3 results of both BuKa3 individuals together with error bars in Fig. 1B, and we present the full results of our f4(Mbuti, UP; BuKa3A, BuKa3C) now in Table S4. We also specifically mention the limitations of outgroup f3 results for both individuals together, and the clade-like results of the f4(Mbuti, UP; BuKa3A, BuKa3C) in the text on
Page 8, line 162-166:
" 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)."

The authors also present a rather complex admixture graph fitted with both BuKa individuals in the context of other IUP and UP individuals. Unfortunately, I could not find any details on how the authors arrived at this fit (as in how many / which other models were tested), nor on the parameters used for qpGraph. As complex models like this are prone to overfitting (see e.g. recent preprint by Maier et al bioRxiv 2022.05.08.491072), I would like to see some results replicated in simpler models and/or further analyses to provide a better understanding if/how many other topologies might also be supported.

We agree that a complex qpGraph with multiple admixture edges may not be the most convincing approach to present our main points in this study. To remedy this, we have decided not to attempt such a bold endeavor with the given data, and have removed this figure and all related discussion. In its place we now present two, much simpler models focused only on modelling the ancestry of each BuKa3 sample separately as Figure 3B. For this revision, we also made use of ADMIXTOOLS 2 to verify the model we selected was optimal (Maier et al paper (https://doi.org/10.7554/eLife.85492). Additionally, for these best-fitting models, we include alternate versions without admixture and their resulting Z-scores as supplemental figure S9, so readers may observe the impact of removing admixture from the models.

Pre-CI gene flow

In this section the authors discuss results indicating the presence of remnant IUP ancestry in BuKa. However, the evidence presented is again weak. While the f4 statistics in Fig 3A do suggest a slight increase in sharing between BuKa3A and Oase over two of the Bacho Kiro individuals, the signal is much weaker or absent when using other IUP samples as comparison. Furthermore, table S5 shows the more direct test of excess Oase ancestry using f4(Mbuti, Oase; BuKa, other UP/IUP). In case of gene flow between Oase and BuKa these statistics would be expected to be significantly >0 when comparing BuKa to other UP indiivudals, but all combinations in S5 show f4=0 which is consistent with Oase being an outgroup symmetrically related to all tested sample pairs.

The f4 results presented in our previous version focusing on allele sharing between BuKa3 and Oase were based primarily on indirect observations without robust statistical support, for example, while, as you mention, for f4(Mbuti, Oase; BuKa3A, UP/IUP), even though the results are all negative for IUP samples, they are not quite statistically different from zero. But if Oase was truly an outgroup, for MUP samples showing affinity for BuKa3A we would expect f4(Mbuti, BuKa3A; Oase, MUP) to be significantly positive, yet for many MUP individuals this is zero with BuKa3A (while positive for BuKa3C). Our explanation was that the allele sharing amount would be low and thus escape robust detection between the low coverage of both of our samples and Oase. We accept that this is not convincing and will have to wait for deeper coverage of these and other UP individuals for these contradictions to be better resolved. Thus, in this revision, we no longer attempt to make a robust case for Oase admixture beyond highlighting some f4 results.

However, in performing extensive f4 analyses in preparation for the current revision, which now includes data from the recently published western European MUP genomes from Posth, et al. 2023, we found better support for pre-CI admixture from a population more closely resembling Zlaty Kun. This was alluded to in our previous version, but we show here that gene flow from Zlaty Kun-like population finds broad statistical support, both for Buran Kaya III and members of the newly published Fournol cluster, with which the Buran Kaya samples are shown to now share the most alleles. We further were able to show Zlaty Kun-like ancestry to be elevated in the Mesolithic CHG cluster. These two findings tie together two disparate observations presented in the previous work, with the Fournol cluster members from southern Europe providing the previously missing information and support. The current version now presents these new results. Regarding the f4 configurations mentioned in the above comment, in contrast to Oase1, we do see a significant amount of allele sharing between BuKa3A and Zlaty kun in different UP and IUP ancestry contexts, for example:

Mbuti	ZlatyKun	BuKa3A	Sunghir3	-0.00138	0.000419	-3.294
Mbuti	ZlatyKun	BuKa3C	Sunghir3	-0.000833	0.000365	-2.281
Mbuti	ZlatyKun	BuKa3A	Kostenki14	-0.001382	0.000637	-2.169
Mbuti	ZlatyKun	BuKa3C	Kostenki14	-0.000903	0.000468	-1.928
Mbuti	ZlatyKun	BuKa3A	Ust_Ishim	-0.001559	0.000466	-3.348
Mbuti	ZlatyKun	BuKa3C	Ust_Ishim	-0.000883	0.000411	-2.146

Additionally, as seen above, we observe a similar consistent trend in BuKa3C, which we now model as receiving ~6% Zlaty Kun-related ancestry, although this f4 result does not quite reach a 2.5 or 3 Z-score cut-off. Our assessment is that these results indicate that the amount of admixture from Zlaty-like ancestry in BuKa3C, although in agreement with f3 and qpGraph analyses, may be too low to cross the significance threshold in our f4 test results.

The Zlaty Kun f3, f4, and qpGraph results are now featured in Figures 1B, 3A and 3B respectively.

Post-CI / Gravettian

The evidence for differential relationship of the Gravettian samples in Vestonice and Krems suffers from similar weaknesses as described in the previous sections. The authors base their conclusions here on the slight differences in outgroup f3 statistics (with their possible ascertainment bias issues) and indirect evidence from comparing two different f4 tests. Instead, the authors should have used the direct test for differential relatedness f4 (Mbuti, BuKa; Vestonice16, Krems), results of which can be found in table S4 and do not support gene flow between Krems and BuKa (f4 not significantly different from 0 with Z-scores -0.25 and 0.98 respectively).

This aspect of our previous work has been changed dramatically in the current revision, due in large part to the release of the Gravettian-associated western European MUP genomes of the Posth et al, 2023 paper. The issue of weak affinities had been due to the fact that our BuKa3 individuals were found to be considerably distant genetically from all other E/MUP high coverage individuals available at the time of our earlier analysis. In our previous work, BuKa3 behaved as nearly an outgroup with nearly all MUP genomes, with few exceptions (e.g. f4(Mbuti, BuKa3C, Krems1, Sunghir). In order to place BuKa3 within the known genomic MUP landscape, we addressed this distance (= the lack of strong statistical allele sharing) by relying on other analyses, such f3 results, inability of Krems1 to differentiate any contemporaries from BuKa3 in f4 (our old Fig. 2B from the previous manuscript), and how only BuKa and Krems1, of all UP individuals tested at that time, showed significant allele sharing with the CHG group when compared to all other E/MUP individuals.

With the recent publication of the Gravettian-associated Fournol cluster from western and southwestern Europe (Posth et al. 2023), we now have identified several individuals who show consistently statistically robust allele sharing with BuKa3 in the specific f4 test referenced above. Some examples with the individual Fournol85 now showing the best supported allele sharing with BuKa3:

Mbuti	BuKa3AFornol85	Sunghir3	-0.005963	0.001243	-4.798
Mbuti	BuKa3CFornol85	Sunghir3	-0.002606	0.000891	-2.925
Mbuti	BuKa3AFornol85	Kostenki14	-0.004817	0.001446	-3.331
Mbuti	BuKa3AFornol85	BK-1653	-0.007042	0.001491	-4.724
Mbuti	BuKa3CFornol85	BK-1653	-0.003546	0.001075	-3.298
Mbuti	BuKa3AFornol85	Vestonice16	-0.005235	0.001818	-2.879
Mbuti	BuKa3AFornol85	Krems1 _I2483-4	-0.005224	0.001752	-2.982
Mbuti	BuKa3AFornol85	Vestonice14	-0.007567	0.001913	-3.955
Mbuti	BuKa3AFornol85	Vestonice15	-0.006208	0.001881	-3.301
Mbuti	BuKa3AFornol85	Ormesson	-0.007278	0.002166	-3.36
Mbuti	BuKa3AFornol85	Paglicci12	-0.003914	0.001489	-2.629
Mbuti	BuKa3AFornol85	Muierii	-0.004406	0.001228	-3.587

The revised text now focuses on these relationships, which are also seen with the Fournol cluster member Serinya. F3 results supporting the relationship between BuKa3 and members of the Fournol cluster from southern France and northeastern Spain are now shown on the new Figures 1A, and 1B, and f4 statistics showing the same are featured in Figures 2A and 2B.

We further show that, like BuKa3, the Fournol members with high allele sharing with BuKa3 also show admixture with a Zlaty Kun-like ancestry, shown in Figure 3A, and also share the affinity with CHG we identified in BuKa3 in the previous manuscript (shown now in Figure 4A and 4B).

With these new results, the previously identified relationship with Krems1 is no longer a central part of the main text. Although it is featured in Figures 2, 3A, and 4, beside the similarly behaving newly published fellow Gravettian-associated sample, Paglicci12 from southern Italy, who was also

found to share some affinities with Zlaty Kun, and CHG. Zlaty Kun gene flow is more convincingly shown with Paglicci12 than with Krems1, as can be seen by f4 results, now in Table S8, such as:

Mbuti	ZlatyKun	Muierii	Krems1_1-2	0.002674	2.963
Mbuti	ZlatyKun	Sunghir3	Krems1_1-2	0.002393	2.743
Mbuti	ZlatýKůň	Vestonice	Paglicci12	0.002058	3.169
Mbuti	ZlatýKůň	Muierii	Paglicci12	0.002853	4.579
Mbuti	ZlatýKůň	Kostenki14	Paglicci12	0.001957	2.739
Mbuti	ZlatýKůň	Sunghir3	Paglicci12	0.00231	3.791
And additionally:					
Mbuti	ZlatýKůň	Krems1_1-2	Paglicci12	0.000253	0.269

Together, this information now allows us to have a specific insight into Gravettian-associated population structure during the MUP, similar to that that escaped strong statistical significance in our previous versions. We still find the individuals at BuranKayaIII in Crimea, archeologically linked both to Gravettian assemblages and the Caucasus, sharing the highest number of alleles with ~5,000 year younger Gravettian-associated individuals, not of the Vestonice cluster of Central Europe as previously reported, but rather certain members of the Fournol cluster, specifically those of the southern range of this cluster in western Europe. This shared ancestry includes admixture with a pre-CI Zlaty Kun-like source, traces of which can also be found in southern Italy and Austria, and which is also elevated in the Caucasus during the Mesolithic. This last aspect is particularly intriguing, and we present some hypotheses in our current text addressing this.

CHG affinity

The final result section discusses genetic affinities between BuKa and palaeolithic hunter-gatherers from the Caucasus. I found this result to be much more well supported and interesting than the previous sections. Both individuals show consistent increase in allele sharing with CHG in the direct test shown in Fig 3C and Table S4, especially when comparing with the geographically and temporally close Sunghir and Kostenki genomes, which additionally don't suffer from shotgun / capture biases in this configuration.

Overall, my take is that the manuscript has some exciting new data, but too much focus is on discussing weak signals based on a few marginally significant f-statistics or a possibly overfitted graph. On the other hand, there is a clear result in the comparison with CHG genomes, but no further exploration of this signal in the context of the other good quality genomes from the region (i.e. Sunghir/Kostenki). Not including a more thorough exploration of these relationships with e.g. more simple and robust admixture graph fit comparisons would in my view constitute a missed opportunity for these exciting new genomes.

Our new analysis explores the CHG connection more deeply, and exposes a link between this and the signal of pre-CI Zlaty Kun-like ancestry we identified in our Buran Kaya III individuals as well as several newly-published MUP Gravettian-associated individuals, which we present in Figure 4B. We additionally include a full range of f4 results concerning pre-LGM populations and CHG ancestry as an additional supplementary table (Table S9), and test several populations, including Kostenki14 and Sunghir3 for relative CHG affinities with our BuKa individuals (Figure 4A). Hypotheses of these ancestry distributions described in this study and how they may reveal interactions of populations found in pre-CI Europe with incoming EUP/MUP ancestries are presented in the discussion.

We thank you for your thoughtful comments and concerns of the last version of our manuscript and have taken them to heart and removed much of the more weakly supported affinities. We have made attempts to explore and quantify the impact of capture-bias on our results, we have removed relationship discussions based solely on f3 results, we have simplified the qpGraphs to be more focused on the ancestry of the samples in this study and to exclude captured data, and we have shifted focus of pre-CI ancestry admixture from Oase1 to the better supported Zlaty Kun. We feel the inclusion of the Fournol cluster data has put these findings in a much more enlightened context and enabled us to give a richer characterization of the Buran Kaya III genomes and surrounding events during the populating of Europe during this time period. We hope that you find our current revision more balanced, convincing, and interesting.

Decision Letter, second revision:

14th July 2023

Dear Eva-Maria,

Thank you for submitting your revised manuscript "Early Upper Paleolithic genomes of Crimea show migration and admixture dynamics of the first modern European ancestries in relation to climatic crisis" (NATECOLEVOL-19087657B). It has now been seen again by the original reviewers and their comments are below. The reviewers find that the paper has improved in revision, and therefore we'll be happy in principle to publish it in Nature Ecology & Evolution, pending minor revisions to comply with our editorial and formatting guidelines.

If the current version of your manuscript is in a PDF format, please email us a copy of the file in an editable format (Microsoft Word or LaTex)-- we can not proceed with PDFs at this stage.

We are now performing detailed checks on your paper and will send you a checklist detailing our editorial and formatting requirements in about a week. Please do not upload the final materials and make any revisions until you receive this additional information from us.

Thank you again for your interest in Nature Ecology & Evolution. Please do not hesitate to contact me if you have any questions.

[REDACTED]

Reviewer #2 (Remarks to the Author):

I enjoyed reading the current version of the manuscript. I especially appreciate the extra work the authors have taken to include newly published data of additional individuals in their analyses and the careful examination of the resulting statistics (as well as the comparisons between the shotgun vs captures), which have made the current version of the manuscript and their findings even more compelling. As for my previous comments, authors have addressed them. Therefore, I believe both the data and the findings to be of great interest to a wider scientific community, and would recommend their publication.

Reviewer #3 (Remarks to the Author):

I appreciate the extensive work the authors have put into further addressing my concerns of some of the results presented in the previous version. In particular, I agree with the authors that the addition of the newly published genomic data from Posth et al provides some intriguing links between the Buran Kaya genomes, the Fournol cluster and pre-CI or CHG ancestry. The newly revised analyses provide more convincing support for these relationships, and as such the manuscript is much improved. I hence don't have any further comments and recommend the manuscript for publication.

nature research

Our ref: NATECOLEVOL-19087657B

26th July 2023

Dear Dr. Geigl,

Thank you for your patience as we've prepared the guidelines for final submission of your Nature Ecology & Evolution manuscript, "Early Upper Paleolithic genomes of Crimea show migration and admixture dynamics of the first modern European ancestries in relation to climatic crisis" (NATECOLEVOL-19087657B). Please carefully follow the step-by-step instructions provided in the attached file, and add a response in each row of the table to indicate the changes that you have made. Please also check and comment on any additional marked-up edits we have proposed within the text. Ensuring that each point is addressed will help to ensure that your revised manuscript can be swiftly handed over to our production team.

We would like to start working on your revised paper, with all of the requested files and forms, as soon as possible (preferably within two weeks). Please get in contact with us immediately if you anticipate it taking more than two weeks to submit these revised files.

When you upload your final materials, please include a point-by-point response to any remaining reviewer comments.

If you have not done so already, please alert us to any related manuscripts from your group that are under consideration or in press at other journals, or are being written up for submission to other journals (see: https://www.nature.com/nature-research/editorial-policies/plagiarism#policy-on-duplicate-publication for details).

In recognition of the time and expertise our reviewers provide to Nature Ecology & Evolution's editorial process, we would like to formally acknowledge their contribution to the external peer review of your manuscript entitled "Early Upper Paleolithic genomes of Crimea show migration and admixture dynamics of the first modern European ancestries in relation to climatic crisis". For those reviewers who give their assent, we will be publishing their names alongside the published article.

Nature Ecology & Evolution offers a Transparent Peer Review option for new original research manuscripts submitted after December 1st, 2019. As part of this initiative, we encourage our authors

to support increased transparency into the peer review process by agreeing to have the reviewer comments, author rebuttal letters, and editorial decision letters published as a Supplementary item. When you submit your final files please clearly state in your cover letter whether or not you would like to participate in this initiative. Please note that failure to state your preference will result in delays in accepting your manuscript for publication.

Cover suggestions

As you prepare your final files we encourage you to consider whether you have any images or illustrations that may be appropriate for use on the cover of Nature Ecology & Evolution.

Covers should be both aesthetically appealing and scientifically relevant, and should be supplied at the best quality available. Due to the prominence of these images, we do not generally select images featuring faces, children, text, graphs, schematic drawings, or collages on our covers.

We accept TIFF, JPEG, PNG or PSD file formats (a layered PSD file would be ideal), and the image should be at least 300ppi resolution (preferably 600-1200 ppi), in CMYK colour mode.

If your image is selected, we may also use it on the journal website as a banner image, and may need to make artistic alterations to fit our journal style.

Please submit your suggestions, clearly labeled, along with your final files. We'll be in touch if more information is needed.

Nature Ecology & Evolution has now transitioned to a unified Rights Collection system which will allow our Author Services team to quickly and easily collect the rights and permissions required to publish your work. Approximately 10 days after your paper is formally accepted, you will receive an email in providing you with a link to complete the grant of rights. If your paper is eligible for Open Access, our Author Services team will also be in touch regarding any additional information that may be required to arrange payment for your article.

Please note that Nature Ecology & Evolution is a Transformative Journal (TJ). Authors may publish their research with us through the traditional subscription access route or make their paper immediately open access through payment of an article-processing charge (APC). Authors will not be required to make a final decision about access to their article until it has been accepted. Find out more about Transformative Journals

Authors may need to take specific actions to achieve compliance with funder and institutional open access mandates. If your research is supported by a funder that requires immediate open access (e.g. according to Plan S principles) then you should select the gold OA route, and we will direct you to the compliant route where possible. For authors selecting the subscription publication route, the journal's standard licensing terms will need to be accepted, including <a href="https://www.nature.com/nature-portfolio/editorial-policies/self-archiving-and-license-to-publish. Those licensing terms will supersede any other terms

that the author or any third party may assert apply to any version of the manuscript.

For information regarding our different publishing models please see our Transformative Journals page. If you have any questions about costs, Open Access requirements, or our legal forms, please contact ASJournals@springernature.com.

Please use the following link for uploading these materials:
[REDACTED]

If you have any further questions, please feel free to contact me.

[REDACTED]

Reviewer #2:
Remarks to the Author:
I enjoyed reading the current version of the manuscript. I especially appreciate the extra work the authors have taken to include newly published data of additional individuals in their analyses and the careful examination of the resulting statistics (as well as the comparisons between the shotgun vs captures), which have made the current version of the manuscript and their findings even more compelling. As for my previous comments, authors have addressed them. Therefore, I believe both the data and the findings to be of great interest to a wider scientific community, and would recommend their publication.

Reviewer #3:
Remarks to the Author:
I appreciate the extensive work the authors have put into further addressing my concerns of some of the results presented in the previous version. In particular, I agree with the authors that the addition of the newly published genomic data from Posth et al provides some intriguing links between the Buran Kaya genomes, the Fournol cluster and pre-CI or CHG ancestry. The newly revised analyses provide more convincing support for these relationships, and as such the manuscript is much improved. I hence don't have any further comments and recommend the manuscript for publication.

Author Rebuttal, first revision:

Reply to the Reviewers

Reviewer #2 (Remarks to the Author):

I enjoyed reading the current version of the manuscript. I especially appreciate the extra work the authors have taken to include newly published data of additional individuals in their analyses and the careful examination of the resulting statistics (as well as the comparisons between the shotgun vs captures), which have made the current version of the manuscript and their findings even more compelling. As for my previous comments, authors have addressed them. Therefore, I believe both the data and the findings to be of great interest to a wider scientific community, and would recommend their publication.

Reviewer #3 (Remarks to the Author):

I appreciate the extensive work the authors have put into further addressing my concerns of some of the results presented in the previous version. In particular, I agree with the authors that the addition of the newly published genomic data from Posth et al provides some intriguing links between the Buran Kaya genomes, the Fournol cluster and pre-CI or CHG ancestry. The newly revised analyses provide more convincing support for these relationships, and as such the manuscript is much improved. I hence don't have any further comments and recommend the manuscript for publication.

We are pleased that we could address the constructive criticism of the reviewers and thank them for their work and input that helped us to improve the manuscript. We are grateful for the final positive statement which is very satisfactory.

Final Decision Letter:

6th September 2023

Dear Eva-Maria,

We are pleased to inform you that your Article entitled "Genome sequences of 36-37,000 year-old modern humans at Buran-Kaya III in Crimea", has now been accepted for publication in Nature Ecology & Evolution.

Over the next few weeks, your paper will be copyedited to ensure that it conforms to Nature Ecology and Evolution style. Once your paper is typeset, you will receive an email with a link to choose the appropriate publishing options for your paper and our Author Services team will be in touch regarding any additional information that may be required

After the grant of rights is completed, you will receive a link to your electronic proof via email with a request to make any corrections within 48 hours. If, when you receive your proof, you cannot meet this deadline, please inform us at rjsproduction@springernature.com immediately.

Due to the importance of these deadlines, we ask you please us know now whether you will be difficult to contact over the next month. If this is the case, we ask you provide us with the contact information (email, phone and fax) of someone who will be able to check the proofs on your behalf, and who will be available to address any last-minute problems . Once your paper has been scheduled for online publication, the Nature press office will be in touch to confirm the details.

Acceptance of your manuscript is conditional on all authors' agreement with our publication policies (see www.nature.com/authors/policies/index.html). In particular your manuscript must not be published elsewhere and there must be no announcement of the work to any media outlet until the publication date (the day on which it is uploaded onto our web site).

Please note that Nature Ecology & Evolution is a Transformative Journal (TJ). Authors may publish their research with us through the traditional subscription access route or make their paper immediately open access through payment of an article-processing charge (APC). Authors will not be required to make a final decision about access to their article until it has been accepted. Find out more about Transformative Journals

Authors may need to take specific actions to achieve compliance with funder and institutional open access mandates. If your research is supported by a funder that requires immediate open access (e.g. according to Plan S principles) then you should select the gold OA route, and we will direct you to the compliant route where possible. For authors selecting the subscription publication route, the journal's standard licensing terms will need to be accepted, including <a href="https://www.nature.com/nature-portfolio/editorial-policies/self-archiving-and-license-to-publish. Those licensing terms will supersede any other terms that the author or any third party may assert apply to any version of the manuscript.

In approximately 10 business days you will receive an email with a link to choose the appropriate publishing options for your paper and our Author Services team will be in touch regarding any additional information that may be required.

If you have any questions about our publishing options, costs, Open Access requirements, or our legal forms, please contact ASJournals@springernature.com

An online order form for reprints of your paper is available at https://www.nature.com/reprints/author-reprints.html. All co-authors, authors' institutions and authors' funding agencies can order reprints using the form appropriate to their geographical region.

We welcome the submission of potential cover material (including a short caption of around 40 words) related to your manuscript; suggestions should be sent to Nature Ecology & Evolution as electronic files (the image should be 300 dpi at 210 x 297 mm in either TIFF or JPEG format). Please note that such pictures should be selected more for their aesthetic appeal than for their scientific content, and that colour images work better than black and white or grayscale images. Please do not try to design a cover with the Nature Ecology & Evolution logo etc., and please do not submit composites of images related to your work. I am sure you will understand that we cannot make any promise as to whether any of your suggestions might be selected for the cover of the journal.

You can now use a single sign-on for all your accounts, view the status of all your manuscript

submissions and reviews, access usage statistics for your published articles and download a record of your refereeing activity for the Nature journals.

To assist our authors in disseminating their research to the broader community, our SharedIt initiative provides you with a unique shareable link that will allow anyone (with or without a subscription) to read the published article. Recipients of the link with a subscription will also be able to download and print the PDF.

You can generate the link yourself when you receive your article DOI by entering it here: <a href="http://authors.springernature.com/share">http://authors.springernature.com/share<a>.

[REDACTED]

P.S. Click on the following link if you would like to recommend Nature Ecology & Evolution to your librarian http://www.nature.com/subscriptions/recommend.html#forms

** Visit the Springer Nature Editorial and Publishing website at www.springernature.com/editorial-and-publishing-jobs for more information about our career opportunities. If you have any questions please click here.**