nature portfolio

Peer Review File

Mid-Pleistocene aridity and landscape shifts promoted Palearctic hominin dispersals

Corresponding Author: Professor Jinbo Zan

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The study hinges on two thoughts: 1) the new and 3.6 Myr long paleoprecipitation record (however, rainfall in numeric values like mm/yr with an uncertainty range is not specified) and 2) a comprehensive analysis of known hominin sites in Europe and Asia. However, the link between the two is not that evident: can the hominin site analysis be done without the new record? The discussion essentially focuses on the MPT and younger times (logically since not many hominin sites older than 1.6 Ma are known outside Africa), so what is the merit of the 3.6 Ma to 1.5 Ma portion of the new records (in terms of hominin dispersal routes)? Note that extended data Fig.2 shows the records starting at the Gauss/Matuyama boundary at ~2.6 Ma. The compilation of hominin sites and their environmental frame is a lot of work. The text is well written but seems to report mainly items that are largely known already. The new aspects remain somehow hidden. The change point analysis in this context is shown but its implications are only worded in very general terms while its necessary technical aspects, how robust, level of significance, etc., remain undiscussed.

Below are more specific line-by-line comments.

Abstract. How high is high resolution? Can authors be more specific?

32-33. “intensive tectonic uplift” reads vague. How much uplift is anticipated/recorded over 0.5-1.0 Myr? Not that much? A fair portion of the Tibetan Plateau was already at significant elevation at ~40 Ma or even before?

41-43. The Palearctic realm is Europe and most of Asia. The meaning of the red dotted line in figure 1 is the Arctic Sub-Realm. However, that is not mentioned in the main text. What is the point of showing it? All sites discussed are (far) to the south of it. In Asia, they seem to be largely confined to the Sino-Japanese realm and the Saharo-Arabian realm.

49. typo, take out “ago” after “Ma”

53. Define MPT again in the main text?

56. Realm, is capitalization required?

62. I wonder about the tectonic factors at about 1 Ma: most of the Eurasian continent had already achieved more or less their present-day elevations.

67. Carbonate-carbon or organic-carbon isotope data? Or both? OK, next sentence organic carbon isotope analysis is done, what about carbonate-carbon isotopes?

77-81. This is the “northern migration route” through Asia, what about the “southern route” through India?

85. Is the entire duration from 3.6 Ma (respectively 2.7 Ma) to Recent sampled? Be explicit here.

86. What is the typical sampling distance? What is the loess/paleosol accumulation rate? And what about the presence of paleosols? There is climatic contrast between them. What if all samples come from loess and none from paleosols? Then

bias could creep in.

94. At other locations the MPT is defined between 1.2 and 0.6 Ma. Something which takes place over 300 kyr is not abrupt.

95. But this shift in orbital expression cannot be resolved in the carbon isotope data given their resolution.

105-110. Again, the expression of paleosols in the record may deserve some attention (some information is provided in 366-371, at one sampling location (Tajikistan) paleosols are described as well developed; in the Tarim basin they are nearly absent. Loess is typically associated with drier conditions. What is the inferred range of paleoprecipitation?

116-121. Nice to know but what is the inferred paleoprecipitation range?

123-202. Most of the discussion relates to the MPT and younger times. What is the point of providing data from 3.6 Ma (or 2.7 Ma) to Recent? From 1.5 Ma to Recent would suffice? It is a good summary of existing knowledge but at the same time rather long. That central Asian climate became increasingly arid since the Pliocene/Pleistocene boundary with a stepwise increase in aridity during the MPT is well known. What is the new record adding to this knowledge base? This remains unclear.

124-128. But this is known for while? It is fine that present data concur with known trends, but the inference is not novel.

139-146. So, the present-day conditions in Tarim Basin prevailed since 0.6 Ma?

147. Here the MPT is implicitly placed between 1.2 and 0.6 Ma.

171-173. This pertains to the northern dispersal route in Asia. Up to now no notion of the southern route via India. The northern route relates to the southern boundary of the Palearctic realm. Hominins went further north since the last 0.4-0.5 Ma or so?

178. Check this tectonic uplift inference. Since 2001 many studies (more than 100?) have returned to the question of when, which parts of, and how much the Tibetan Plateau was uplifted. Current mainstream thinking is that most of the plateau reached more or less its current elevation at the Pliocene/Pleistocene boundary or even before.

207, 208. Provide the ages for MIS11 and MIS13.

213. A lower sea level is not “felt” by a river beyond its coastal plain. In the upstream region of interest here the direct impact of a lower sea level on rivers is marginal at best.

215. Figure 3. What is the basis for allocating a terrace sequence to climate-controlled or tectonic-climate-controlled? How are these two distinguished?

227. Serbia is more of southeastern Europe; northern Europe is Scandinavia?

231. And in Atapuerca? There hominin occupation is argued to be continuous since ~1.3 Ma?

251-265. Reads somehow repetitive?

266-275. Well-worded text but the statements are generic.

276. A distinct jump in reasoning? What are “our observations”?

279-284. Can this be rhymed with younger migration waves as analyzed in Ao et al. (2024) Proceedings of the National Academy of Sciences?

292. How is the “ready availability of fresh water” to be placed in an environment that features increased aridity?

296. In the very southern margin of the Palearctic realm.
Figure 1. The light blue on greenish background does not read well, pick another color combination.

Figure 3 and 4. The methodology behind the analyses deserves more mention in the main text since this represents a new twist to existing knowledge. How robust are the inferences made?

358. Why in Europe? That lies up-wind of the Tarim basin loess source areas?

364-366. So, authors rely on the age model provided by reference 14. How robust is that age model? It is robust, authors should provide some more information concerning sampling resolution and paleomagnetic stability here (is actually done in 374-378).

372-374. Superfluous, leave out?

384. Is a winter monsoon/westerlies proxy, specify that?

385. Is a summer monsoon proxy, specify that?

392. Strict adherence to GTS2020 nomenclature would imply kyr here.

393-394. But are they from loess, paleosol, or both per couplet?

397. At which temperature? Room temperature?

404-406. Also mentioned in the main text, however without much context over there.

421. Typos, reconstructions; appear

425. Numbers become less negative, is that a reduction?

Change point analysis section.
It may be wise to word the concept of this analysis and these tests with reference to a textbook. Extended data Figure 3 (and elsewhere). What is the meaning of the deviation from the zero line (always negative) in the red lines (brown-ish in my print) and green lines? What is the meaning of log(p)? What is implied with Central Tendency and Dispersion?

459. Asia is probably better here? Not a single record shown in extended data Figure 3 comes from Europe.

461. Reference for MatLab2016b?

466-473. These bits are not that helpful for a general readership. What is the outcome of this analysis? Where are the corresponding figures?

477. Dating values: Is this the number of sites compiled here?

479-480. What is inferred with this sentence?

482. No new hominin finds are reported here, what is implied with 'new discoveries'? How many new entries are included in comparison to previous compilations?

483. Is all duly referred to, how good are the updates? No new age assignments were obtained in the present study. How many sites/data entries were updated?

488. How many? Which key sites? Be specific. I presume the examples provided are not exhaustive.

503-506. Start this section with these notions. They could also be mentioned in the main text.

509-510. You mean that the methods are suited for the age range considered here? As worded it comes across as not overly informative.

510. Why 'residual'? Residual of what?

510-512. A histogram of the apparently more or less even distribution through time would be insightful.

512. And distribution through time was more or less even (at least it reads like that to me)? What is the cause of the inflection points?

513-517. What is the purpose of the notions made here?

521. How is this done and how is somebody supposed to read the extended data figures 11 and 12?

530. What are CAM plants? Explain the acronym.

554-603. How climatically controlled, tectonic-climate, and uplift controlled river terraces are being distinguished and how robust that distinction would be is not discussed.

564-567. The relation to the present study may become rather thin. What information would be lost if this was omitted?

566. Presumably dust fluxes are to those seas and oceans, since they cannot deliver dust.

579. dating values = ages?

590-593. All correct but what is the relevance since hominin dispersal is being discussed between 0.1 and ~1.3 Ma?

Extended data Figure 1. Are these figures obtained in this study? If so, provide some technical details in the caption. Which model was used, which years were integrated, etc. If not, properly refer to the source(s).

Extended data Figure 3. Explain log(p) in the caption. Mean and variance are being mentioned in the caption while central tendency and dispersion are being displayed on the figure panels. What is what?

878. Are the changes observed significant? At which level?

Extended data figure 4. How robust are those cluster analysis results? From the text it seems that authors have carried out that cluster analysis. Which input parameters were used and how was the stratigraphic control achieved?

Extended data figure 6. The age distinction seems rather distinct. How robust are the ages involved in this analysis?

Extended data figure 8l. The temperature anomaly is with respect to what?

917. explain acronym IRM@AF.

Reviewer #2

(Remarks to the Author)
The rationale for this study is provided in the second paragraph of the main text:
"The climatic and environmental histories of Europe and Asia are usually discussed separately. However, they are parts of the same biological Realm…As phylogenetic compositions are driven by regional climatic and ecologic factors 11, the histor[ies] of both regions need to be considered together when examining the drivers of hominin dispersals out of Africa".

This is an important point, but also a substantial undertaking. In this work, the authors start by presenting new δ13C data from loess deposits in Central Asia and northwest China over the last 3.6 Myr. The evidence points to a significant climatic shift during the so-called Middle Pleistocene Transition (MPT; 1.2-0.7 Ma). This is then expanded to a synthesis of palaeoenvironmental and palaeoanthropological/archaeological data to establish the transformation of landscapes across the Eurasian Palaearctic during the MPT, with implications for early human evolution and dispersal.

This is an impressive collection of new and published information on a grand scale that is worthy of publication. I have two main general points:

1. While the MPT 'sensu lato' is considered in general terms, the authors tend to focus more specifically on climate changes commencing at 0.9 Ma. Indeed, the so-called '0.9 Ma event' is traditionally taken to represent the first major Northern Hemisphere glaciation of MIS 22 (e.g. Clark et al.., 2006, QSR). However, closer inspection of the Tarim Basin loess δ13C record shows an earlier shift (~1.1 Ma) and this is also observed in certain other records in Fig. 2 and Ext. Data Fig. 8. Margari et al. (2023, Science) presented evidence for an extensive glaciation during the 30kyr-long MIS 34 (possibly associated with the Menapian cold stage of NW Europe), culminating in an extreme terminal stadial in the eastern North Atlantic. Climate envelope model simulations suggest that the intensity of the declines in temperature and net primary productivity would have led to the extirpation of early human populations in Europe and SW Asia. A critical review of archaeological record further suggested a paucity of European and SW Asian early hominin sites after that, possibly until MIS 21 (Margari et al., 2023). This may be worth considering in the present work.

2. The synthesis part of the paper represents a large collection of information, but space restrictions limit a fuller presentation and discussion of the evidence. For example, the river terraces dataset of Ext Data Table 2 and Fig. 3 are divided into tectonic-climate interactions, climatically-controlled and uplift-control, but their attribution and significance are not explained in the text or in the methods. In the same vein, the archaeological and human fossil evidence of Fig. 4 merit a fuller discussion of the granularity contained therein. More generally, the broad spatio-temporal canvases of Fig. 3 and 4 require a more expanded description.

In addition, the following specific points need attention:

l. 110. "the stable pCO2 levels…associated with the intensification of NH glaciation": While this may be in line with reconstructions based on carbonated precipitated in palaeosols (Da et al., 2019), it is not in line δ11B reconstructions showing drops in pCO2 after 2.8 and 2 Ma (Bartoli et al., 2011); Martinez-Boti et al., 2015). This may have to be considered in the δ13C changes of the loess records.

l. 132. "a pronounced shift in dominant periodicity". This is based on earlier work, but have the authors attempted to explore the periodicities embedded in their new δ13C records?

l. 166-170. "shift in vegetation from closed forests or woodland to steppe and grasslands is evident at ~0.9-0.6 Ma": this is not correct. While European pollen records show a serial extinction of moisture-requiring taxa (the so-called Tertiary relicts) over the MPT, interglacial stages were still characterized by forests (though less diverse). In the same vein, the inferences based on mammal evidence of a shift from closed forests to open woodland in Europe (Ext Data Fig. 6) are also tenuous, as they may represent local contexts (e.g. flood plains) and not supported by other lines of evidence (beetles, molluscs).

L. 230-231. “stable occupation of hominins in Europe was possibly since ~0.9 Ma”. It is worth adding here that Margari et al. (2023) suggested that repopulation of Europe after 0.9 Ma may have been by a more resilient species (e.g. *H. antecessor*) that was able to withstand the increasing intensity of glacial conditions.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The new version rightfully downplays the impact of tectonics. However, the river terraces are still categorized in the same three groups as before. The basis for their allocation is still pretty unclear (it seems to be taken at face value from the respective literature sources). Authors mention some generic geological approaches, but which criteria were used to allocate a river terrace to a certain group is not mentioned. What is the point of making the allocation in the first place since the tectonic impact is argued to be not that relevant?

To what extent the present analysis focusing on the MPT adds to the review by Ao et al. (2020) Earth Science Reviews? OK, data from Europe are not considered in that paper but climate change is dealt with in quite some detail. The final sentence of their abstract: “Combined with environmental theories of human evolution, this chronological coincidence suggests a potential link between mid-Pleistocene environmental changes and human evolution.“ A lot seems to have happened during the 400-500 ka age interval, i.e. formally after the MPT range discussed here?

The reason for keeping the long paleoclimate records as they are, is explained in the rebuttal letter (providing a kind of base line to better appreciate the changes in the MPT interval); in the revised text this is less clear.

50. put Spain in between brackets or commas: (Spain) or: , Spain,

52. Provide the age range of the MPT?

88-150. OK, the idea is to use 13C for reconstructing a paleoprecipitation record. All is featuring low precipitation but why is the analysis needed? Hominin evolution aspects are introduced before but it seems to disappear somehow here (at least in the text). Perhaps making it a little more specific in 82-86 would be helpful.

104. is 0.9-0.6 Ma the MPT in its entirety or a part of it?

158. Were the early Brunhes glaciations the strongest?

194. Oriental Realm sufficiently clear in the present context?

247-248. Earlier on emphasis is put on glacial and arid conditions, here interglacials are also stated to be windows of opportunity for hominin migration. This may require some further clarification.

258. Base level lowering is probably more relevant than sea level lowering for river segments 1000s of km upstream from their respective river mouths. The area is partially an endorheic basin with no outlet to any ocean?

259. Apparently, we now shift to NW Europe and northern China. Away from Central Asia?

308-309. Sentence complete?

339-341. What would be the underlying reason that hominins migrated mainly during humid periods in the Late Pleistocene while dry periods would be required earlier on?

350. Our analysis, the data comes from various sources.

376. When is a 0.1 proportion window (meaning?) used and when 0.2?

Figure 2. In the text mostly 0.9-0.6 Ma is being used for the MPT while in the figure 0.6-1.2 Ma is indicated. Why is that different? Why is the emphasis on the 0.9-0.6 Ma portion of the MPT?

Figure 3. What is the difference between tectonic-climate interactions (vague wording?) and uplift-controlled? The text in 388-391 is not that helpful in clarifying. Methods used by the references are mentioned but their criteria to allocate terraces to a certain category are not. Tectonic trends typically have time scales of millions of years beyond the age range analyzed here. Climate change may operate at orbital time scales.

Figures 3 and 4. It looks like that beyond 1.0-1.5 Ma the number of input data is small compared to the 0.0-1.0 Ma interval. How does that impact the analyses performed?

401-403. This should also be explained in the caption to figure 3.

406-409. This reads cryptic to me. OK, focus here is on the (younger portion of) the MPT. However, dating is anticipated to be more precise for the most recent 100 kyr? Also the number of sites probably is high in comparison to older hominin sites.

419. Which regions in Europe? It lies (mostly) upwind of the loess source area? This may need some further clarification.

425. You mean the 2x2x2cm cubes? A “block sample” typically is a much larger hand sample.

437. this = the present? (or that?) (this is not ref 70?)

459. the averaging = averages of

475. opposing the temperature controls = unclear wording, reword for clarity.

478-483. Somehow wordy text to argue for a dominant precipitation control on d13C.

484-485. for the plants under C3 vegetation = unfortunate wording, some rewording is appropriate.

485-487. We know that already, leave out/merge with foregoing?

490. take out ‘the’

498 ff. a jump in the reasoning. Some more lead-in may be required.

519-520. “without requiring a normality assumption for the underlying population” reads somehow unfortunate, is it the correct statistical terminology?

525. Why is the age window selected? Why not a shorter or longer window?

533. Really so? If there are no data the residual is undefined? With few data it is subject to a large uncertainty?

534-538. With which software package the analysis was performed?

572. take out ‘ago’

597. A kind of to be expected? Age uncertainties are probably similar, the minimum and maximum age diagrams would imply shifting the mean diagram downward and upward?

599. linguistically: can a distribution increase?

600. How is that residual determined? It is the residual of what (number of sites?) with respect to what (time?)?

602-603. Refer to the relevant figure(s) here. What is the reader supposed to see?

605-606. So, how robust are such estimates for data-poor times?

606-613. So, apparently number of sites and age uncertainty effects play up. Earlier on this was downplayed. This makes me worrying.

615-616. “broadly increasing trends in the distribution of hominin dispersals” is this wordy for increasing hominin presence?

620. you may wish to explain concisely ‘Crassulacean acid metabolism plants’

622. carbonate

625. organic carbon, carbonate or both? Be more specific?

629-630. “analysis of the astronomically tuned timescale” reads cryptic, records are tuned to that scale, the scale itself is not analyzed.

631-632. Nothing new here, at least for the CLP.

633. take out ‘the’

637. and when the ‘open forest steppe’ vegetation type did start?

648. this is textbook? For loess deposits one needs an eolian source, isn’t it?

649. take out ‘the’

651. were these regions also sub-humid (correct term?) and humid at the time of loess expansion? From the text further on it does not seem that way. So, why the emphasis on humid is required?

667-669. This pertains to glacial times! In the current interglacial in by far most of Europe there is no active loess accumulation. Loess-paleosol profiles are typically featuring ‘telescoping’, the downward migration of soil development with time.

670 ff. A jump in reasoning, suddenly a switch to river terraces is made. What is the point of this paragraph? Older river terraces have much more potential to be eroded since river terraces are something dynamic (for example meandering rivers cut away riverine deposits down-stream with time).

971-972. “Average wind data at 850 hPa is reached from ERA5 dataset” reads cumbersome. Did authors download the data and plot them with their favorite plotting package?

972. insert ‘the’ before ‘European’

Extended data Fig. 2 here the MPT is indicated from 1.2 to 0.8 Ma. Make consistent throughout text and figures.

991-1001. Apparently shifts in central tendency and dispersion are something independent since they change at different times.

994. I can’t see those indicated.

994-997. OK but what is new here? Many things have changed during the MPT.

1008-1010. Was this cluster analysis carried out by the present authors or by the studies referred to? Be explicit here.

Extended data figures 5 and 6 seem compilations of literature data. Is that sufficiently clear from the wording?

1010-1011. What variables are clustered? Some words on the premises and difference/coherence measure of that cluster analysis type would add.

1019. typo, take out ‘the’ at the end of the sentence.

Extended data figure 8. Note that the MPT (yellow shading) is differently indicated as appears from the main text.

Extended data Figures 9 and 10. These are all literature data. Refer to the sources explicitly.
What is the point of extended data figure 11? The caption is not overly helpful.

The extended data tables 1, 2 and 3, duly refer each entry. Where are these references provided in full? Do they correspond to the references in the main text? Probably only in part. There is some file named supp_refs but it is not specified in table headings.

Reviewer #2

(Remarks to the Author)

I thank the authors for addressing the majority of comments/suggestions by reviewers. Only a few matters remain outstanding.

1. In response to a question by Reviewer 1on the inferred range of palaeoprecipitation, the authors have attempted to do this:
“Using previously established linear correlations between the MAP and the δ13Corg of surface soils in Central Asia and the NW China, paleoprecipitation changes in south Tajikistan and the Tarim Basin can be estimated. These correlations suggest the precipitation in Tajikistan decreased by approximately 100-200 mm after the Mid-Pleistocene. In the Tarim Basin, precipitation decreased from ~200 mm during the late Pliocene to ~100 mm after the MPT”

However, closer inspection shows that the data set used for the transfer function applied to the Tajikistan record has been developed from surface soil samples in northern Iran, in the western part of arid central Asia (Wang et al, 2019), while that applied to the Tarim Basin record is based on a surface soil samples from a transect in central East Asia from the northern slopes of Qinling Mountain to near the border of Mongolia and Russia (Lee et al., 2006). Since the relationship between

δ13Corg and precipitation appears to vary geographically, the lack of modern datasets for Tajikistan and Tarim Basin suggest that the inferred palaeoprecipitation values need to be treated with caution. This uncertainty should be discussed more clearly in the text.

2. lines 204-213. I still disagree with the statement “the replacement of woodland and closed forest by steppe and woodland-steppe environments across most of Palearctic Eurasia during the Mid-Pleistocene”. Forest biomes occurred less frequently, because of the shift from 41-kyr to ~100-kyr glacial cycles, but interglacial forests did not turn to steppe-woodland or open forests in most places, only in the lower precipitation areas. Some of the moisture-requiring tree taxa disappeared from interglacial forests, but what mainly changed was the character of the glacial periods with some tree presence during the Early Pleistocene, but then mostly absence later on (see for example Zhao et al 2020 Science Advances; Donders et al. 2021 PNAS). The same correction is needed for the caption of Ext. Fig. 5, line 1020, the herb values (pollen percentages not pollen concentrations) increased during the glacial periods (not the interglacials).
(Incidentally, the attribution of the serial extinction of Tertiary relicts (line 211, ref 42) is incorrect, it should be Tzedakis et al. 2006 QSR.)

3. Fig. 4b,c,d: the density diagrams appear to show a low value at ~2.4 Ma near 0° longitude. Since the data refer only to Europe, is this an artefact of the analysis?

4. Extended Data Table 2:
“Loir” should be “Loire”
The Voidomatis River in Greece is in the NW not NE. Also, while I agree that both tectonic and climate influences are present in that basin, the reference provided (Hughes 2010) does not discuss that. So where was the information taken from?

5. Finally, a more general comment: the main thrust of this work is that Mid-Pleistocene enhanced aridity and landscape shifts “drove” Palaearctic hominin dispersals. This is plausible, but the confluence of the climate shift over the MPT and human dispersals does not necessarily mean a direct causal relationship (since other factors are also important). I wonder, therefore, whether this could be phrased in a more circumspect manner (e.g. may have “promoted” or “influenced”).

Provided that that the authors address/consider these comments, I am happy recommend publication.

Version 2:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The re-revision has addressed all my comments to a sufficient extent. I recommend acceptance for publication. Usage of articles may require some attention (a few times 'the' seems superfluous). AF in line 1096 stands for Alternating Field; inconsistencies in the reference list can be remedied at the proof-reading stage.

Reviewer #2

(Remarks to the Author)
The authors have addressed the points I raised and I am happy to recommend publications.

In the final MS, two more points require minor modification:

l. 218 After reference “42” please add the following reference: MAgri et al., Quaternary disappearance of tree taxa from Southern Europe: Timing and trends. Quaternary Science Reviews 163 23-55 (2017).

l. 648-650 and Extended Data Fig. 4: “Vegetation types in East Asia show a remarkable shift from forest to open forest at 1.2-1 Ma, and subsequently changed from open forest steppe to steppe at ~0.9-0.6 Ma (Extended Data Fig. 4)”

With respect to Zoige Basin, this sentence is not correct: after 0.62 Ma, interglacials were still characterized by tree population expansions (mainly Picea), but with presence of alpine meadows; glacial periods were longer and dominated by steppe and meadows. Please correct this in the text and also add some spruce trees in the Ext Fig 4 in the interval 0-0.62 Ma and even more for the interval 0.62-0.9 Ma

Reply to Reviewer: 1

The study hinges on two thoughts: 1) the new and 3.6 Myr long paleoprecipitation record (however, rainfall in numeric values like mm/yr with an uncertainty range is not specified) and 2) a comprehensive analysis of known hominin sites in Europe and Asia. However, the link between the two is not that evident: can the hominin site analysis be done without the new record? The discussion essentially focuses on the MPT and younger times (logically since not many hominin sites older than 1.6 Ma are known outside Africa), so what is the merit of the 3.6 Ma to 1.5 Ma portion of the new records (in terms of hominin dispersal routes)? Note that extended data Fig.2 shows the records starting at the Gauss/Matuyama boundary at ~2.6 Ma. The compilation of hominin sites and their environmental frame is a lot of work. The text is well written but seems to report mainly items that are largely known already. The new aspects remain somehow hidden. The change point analysis in this context is shown but its implications are only worded in very general terms while its necessary technical aspects, how robust, level of significance, etc., remain undiscussed.

Response: Thanks for the critical comments. In the revised manuscript, a rough estimation of paleoprecipitation changes in south Tajikistan and the Tarim Basin were conducted, which was based on the linear correlation between the mean annual precipitation and the $\delta^{13}C_{org}$ of surface soils in Central Asia and the NW China. Further, we have added several new paragraphs to extend the knowledge about the vegetation cover and paleoclimatic changes as well as their implications for hominin-environmental interactions between 3.6 Ma and 1.5 Ma. Combined with further syntheses of Pleistocene paleontological and carbon isotope records from the SW China and the Southeast Asia, we have provided some new discussions about the southern migration routes of early hominins across Asia in the revised manuscript. In addition, we provided more details about the change point analysis and the kernel density estimation diagrams in the text and the supplementary information. As outlined by Reviewer 2, the main contribution of our manuscript is a consideration of Europe and Asia together, something that is rarely explicitly considered in studies exploring drivers of hominin dispersals, and using both new environmental data and a large compilation of archaeological and hominin records to determine how these may be related.

1. Abstract. How high is high resolution? Can authors be more specific?

Response: We have provided more details about the sample resolution in the abstract.

2. 32-33. "intensive tectonic uplift" reads vague. How much uplift is anticipated/recorded over 0.5-1.0 Myr? Not that much? A fair portion of the Tibetan Plateau was already at significant elevation at ~40 Ma or even before?

Response: Thanks for the critical comments. We agree that the influences of tectonic uplift on the Mid-Pleistocene aridity and landscape shifts are overestimated. The discussions about the tectonic forcing have been abandoned in the revised manuscript.

3. 41-43. The Palearctic realm is Europe and most of Asia. The meaning of the red dotted line in figure 1 is the Arctic Sub-Realm. However, that is not mentioned in the main text. What is the point of showing it? All sites discussed are (far) to the south of it. In Asia, they seem to be largely confined to the Sino-Japanese realm and the Saharo-Arabian realm.

Response: We have removed the red dotted line in Figure 1.

4. 49. typo, take out "ago" after "Ma"

Response: Done

5. 53. Define MPT again in the main text?

Response: Done

6. 56. Realm, is capitalization required?

Response: Thanks, we have corrected it.

7. 62. I wonder about the tectonic factors at about 1 Ma: most of the Eurasian continent had already achieved more or less their present-day elevations.

Response: Thanks for the critical comments. We agree that the influences of tectonic uplift on the Mid-Pleistocene aridity and landscape shifts are overestimated. The discussions about the tectonic forcing have been abandoned in the revised manuscript.

8. 67. Carbonate-carbon or organic-carbon isotope data? Or both? OK, next sentence organic carbon isotope analysis is done, what about carbonate-carbon isotopes?

Response: Given that the paleosol layers in the loess deposits of the Tarim Basin, NW China, are very weakly developed due to low precipitation, and pedogenic carbonate is difficult to distinguish from the loess layers. Thus, the carbonate-carbon isotopes were not analysed in the present study.

9. 77-81. This is the "northern migration route" through Asia, what about the "southern route" through India?

Response: Thanks for the good suggestions. We have provided more information about this question in lines 82-86:

"Our novel carbon isotope results, combined with syntheses of Pleistocene paleontological and carbon isotope records from the SW China and the Southeast Asia, provide holistic insights into the regional vegetation cover and paleoclimatic conditions for the northern and southern migration routes of early hominins through Asia."

10. 85. Is the entire duration from 3.6 Ma (respectively 2.7 Ma) to Recent sampled? Be explicit here.

Response: Yes, we have revised the manuscript accordingly in lines 96-99:

"A total of 286 and 244 loess samples, spanning 3.6 and 2.7 Ma to the present with a temporal resolution of ~5-10 kyr and ~10-20 kyr, respectively, were collected from Tajikistan and the Tarim Basins."

11. 86. What is the typical sampling distance? What is the loess/paleosol accumulation rate? And what about the presence of paleosols? There is climatic contrast between them. What if all samples come from loess and none from paleosols? Then bias could creep in.

Response: Thanks for the critical comments. In the revised manuscript, we added a new paragraph in lines 89-100 to address these issues:

"Eolian dust deposition and pedogenesis in Central Asia and NW China are continuously competing processes, with dust accumulation rates usually showing much higher values in loess units than those in paleosol layers. In southern Tajikistan, paleosols are well developed due to their location on the windward slope of the Westerlies-influenced areas, and lithological characteristics of the loess deposits resemble those of contemporaneous loess-paleosol units in the CLP. In contrast, paleosols in the Tarim Basin are very weakly developed due to low precipitation, and the carbonate nodular (Bk) horizons are difficult to define in most levels. A total of 286 and 244 loess samples, spanning 3.6 and 2.7 Ma to the present with a temporal resolution of ~5-10 kyr and ~10-20 kyr, respectively, were collected from Tajikistan

and the Tarim Basins. Samples for the carbon isotope analysis of organic matter were collected from all loess and paleosol units (Extended Data Fig. 2)."

12. 94. At other locations the MPT is defined between 1.2 and 0.6 Ma. Something which takes place over 300 kyr is not abrupt.

Response: Thank you for pointing this out. We have reorganized the text accordingly and change the word "abrupt" to "pronounced".

13. 95. But this shift in orbital expression cannot be resolved in the carbon isotope data given their resolution.

Response: Yes, the shift in orbital expression cannot be resolved in the carbon isotope data. Nevertheless, we provide the grain size evidence of loess deposits in the Tarim Basin and Tajikistan to show the orbital shift after 0.9-0.6 Ma in lines 159-162.

14. 105-110. Again, the expression of paleosols in the record may deserve some attention (some information is provided in 366-371, at one sampling location (Tajikistan) paleosols are described as well developed; in the Tarim basin they are nearly absent. Loess is typically associated with drier conditions. What is the inferred range of paleoprecipitation?

Response: Thanks for the critical comments. In the revised manuscript, we have added two new paragraphs in the main text and the "Methods" to show the inferred range of paleoprecipitation.

Lines 136-144:

"Using previously established linear correlations between the MAP and the $\delta^{13}C_{org}$ of surface soils in Central Asia and the NW China, paleoprecipitation changes in south Tajikistan and the Tarim Basin can be estimated. These correlations suggest the precipitation in Tajikistan decreased by approximately 100-200 mm after the Mid-Pleistocene. In the Tarim Basin, precipitation decreased from ~200 mm during the late Pliocene to ~100 mm after the MPT. Generally, the absence of paleosols in Chinese eolian deposits indicates that the MAP is less than 200-300 mm. These findings are consistent with lithological observations that paleosols in the Tarim Basin are very weakly developed since the late Pliocene (see Methods for further details)."

And in lines 488-497:

"Previous studies demonstrate that a linear correlation exists between the mean annual precipitation and the $\delta^{13}C_{org}$ of surface soils in Central Asia (Precipitation = (δ^{13}Corg+22.84)/(-0.0079); n=44, R^2=0.44) and the NW China/Mongolia (Precipitation = $-58 \times \delta^{13}$Corg $-$ 1266.5; n=19, R^2=0.9). Following these observations, a rough estimation of paleoprecipitation changes in south Tajikistan and the Tarim Basin was conducted based on the temporal variations in $\delta^{13}C_{org}$ values. The results demonstrate that the precipitation in Tajikistan was >500 mm during the early Pleistocene and decreased to ~400-300 mm after the Mid-Pleistocene. In the Tarim Basin, the precipitation decreased from ~200 mm during the late Pliocene to ~100 mm after the MPT."

15. 116-121. Nice to know but what is the inferred paleoprecipitation range?

Response: Please see the replies to the question 14.

16. 123-202. Most of the discussion relates to the MPT and younger times. What is the point of providing data from 3.6 Ma (or 2.7 Ma) to Recent? From 1.5 Ma to Recent would suffice? It is a good summary of existing knowledge but at the same time rather long. That central Asian climate became increasingly arid since the Pliocene/Pleistocene boundary with a stepwise increase in aridity during the MPT is well known. What is the new record adding to this knowledge base? This remains unclear.

Response: Thanks for the critical comments. While we agree that there has been a general appreciation of the increased aridity of the central Asian climate since the Pliocene, this knowledge has been either broadbrush in approach or rarely focused on our area of study. Our results provide a significant new and detailed record of the dynamics of this increase in Tajikistan and the Tarim Basin. In providing a longer (Pliocene) record, we can highlight just how unique the changes centred around the MPT are for the Plio-Pleistocene. In the revised manuscript, we have therefore added a new paragraph to extend the knowledge about the vegetation cover and paleoclimatic changes during the older times.

In lines 176-187: "Tajikistan sits on a crucial migration pathway of hominins dispersing through Eurasia. The Palaeolithic sites in the loess-palaeosol sections of southern

Tajikistan provide the most spectacular picture of the loessic Palaeolithic across Eurasia. To date, little information is known about the regional vegetation cover and paleoclimatic changes prior to the Mid-Pleistocene in this area. Our results demonstrate that although vegetation oscillations existed during glacial-interglacial cycles in Tajikistan, the vegetation community was relatively stable during the late Pliocene and early Pleistocene. A prominent shift in vegetation cover and vegetation communities is recorded only around the MPT. Similar patterns of vegetation changes were also observed in the Tarim Basin (Fig. 2b-c). Collectively, these observations demonstrate that a marked ecological shift towards more open habitats is coeval with the MPT along the northern migration route of early hominins through Eurasia."

17. 139-146. So, the present-day conditions in Tarim Basin prevailed since 0.6 Ma?
Response: Yes, the $\delta^{13}C_{org}$ records of the Lop Nor lake and the loess deposits (this study) and eolian dust flux records from the Tarim Basin all demonstrate that the present-day conditions in Tarim Basin prevailed since the MPT, and a permanent presence of the Taklimakan Desert has been identified since 0.7-0.6 Ma (Fang et al., 2020; Liu et al., 2020).

18. 147. Here the MPT is implicitly placed between 1.2 and 0.6 Ma.
Response: In fact, a prominent increase in $\delta^{13}C_{org}$ and $\delta^{13}C_{carb}$ records from the monsoon-controlled areas in the Chinese Loess Plateau and the adjacent regions also mainly occurred after 0.9-0.6 Ma, which is consistent with the results from the Tarim Basin and southern Tajikistan. However, we note that the $\delta^{13}C_{carb}$ records in the Qaidam Basin showed a slight increase at 1.2 Ma, and thereafter increased significantly at ~0.6 Ma. Thus, here the MPT is placed between 1.2 and 0.6 Ma. Nevertheless, these results are not contradictory to our main conclusions.

19. 171-173. This pertains to the northern dispersal route in Asia. Up to now no notion of the southern route via India. The northern route relates to the southern boundary of the Palearctic realm. Hominins went further north since the last 0.4-0.5 Ma or so?
Response: Thanks for the critical comments. Combined with further syntheses of Pleistocene paleontological and carbon isotope records from the SW China and the

Southeast Asia, we have provided some new discussions about the southern migration routes of early hominins across Asia in the revised manuscript. In addition, our syntheses of Pleistocene archaeological records suggest that hominins went further north since the last 0.6-0.5 Ma.

Lines 193-199: "In addition, $\delta^{13}C$ of fossil tooth enamel suggests that the spread of grazers and extinction of browsers occurred synchronously in the Oriental Realm since the Mid-Pleistocene due to the ecosystem conversion from forests to savannahs, providing support for reduced precipitation in the Indian summer monsoon-controlled areas. This observation suggests that a marked ecological shift towards more open habitats due to decreasing precipitation also occurred along the southern migration route of early hominins in Asia since the Mid-Pleistocene."

20. 178. Check this tectonic uplift inference. Since 2001 many studies (more than 100?) have returned to the question of when, which parts of, and how much the Tibetan Plateau was uplifted. Current mainstream thinking is that most of the plateau reached more or less its current elevation at the Pliocene/Pleistocene boundary or even before.

Response: Thanks for the critical comments. We agree that the influences of tectonic uplift on the Mid-Pleistocene aridity and landscape shifts are overestimated. The discussions about the tectonic forcing have been abandoned in the revised manuscript.

21. 207, 208. Provide the ages for MIS11 and MIS13.

Response: Done.

22. 213. A lower sea level is not "felt" by a river beyond its coastal plain. In the upstream region of interest here the direct impact of a lower sea level on rivers is marginal at best.

Response: Thank you for pointing this out. We reorganized the text accordingly in lines 257-258: "intense river incision along the river margins in the upstream regions occurred in response to lowering sea-levels and base level"

23. 215. Figure 3. What is the basis for allocating a terrace sequence to climate-controlled or tectonic-climate-controlled? How are these two distinguished?

Response: All these climate-controlled or tectonic-climate-controlled terraces are from the previously published papers. The authors distinguished these two terrace sequences

through examining the stratigraphy (layering) of the terraces, studying the regional geology and tectonic history and using radiometric dating techniques to determine the ages of the terraces, thereby correlating them with known climatic and tectonic events. We have provided more discussion about this question in lines 388-391.

24. 227. Serbia is more of southeastern Europe; northern Europe is Scandinavia?
Response: Yes, Serbia is more of southeastern Europe and UK is more of western Europe. We corrected them accordingly

25. 231. And in Atapuerca? There hominin occupation is argued to be continuous since ~1.3 Ma?
Response: Yes, the earliest secure dates from Atapuerca and the Orce Basin Spain are ca. 1.4 – 1.2 Ma (Bermúdez de Castro et al., 2011; Oms et al., 2011). However, Muttoni et al. (2018) argued that although Europe may have become initially populated during the early Pleistocene, the stable occupation of hominins in Europe was possibly only since ~0.9 Ma.

26. 276. A distinct jump in reasoning? What are "our observations"?
Response: We have reorganized the text accordingly in lines 328-331:
"Our results indicate that enhanced aridity and pronounced landscape shifts drove Palearctic hominin dispersals during the Mid-Pleistocene, suggesting that the environmental drivers of early hominin dispersals in Eurasia were quite different from those in Africa (deMenocal, 1995; Levin, 2015; Potts et al., 2020)."

27. 279-284. Can this be rhymed with younger migration waves as analyzed in Ao et al. (2024) Proceedings of the National Academy of Sciences?
Response: We have provided more discussion about this question in lines 339-341. Our results demonstrate the dominant influences of a prominent shift in vegetation and landscape on Palearctic hominin dispersals at 0.9-0.6 Ma. Nevertheless, it is reasonable to expect that hominin dispersals in Eurasia were likely facilitated by humid periods when arid climate and steppe/grassland prevailed during the late Pleistocene (Ao et al., 2024).

28. 292. How is the "ready availability of fresh water" to be placed in an environment that features increased aridity?

Response: Here we focused on the critical role of increasing drainage density in the availability of fresh water. We have reorganized the text accordingly in lines 347-348.

29. 296. In the very southern margin of the Palearctic realm.
Figure 1. The light blue on greenish background does not read well, pick another color combination.
Response: Done.

30. Figure 3 and 4. The methodology behind the analyses deserves more mention in the main text since this represents a new twist to existing knowledge. How robust are the inferences made?
Response: We have done so as suggested in lines 401-403 and lines 516-546.

31. 358. Why in Europe? That lies up-wind of the Tarim basin loess source areas?

Response: Yes, it is because the Europe lies up-wind of the loess source areas in Tajikistan and the Tarim Basin.

32. 364-366. So, authors rely on the age model provided by reference 14. How robust is that age model? It is robust, authors should provide some more information concerning sampling resolution and paleomagnetic stability here (is actually done in 374-378).

Response: We have provided more information about this question in lines 425-428:

"Paleomagnetic block samples were taken at 50-cm intervals in the upper 80 m of the Tarim Basin loess core and at 25-cm intervals in the lower 591 m. A total of 2,331 samples gave reliable characteristic remanent magnetizations to determine the magnetostratigraphy."

And in lines 431-446:

"Based on chronological constraints of the magnetostratigraphic and OSL analyses, an astronomically tuned age model was established for the 671 m-deep loess core in the Tarim Basin with a temporal resolution of ~0.5 kyr, which has been achieved through tuning the grain size record to theoretical variations in the Earth's orbital parameters (Fang et al., 2020). In addition, an astronomically tuned age model of the Tajikistan loess has been obtained previously through tuning the magnetic susceptibility record to

theoretical variations in the Earth's orbital parameters (Ding et al., 2002). In this study, the timescale of the Tajikistan loess profile, with a temporal resolution of ~5-10 kyr, was established based on magnetostratigraphic constraints and the correlation of the magnetic susceptibility record with the marine oxygen isotope results (Extended Data Fig. 2). Previous studies confirmed that the grain size records of the Tarim Basin loess and the magnetic susceptibility records of the Tajikistan loess both show a dominant periodicity shift from 41-kyr cyclicity to 100-kyr cyclicity between 1.2 Ma and 0.6 Ma (Ding et al., 2002; Fang et al., 2020). These features are consistent with global ice volume variability during the MPT (Extended Data Fig. 2), highlighting the reliability of the chronological framework."

33. 372-374. Superfluous, leave out?
Response: We have done so as suggested.

34. 384. Is a winter monsoon/westerlies proxy, specify that?
Response: We have done so as suggested. That is, the grain size records represent the intensity of westerlies.

35. 385. Is a summer monsoon proxy, specify that?
Response: We have done so as suggested. That is, the magnetic susceptibility records reflect the precipitation variability.

36. 392. Strict adherence to GTS2020 nomenclature would imply kyr here.
Response: We have done so as suggested.

37. 393-394. But are they from loess, paleosol, or both per couplet?
Response: The samples for the carbon isotope analysis of organic matter were collected from all the loess and paleosol units. We have reorganized the text accordingly.

38. 397. At which temperature? Room temperature?
Response: Yes, we have provided this information in the text.

39. 404-406. Also mentioned in the main text, however without much context over there.
Response: We have deleted these sentences in the revised text.

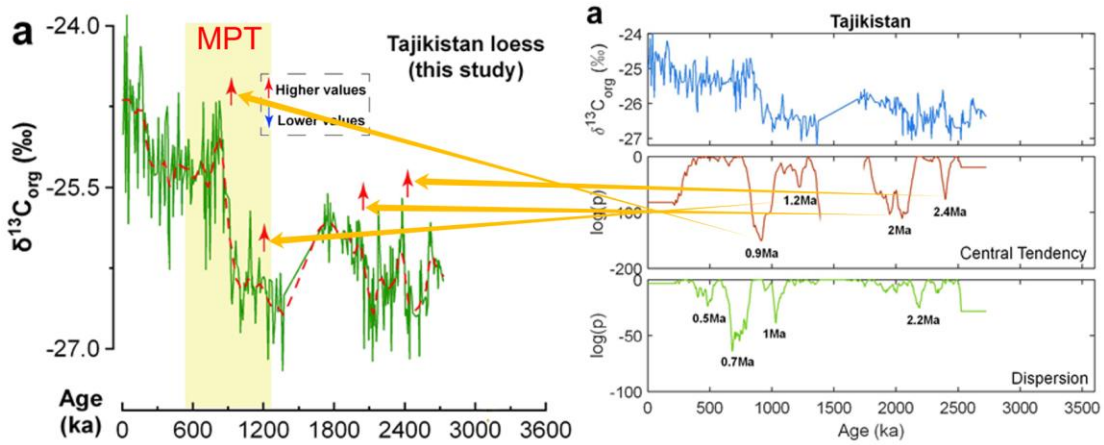
40. 421. Typos, reconstructions; appear
Response: We have corrected it.

41. 425. Numbers become less negative, is that a reduction?
Response: We have reorganized the text accordingly.

42. Change point analysis section. It may be wise to word the concept of this analysis and these tests with reference to a textbook. Extended data Figure 3 (and elsewhere). What is the meaning of the deviation from the zero line (always negative) in the red lines (brown-ish in my print) and green lines? What is the meaning of log(p)? What is implied with Central Tendency and Dispersion?
Response: We have provided more discussion about these questions in lines 517-522:
"The change point analysis of Mann–Whitney and Ansari–Bradley tests can provide crucial constraints on the rhythms, events and trends in Plio-Pleistocene climate changes of Asia (Foerster et al., 2022; Trauth et al., 2009). These statistical approaches, without requiring a normality assumption for the underlying population, are non-parametric hypothesis tests (Lepage, 1971). In this study, we use the *ranksum* function embedded in Matlab 2016b to perform the Mann-Whitney and Ansari-Bradley tests (Trauth, 2020)."

And in the captions of Extended Data Fig. 3, 9-10, 12:
"The log(p) denotes the logarithm of the p value of the Mann-Whitney and Ansari-Bradley tests at the default 5% significance level. The downward prominent peaks indicate the position where the central tendency and dispersion change significantly determined by the Mann-Whitney and Ansari-Bradley tests, respectively."

At the same time, we can combine visual and statistical results to determine whether the central tendency and dispersion are increasing or decreasing. For example:



Left picture: paleoclimate time series. **Right picture**: The Mann-Whitney (middle) and Ansari-Bradley test (bottom) of time series.

43. 459. Asia is probably better here? Not a single record shown in extended data Figure 3 comes from Europe.

Response: We have done so as suggested.

44. 461. Reference for MatLab2016b?

Response: Done.

45. 466-473. These bits are not that helpful for a general readership. What is the outcome of this analysis? Where are the corresponding figures?

Response: This analysis can avoid the influence of sampling bias/burial loss of older archaeological records on the temporal distributions of fossil hominins. We have provided the corresponding figure in the revised text (Fig. 4h-j).

46. 477. Dating values: Is this the number of sites compiled here?

Response: Not exactly, it represents the age number of Pleistocene archaeological and fossil hominin records. For some archaeological sites, fossil hominin records can be found over varying time periods and at different ages.

47. 479-480. What is inferred with this sentence?

Response: We have deleted this sentence in the revised manuscript.

48. 482. No new hominin finds are reported here, what is implied with 'new discoveries'? How many new entries are included in comparison to previous compilations?

Response: Thank you for pointing this out. In the revised manuscript, we have deleted

the words “new discoveries” and reorganized the text accordingly.

49. 483. Is all duly referred to, how good are the updates? No new age assignments were obtained in the present study. How many sites/data entries were updated?

Response: Thank you for pointing this out. In the revised manuscript, we have reorganized the text accordingly.

50. 488. How many? Which key sites? Be specific. I presume the examples provided are not exhaustive.

Response: We have provided more discussions about these questions in the section of “Synthesis of archaeological and fossil hominin sites”

51. 503-506. Start this section with these notions. They could also be mentioned in the main text.

Response: We have done so as suggested. In addition, these notions were also mentioned in the captions of Figure 4.

52. 509-510. You mean that the methods are suited for the age range considered here? As worded it comes across as not overly informative.

Response: We have deleted this sentence in the revised manuscript.

53. 510. Why ‘residual’? Residual of what?

Response: To avoid the influence of sampling bias/burial loss of older archaeological and fossil hominin sites, we perform change point analysis on the residual value of the site number regressed against age. These residual values should be, in a sense, “free” of the tendency for an increase in the preservation and documentation of sites closer in time to the present (Maxwell et al., 2018). We have provided more information about this question in the section of methods and the main text accordingly.

54. 510-512. A histogram of the apparently more or less even distribution through time would be insightful.

Response: We have done so as suggested and a histogram was added in Figure 4.

55. 512. And distribution through time was more or less even (at least it reads like that to me)? What is the cause of the inflection points?

Response: In Fig. 4 and Extended Data Fig. 11, we supplement the histograms of *Homo*

fossils, handaxes and core-flakes site numbers over time. The raw histogram data show clearly that the site number of *Homo* fossils, handaxes and core-flakes increased after the Middle Pleistocene. Statistically, the Mann-Whitney test (red arrows) also show that the number of archaeological and *Homo* fossils sites changed significantly since 0.9-0.6 Ma. However, to avoid the influence of sampling bias/burial loss of older archaeological and fossil hominin sites, we further performed change point analysis on the residual value of the site number regressed against age (fig. 4h-j). These residual values should be, in a sense, "free" of the tendency for an increase in the preservation and documentation of sites closer in time to the present, which may account for the increased trends observed in the histograms. In Extended Data Fig. 12, the inflection points show downward peaks in both central tendency and dispersion of the data at 0.9-0.6 Ma. As a correlated change in the central tendency and dispersion of a dataset is related to changing the scale of the data, these results suggest that examining the residual values of our dataset against time has scaled but not shifted the significant relationship between variables, thus supporting our inferences based on the raw counts.

56. 513-517. What is the purpose of the notions made here?
Response: Here, these notions are aimed to exclude the sampling bias/burial loss of older archaeological and fossil hominin sites on the spatial and temporal distributions of early hominins in Europe and Asia.

57. 521. How is this done and how is somebody supposed to read the extended data figures 11 and 12?
Response: We have provided more details about the change-point analysis and the kernel density estimation diagrams in lines 517-547 and lines 592-616.

58. 530. What are CAM plants? Explain the acronym.
Response: Done.

59. 554-603. How climatically controlled, tectonic-climate, and uplift controlled river terraces are being distinguished and how robust that distinction would be is not discussed.
Response: All these climate-controlled or tectonic-climate-controlled terraces are from the previously published papers. The authors distinguished these two terrace sequences

through examining the stratigraphy (layering) of the terraces, studying the regional geology and tectonic history and using radiometric dating techniques to determine the ages of the terraces, thereby correlating them with known climatic and tectonic events. We have provided more discussion about this question in lines 388-391.

60. 564-567. The relation to the present study may become rather thin. What information would be lost if this was omitted?

Response: We use this evidence to demonstrate that increasing dust fluxes of marine sediments from the Mediterranean Sea, the Arabian Sea, and the North Pacific Ocean provide further supports for the occurrence of continental aridification and expansion of the Eurasian loess belt during the MPT, which led to a pronounced increase in eolian sediment supply from the terrestrial dust sources to the downwind oceans. We have revised these sentences accordingly.

61. 566. Presumably dust fluxes are to those seas and oceans, since they cannot deliver dust.

Response: Yes. We revised the text accordingly.

62. 579. dating values = ages?

Response: Yes. We change the words "dating values" to "ages" in the revised manuscript.

63. 590-593. All correct but what is the relevance since hominin dispersal is being discussed between 0.1 and ~1.3 Ma?

Response: We deleted these sentences in the revised manuscript.

64. Extended data Figure 1. Are these figures obtained in this study? If so, provide some technical details in the caption. Which model was used, which years were integrated, etc. If not, properly refer to the source(s).

Response: Yes, these figures are obtained in this study. We have provided more information about the technical details in the captions.

65. Extended data Figure 3. Explain log(p) in the caption. Mean and variance are being mentioned in the caption while central tendency and dispersion are being displayed on the figure panels. What is what?

Response: We have provided more information about these questions in the revised

manuscript. The log(p) denotes the logarithm of p value of the Mann-Whitney and Ansari-Bradley tests. When log(p) is less than -1.3 (corresponding to the p value < 0.05), the difference is clear and the changes have statistical significance. In statistics, the central tendency of time series is usually expressed as mean and median mean, median, mode; the dispersion of time series is commonly expressed by range, variance, quantiles (Trauth, 2020).

66. 878. Are the changes observed significant? At which level?

Response: The embedded *ranksum* function in Matlab 2016b were used for Mann-Whitney and Ansari-Bradley tests. The default significance level (α) of *ranksum* function is 0.05. When log(p) is less than -1.3 (corresponding to the p value < 0.05), the difference is clear and the changes have statistical significance.

67. Extended data figure 4. How robust are those cluster analysis results? From the text it seems that authors have carried out that cluster analysis. Which input parameters were used and how was the stratigraphic control achieved?

Response: Yes. The cluster analyses have been carried out by the published papers. We have provided more details about these questions in the figure captions.

68. Extended data figure 6. The age distinction seems rather distinct. How robust are the ages involved in this analysis?

Response: We have provided more information about the ages in the caption involved in this analysis: The ages of large mammal assemblage were constrained by paleomagnetic analyses and the correlations with typical European Mammal Zonation during the late Neogene and Quaternary (MNQ).

69. Extended data figure 8l. The temperature anomaly is with respect to what?

Response: Temperature anomaly indicates the temperature deviation (in °C) from present (average over 0–5 ka). Such information has been provided in the revised manuscript.

70. 917. explain acronym IRM@AF.

Response: Done. It indicates an isothermal remanent magnetization (IRM) that had

been AF (Alternating demagnetization)-demagnetized at 120 mT, representing the eolian input of hematite.

Reviewer #2 (Remarks to the Author):

The rationale for this study is provided in the second paragraph of the main text: "The climatic and environmental histories of Europe and Asia are usually discussed separately. However, they are parts of the same biological Realm…As phylogenetic compositions are driven by regional climatic and ecologic factors 11, the histories of both regions need to be considered together when examining the drivers of hominin dispersals out of Africa".

This is an important point, but also a substantial undertaking. In this work, the authors start by presenting new δ13C data from loess deposits in Central Asia and northwest China over the last 3.6 Myr. The evidence points to a significant climatic shift during the so-called Middle Pleistocene Transition (MPT; 1.2-0.7 Ma). This is then expanded to a synthesis of palaeoenvironmental and palaeoanthropological/archaeological data to establish the transformation of landscapes across the Eurasian Palaearctic during the MPT, with implications for early human evolution and dispersal.

This is an impressive collection of new and published information on a grand scale that is worthy of publication.

Response: We greatly appreciate your very positive responses and insightful comments to our manuscript. We are very happy to revise our manuscript in light of these helpful and insightful comments.

1. While the MPT 'sensu lato' is considered in general terms, the authors tend to focus more specifically on climate changes commencing at 0.9 Ma. Indeed, the so-called '0.9 Ma event' is traditionally taken to represent the first major Northern Hemisphere glaciation of MIS 22 (e.g. Clark et al.., 2006, QSR). However, closer inspection of the Tarim Basin loess δ13C record shows an earlier shift (~1.1 Ma) and this is also observed in certain other records in Fig. 2 and Ext. Data Fig. 8. Margari et al. (2023, Science) presented evidence for an extensive glaciation during the 30kyr-long MIS 34 (possibly associated with the Menapian cold stage of NW Europe), culminating in an extreme terminal stadial in the eastern North Atlantic. Climate envelope model simulations suggest that the intensity of the declines in temperature and net primary productivity would have led to the extirpation of early human populations in Europe and SW Asia. A critical review of archaeological record further suggested a paucity of European and SW Asian early hominin sites after that, possibly until MIS 21 (Margari et al., 2023). This may be worth considering in the present work.

Response: Thanks for the critical comments. We have added these new findings to our

manuscript in lines 275-282:

"Paleoenvironmental reconstruction and climate envelope model simulations demonstrate that an extensive glaciation during MIS 34 (~1.15 to ~1.12 Ma) may have resulted in the depopulation of early human populations in Europe and SW Asia due to the declines in temperature and net primary productivity, and that reoccupation of Europe may have been delayed until during MIS 21 (~0.81 to ~0.87 Ma) (Margari et al., 2023). The repopulation of Europe after 0.9 Ma may have been made only by a more resilient species (e.g. *H. antecessor*) that was able to withstand the increased intensity of glacial conditions (Margari et al., 2023)."

2. The synthesis part of the paper represents a large collection of information, but space restrictions limit a fuller presentation and discussion of the evidence. For example, the river terraces dataset of Ext Data Table 2 and Fig. 3 are divided into tectonic-climate interactions, climatically-controlled and uplift-control, but their attribution and significance are not explained in the text or in the methods.

Response: Thanks for the critical comments. All these climate-controlled or tectonic-climate-controlled terraces are from the previously published papers. The authors distinguished these two terrace sequences through examining the stratigraphy (layering) of the terraces, studying the regional geology and tectonic history and using radiometric dating techniques to determine the ages of the terraces, thereby correlating them with known climatic and tectonic events. We have provided more discussion about this question in lines 388-391.

3. In the same vein, the archaeological and human fossil evidence of Fig. 4 merit a fuller discussion of the granularity contained therein. More generally, the broad spatio-temporal canvases of Fig. 3 and 4 require a more expanded description.

Response: We have provided a more detailed discussion of the implications of these results in the following lines 600-616:

"In contrast, the residual value of the site number of fossil hominins and artefacts in the current dataset does not show a significant increase through time during the Early and Middle Pleistocene. However, marked inflection points in these distributions are consistently observed around the MPT. This commonality suggests they aren't artefacts of any biases present in our dataset but rather represent genuine changes in the

distribution of hominin, handaxe, and archaeological sites. However, these data also suggest that these changes may not have occurred synchronously. Determining the relative timings of increases in archaeological sites and/or handaxe occurrences versus records of hominin remains is a geological question that will require much finer-scale considerations of how taphonomic factors control preservation than we currently have available. As such, our results echo the cautionary note of researchers working on the African hominin records regarding the need for better confidence intervals on the stratigraphic ranges of hominins and their proxies, and stronger efforts to identify, discuss, and correct fossil sampling biases. Nevertheless, the current dataset represents the largest compilation of archaeological records in Europe and Asia that suggests, despite taphonomic and sampling limitations, broadly increasing trends in the distribution of hominin dispersals in Eurasia since the MPT (see Discussion)."

4. 110. "the stable pCO2 levels…associated with the intensification of NH glaciation": While this may be in line with reconstructions based on carbonated precipitated in palaeosols (Da et al., 2019), it is not in line δ11B reconstructions showing drops in pCO2; Martinez-Boti et al., 2015). This may have to be considered in the δ13C changes of the loess records.

Response: We have reorganized the text accordingly in lines 126-132:

"Planktic foraminiferal $\delta^{11}B$ ratios demonstrate that while a significant decrease in atmospheric CO_2 occurred mainly after 2.8 and 2 Ma, pCO_2 levels were relatively stable or increased slightly since the MPT. Meanwhile, a colder climate associated with the intensification of Northern Hemisphere glaciations since the MPT favors the growth of C_3 over C_4 plants, thereby leading to a significant reduction in $\delta^{13}C_{org}$. These observations suggest that changes in temperature and pCO_2 levels cannot account for the Mid-Pleistocene increase of $\delta^{13}C_{org}$ records in Central Asia and NW China."

5. 132. "a pronounced shift in dominant periodicity". This is based on earlier work, but have the authors attempted to explore the periodicities embedded in their new δ13C records?

Response: The shift in orbital expression cannot be examined in the carbon isotope data given their resolution. Nevertheless, we provide the grain size evidence (representing

the intensity of westerlies) of loess deposits in the Tarim Basin and Tajikistan to show the orbital shift after 0.9-0.6 Ma in lines 159-162.

6. 166-170. "shift in vegetation from closed forests or woodland to steppe and grasslands is evident at ~0.9-0.6 Ma": this is not correct. While European pollen records show a serial extinction of moisture-requiring taxa (the so-called Tertiary relicts) over the MPT, interglacial stages were still characterized by forests (though less diverse). In the same vein, the inferences based on mammal evidence of a shift from closed forests to open woodland in Europe (Ext Data Fig. 6) are also tenuous, as they may represent local contexts (e.g. flood plains) and not supported by other lines of evidence (beetles, molluscs).

Response: Thanks for the critical comments. We have revised the manuscript accordingly in lines 209-213:

"Pollen records from Spain, France, Greece, and Italy suggest that an ecological shift towards more open habitats, together with a serial extinction of moisture-requiring taxa (the so-called Tertiary relicts), is evident since ~0.9-0.6 Ma (Extended Data Fig. 5). Although the principal vegetation biomes in Europe during interglacial stages were still characterized by forests after the MPT, vegetation nevertheless became less diverse."

For the Extended Data Fig. 6, we provide an additional note that local contexts (e.g. flood plains) may lead to some uncertainties on the vegetation reconstruction.

7. 230-231. "stable occupation of hominins in Europe was possibly since ~0.9 Ma". It is worth adding here that Margari et al. (2023) suggested that repopulation of Europe after 0.9 Ma may have been by a more resilient species (e.g. H. antecessor) that was able to withstand the increasing intensity of glacial conditions.

Response: Thanks for the critical comments. We have added this finding to the revised manuscript.

Reviewer #1:

We are extremely grateful that you could take the time to conduct the second review of our paper. Based on these insightful comments and suggestions, we revise the manuscript carefully. In the revised manuscript, we abandon the river terrace classification and put all the terrace data together. We also provide more details about the statistical analysis software in the methods. In addition, we have carefully proofread the manuscript and corrected all grammatical errors as suggested by you.

1. The new version rightfully downplays the impact of tectonics. However, the river terraces are still categorized in the same three groups as before. The basis for their allocation is still pretty unclear (it seems to be taken at face value from the respective literature sources). Authors mention some generic geological approaches, but which criteria were used to allocate a river terrace to a certain group is not mentioned. What is the point of making the allocation in the first place since the tectonic impact is argued to be not that relevant?

Response: We thank the reviewer for pointing out this issue. We agree that there is no need to categorize the river terraces into three groups as before. We abandon the river terrace classification in the revised manuscript and put all the terrace data together.

2. To what extent the present analysis focusing on the MPT adds to the review by Ao et al. (2020) Earth Science Reviews? OK, data from Europe are not considered in that paper but climate change is dealt with in quite some detail. The final sentence of their abstract: "Combined with environmental theories of human evolution, this chronological coincidence suggests a potential link between mid-Pleistocene environmental changes and human evolution." A lot seems to have happened during the 400-500 ka age interval, i.e. formally after the MPT range discussed here?

Response: In fact, the review paper by Ao et al. (2020) was focused on the precipitation reconstructions in East Asia (mainly in the Chinese Loess Plateau) over the past 0.8 Ma. The paleoclimatic records are mostly derived from the rock magnetic results. In our

study, we report high-resolution (up to ~5-10 kyr sample interval) carbon isotope time series of loess deposits in Central Asia and northwest China, providing unique paleoecological and paleoclimatic records spanning the past 3.6 Ma. These isotopic data, combined with novel syntheses of Pleistocene paleontological and archaeological records and spatio-temporal distributions of Eurasian eolian deposits and river terraces, provide a comprehensive picture about the critical role of aridity and landscape shifts in promoting Palearctic hominin dispersals.

3. The reason for keeping the long paleoclimate records as they are, is explained in the rebuttal letter (providing a kind of base line to better appreciate the changes in the MPT interval); in the revised text this is less clear.

Response: In the revised manuscript, the reason for keeping the long paleoclimate records was explained in the introduction in lines 80-85:

"Thus, $\delta^{13}C_{org}$ records of these loess deposits can provide critical information on precipitation and precipitation-controlled vegetation density as well as their impacts on early hominin dispersals in the Westerlies-influenced regions since the late Pliocence. Moreover, in providing these long paleoclimate records, we can highlight just how unique the changes centred around the MPT are for the Plio-Pleistocene."

4. 88-150. OK, the idea is to use 13C for reconstructing a paleoprecipitation record. All is featuring low precipitation but why is the analysis needed? Hominin evolution aspects are introduced before but it seems to disappear somehow here (at least in the text). Perhaps making it a little more specific in 82-86 would be helpful.

Response: We have provided more information about this question in lines 80-83:

"Thus, $\delta^{13}C_{org}$ records of these loess deposits can provide critical information on precipitation and precipitation-controlled vegetation density as well as their impacts on early hominin dispersals in the Westerlies-influenced regions since the late Pliocence."

5. 104. is 0.9-0.6 Ma the MPT in its entirety or a part of it?

Response: Yes, the MPT covers the time interval between 1.2-0.7 Ma. Here, we

emphasized that the pronounced shifts towards more positive δ^{13} Corg values primarily occurred at the end of the MPT, i.e. ~ 0.9-0.6 Ma. We reorganized the text accordingly.

6. 158. Were the early Brunhes glaciations the strongest?

Response: In fact, global ice volume began to increase significantly between 0.9 and 0.6 Ma, and thereafter a further increase occurred during the early Brunhes. We reorganized the text accordingly.

7. 194. Oriental Realm sufficiently clear in the present context?

Response: The term "Oriental Realm" has been replaced with " Southeast Asia".

8. 247-248. Earlier on emphasis is put on glacial and arid conditions, here interglacials are also stated to be windows of opportunity for hominin migration. This may require some further clarification.

Response: We refer here specifically to periods in interglacials that have been identified based on local geological information as particularly humid. We thus further argue that it was not just the increased aridity that drove dispersals, but that establishment of river terraces was essential to dispersals under drier conditions (as opposed to sporadic migrations during interglacials prior to the establishment of river terraces). This has now been made explicit in the text.

9. 258. Base level lowering is probably more relevant than sea level lowering for river segments 1000s of km upstream from their respective river mouths. The area is partially an endorheic basin with no outlet to any ocean?

Response: We reorganized the text accordingly and emphasized the influences of lowering Base level.

10. 259. Apparently, we now shift to NW Europe and northern China. Away from Central Asia?

Response: Given that the river terrace data in Central Asia is relatively limited, here we

have to shift to NW Europe and northern China.

11. 339-341. What would be the underlying reason that hominins migrated mainly during humid periods in the Late Pleistocene while dry periods would be required earlier on?

Response: We emphasized that the climate in most areas of Eurasia switched from humid to arid conditions at the end of the MPT. During this transition period, dry climate led to remarkable landscape shifts (including the ecosystem conversion), thereby promoting Palearctic hominin dispersals. In contrast, a typical arid climate and steppe/grasslands prevailed in most areas of Eurasia during the Late Pleistocene. Thus, hominin dispersals in Eurasia during the Late Pleistocene may have been better facilitated by humid periods like in northern Africa.

12. 376. When is a 0.1 proportion window (meaning?) used and when 0.2?

Response: The length of sliding window was determined according to the data resolution. When the data resolution is high, a longer window was used; when the data resolution is low, a shorter window was used.

13. Figure 2. In the text mostly 0.9-0.6 Ma is being used for the MPT while in the figure 0.6-1.2 Ma is indicated. Why is that different? Why is the emphasis on the 0.9-0.6 Ma portion of the MPT?

Response: Yes, the MPT covers the time interval between 1.2-0.7 Ma. Here, we emphasized that the pronounced aridity and landscape shifts primarily occurred at the end of the MPT, i.e. ~ 0.9-0.6 Ma. We reorganized the text accordingly.

14. Figure 3. What is the difference between tectonic-climate interactions (vague wording?) and uplift-controlled? The text in 388-391 is not that helpful in clarifying. Methods used by the references are mentioned but their criteria to allocate terraces to a certain category are not. Tectonic trends typically have time scales of millions of years beyond the age range analyzed here. Climate change may operate at orbital time scales.

Response: We thank the reviewer for pointing out this issue. We agree that there is no need to categorize the river terraces into three groups as before. We abandon the river terrace classification in the revised manuscript and put all the terrace data together.

15. Figures 3 and 4. It looks like that beyond 1.0-1.5 Ma the number of input data is small compared to the 0.0-1.0 Ma interval. How does that impact the analyses performed?

Response: To avoid the influence of sampling bias/burial loss of archaeological and fossil hominin sites, we perform change point analysis on the residual value of the site number regressed against age (Fig. 4h-j). These residual values should be, in a sense, "free" of the tendency for an increase in the preservation and documentation of sites closer in time to the present.

16. 406-409. This reads cryptic to me. OK, focus here is on the (younger portion of) the MPT. However, dating is anticipated to be more precise for the most recent 100 kyr? Also the number of sites probably is high in comparison to older hominin sites.

Response: The reviewer is correct, and these properties mean that if sites younger than 100 ka are added to the dataset, the relationship between site number and date becomes even more exponential than currently modelled, with the asymptote at the young end of the scale. Such a relationship is heavily skewed towards the youngest sites and may obscure any patterns in the flatter (i.e. mid-Pleistocene) part of the exponential, even when using log-transformed values.

17. 478-483. Somehow wordy text to argue for a dominant precipitation control on $d^{13}C$.

Response: We reorganized the text accordingly in lines 488-494.

18. 498 ff. a jump in the reasoning. Some more lead-in may be required.

Response: A new sentence has been added at the beginning of this paragraph: "The $\delta^{13}C$ ratios of carbonate from paleosols and fluvial-lacustrine deposits can provide additional information about the paleoenvironmental changes."

19. 525. Why is the age window selected? Why not a shorter or longer window?

Response: In most cases, trends and events in the time and frequency domain are detected by computing the statistical parameters of the data (e.g., measures of central tendency and dispersion). The precision of these parameters depends on the length of the window. According to the data resolution and research purpose, the appropriate sliding window are selected.

20. 533. Really so? If there are no data the residual is undefined? With few data it is subject to a large uncertainty?

Response: The residual values are calculated by subtracting the predicted value from the observed value, so if there are no values at all then there are no residuals. However, if the observed value is zero but this is drawn from a valid statistical sample, then a residual will exist. The degree of certainty attached to the residual is only as good as the underlying dataset describing the phenomena of interest. We have compiled as extensive a dataset as possible, thereby seeking to minimize uncertainties.

21. 534-538. With which software package the analysis was performed?

Response: The SMATR package was used for the data analysis. We added this information in the revised manuscript.

22. 597. A kind of to be expected? Age uncertainties are probably similar, the minimum and maximum age diagrams would imply shifting the mean diagram downward and upward?

Response: As commented by the reviewer, the minimum and maximum age diagrams respectively shift the mean diagram downward and upward. Nevertheless, an expansion in the spatial distribution of fossil hominins and artefacts in Eurasia is evident at ~0.9–0.6 Ma. Moreover, marked inflection points in these distributions are consistently observed around the MPT.

23. 600. How is that residual determined? It is the residual of what (number of sites?) with respect to what (time?)?

Response: The residual value of the site number regressed against time is obtained for further change point analysis. To obtain the residual value, site number was transformed to non-zero values by adding 1, and a reduced major axis (RMA) regression analysis was performed between log (age) and log (site number + 1).

24. 605-606. So, how robust are such estimates for data-poor times?

Response: Indeed, the residual value of the site number primarily reflects the long-term trends in the distribution of hominin dispersals rather than fine-scale spatial patterns.

25. 615-616. "broadly increasing trends in the distribution of hominin dispersals" is this wordy for increasing hominin presence?

Response: Here we focused on the expansion in spatial distribution of hominins rather than the increase in hominin populations. We reorganized the text accordingly.

26. 637. and when the 'open forest steppe' vegetation type did start?

Response: According to our novel syntheses of Pleistocene paleontological records, the vegetation type of open forest steppe in East Asia appeared around 1.2-1 Ma.

27 651. were these regions also sub-humid (correct term?) and humid at the time of loess expansion? From the text further on it does not seem that way. So, why the emphasis on humid is required?

Response: Thanks for pointing out this issue. In the revised manuscript, we deleted these words.

28. 667-669. This pertains to glacial times! In the current interglacial in by far most of Europe there is no active loess accumulation. Loess-paleosol profiles are typically featuring 'telescoping', the downward migration of soil development with time.

Response: We have deleted the word "continuous".

29. 670 ff. A jump in reasoning, suddenly a switch to river terraces is made. What is the point of this paragraph? Older river terraces have much more potential to be eroded since river terraces are something dynamic (for example meandering rivers cut away riverine deposits down-stream with time).

Response: In this paragraph, we aim to provide some background information for the novel syntheses of river terraces. A new sentence has been added at the beginning of this paragraph: "A widespread occurrence of fluvial terraces could exert significant impacts on the availability of fresh water and lithic raw materials (fluvial gravels) for early hominins in Palearctic Eurasia."

In addition, we agree that older river terraces have much more potential to be eroded in the down-stream areas, especially for tectonically active regions. Nevertheless, growing evidence suggest that river terraces in Central Asia and NW China are preserved well due to the arid climate and relatively stable tectonic setting during the Pleistocene. Thus, the spatio-temporal patterns of river terraces in these regions are discernable and believable.

30. Extended data Fig. 2 here the MPT is indicated from 1.2 to 0.8 Ma. Make consistent throughout text and figures.

Response: We have done so as suggested.

31. 991-1001. Apparently shifts in central tendency and dispersion are something independent since they change at different times.

Response: In statistics, the central tendency of time series is usually expressed as mean and median mean, median, mode; the dispersion of time series is commonly expressed by range, variance, quantiles (Trauth, 2020). Thus, the shifts in central tendency and dispersion are something independent.

32. 1008-1010. Was this cluster analysis carried out by the present authors or by the studies referred to? Be explicit here.

Response: The cluster analysis was carried out by the previous studies. We provided this information in the revised manuscript.

33. Extended data figures 5 and 6 seem compilations of literature data. Is that sufficiently clear from the wording?

Response: We have provided more information in the revised manuscript.

34. 1010-1011. What variables are clustered? Some words on the premises and difference/coherence measure of that cluster analysis type would add.

Response: The relative abundance and pollen assemblage were used for cluster analysis. We have added some worlds on the premises and difference/coherence measure: "Prior to pollen cluster analysis, the premises and difference/coherence measure, including pollen representativeness and source stability, statistical independence and environmental coherence, have been evaluated."

35. Extended data figure 8. Note that the MPT (yellow shading) is differently indicated as appears from the main text.

Response: We revised the figure accordingly.

36. Extended data Figures 9 and 10. These are all literature data. Refer to the sources explicitly.

Response: We have done so as suggested.

37. What is the point of extended data figure 11? The caption is not overly helpful.

Response: To clarify the potential influences of dating uncertainties on the spatial and temporal distributions of archaeological records, the kernel density estimation diagrams and the breakpoints of archaeological and fossil hominin sites with the maximum and minimum age estimations were investigated (Extended Data Fig. 11). The figure caption has been revised accordingly in lines 1121-1126:

"To clarify the potential influences of dating uncertainties on the spatial and temporal

distributions of archaeological records, the kernel density estimation diagrams and the breakpoints of archaeological and fossil hominin sites with the maximum and minimum age estimations were investigated. An expansion in the spatial distribution of fossil hominins and artefacts in Eurasia are consistently observed around the MPT."

38. The extended data tables 1, 2 and 3, duly refer each entry. Where are these references provided in full? Do they correspond to the references in the main text? Probably only in part. There is some file named supp refs but it is not specified in table headings.

Response: Given that a vast amount of literature is available, we thus provided these references in the supplementary information. In the supplementary file, table headings are specified.

Reviewer #2:

I thank the authors for addressing the majority of comments/suggestions by reviewers. Only a few matters remain outstanding.

We are extremely grateful that you could take the time to conduct the second review of our paper. Based on these insightful comments and suggestions, we revise the manuscript carefully.

1. In response to a question by Reviewer 1on the inferred range of palaeoprecipitation, the authors have attempted to do this: "Using previously established linear correlations between the MAP and the δ13Corg of surface soils in Central Asia and the NW China, paleoprecipitation changes in south Tajikistan and the Tarim Basin can be estimated. These correlations suggest the precipitation in Tajikistan decreased by approximately 100-200 mm after the Mid-Pleistocene. In the Tarim Basin, precipitation decreased from ~200 mm during the late Pliocene to ~100 mm after the MPT". However, closer inspection shows that the data set used for the transfer function applied to the Tajikistan record has been developed from surface soil samples in northern Iran, in the western part of arid central Asia (Wang et al, 2019), while that applied to the Tarim Basin record

is based on a surface soil samples from a transect in central East Asia from the northern slopes of Qinling Mountain to near the border of Mongolia and Russia (Lee et al., 2006). Since the relationship between δ13Corg and precipitation appears to vary geographically, the lack of modern datasets for Tajikistan and Tarim Basin suggest that the inferred palaeoprecipitation values need to be treated with caution. This uncertainty should be discussed more clearly in the text.

Response: Thanks for the critical comments. We agree that the relationships between δ^{13}Corg and precipitation vary geographically, and the lack of modern datasets for Tajikistan and Tarim Basin suggest that the inferred palaeoprecipitation values need to be treated with caution. In the revised manuscript, we have provided more discussion about this question in lines 148-150:

"The relationships between $\delta^{13}C_{org}$ and precipitation vary geographically, and our inferred palaeoprecipitation values require validation from modern datasets from Tajikistan and Tarim Basin, which are currently lacking."

2. lines 204-213. I still disagree with the statement "the replacement of woodland and closed forest by steppe and woodland-steppe environments across most of Palearctic Eurasia during the Mid-Pleistocene". Forest biomes occurred less frequently, because of the shift from 41-kyr to ~100-kyr glacial cycles, but interglacial forests did not turn to steppe-woodland or open forests in most places, only in the lower precipitation areas. Some of the moisture-requiring tree taxa disappeared from interglacial forests, but what mainly changed was the character of the glacial periods with some tree presence during the Early Pleistocene, but then mostly absence later on (see for example Zhao et al 2020 Science Advances; Donders et al. 2021 PNAS). The same correction is needed for the caption of Ext. Fig. 5, line 1020, the herb values (pollen percentages not pollen concentrations) increased during the glacial periods (not the interglacials). (Incidentally, the attribution of the serial extinction of Tertiary relicts (line 211, ref 42) is incorrect, it should be Tzedakis et al. 2006 QSR.)

Response: Thanks for the critical comments. We agree that the replacement of woodland and closed forest by steppe and woodland-steppe environments mainly

occurred during the glacial periods. We reorganized the text accordingly in lines 209-218 and in the caption of Ext. Fig. 5:

"The first was the replacement of woodland and closed forest by steppe and woodland-steppe environments across most of Palearctic Eurasia during the glacial periods of the Mid-Pleistocene. In northern China and southeast Tibet, the principal vegetation biomes during the early Pleistocene glacial periods were characterized by coniferous or broadleaf forests, and then show a glacial change towards open forest steppe at the onset of the MPT (Extended Data Fig. 4). Pollen records from Spain, France, Greece, and Italy suggest that an ecological shift towards more open habitats, together with a serial extinction of moisture-requiring taxa (the so-called Tertiary relicts), is evident during the glacial periods since ~0.9-0.6 Ma (Tzedakis et al., 2006)."

3. Fig. 4b,c,d: the density diagrams appear to show a low value at ~2.4 Ma near 0° longitude. Since the data refer only to Europe, is this an artefact of the analysis?

Response: We thank the reviewer for pointing out this issue. Yes, this data is an artefact of the analysis, which is aimed to set a suitable scale to show the density diagrams. We are really sorry for the mistake and remove this data in the revised figures.

4. Extended Data Table 2:"Loir" should be "Loire". The Voidomatis River in Greece is in the NW not NE. Also, while I agree that both tectonic and climate influences are present in that basin, the reference provided (Hughes 2010) does not discuss that. So where was the information taken from?

Response: Thanks for pointing out this. We correct these mistakes accordingly. The information about the river terrace classification is from two additional references by Lewin et al. (1991, Quat. Res.) and Hamlin et al. (2000). As commented by the reviewer 1, there is no need to categorize the river terraces into three groups as before. Hence, we abandon the river terrace classification in the revised manuscript and put all the terrace data together.

5. Finally, a more general comment: the main thrust of this work is that Mid-Pleistocene

enhanced aridity and landscape shifts “drove” Palaearctic hominin dispersals. This is plausible, but the confluence of the climate shift over the MPT and human dispersals does not necessarily mean a direct causal relationship (since other factors are also important). I wonder, therefore, whether this could be phrased in a more circumspect manner (e.g. may have “promoted” or “influenced”).

Response: We thank the reviewer for pointing out this issue. In the revised manuscript, the main thrust of this work has been phrased in a more circumspect manner and we also change the word "drove" to "promoted" in the title.

Reviewer #1:

The re-revision has addressed all my comments to a sufficient extent. I recommend acceptance for publication. Usage of articles may require some attention (a few times 'the' seems superfluous). AF in line 1096 stands for Alternating Field; inconsistencies in the reference list can be remedied at the proof-reading stage.

Response: We are extremely grateful that you could take the time to conduct the third review of our paper. In the revised manuscript, we have carefully proofread the manuscript and corrected all grammatical errors as suggested by you.

Reviewer #2

Response: We are extremely grateful that you could take the time to conduct the third review of our paper. In the revised manuscript, we have reorganized the text accordingly.

l. 218 After reference "42" please add the following reference: MAgri et al., Quaternary disappearance of tree taxa from Southern Europe: Timing and trends. Quaternary Science Reviews 163 23-55 (2017).

Response: We have done so as suggested.

l. 648-650 and Extended Data Fig. 4: "Vegetation types in East Asia show a remarkable shift from forest to open forest at 1.2-1 Ma, and subsequently changed from open forest steppe to steppe at ~0.9-0.6 Ma (Extended Data Fig. 4)" With respect to Zoige Basin, this sentence is not correct: after 0.62 Ma, interglacials were still characterized by tree population expansions (mainly Picea), but with presence of alpine meadows; glacial periods were longer and dominated by steppe and meadows. Please correct this in the text and also add some spruce trees in the Ext Fig 4 in the interval 0-0.62 Ma and even more for the interval 0.62-0.9 Ma

Response: We have done so as suggested.