Response to comments of Anonymous Referee #1

General comments

The manuscript has been worked on by the authors but there are still most questions left unanswered mentioned already in my first review.

Response: Thanks for all the constructive comment. We have revised the manuscript carefully and our point-by-point responses are provided below.

Q1

Abstract: "around 6000 cal. a BP, coinciding with the Holocene optimum….". The Holocene optimum in the 6th millennium was cal BC not BP ((the latter is around 4000 BC)). In addition, the cultivated plants (*Triticum* species, *Hordeum*, *Secale*, pulses, etc.) in South-West Asia (Near East) are millennia older.

Response: Thanks for the helpful comment.

1) We have revised the age of Holocene optimum as intervals between 8000 and 6000 cal. a BP. Further clarification of the Holocene optimum, defined as maximum combination of precipitation and temperature (Chen et al., 2015; Dong et al., 2022), was added in lines 302–304.

2) Yes, we acknowledged that the domestication of crops in Southwest Asia occurred in the early Holocene (~9000 cal. a BP), which was millennia older than the millets in East Asia.

References:

Chen, F. H., Xu, Q. H., Chen, J. H., et al.: East Asian summer monsoon precipitation variability since the last deglaciation, Sci. Rep., 5, 11186, 2015.

Dong, Y. J., Wu, N. Q., Li, F. J., et al.: The Holocene temperature conundrum answered by mollusk records from East Asia, Nat. Commun., 13, 5153, 2022.

Q2

Fig. 1: Different colored dots are entered in the map, which seem to correspond to regions. Are the dots of one color the same age? Otherwise, it would make more sense to present the points according to the archaeological epochs. Looking at the chronology table below, archaeological cultures are already present well before 6000 BP (see also in the text rows 80-90). Do we know nothing about these cultures archaeobotanically? Is this really a Neolithic (by what proved?) or are these hunter-gatherer cultures? What characterizes and differentiates the different archaeological cultures that are compared in Figs 2 and 3?

Response: Thanks for the suggestion.

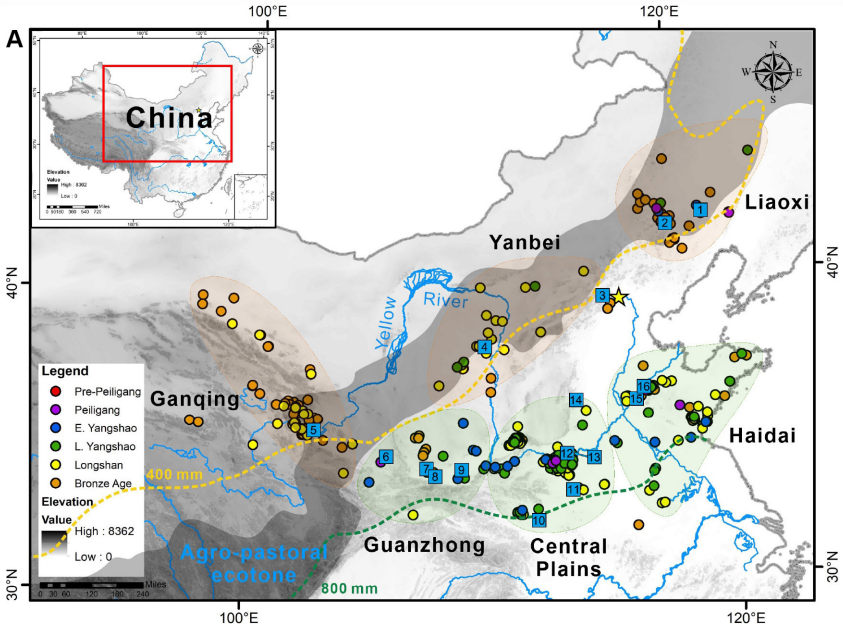
1) These dots of one color in the original Fig. 1A were not the same age and all the dots have been redrawn according to the archaeological epochs, which corresponded to the chronology table in Fig. 1B.

2) The archaeological cultures before 6000 BP have also been archaeobotanically investigated and exhibited as "Pre-Peiligang, Peiligang and Early Yangshao" periods in Figs. 2 and 3.

3) These cultures were proved to be Neolithic mainly based on the sedentary settlements, possible agriculture, pottery, and ground stone tool (Liu and Chen, 2012), and the characterizes and differentiates of these cultures were briefly summarized in lines 85–92.

References:

Liu, L. and Chen, X. C.: The archaeology of China: from the late Paleolithic to the early Bronze Age, Cambridge University Press, New York, 2012.



Q3

Line 100: the archaeozoological NISP method used for counting crop fragments is not understandable here. This should be clarified.

Line 102ff.: Perhaps the different size of the grains of *Triticum*, *Panicum*, *Setaria* etc. is meant here? This should be clarified.

Response: The definition of archeozoological NISP and possible effects of different size of crop grains in identification have been clarified in lines 111–115.

Line 111–115

…(NISP)" in zooarchaeology (Grayson, 2014), defined as the number of identified specimens for a specific site or skeleton. According to this criterion, fragments of unidentified crop seeds were not counted, while each fragment of identifiable large crop seeds, such as wheat and rice, that retained more than half of intact seed were counted as an intact seed; different parts of the crop seeds, i.e. diagnostic grains and spikelet bases of wheat, barley, and rice, retrieved from the same context were added up to denote the total numbers of crop seeds.

Q4

Fig. 2 and 3: the dashed lines are mathematically not correct (see my last review), the percentages of counts and weights show the same trend.

The archaeobotanical results cannot be understood without the following (citation from my first review already):

"The temporal-quantitative evaluation is not comprehensible if the authors do not disclose how many sites (features, samples) they have per region and per time slice or archaeological culture. Only then is it clear whether the quantitative changes are not artifacts. According to

page 4, they have 487 flotation results (are these samples?) from 349 sites. That is, less than 2 samples per site on average? Maybe also a few sites (which epochs) with many samples? Therefore, the representativeness of the data is not clear. What about the earliest time slice (e.g. Fig. 5 above): is there nothing investigated, or is it investigated, but nothing found?"

Response: Thanks for the insightful comment.
1) The dashed lines in Figs. 2 and 3 have been completely deleted in the manuscript.
2) Numbers of flotation results (sites or cultural phases) applied to illustrate the cropping patterns per region and per time slice have been added below each period in revised Figs. 2 and 3. Further detailed information on the spatiotemporal composition of flotation results have been added in Table 1 and lines 99–104.
3) 78 of 349 sites contained more than one cultural phase, while the rest contained one single cultural phase.
4) Only one site with millets were investigated in the earliest time slice of Fig. 5 (11000–9000 cal BP), i.e., Donghulin site (Fig. 3B), and thus were not illustrated again.

Line 99–104
Each flotation result addressed here indicated a compilation of original samples floated from the same cultural phase of an archaeological site. The original sample numbers investigated and compiled per cultural phase of each site ranged from 1 to 1082, with an average of approximate 30 samples. The 487 flotation results from 349 sites were classified into six geographical regions and six cultural periods (Table 1), with 78 sites containing more than one single cultural phase. The number of flotation results across the six regions during each period was generally even, except for the common lack of flotation results in the agro-pastoral ecotone before Longshan period.

Table 1. *Spatiotemporal composition of flotation results in northern China*

Period	*Mid-lower Yellow River (MLY)*			*Agro-pastoral ecotone (APE)*			
	Guanzhong region	*Central Plains*	*Haidai region*	*Ganqing region*	*Yanbei region*	*Liaoxi region*	*Sum*
Pre-Peiligang	*0*	*0*	*0*	*0*	*1*	*0*	*1*
Peiligang	*2*	*4*	*4*	*0*	*0*	*3*	*13*
Early Yangshao	*11*	*6*	*3*	*0*	*0*	*2*	*22*
Late Yangshao	*15*	*50*	*11*	*0*	*3*	*1*	*80*
Longshan	*11*	*72*	*29*	*22*	*26*	*0*	*160*
Bronze Age	*18*	*71*	*30*	*52*	*6*	*34*	*211*
Sum	*57*	*203*	*77*	*74*	*36*	*40*	*487*

Q5
The phytolith picture of millets (Fig. 6 below to the left) is not good enough to be distinguished from the one to the middle and they have to be named by the species name. The grains have to be turned (embryo has to be below) and the pictures must be larger, to see if the curves of the embryos are typical for *Panicum* and *Setaria* respectively which are different.

Response: Both the images of charred millet seeds and phytoliths have been rearranged, which were large and clear to be identified with diagnostic characteristics, i.e., the grain embryos of millets have been turned to be below (E and F) and the Ω- and η- undulated epidermal long cell walls (H and I) have been illustrated. Besides, they have all named by the Latin species name.