Response to comments of Anonymous Referee #1

General comments

The essay on the development of millet cultivation in China includes an interesting and large collection of data well suited for publication.

Specific comments

1. 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 helpful comment. Sample numbers had been added to the top-right corner of figures per region (Fig. 2 and 3) and per time slice (Fig. 4). The 487 flotation results addressed in the manuscript indicate compilations of samples from the same phase of an archaeological site rather than original samples floated, and sites with less than 2 or many samples just indicate the number of cultural phases per site. Original sample numbers and volumes investigated per site have been added in the new version of the data tables (https://doi.org/10.5281/zenodo.6669730). The earliest time slice of Fig. 5 (11000–9000 cal BP) only consists of one site, i.e. Donghulin site (Fig. 3B), and thus is not illustrated again.

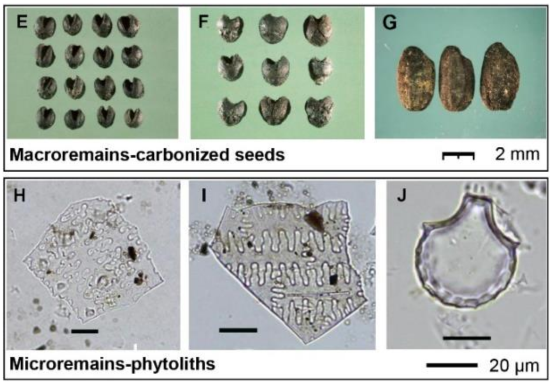
2. An image of selected macro-remains and phytoliths of the millets is missing. What are the criteria to distinguish the millets on the basis of their phytoliths? Is this possible? In any case, the comparison of grain numbers and phytoliths is not useful. If there has not been threshing and dehusking within the village, phytoliths will be hardly found on site.

Response: Images of selected macro-remains and phytoliths of the millets have been added in Fig. 6. Discrimination of millets phytoliths was based on five key diagnostic characteristics in phytolith morphology of inflorescence bracts, especially the Ω-type and η-type (Fig. 6H and I), which had been published by our research team (Lu et al., 2009). The divergence between grain numbers and phytoliths may attribute to several factors, and the possible effect of the processing method had been added to the discussion.

Lu, H. Y., et al. Phytoliths analysis for the discrimination of Foxtail millet (*Setaria italica*) and Common millet (*Panicum miliaceum*), PLoS One, 4, e4448, 2009a.

Lines 206–208

…such as the different sources of crop grains and phytoliths or biases in the representativeness of quantity. Diagnostic phytoliths of millets were derived from the inflorescence bracts (Figures 6G and I), which may have been discarded during threshing and dehusking. …



3. Important would still be the climate discussion: around 6000 BP there is a climate deterioration in Central Europe, what it's like in China? The authors write warm/humid. How can they read this from the pollen data?

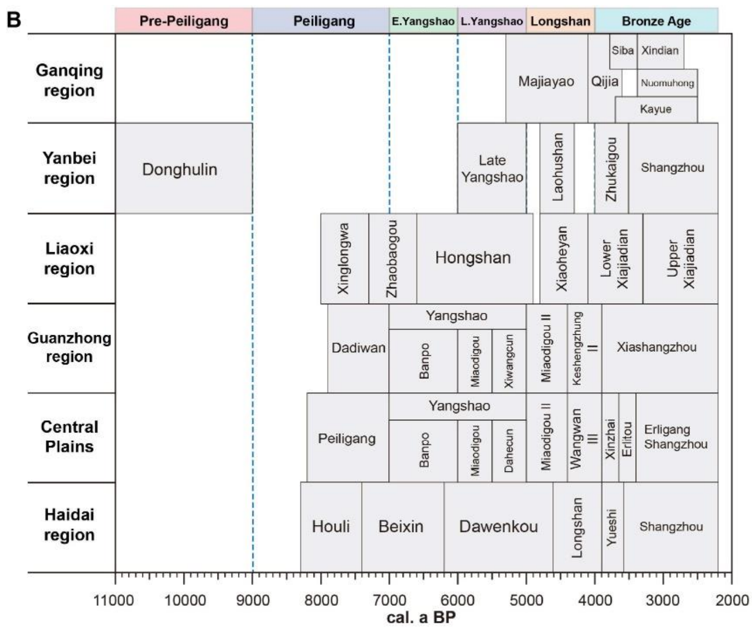
Response: According to the high-resolution, pollen-based quantitative precipitation reconstructed from the Gonghai Lake in northern China (Fig. 7A), the annual precipitation was 30% higher than present from ~7800–5300 ka, and no climate deterioration had been observed around 6000 cal BP in northern China. Besides, similar results could also be observed in pollen records from the Daihai and Hulun Lake.

4. Absolutely necessary is a chronology table, broken down by regions and millennia (better centuries), otherwise the arguments and data are not understandable. And it would be important to have a brief summary of what characterizes these archaeological cultures about which they are writing, see also Fig. 1. Are these comparable settlement types and types of findings?

Response: A chronology table broken down by regions and centuries has been added as Fig. 1B, and a brief summary of these archaeological cultures has been added in section 2.

Lines 83–90

During the Pre-Peiligang period, only a few archaeological sites were scattered in the Yanbei region, Central Plains, and Haidai region. Subsequently, four archaeological cultures, i.e. Xinglongwa, Dadiwan, Peiligang, and Houli, formed almost synchronously across Liaoxi, Guanzhong, Central Plains, and Haidai regions, which were widely regarded as sedentary settlements and the origin of millet agriculture. During the early and late Yangshao period, settlements increased dramatically and spread wildly across all the six regions, especially Miaodigou culture, with remarkable signals of social hierarchy emergence. The Longshan period witnessed high population densities and the rise and fall of early complex society, and eventually the early states—Erlitou culture formed in the Central Plains during the Bronze Age.



5. That more existing 14C data of settlements means an increase in population is an old idea but not convincing. The amount of 14C data depends among others on how much money the archaeologists spend on it. It is enough argument that there are more sites. But for that you would have to know whether the fewer, older sites are just as easy to find. If they have left fewer traces (e.g. block construction of houses, no pits), then you will also find less. Are all epochs sampled and examined equally (see above?).

Response: Thanks for the comment. One key process in the use of 14C data of settlements dealing with the inter-site sampling intensity was aggregating samples from the same site before calculating summed probability distribution, and thus may diminish possible anthropogenic biases and reflect the fluctuation of population. The taphonomic loss was an inevitable question in the flotation results, and feature numbers (mean, standard deviation) (Excel data, MLY_Fig.2 and APE_Fig.3 sheets) had been applied to infer changes in all epochs.

6. There is also a lack of inclusion of other crops. For example, a change from dry to wet rice cultivation from the Neolithic to the Bronze Age can be expected, etc. This would clearly substantiate the author's thesis of increasing effectiveness. In addition, the archaeological background would be interesting, what do we know: settlement concentration? Which raw materials? Already metal? Trade?

Response: Thanks for your suggestion. Due to limited research on the determination of rice arable systems, whether the paddy fields in northern China were wet (e.g. Huizui site) or dry (e.g. Zhuzhai site) remains controversial and a shift from dry to wet rice cultivation had not been observed so far. The increasing effectiveness was supposed to be mainly focused on millet agriculture in northern China as rice only played a minor role in the cropping patterns. The key change in cropping patterns from the Neolithic to the Bronze Age was the introduction of wheat and barley through the trans-Eurasian exchange, as illustrated in Figs. 2 and 3.

7. There is something strange with nitrogen and loess on p. 15 below. The black soils from loess have been the most fertile soils ever, which probably did not have to be fertilized for the first 1-2 millennia.

Response: Though the black soils from loess were fertile, loess soils were vulnerable to the loss of organic matter due to the little content of clay in loess and thus may not be able to sustain continuous intensive cultivation. Besides, a millet–pig system illustrating the fertilization of millet fields with pig or human dung was supposed to be in practice at Dadiwan around 5500 cal BP in a recent study (Yang et al., 2022).

Yang, J. S., et al. Sustainable intensification of millet–pig agriculture in Neolithic North China, Nature Sustainability, https://doi.org/10.1038/s41893-022-00905-9, 2022.

Lines 297–299

Besides, a recent study in the Dadiwan site also suggested an intensive crop–livestock system was in practice around 5500 cal BP (Yang et al., 2022)

8. In the case of bar charts, the dashed lines must be removed or consistently applied to everything. Mathematically, they are strange, because these are different times and data sets. Why are some bars missing in the charts? The number of the figures has to be checked.

Response: The horizontal dashed lines in Fig. 3 have been removed and the green and blue dashed lines have been plotted consistently to illustrate the diachronic trends of increase or decrease across different periods. The missing bars in the charts indicate the absence of certain crops or lack of flotation results during certain periods (Fig. 3). The number of Fig. 4 was mistakenly written as Fig. 2 and has been revised.

9. Something important for the calculations is the differentiation of mass finds (charred storage finds) versus normal settlement waste. If such mass finds, which are singular events, are included in the calculations, they confuse the results, e.g. by pretending an "increase".

Response: Thanks for the comment. Two rules applied in this study could ensure the representativeness of these data. Firstly, most of the flotation results compiled here were a mixture of contemporaneous samples recovered from cultural layers or pits of the same sites, with more than 74% containing 2 or more samples. Secondly, the absolute numbers of crops from each site were transferred to percentage first and then calculated for the mean values and standard deviations of each period, which diminished possible representativeness bias induced by samples with mass finds.

Technical corrections

10. As for the data tables (Excel):

They are not understandable for people not involved in the project.

The feature numbers, sample numbers and sample volumes investigated per site are lacking.

Are all items preserved charred?

The order of the data is unclear.

The archaeological period and the publication per site are lacking.

Response: As for the data tables (Excel): A new version of V1.0 has been uploaded, including a revised Excel file and a new KML file for better visualization of these data. Sample numbers and volumes of each flotation results have been added and feature numbers of each period (MLY_Fig.2 and APE_Fig.3 sheets) have also been illustrated. All staple crops are preserved charred. The order of the data has been rearranged. The archaeological period and the publication per site have also been added (https://doi.org/10.5281/zenodo.6669730).

Response to comments of Anonymous Referee #2

General comments

Millet agriculture were initially domesticated in northern China and played an important role in early agriculture evolution and the formation of the Chinese civilization. The manuscript reports a dataset of archaeobotanical macroremains spanning the Neolithic and Bronze Ages in northern China. Authors also suggest a significant spatiotemporal divergence of millet agriculture, discuss the past human-environment interaction, and provide a valuable perspective of agricultural sustainability for the future. This manuscript meets the scope of Earth System Science Data and could arise a wide audience as well. I would like to suggest a publication after a moderate revision.

Response: Thanks for the helpful comment. Our point-by-point responses are provided below.

Specific comments

1. In Introduction Part and Fig. 1. I suggest that authors check the names of different regions in North China, which belong to geographical division or archaeological culture division. For example, many archaeological sites distribute in Loess Plateau and are not in Guanzhong basin. Yanbei region is not clear.

Response: The reference for the names of different regions has been revised, and a brief introduction to the geographical and archaeological division of the six subregions has been added in Lines 73–79. Archaeological sites distributed on the Chinese Loess Plateau were divided into the Yanbei region in the north and Guanzhong region in the south by bondary of 37°N.

Lines 73–79

The Liaoxi region is situated in the west of Liaoning Province and southeast of Inner Mongolia Autonomous Region; to the west, the Yanbei region is situated in the central-south of Inner Mongolia Autonomous Region and north of Shaanxi, Shanxi, and Hebei Provinces, northern parts of Chinese Loess Plateau; in the westmost, the Ganqing region is situated in the northeast of Qinghai Province and central-south of Gansu Province. The Guanzhong region is located in the southeast of Gansu Province and south of Shaanxi Province, which is merged into the Central Plains in some studies; to the east, the Central Plains is located in the south of Shanxi Province, and the bulk of Henan Province; the Haidai region is mainly located in the Shandong Province.

References

Yan, W. M.: Cradle of Oriental Civilization. In: The origins of agriculture and rise of civilization, Yan, W. M. (Ed.), Science Press, Beijing, 2000.

2. A reference (Zhou et al., 2011) need to be cited which has discussed the significant divergence of millet west Loess Plateau around 5500 BP.

Response: A new reference (Zhou et al., 2011) has been added in Line 40.

References

Zhou, X. Y., Li, X. Q., Zhao, K. L., Dodson, J., Sun, N., and Yang, Q.: Early agricultural development and environmental effects in the Neolithic Longdong basin (eastern Gansu), Chinese Sci. Bull., 56, 762, https://doi.org/10.1007/s11434-010-4286-x, 2011.

3. In Discuss Part.

Line 170-175

The description on "The spread of millet intensified from the late Yangshao to Longshan periods in two directions (Figures 5C–D and H–I): westward routine to the Ganqing region and northward routine to the Yanbei region" and Fig. 5 need more evidences and the references to support

Response: More evidence of routes and dates in the two directions had been added in Lines 188–192, and new references supporting the spread of millet agriculture had also been added.

Lines 188–192

The discontinuous agricultural region linked together at stage II and expanded to marginal areas without solid evidence of agriculture prior to 6000 cal BP. The spread of millet intensified from the late Yangshao to Longshan periods in two directions (Figures 5C–D and H–I): westward routine to the Ganqing region along the Wei River and Hexi corridor around 5400 cal BP (Leipe et al., 2019); and northward routine to the Yanbei region along the middle Yellow River around 5200 cal BP (Bao et al., 2018).

References

Bao, Y. G., Zhou, X. Y., Liu, H. B., Hu, S. M., Zhao, K. L., Atahan, P., Dodson, J., and Li, X. Q.: Evolution of prehistoric dryland agriculture in the arid and semi-arid transition zone in northern China, PLoS One, 13, e0198750, https://doi.org/10.1371/journal.pone.0198750, 2018.

Leipe, C., Long, T. W., Sergusheva, E. A., Wagner, M., and Tarasov, P. E.: Discontinuous spread of millet agriculture in eastern Asia and prehistoric population dynamics, Sci. Adv., 5, eaax6225, https://doi.org/10.1126/sciadv.aax6225, 2019.

4. Line 195-210

The discussion on the possible biases of archaeobotanical macroremains and the reason of divergence of the foxtail and broomcorn millet need to add some information on the different ecological habits and the way of seed yield from the foxtail and broomcorn millet. I think that the discussion of phytolith and Fig. 6 are not necessary, which can't support the changes and divergence of the foxtail and broomcorn millet during the Neolithic.

Response: Different ecological habits and the seed yield of foxtail and broomcorn millets have been added in Lines 227–234. Though discussion of phytolith (Fig. 6) didn't support the transition from broomcorn to foxtail millet, it could also provide a valuable perspective for comparison to that of charred seeds.

Lines 227–234

Furthermore, the way of seed yield may also affect the biases between foxtail and broomcorn millet. Morphological data from the northern Chinese Loess Plateau showed that the size of broomcorn millet increased significantly during 5500–4000 cal BP, while that of foxtail millet didn't exhibit an obvious increasing trend (Bao et al., 2018). Given the contemporaneous increasing human population, the crop yield was supposed to increase to feed a large population. Nevertheless, the increasing yields of broomcorn millet may depend on the increase in the seed size, while that of foxtail millet may result from the increase in the number of seeds per plant. Thus, the different ways of seed yield may also be one cause of the increase in the proportion of foxtail millet within the total seed number of seeds. In brief, both biases in the carbonization process and the way of seed yield may have exaggerated the proportion of foxtail millet.

5. Line 259-269

The manuring enhanced the crop yields and provide the possible reasons that human adapt the environmental changes and can't well understand the divergence of the foxtail and broomcorn millet around 6000 BP. Authors need to more discussions on the driving factors.

Response: More discussions about the relationship between field manuring practices and different crop yields have been added in Lines 297–301. Given the yield of foxtail millet was typically twice that of broomcorn millet, the ancient human may choose to cultivate more foxtail millet to maximize crop yield on the condition of enough fertilizer.

Lines 297–301

Besides, a recent study in the Dadiwan site also suggested an intensive crop–livestock system was in practice around 5500 cal BP (Yang et al., 2022). Considering the yield of foxtail millet was typically twice that of broomcorn millet, the dramatic increase in soil fertilization may promote the conscious choice of high-yield foxtail millets to maximize agricultural productivity.

References

Yang, J. S., Zhang, D. J., Yang, X. Y., Wang, W. W., Perry, L., Fuller, D. Q., Li, H. M., Wang, J., Ren, L. L., Xia, H., Shen, X. K., Wang, H., Yang, Y. S., Yao, J. T., Gao, Y., and Chen, F. H.: Sustainable intensification of millet–pig agriculture in Neolithic North China, Nat. Sustain. https://doi.org/10.1038/s41893-022-00905-9, 2022.