

For research article

Response to Reviewer 1 Comments

1. Summary

Thank you very much for taking the time to review this manuscript. Please find the detailed responses below and the corresponding revisions in track changes in the re-submitted files.

2. Questions for General Evaluation	Reviewer's Evaluation	Response and Revisions
Does the introduction provide sufficient background and include all relevant references?	Yes	No
Are all the cited references relevant to the research?	Yes	No
Is the research design appropriate?	Yes	No
Are the methods adequately described?	Yes	No
Are the results clearly presented?	Can be improved	See in point-by-point response
Are the conclusions supported by the results?	Yes	No

3. Point-by-point response to Comments and Suggestions for Authors

Authors analyzed the influence of presence/absence variations (PAVs) in 141 rice varieties genome on the five selected traits (days to heading, plant height, panicle number, flag leaf length and flag leaf width).

Abstract and introduction sections are of good quality and present necessary information. In the materials and methods I propose the justification of values described in line 114.

In the results section Authors presented the average sequencing depth and coverage as well as the number of PAVs (33220). The number of observed PAVs was inversely correlated with their length. Most of observed PAVs (67,4 %) occurred in gene sequences.

The number of PAVs per chromosome was also calculated to find their hotspots for example on chromosome 11.

After performing phenotype statistics studies Authors confirmed the continuous and

approximately normal distribution for five traits, suggesting that each of them is quantitative and controlled by multiple genes.

To analyse the role of particular PAVs in rice improvements, Authors applied 33220 PAVs as genomes and performed a genome-wide association analysis for five traits. As a results, were found PAVs that are significantly associated with flowering time (29), plant height (9), panicle number (37), flag leaf length (82), and flag leaf width (29). Finally, the annotation informations were used to find genes corresponding to described PAVs. Available references were applied to support the role of observed genes in modification of five traits.

The Discussion section is of good quality.

However, the explanation of results presented on Fig S2 and correction of typo in line 303 was proposed.

Authors selected presence/absence variations (PAVs) in rice using existing sequencing data from 148 improved rice varieties in Asia. Moreover, the genome-wide association study (GWAS) using PAVs and phenotypic data was performed for five traits (flowering time, plant height, flag leaf length, flag leaf width, and panicle number). As a results, 186 significantly associated PAVs were identified. Obtained information could be used to further improve the rice cultivars.

Results could be interesting to researchers working on breeding rice varieties. The study is properly planned and performed. Obtained results support conclusions. Figures are of good quality. In my opinion following minor comments should be addressed to improve the manuscript.

Response: We appreciated the reviewer for the valuable comments and have revised the manuscript as your suggestion.

Comments 1: Line 114 Describe why presented values were applied in the study: (1) minor allele frequencies (MAFs ≥ 0.01); (2) missing rate $< 40\%$.

Response 1: The variant with MAF < 0.01 were defined as rare variants and we filtered them to obtain polymorphic locus. Excessive missing data at loci can affect the accuracy of GWAS analysis. Therefore, we conducted the missing rate $< 40\%$ to filter the genotype data accordingly. The filtering criteria were referenced from some literature [1, 2] that we have supplemented in the Line 114.

Comments 2: Figure S2-Describe in the text how values of PC-9,45 % and 3,96 % were obtained/calculated.

Response 2: The proportion of variance explained by each component is computed by

dividing the variance of each component (eigenvalue) by the total variance of the dataset. We have added it in the section of Materials and Methods in Line 130. We also revised the Supplementary Materials in Line 322.

Comments 3: Line 303-*Brassica napus*-should be written in italice.

Response 3: We have revised the *Brassica napus* in italic and modified the other errors in the manuscript.

Reference

1. Zhao, G.; Lian, Q.; Zhang, Z.; et al. A comprehensive genome variation map of melon identifies multiple domestication events and loci influencing agronomic traits. *Nat. Genet.* 2019, 51(11), 1607-1615.
2. Zheng, X.; Zhong, L.; Pang H.; et al. Lost genome segments associate with trait diversity during rice domestication. *BMC Biol.* 2023, 21(1), 1-15.