nature portfolio

Peer Review File

Allohexaploid wheatgrass (Elymus nutans) genome reveals its Y haplome origin in Triticeae and high-altitude adaptation

Corresponding Author: Professor Xiao Ma

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 manuscript reports a study focused on the genome structure and sequence of the allohexaploid (StStYYHH) wheatgrass species Elymus nutants. This is an important forage that grows within a broad altitude range in the Qinghai-Tibet Plateau. The manuscript is very well written with simple and clear explanations of the results. The methods used to generate the datasets are also well described. I believe this manuscript will attract the attention of the scientific community.

A key novel result reported in this manuscript is the phylogenetic reconstruction of the different haplomes that gave rise to E. nutants. The manuscript also reports a genome-wide association study that is based on the re-sequencing of several lines that grow at different altitudes. This highlights a novel regulatory element that is associated with the expression of the gene MAPKKK1. The reduction in the expression of this gene enhances the adaptation to UV-B radiation that is expected in higher altitudes.

The results about the phylogenetical history of the haplomes is very detailed and contain several new discoveries. This is a significant outcome of the study that is relevant to the research community, in particularly, crop scientists and breeders. The understanding of the evolutionary drivers and constraints as described in the manuscript could be a great tool to enhance wheat and barley breeding. As it is the case with other grasses, the genome structure preserves the number of chromosomes and gene synteny.

The authors also report the analysis of gene content between the three haplomes of E. nutants. Remarkably, the expression between genes that are highly conserved between the haplomes also display a similar level of expression in response to stress. This raises several questions that I would like to posed to the authors:

Is the number of triads of syntenic genes a relatively small number? How this compares to the sequenced hexaploidy bread wheat genome?

How many pairs of genes can be found between the haplomes? Is there any pair of haplomes that are genetically closer?

Finally, for those pairs of syntenic genes, is the third missing copy a pseudogene? If yes, in how many cases.

Another novel result reported in this manuscript is the close relationship of the Y-haplome in E. nutants and the V-haplome in D. villosum and Jv-haplome in Thinopyum intermedium which challenges previous assumptions that the St and Y haplome had a common recent ancestor.

The molecular assays and data analysis tools used to support the conclusions of the manuscript are well established and suitable for type of results. The description of the methods is thorough. The only comment I have regarding the data analysis is the missed opportunity to delve deeper into the annotated genes (as for instance, checking for signatures of subfunctionalisation in the homeologous genes analysis).

Regarding the data I could not validate the access to the raw sequence data that is deposited in China National GeneBank DataBase. The accession numbers for the datasets are visible but result in a message indicating that the raw sequencing data will not be available until 25 November 2025. Unfortunately, these datasets will need to be available for public access before I could recommend this manuscript for publication.

Typos

Line 117 “sytenic relationships” should be “syntenic relationships”

Line 174 “Dasypyrum Villosum” should be “Dasypyrum villosum”

Line 309 “clod stress” should be “cold stress”

(Remarks on code availability)

Reviewer #2

(Remarks to the Author)
This manuscript offers a multi-dimensional presentation of the genome, evolutionary history, and genomic adaptation of Elymus nutans, an allohexaploid composed of three subgenomes. The paper describes a genome assembly for the hexaploid Elymus nutans, pinpoints the evolutionary position of two of the three subgenomes of E. nutans (Y and H), provides evidence of reticulate allopolyploidization histories of Elymus, explores population structure and demographic history of E. nutans in Central Asia and the Qinhai-Tibet Plateau, and then documents genetic signatures of environmental adaptation and GWAS to identify genes underlying UV-B resistance.

This manuscript is one of the most intensive studies of a large polyploid genome I have seen. It is a dense, comprehensive, multi-faceted, generally well-written manuscript that provides deep insights into a large hexaploid genome. It was interesting that although there are some minor differences in expression among subgenomes, overall expression across subgenomes is relatively balanced. In looking at the evolutionary history of the Y and H haplomes, I appreciated that the authors were careful in their wording (e.g., “the ancestral donor of the Y haplome in Elymus s.l. may be phylogenetically proximate to Dasypyrum and Thinopyrum….”). In this part of the work, the authors have not identified the donor of the Y genome, but they have an idea of the general area within the Triticeae where it may have originated. The subsequent sections that use resequencing data from other Psudoroegneria, Hordeum, Roegneria, and Elymus and chloroplast sequence data provide convincing evidence of multiple origins of polyploidization and reticulation. The model showing reticulate evolution in Elymus (Fig 3C) is difficult to interpret. At the very least, a more detailed caption is needed. And it would be even better to have more clarity within the figure.

While the data are solid, I felt the introduction and discussion, and to a lesser extent the results, failed to make the most of the work presented. In some cases, the writing is a little choppy (“perennial” is used three times in the first five sentences). I think there is a gap between the genome assembly/evolutionary analyses , and the local adaptation/GWAS studies. The introduction offers an opportunity to highlight the novelty of combining these approaches in a single study, and the discussion presents an opportunity to more deeply consider the role of polyploidization in high-altitude evolution. For example, specifically which subgenomes within E. nutans were most important in the adaptation to UV-B and other aspects of the environment? Perhaps this is buried in results, but it could be more pronounced. How has reticulate evolution facilitated local adaptation? Were the same gene regions underlying stress tolerance selected following distinct allopolyploid origins, or different ones? If this is beyond the scope of this paper, that should be more clearly stated.

All in all, I enjoyed this paper and think it represents a novel and important contribution.

(Remarks on code availability)

Reviewer #3

(Remarks to the Author)
I was asked to comment on the UV-B part of this manuscript. Therefore, I specifically focus my review on this part. This work identified MAPKKK18 as a new gene with negative roles in UV-B tolerance in Elymus nutans through phenotype-environmental association and common garden analyses. The authors show that allelic mutations in the promoter region of MAPKKK18 that decrease its gene expression, facilitating adaptation to UV-B radiation in the high-latitude populations. The authors could confirm the negative role of MAPKKK18 in UV-B stress tolerance in Arabidopsis with an Arabidopsis mutant line and overexpression lines. The data of UV-B stress tolerance appear to be convincing. I raise several issues related to UV-B tolerance.

1. In Fig S19, flavonoids and total anthocyanins should also be quantified since both are important for UV-B stress tolerance.
2. Photoreceptor UVR8-mediated UV-B acclimation pathway promotes UV-B stress tolerance in Arabidopsis. Is there any relationship between UVR8 pathway and the reported MAPKKK18 gene in UV-B stress tolerance?
3. The UV-B conditions used in Figure 6h is not clear. Please provide more detailed information here.
4. OE-MAPKKK18 lines show very low chlorophyll contents as shown in Fig S19 A-B, however, the OE lines shown in

Figure 6h appear to develop green leaves as WT Arabidopsis plants. Please comment on these observations.

(Remarks on code availability)

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
I thank the authors for responding to my questions as well as the ones raised by the other reviewers. I am satisfied the responses and further analysis covered all the points I raised in my original review. In my opinion the manuscript makes a substantial contribution to our understanding of the evolutionary history of the haplomes within the Triticeae tribe. It also highlights the role of the gene MAPKKK18 in the context of UV-B radiation adaptation. I also validated that the sequences are now available in the China National GeneBank DataBase.

(Remarks on code availability)

Reviewer #2

(Remarks to the Author)
I reviewed the revised manuscript. The authors have done a nice job responding to reviewer comments from a previous version of this manuscript. For the most part, the result is a dense, thorough contribution expanding current understanding of the evolutionary history of complex hybrids within Elymus. Improvements to Figure 3 are particularly appreciated.

A couple of quick comments – on line 488 the authors use the term “pseudogeneticization” – I’m not sure this is a term. Perhaps change to “formation of pseudogenes.” In that same paragraph, there are several grammatical issues. This recently edited section needs to be carefully re-written.

(Remarks on code availability)

Reviewer #3

(Remarks to the Author)
My previous comments and questions have been properly addressed in the revised manuscript.

(Remarks on code availability)

We sincerely thank the editor and all reviewers for their valuable comments, which we have used to improve the quality of the manuscript. Reviewers' comments are listed in black font and changes/additions to the manuscript are listed in bold blue font. The changes in the manuscript were marked with red font.

REVIEWER COMMENTS

Reviewer #1

1. Is the number of triads of syntenic genes a relatively small number? How this compares to the sequenced hexaploidy bread wheat genome?

Reply: We have identified the 1:1:1 syntenic genes among the three haplomes, retrieving a total of 15,875 triads of syntenic genes in *Elymus nutans* (Extended Data Table 1). This number is nearly equivalent to that of hexaploid bread wheat, which contains 17,765 triads (*Ramírez-González et al., 2018*), suggesting that these figures reflect normal conditions. We have included the detailed number in the revised manuscript (Line 165).

Ramírez-González, R.H. et al. The transcriptional landscape of polyploid wheat. Science 361, eaar6089 (2018). DOI:10.1126/science.aar6089

2. How many pairs of genes can be found between the haplomes? Is there any pair of haplomes that are genetically closer?

Reply: A similar number of syntenic gene pairs was identified among the different haplomes of *Elymus nuans*, with 22,427, 19,176, and 21,750 syntenic gene pairs found between H and St, H and Y, and St and Y, respectively (Extended Data Table 1). The Ks values for each gene pair were calculated, revealing that the St and Y haplomes exhibited smaller average Ks values (Fig S5b). This finding is consistent with the phylogenetic analyses (Fig 2a), which indicate a closer relationship between St and Y compared to the H haplome. All the results have been included in the revised manuscript (Lines 131-133).

3. Finally, for those pairs of syntenic genes, is the third missing copy a pseudogene? If yes, in how many cases.

Reply: The number of genes absent in the third subgenome was counted based on the 1:1 syntenic gene pairs mentioned above, which revealed that 60.18%, 51.90% and 51.52% of the missing genes in the St, H, and Y subgenomes, respectively, were pseudogenized (Extended Data Table 2). Fig S7 shows the missing for the syntenic genes in St07 (*Enu.ptg000117l.G1432* and *Enu.ptg000117l.G1430*) and Y07 (*Enu.ptg000016l.G733* and *Enu.ptg000016l.G730*) that have undergone pseudogeneticization. All the results are included in the revised manuscript (Lines 167-172).

4. The only comment I have regarding the data analysis is the missed opportunity to delve deeper into the annotated genes (as for instance, checking for signatures of subfunctionalisation in the homeologous genes analysis).

Reply: The ka/ks ratios within syntenic gene pairs in *Elymus nutans* were calculated and showed no significant differences in the values between St vs H (0.325), Y vs H (0.318) and St vs Y (0.332) (Extended Data Table 4; Fig S16). In addition, more than 98% of gene pairs have a ka/ks ratio less than 1, indicating purifying selection. The gene ontology (GO) enrichment result showed that gene pairs with a ka/ks ≤ 1 exhibited functional conservation, primarily enriched in processes related to protein synthesis, processing and transport (Extended Data Table 5; Fig S17). However, for gene pairs with ka/ks > 1, suggesting positive selection, the genes in St vs H, St vs Y, and Y vs H exhibited varied gene function. Characterization of the classical star DEGs (with 1:1:1 gene triads) known to correspond to UV-B, cold and drought stress conditions revealed that most of these gene triads showed a divergent expression pattern (sub-functionalization). Interestingly, the "homologous gene function change (neo-functionalization)" was also found on specific star DEGs triads in response to these stress conditions (Extended Data Table 6). Above results indicate a preliminary sub-functionalization and neo-functionalization of genes within the *E. nutans* genome, which may have played an important role in the evolution and adaptation of *E. nutans*. All the results have been included in the revised manuscript (Lines 181-186, 488-507).

5. Regarding the data I could not validate the access to the raw sequence data that is deposited in China National GeneBank DataBase. The accession numbers for the datasets are visible but result in a message indicating that the raw sequencing data will not be available until 25 November 2025. Unfortunately, these datasets will need to be available for public access before I could recommend this manuscript for publication.

Reply: We have released the uploaded data in China National GeneBank DataBase, now all sequencing raw data and the assembled genome sequence are available.

6. Typos

Line 117 "sytenic relationships" should be "syntenic relationships"

Line 174 "Dasypyrum Villosum" should be "Dasypyrum villosum"

Line 309 "clod stress" should be "cold stress"

Reply: We have amended "sytenic relationships" into "syntenic relationships" at line 127, "*Dasypyrum Villosum*" into "*Dasypyrum villosum*" at line 208, "clod stress" into "cold stress" at line 352, respectively.

Reviewer #2

1. The model showing reticulate evolution in *Elymus* (Fig 3C) is difficult to interpret. At the very least, a more detailed caption is needed. And it would be even better to have more clarity within the figure.

Reply: Thank you for your question. We have revised the reticulated evolutionary model of *Elymus* allopolyploid species involving the St, Y, and H subgenomes in Fig 3C. Specifically, different proportions of the background color in each rectangle represent lineages derived from *Hordeum* (H), *Pseudoroegneria* (St), or an unknown (Y) diploid donor species. The color of the rectangular border corresponds to the genome composition and ratio of the different polyploid *Elymus* species, including *Elymus* s.s. (StH, StStH and StHH), *Roegneria* (StY, StStY and StYY), and *Campeiostachys* (StYH). The arrow points from the donor species to the recipient species, with the dotted arrow representing the undetermined donor species. All the results have been included in the revised manuscript (Lines 288-294).

2. While the data are solid, I felt the introduction and discussion, and to a lesser extent the results, failed to make the most of the work presented. In some cases, the writing is a little choppy ("perennial" is used three times in the first five sentences). I think there is a gap between the genome assembly/evolutionary analyses, and the local adaptation/GWAS studies. The introduction offers an opportunity to highlight the novelty of combining these approaches in a single study, and the discussion presents an opportunity to more deeply consider the role of polyploidization in high-altitude evolution. For example, specifically which subgenomes within *E. nutans* were most important in the adaptation to UV-B and other aspects of the environment? Perhaps this is buried in results, but it could be more pronounced. How has reticulate evolution facilitated local adaptation? Were the same gene regions underlying stress tolerance selected following distinct allopolyploid origins, or different ones? If this is beyond the scope of this paper, that should be more clearly stated.

Reply: We appreciate your thoughtful review and constructive feedback. We agree with your suggestions and have revised the entire manuscript including the introduction (Lines 81-83, 86-90), results (Lines 334-336) and discussion (line 488-507) to makes the logic smoother.

We have performed extensive analyses to assess the contribution of each subgenome to the environmental adaptation. First, the transcriptome results showed that the number of gene triads in the seven homologous expression patterns demonstrated a negligible difference before and after stress (Fig S12a). Further analysis showed that the majority of the gene triads showed balanced expression (54.34%~62.69%), while genome dominance and repression effects between subgenomes were not significant (Fig S8-S13; Table S5-S10). In addition, no significant difference ($P > 0.05$, chi-square test) was detected in the number of up- and down-regulated genes for each subgenome across all abiotic stress conditions (Fig S12b). The GO enrichment result based on the differentially expressed genes (DEGs) also showed no clear functional bias between the subgenomes under each stress condition (Fig S14). What's more, the co-expression network did not detect a unique module divergence of a specific subgenome under each stress, meaning that co-expressed homeologues were evenly distributed across gene triads (Fig S15; Extended Data Table 3). We also assessed the average ka/ks values for these gene triads, which revealed no significant differences in the ka/ks values between the subgenomes of *E. nutans* (Fig S16). In summary, the above results suggest that the expression profiles of the three subgenomes in *E. nutans* are relatively balanced with no clear subgenome

dominance. Similarly, the observed non-subgenome dominance pattern was also observed in the alloploid species such as *Avena sativa* (Kamal et al., 2022), *Brassica napus* (Chalhoub et al., 2014), and *Triticum aestivum* (Pfeifer et al., 2014; Harper et al., 2015) using the same homoeologue method.

We have characterized the classical star DEGs (with 1:1:1 gene triads) known to correspond to UV-B, cold and drought stress conditions to explore how reticulate evolution may have facilitated local adaptation in *E. nutans* (Extended Data Table 6). The result showed that most of these DEGs triads showed a divergent expression pattern. Interestingly, we observed the "homologous gene function change" phenomenon on specific gene triads such as the *RUG3-UVR8-HERC2* (St-H-Y) shift, the *BHLH56-PIF3-BHLH119* (St-H-Y) shift, and the *FACR5-FAR-FACR5* (St-H-Y) shift in response to UV-B, cold and drought, respectively. These results suggest that the sub-functionalization and neo-functionalization appear to be occur between the three subgenomes. In addition, most of the associated genes or DEGs underlying stress tolerance were located in distinct gene (syntenic) regions (i.e. these genes usually don't belong to the 1:1:1 gene triads) (Extended Data Table 7). In short, all subgenomes in *E. nutans* play an important role in coping with various adversities, and their systemic cooperation formed during reticulate evolution, has facilitated its local adaptation to high altitudes. All the results have been included in the revised manuscript (Lines 131-133, 187-199, 405-407, and 488-507).

Kamal, N. et al. The mosaic oat genome gives insights into a uniquely healthy cereal crop. Nature 606, 113–119 (2022). https://doi.org/10.1038/s41586-022-04732-y

Chalhoub, B. et al. Early allopolyploid evolution in the post-Neolithic Brassica napus oilseed genome. Science 345, 950-953 (2014). DOI:10.1126/science.1253435

Pfeifer, M. et al. Genome interplay in the grain transcriptome of hexaploid bread wheat. Science 345, 1250091 (2014). DOI:10.1126/science.1250091

Harper, A.L. et al. Genome distribution of differential homoeologue contributions to leaf gene expression in bread wheat. Plant Biotechnology Journal 14, 1207-1214 (2015). https://doi.org/10.1111/pbi.12486

Reviewer #3

1. In Fig S19, flavonoids and total anthocyanins should also be quantified since both are important for UV-B stress tolerance.

Reply: We agree with your suggestions and have measured these two indicators flavonoids (Flavo) and total anthocyanins (TA). The results are presented in Fig S24 and Extended Data Table 11. All the results have been included in the revised manuscript (Lines 429-430).

2. Photoreceptor *UVR8*-mediated UV-B acclimation pathway promotes UV-B stress tolerance in *Arabidopsis*. Is there any relationship between UVR8 pathway and the reported *MAPKKK18* gene in UV-B stress tolerance?

Reply: Thank you for your question. Indeed, the photoreceptor *UVR8*-mediated UV-B resistance is one of the most important regulatory pathways in plants. However, the *MAPK* signaling pathway (for example, the *MKP1* - *MPK3*/MPK6 pathway), regulated by its upstream *MAP3Ks* and *MAP2Ks*, is another conserved UV-B response pathway that is independent of *UVR8* in the plant response to acute UV-B stress (Chen et al., 2022; Gonzάlez Besteiro et al., 2011). In addition, Gonzάlez Besteiro et al. (2011) showed that the MAPK and *UVR8* pathways can be genetically separated: the MAPK pathway is hypersensitive to acute UV-B stress but *UVR8* is not; *UVR8* is impaired in the UV-B acclimation response, but MAPK pathway is not. In agreement, *MAPKs* is not altered in the process of UV-B-induced photomorphogenesis, and *UVR8* does not directly affect the stress-induced MAPK pathway (Gonzάlez Besteiro et al., 2011). We have applied the *Arabidopsis thaliana* wild-type (WT) and *MAPKKK18* over-expressed (OE) lines to UV-B stress and measured their gene expression levels by RNA sequencing under 0 day (CK), 3 days, and 10 days with three biological duplications. The result showed that there is no significant difference for the expression levels of *UVR8* genes between the OE lines and WT in all treatments, indicating that the regulatory pathway of *MAPKKK18* might be independent of that of *UVR8* (Fig S25; Extended Data Table 11). All the results have been included in the revised manuscript (Lines 436-439 and 851-852.

Chen, Z. et al. Plant responses to UV-B radiation: signaling, acclimation and stress tolerance. Stress Biology 2, 51 (2022). https://doi.org/10.1007/s44154-022-00076-9

Gonzάlez Besteiro, M.A. et al. Arabidopsis MAP kinase phosphatase 1 and its target MAP kinases 3 and 6 antagonistically determine UV-B stress tolerance, independent of the UVR8 photoreceptor pathway. The Plant Journal 68, 727-737 (2011). https://doi.org/10.1111/j.1365-313X.2011.04725.x

3. The UV-B conditions used in Figure 6h is not clear. Please provide more detailed information here.

Reply: The UV-B treatment condition used for the WT, mutant and overexpression lines of *Arabidopsis thaliana* was under a wavelength of 311 nm and an intensity of 1,200 • mw • m^{-2}. We have added the detailed information in the "Methods: Overexpression and mutant lines construction of *M3K18* in *Arabidopsis thaliana*" section (Lines 847-850).

4. OE-*MAPKKK18* lines show very low chlorophyll contents as shown in Fig S19 A-B, however, the OE lines shown in Figure 6h appear to develop green leaves as WT *Arabidopsis* plants. Please comment on these observations.

Reply: As shown in Fig 6h, there are significant differences in the degree of yellowing and the proportion of green leaves between WT and OE lines after 20 days of UV-B stress. The Fig S24a-e showed the indicators related to UV-B response after 20 days of UV-B treatment, corresponds to the growth status in the bottom left of Figure 6h. After UV-B treatment, although the old leaves still retained a light green color (bottom left of Fig 6h), by this time the OE lines were already wilting and unable to grow young new leaves. In contrast, after 20 days of UV-B treatment, the WT line was still keep in a relatively robust state and showed no significant wilting or yellowing (bottom left of Fig 6h). As a result, we observed significant differences in chlA and chlB between WT and OE lines after UV-B stress (Fig S24a, b). We have added the measurement status of the indicator (i.e. the measurement after 20 days of UV-B exposure) in the legend of Fig S24 to make it easier for the reader to understand.

We sincerely thank the editor and all reviewers for their valuable comments, which we have used to improve the quality of the manuscript. Reviewers' comments are listed in black font and changes/additions to the manuscript are listed in bold blue font. The changes in the manuscript were marked with red font.

REVIEWER COMMENTS

Reviewer #2

1. A couple of quick comments – on line 488 the authors use the term "pseudogeneticization" – I'm not sure this is a term. Perhaps change to "formation of pseudogenes." In that same paragraph, there are several grammatical issues. This recently edited section needs to be carefully re-written.

Reply: We have revised the incorrect term "pseudogeneticization" to the accurate term "pseudogenization" in line 514. The entire paragraph has been thoroughly reviewed and carefully refined to correct any spelling or grammatical errors, ensuring clarity and precision in our presentation.