nature portfolio

Peer Review File

Genomes of wheat wild relatives Thinopyrum intermedium and Roegneria kamoji reveal different polyploid evolution paths

Corresponding Author: Dr Silong Sun

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)
T. intermedium and R. kamoji are two important species that breeders frequently used to transfer important genes for disease resistance and other traits into wheat through distant hybridization to enrich wheat gene pool. However, their genome compositions and evolution relationships with other members in Triticeae family remain ambiguous. This manuscript addressed these issues and found that T. intermedium possesses StStVVDD genome and R. kamoji varies StStVVHH. The authors conducted genome assembly by use of the long-read sequencing technology and confirmed T. intermedium does not carry E genome as previously reported. Although both T. intermedium and R. kamoji carry StStVV genomes, but they might have different origins. They found that both species carry Fhb7 homologs that showed resistance to FHB. Using whole genome sequence analysis to determine genome composition of wheat relatives is very interesting. Comparative analysis of genome sequences among different spaces can provide more accurate information on their genome composition and evolution history than using classical cytological methods. The new findings in this manuscript provide further insight into the genome evolutions of the two species. The genome assemblies for the two species will provide important genetic resources for identification and cloning of the disease resistance genes from the two species, therefore the information reported in this paper will benefit cereal crop researchers worldwide. The paper is well written, and the experiment was well designed and conducted using appropriate materials and methods. Data analysis and interpretation are appropriate. Methods were described with necessary detail.
One suggestion for title: High-quality genome assemblies of wheat wild relatives Thinopyrum intermedium and Roegneria kamoji revealed different polyploid genome evolution paths

Reviewer #2

(Remarks to the Author)
To explore the genomic composition of Thinopyrum intermedium and Roegneria kamoji, theoretically all known diploid species should be included except the genome used in the article, e.g. Australopyrum species (W genome), Agropyron species (P genome), Psathyrostachys species (Ns genome), Peridictyon species (Xp genome), Taeniatherum species (Ta genome). If the research does not include all the diploid genomes of Triticeae, the conclusions drawn may be incorrect.
In Figure 2C, the V01-07 mapping signals of V mapping to R. kamoji are not very strong, while the mapping signals of other genomes are strong. How did the author obtain the conclusion that R. kamoji contain the V genome?
In conclusion, the study reports a genome constitution of StStVVDD for Th. intermedium and StStVVHH of Roegneria kamoji. The genome reports of these two species are new and different from previous reports. It is suggested that the authors need more experiments and data to support them.

Reviewer #3

(Remarks to the Author)
In this study, Sun and colleagues present high-quality assemblies of two Triticeae species, Thinopyrum intermedium and Roegneria kamoji. Both species have been used in wide crosses with wheat to introduce novel traits. The two assemblies have been generated with the latest sequencing technologies (PacBio HiFi and Hi-C), matching the expected standards in

the field. Phylogenetic analyses allowed sub-genome designations for the two hexaploid species. In addition, the authors identify and validate novel alleles of the Fhb7 Fusarium resistance gene that was initially described in Thinopyrum elongatum. The manuscript is well written, and the results are presented in a logical order. Below I list some comments that the authors should consider:

Major comments:

• Selection of accessions: How were the two accessions selected? As with any wild plant material, a major issue is to validate that the species designations in the genebanks were correct? It is widely accepted that about 10% of wild genebank material has been wrongly classified on the species level. This is particularly relevant for this study because the authors report some 'surprises' in the subgenome compositions of Th. intermedium and R. kamoji. How did the authors make sure that the two selected accessions represent 'typical' representatives of their species? Critical information about accession selection and species validation is lacking in the methods section. I could not find any information on R. kamoji accession Pr 87-88 353 in Genesys. From which genebank does this accession originate? From the methods section, it appears that the sequenced accessions were not obtained from the respective genebanks directly, but through intermediary labs. This always involves the risks of seed mix-up. How can seeds of the two sequenced accessions be obtained by the international research community?
• Heterozygosity rate: Thinopyrum intermedium is described as a self-sterile, outcrossing species. The heterozygosity rate of 1.35% (Figure S1) seems to be low for an obligate outcrossing species. Do the authors have any explanation for this?
• Fhb7: There is a major discrepancy regarding the origin/distribution of the Fhb7 gene with previous literature. In their 2020 paper (Wang et al. (2020), Science 368, eaba5435), the authors concluded 'this FP-HGT event apparently occurred after the divergence of the E genome from Triticum sp. but before the formation of the decaploid Th. ponticum'. This conclusion was based on the reported absence of Fhb7 orthologs across Triticum species. Here, the authors state 'These findings support the hypothesis that the horizontal gene transfer of Fhb7 occurred prior to the divergence of the Triticeae family, resulting in the integration of a small genomic fragment into the plant genome.' This conclusion is very different from the 2020 paper, implying that Fhb7 was introduced much earlier and occurs across a wide range of species. This discrepancy needs to be addressed. Why did the authors miss the Fhb7 orthologs in their 2020 study? Have the authors validated that Fhb7 orthologs are absent from Pooideae tribes other than Triticeae?
• Figure 3B: The Fielder spike looks resistant to me, while the T3_Fhb7RkaH looks susceptible.
• Data availability: The two accessions numbers provided in the manuscript are not active yet, so I was unable to validate the deposited data. The data availability statement indicates that only assembly and annotation data have been deposited. What about the raw PacBio and Illumina reads? These also need to be deposited and made available. I also suggest to deposit the data in NCBI or ENA.

Minor comments:
• Line 57: The term 'elite gene' sounds strange to me. Maybe 'superior allele' might be more appropriate.
• Figure 3, legend: invocated or inoculated?

Reviewer #4

(Remarks to the Author)
The manuscript by Sun et al. presents the whole-genome sequence of two allohexaploid grass species from the tribe Triticeae, including an outline of the possible origin of these species. They also discuss the analysis of the FHB resistance of both species, which is controlled by the Fhb7 gene.
Overall, my assessment of the manuscript is positive, as it contains interesting and important information backed by a large amount of work and data. The data are plausible and their interpretation is, in my opinion, mostly correct. The main output of the work, the whole genome sequences of these two allohexaploid species, will be a significant step forward in wheat breeding.
What I have a little problem with is the selection of genotypes for sequencing and their characterization. It would be good to characterize the reference genotypes in more detail. Where the plants were collected, what the climatic conditions are there, whether they have been used in previous work and whether they have been characterized in any way, either morphologically or genetically. Both species have large ranges and are certainly not genetically uniform, so it would be good to know which populations the reference genomes actually represent. I also wonder about the selection of genotypes for resequencing. Why were genotypes used from the USA (5 genotypes) and Australia (1), where the species is not native? I could still understand it if (1) these were genotypes used in wheat breeding and (2) the primary goal of the work was the whole genome sequence of genotypes used in wheat breeding. Please explain.
Another problem is the interpretation of the genomic composition of the two hexaploids. Phylogenetic analysis depends on the input data, especially the representation of diploid Triticeae species. This was far from complete in this case. I understand that in this case it was necessary to use only those species for which whole-genome sequences are available, because otherwise it would probably have been impossible to obtain sequence data for 129 genes. But then you have to take into account the incomplete sampling and be very cautious in interpreting the data.
Furthermore, I wonder whether the authors also performed a chloroplast genome assembly. If so, identification of the maternal genome would provide an important clue to the origin of the two hexaploids.
Another comment is more technical, namely that the figures (especially the legends and axis labels) are harder to read throughout the manuscript and the legends are not self-explanatory.
Further specific comments are:
l. 48. In species names, authority should be given at first mention in the manuscript. Please check this throughout the manuscript (l. 60, 61 etc.).

l. 73. ‘Genomic in situ hybridization analyses...’ This sentence is a bit misleading because if GISH is the only marker used (is that the case with ref. 14?), it does not give independent results (the results depend on the probes used). It would be more accurate to write that the study, based solely on in situ hybridization, have designated the genomic formula... In fact, GISH has also been used in other papers to confirm hypotheses generated from sequencing data, and has shown different results to those mentioned by the authors in this sentence.
Btw, journal title is missing in ref. 14 (l. 796).
l. 85. The sentence ‘While Pseudoroegneria is a well-known diploid ancestor...’ does not describe the situation. How about 'While the involvement of Pseudoroegneria in the origin of Th. intermedium has been demonstrated, the origin of the other subgenomes has not been reliably identified'?
l. 90. If species distribution is mentioned for Roegneria, then it would be good to mention it for Thinopyrum as well.
l. 92-93. It might be good to mention that Roegneria is exclusively a polyploid species (as opposed to Thinopyrum where diploids occur).
l. 92. In references 20, 23, 25 and 26 the representatives of the genus Elymus are mentioned. Is Elymus an older name for current Roegneria? It would be useful to briefly explain what the relationship between the genera Roegneria and Elymus is (in terms of taxonomy), and which designation is actual.
l. 93-94. How many tetraploids and hexaploids are there?
l. 112-119. This part starting with ‘Our findings...’ can be omitted here as it contains results. It should be replaced by hypotheses and aims of the study.
l. 122. Is there a particular reason for choosing these genotypes? Both species are characterized by a large geographic range and, we can assume, corresponding genetic diversity. In Th. intermedium, six accessions were used for resequencing, but none in R. kamoji. The question arises as to what parameters were used to select the reference genomes for sequencing and to what extent they represent the genetic diversity of the species. Why were additional genotypes not used for resequencing in Roegneria?
l. 123. ‘genome size’ is not the correct term here. DNA content per 1C should be given instead. In addition, the value given for Th. intermedium is quite low. The 2C-DNA database gives higher values (from 13 to 13.5 pg, which is equivalent to between 12.7 and 13.2 GB). How was the genome size estimated? No details are provided in M&M section.
l. 124. Figures S1 and S2 refer to the morphology of the species, but morphological details are not visible in them. I suggest mentioning the key morphological differences in the text of the Introduction.
l. 131. 97.7 and 99.4 % of sequence data were assambled into pseudomolecules. It would be useful to mention what amount of data (in MB or GB) was left as unplaced contigs.
l. 135 (744). Table S4. What do the letters C, S, D etc. stand for? Please explain in table heading.
l. 145 (664). The framed legend is illegible. It should be enlarged.
l. 150 (666 and 673), Figure S6 and S7. Legends (colour scales) are missing.
l. 151 (751). Table S8. Pse. Libanotica in table heading should be corrected to Pse. libanotica. Same in Ae. Tauschii and Pse. Libanotica in the table (first row).
l. 155 (745). The header should explain what LAI means and how it is calculated.
l. 157-159. It’s not accurate to talk about potential progenitors here. This designation is based on the phylogenetic analysis, which depended on the reference genomes used (Fig. 2A). Only a few species were represented in each genus. For example, Hordeum is represented by only two species, while the entire genus contains 31 diploids (with 4 distinct genomes). Therefore, it would be more accurate to say that the TE annotation was also performed on selected representatives of genera whose members are likely to have contributed to the origin of the hexaploids studied. The result of the comparative analysis should also be briefly commented on (not just refer to table S8).
The other thing is that since the possible progenitors of the polyploids Thinopyrum and Roegneria are mentioned here, but the identification of the individual subgenomes is not described until the next chapter, it might be more logical to place the Molecular phylogeny chapters (both) earlier, after line 142. A paragraph beginning with the sentence 'Transposable element...' would follow the Molecular phylogeny chapters.
l. 167 (754). Table S10. ‘Ratio’ on the third row could be replaced with ‘[%]’ or ‘Ratio [%]’
l. 179-182. Mapping of St reads on Thinopyrum chromosomes shows gaps at the right end of the chromosomes. While reads from other diploids map end-to-end on the corresponding chromosomes, this is not the case for St reads on the St genome. Is there an explanation for this pattern?
l. 182 (696 and 700). Fig S10. replace Ae. Tauschii with Ae. tauschii (twice).
l. 182. not Figure 2B but Figure 2C includes that analysis.
l. 182-184. ‘Notably, the ThinD subgenome was clustered between wheat A lineages and other Aegilops-clade species, diverging approximately 3.4 million years ago.’ In phylogenetics, it is not appropriate to say that a species is clustered between some clades. To be correct, the ThinD subgenome is sister to the D- + S-genome Aegilops species (according to what I quickly found in Feldman's classification for Ae. searsii and other species. Feldman, M., Levy, A.A. 2023. Aegilops L. In: Wheat Evolution and Domestication. Springer, Cham. https://doi.org/10.1007/978-3-031-30175-9_9.). To be more accurate in identifying the ThinD subgenome, a more comprehensive representation of Aegilops species would be needed. Therefore, the designation of this subgenome as ThinD should be considered tentative. The text should be corrected.
l. 184-189. Since the closest relatives of the ancestors of the hexaploids are not known, the comparison of the TE content of the hexaploids with a particular diploid (shown in Fig. S9) makes little sense. I would omit this analysis except for Thinopyrum vs. Roegneria comparison.
l. 190-192. Figure 2B. How was the synteny analysis done? There are no details provided in methods section.
l. 196. ...‘previously proposed as an ancestral donor.’ A reference(s) should be added.
l. 200 (756). Table S11. correct Th. Intermedium to Th. intermedium. Which gene bank did these accessions come from?
l. 213-221. Based on the analyses performed, it appears that the ancestor is indeed Dasypyrum. However, it should be noted that the phylogenetic analysis is far from including all possible representatives of Triticeae. It is therefore theoretically possible that the ancestor is another species related to Dasypyrum that was not included in the analyses. More detailed analyses in this direction would be useful in the future, and it would also be fair to discuss this possibility in the text.

l. 220 (710). Fig S12. not blue but green lines represent the translocation blocks.
l. 232-234. It is also possible that each polyploid arose from a different (not sampled here) St ancestor.
l. 235. ... 'a specific inversion in the V03 subgenome'...why V03? Wouldn't be enough V subgenome?
l. 237. Is R. ciliaris the only known tetraploid, or are there more candidates?
l. 240. The involvement of genomic in situ hybridization would confirm the identity of the subgenomes and the presence/absence of translocations, and greatly increase the credibility of the results.
l. 244-246. ...'Because St genome...shares a closer evolutionary relationship with Dasypyrum than with Aegilops, it is likely that the ThinSt and ThinV subgenomes first combined to form a tetraploid species.' I don't understand what is meant by the 'closer evolutionary relationship with Dasypyrum than with Aegilops'.
l. 260 (714 and 717). Fig S13. Instead of the term comparison, the term alignment would be more appropriate.
l. 260. I think that synteny is not the right term for an analysis that deals with a structure of a single gene. It is just a comparison of the gene region sequences.
l. 262 (394). Fig 3A. 'Infected' instead of 'invocated'? Next, in panels 3B and C, it would be good to indicate which spikes are supposed to be sensitive and which insensitive (transgenic).
l. 266. Transcriptome with lowercase t.
l. 268 (720). Fig S14. Please explain what is on y axis.
l. 271 (723). Fig S15. The figure as it stands doesn't say much and could be omitted.
l. 279. Could you briefly describe what deoxynivalenol-glutathione is good for? What is the principle of the resistance?
l. 284. please correct Fhb7RkaSt (missing t)
l. 293 (393). Fig 3E. The figure legend should be self-explanatory. It should be explained what is on x axis etc.
l. 305. The expression...'with contributions from the present-day genera'... may give the false impression that the hexaploid was formed recently with the contribution of present-day species. It would probably be better to write 'with contributions from progenitors corresponding to present-day genera...'
l. 307. ...suggesting that DIPLOID Thinopyrum species...
l. 309. The contribution from Aegilops have been suggested in other papers. It would be good to add reference(s).
l. 325. How does genomic stability relate to the hybridization ability with wheat? Many genomes are stable but don't cross with wheat.
l. 332-333. ...'but not in annual species'... Have the annual species been tested? If not, then this statement is unsupported.

Materials and Methods
Overall, it seems to me that the methodology could be more detailed.

l. 409. Please add a reference for CTAB method.
l. 433. subjected instead of subjection
l. 468-469. Accession numbers for the diplois should be provided.
l. 476. What substitution model was used to calculate divergence K?
l. 480. Please add a reference paper to the TE-specific mutation rate used.
l. 523. A sentence explainig why the phylogenetic analysis was performed would be useful. The same should be stated for the other chapters.
l. 529. Please add accession number for Ps. stipifolia. What tissues were used for RNA isolation? Please add details on how the RNA seq was done.
l. 532. In total, 129 single-copy genes were identified... I wonder if only 129 common genes were found for the selected taxa, or if there were more and 129 represents only a subset (in that case I would be interested in the selection criteria).
l. 537 and 553. As noted above, a sentence explaining the purpose of the analyses should be added to each chapter.
l. 565. Please provide the primers' sequences.
l. 570-571. What cultivar of wheat was used? How many plants of each line were used for the experiment?
l. 584. Which genotypes of Thinopyrum and Roegneria were used? Same as for the sequencing? Regarding wheat lines, were both Fusarium susceptible and transgenic lines used? Please add details in the text.
l. 586. I don't understand the term one-heart stage.
l. 593. Add a sentence describing the purpose of the analysis.

Version 1:

Reviewer comments:

Reviewer #2

(Remarks to the Author)
The author carefully revised the reviewer's questions in the revised manuscript, and I have no other questions. Please supplement and improve a few precautions:
1. The origin time of Thinopyrum intermedium and Roegneria kamoji should be predicted based on the results of whole genome sequencing in this paper.
2. In supplementary Fig. 10, Where does the data for Pse. stipifolia?
3. In Genome in situ hybridization experiments, if a species genome is used as a probe, the probe name should be V-genome instead of Oligo-pDb12H. Suggest making modifications in the text and images.
4. How the Oligo-Ae584 probe was obtained should be explained in the methods section.
5. Triticeae does not italicize in the text.

6. In supplementary Fig. 14, “Ps. Spicata” should be “Pse. spicata”
7. In Line 49, “Triticum aestivum Lam.” should be “Triticum aestivum L.”
8. Genomic symbols, if not bolded, the entire text will not be bolded. Please check.

Reviewer #3

(Remarks to the Author)
The authors have adequately addressed my comments. I particularly appreciate the additional information on the origin, management, and validation of the sequenced accessions.

Reviewer #4

(Remarks to the Author)
The revision of the manuscript increased the credibility of the results, especially the identification of the ancestors of the studied hexaploids by phylogenetic analyses. I welcome the use of cytogenetic methods to verify genomic composition, but I have reservations about this section, both in terms of the methodology used and the interpretation. Specific comments are mentioned below.
l. 46. The authority for Triticum aestivum is L. Please check the International Plant Names Index (IPNI) for species names and their authorities.
l. 62. According to IPNI, Roegneria kamoji (Ohwi) Ohwi ex Keng is the correct name for R. kamoji.
l. 72-76. What source is the species‘ distribution from? Th. intermedium is also native to a major part of Europe (ca from Spain to the Ural mts.). Please check and add reference(s). Add a reference for the morphological description, too.
l. 92. Dasypyrum breviaristatum (H. Lindb.) Fred.
l. 144-148. Keep in mind that these are 1C values. It should be specified, e.g., on l. 147 as follows:...‘flow cytometry measurements (11.84 Gb and 11.16 Gb per 1C, respectively)’...
l. 169-170. The sentence sounds weird. I suggest rephrasing, e.g., Transposable element (TE) annotation revealed that TEs occupy 77.15% and 78.49% of the genome sequences in Th. intermedium and R. kamoji, respectively.
In this context, I don’t understand the category ‘Transposons’ in supplementary Figs. S6 and S7 (first line and first column in the tables). It has no values and could be omitted in both tables.
l. 189. replace ‘transposable elements’ with ‘TEs’
l. 212-213. This interpretation is incorrect. The Thinopyrum sequence does not cluster with the D-clade of Aegilops. Instead, it is a sister species to a clade comprising species ranging from Ae. sharonensis to Ae. tauschii. As the clade can rotate at the node, ThinJ is as closely related to Ae. tauschii as to Ae. sharonensis. The analysis shows that ThinJ is not closely related to any of the Aegilops species present today (as expected), but it is likely to come from one of the species included in the aforementioned sister group comprising various Aegilops species (node XI). If we looked deeper into the three, we would see that the ThinJ sequence falls within the paraphyletic clade comprising the Aegilops and Triticum genomes (node IX) . In conclusion, I suggest rephrasing the sentence as follows (or something similar):
...‘yielded concordant topologies. They revealed that the subgenomes of Th. intermedium clustered with the St-clade of Pseudoroegneria, the V-clade of Dasypyrum, and a heterogeneous clade of Aegilops species, which we designate as the ThinSt, ThinV, and ThinJ subgenomes, respectively.

Finally, it would be good to include more information in the legend of Fig. 2a. For example, it would be useful to specify what type of phylogenetic analysis the tree shows. It would also be desirable to add bootstrap values to the tree. The same applies to Supplementary Fig. 10.

L. 240–251. I welcome the inclusion of cytogenetic verification. However, I am not convinced by it. Specifically, I am missing information about how the probes were chosen and how their specificity was verified.
Th. intermedium: I am no expert in cytogenetics, but the assignment of individual chromosomes to particular progenitors seems somewhat arbitrary. Have you tried hybridizing genomic DNA probes from Aegilops and Dasypyrum as well? Mapping of reads (Fig. 2c) suggests that GISH with these probes could work (since both mapping of reads and GISH show repeatome homology).

R. kamoji: GISH (Fig. 15a) – Why not hybridize a total genomic DNA probe from Hordeum? Mapping of Hordeum reads shows the greatest homology (suggesting Hordeum is indeed a progenitor), so this experiment would help to identify the Hordeum chromosomes first. The remaining chromosomes could then be analyzed using oligo-specific (or genomic) probes.
Fig. 15b and d. Are you sure that the oligo-pTa535 is Hordeum-specific as stated in the legend? Its name suggests that it is Triticum aestivum-specific. If it was Hordeum-specific, both probes should hybridize on the same chromosomes.
Fig. 15c. The figure is confusing. Why is there a signal on H chromosomes when the probes are St- and V-specific??
I suggest that the authors clarify which probes were used, what their specificity was, and what sequence they target. To make the use of probes reliable, a reciprocal cross-hybridization test should be performed on all diploids (St, V, J).

l. 257. Please add the bootstrap values to the tree in Fig. 2a. If clade XII is supported, the V-genome ancestry conclusion is correct.

l. 297 (Suppl. Fig. 17). Pleaes add all necessary details to the phylogenetic analysis so that the legend is self-explanatory.

l. 304-305. No, there is no such close relationship. Pseudoroegneria is clearly the maternal lineage. Omit this part from l. 304 to 311 and do not complicate the story.

l. 396-400. In line with the previous comment, I suggest deleting this part (starting with ‘whereas...’ on l. 395).

l. 515. petri dish with uppercase P

l. 650. Aegilops caudata

l. 655. Pse. cognata

l. 663. Secale cereale

l. 686. high-molecular with lowercase h

l. 911. predicted instead of predicated

Version 2:

Reviewer comments:

Reviewer #2

(Remarks to the Author)
The authors have made careful revisions based on the reviewer's questions, and I have no further comments.

Reviewer #4

(Remarks to the Author)
In terms of cytogenetic analyses, it is evident that identifying chromosomes based on in situ hybridization is more challenging than we would like. This may be because both polyploid species have undergone so many chromosomal changes compared to the current diploid species that unambiguous identification is no longer possible. I have no further comments on the revised version of the manuscript and wish the authors every success in their future research.

-delete multiple refs. [42] on l. 633

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

T. intermedium and R. kamoji are two important species that breeders frequently used to transfer important genes for disease resistance and other traits into wheat through distant hybridization to enrich wheat gene pool. However, their genome compositions and evolution relationships with other members in Triticeae family remain ambiguous. This manuscript addressed these issues and found that T. intermedium possesses StStVVDD genome and R. kamoji varies StStVVHH. The authors conducted genome assembly by use of the long-read sequencing technology and confirmed T. intermedium does not carry E genome as previously reported. Although both T. intermedium and R. kamoji carry StStVV genomes, but they might have different origins. They found that both species carry Fhb7 homologs that showed resistance to FHB. Using whole genome sequence analysis to determine genome composition of wheat relatives is very interesting. Comparative analysis of genome sequences among different species can provide more accurate information on their genome composition and evolution history than using classical cytological methods. The new findings in this manuscript provide further insight into the genome evolutions of the two species. The genome assemblies for the two species will provide important genetic resources for identification and cloning of the disease resistance genes from the two species, therefore the information reported in this paper will benefit cereal crop researchers worldwide. The paper is well written, and the experiment was well designed and conducted using appropriate materials and methods. Data analysis and interpretation are appropriate. Methods were described with necessary detail.

Response: We sincerely appreciate the reviewer's comprehensive evaluation and encouraging comments on our manuscript. We are grateful for the acknowledgment of our work in clarifying the genome compositions and evolutionary relationships of *Thinopyrum intermedium* and *Roegneria kamoji*, which have long remained ambiguous despite their significance in wheat breeding. We are also delighted that the reviewer emphasized the value of identifying *Fhb7* homologs in these species, which highlights their potential as valuable genetic resources for enhancing *Fusarium* head blight resistance in wheat. We hope the genome assemblies of these species will serve as critical resources for the global cereal research community, facilitating the cloning and functional characterization of agriculturally important genes.

One suggestion for title: High-quality genome assemblies of wheat wild relatives *Thinopyrum intermedium* and *Roegneria kamoji* revealed different polyploid genome evolution paths

Response: Thanks for the helpful comment, we have changed the title as suggested.

Reviewer #2 (Remarks to the Author):

To explore the genomic composition of *Thinopyrum intermedium* and *Roegneria kamoji*, theoretically all known diploid species should be included except the genome used in the article, e.g. *Australopyrum* species (W genome), *Agropyron* species (P genome), *Psathyrostachys* species (Ns genome), *Peridictyon* species (Xp genome), *Taeniatherum* species (Ta genome). If the research does not include all the diploid genomes of Triticeae, the conclusions drawn may be incorrect.

Response: We sincerely appreciate the reviewer's helpful comments regarding the necessity of including all relevant diploid *Triticeae* species in elucidating the genomic composition of *Thinopyrum intermedium* and *Roegneria kamoji*.

Following Löve's *Triticeae* taxonomic system (Löve, Á. 1984. Conspectus of the *Triticeae*. *Feddes Repertorium* 95: 425–521), we have now incorporated sequencing data from 21 of the 24 recognized diploid/polyploid genera (A to Y) in *Triticeae* in our analysis. For genera without assembled genome sequences, we collected transcriptome or exome datasets from public databases which includes: *Aegilops* spp. (C, M, T and U), *Pseudoroegneria* spp. (St), *Eremopyrum triticeum* (F), *Crithopsis delileana* (K), *Psathyrostachys juncea* (Ns), *Henrardia persica* (O), *Agropyron cristatum* (P), *Heteranthelium piliferum* (Q), *Taeniatherum caput-medusae* (Ta), *Australopyrum retrofractum* (W), *Peridictyon sanctum* (Xp), and multiple *Hordeum* species (H, I and Xu genomes) (Supplementary Table 11). Some of these datasets were also previously used for *Triticeae* phylogeny analysis (Mason-Gamer, R. J., and D. M. White. 2024. *American Journal of Botany* 111(10): e16404). Two exceptions include the G genome in *Triticum timopheevii* (AG) and the N genome of *Aegilops uniaristata*, which were reported to exhibit close phylogenetic affinities with the wheat B-linage and the M genome of *Ae. comosa* respectively, were not included in the phylogenetic analyses. The third exception involves the *Festucopsis* (L) genus, for which no genomic datasets were publicly accessible.

The reconstructed phylogenetic tree based on genome-wide shared SNP loci among these *Triticeae* species, demonstrates a general consistency with prior single-copy orthologous genes based phylogenetic reconstructions regarding the genomic composition of *Th. intermedium* and *R. kamoji* (Fig. 2a). The Y subgenome in *R. kamoji* exhibited phylogenetic adjacency to V-clade of *Dasypyrum*, whereas one subgenome in *Th. intermedium* displayed closer affinity to *Aegilops* genus, indicating the absence of ancestral diploid donors for these two subgenomes within the reconstructed phylogeny. This conclusion is further supported by the genomic *in situ* hybridization (GISH) experiments utilizing V-specific probes that failed to yield definitive chromosomal differentiation in *R. kamoji*, reinforcing the inadequacy of the "V" genome designation for this evolutionarily distinct subgenome (Supplementary Fig. 15). Similarly, no clear GISH signals were observed for the *Aegilops*-related subgenome when using *Ae. tauschii* (D) as genomic probe, while clear FISH results were obtained using the repeat sequence 'Oligo-Ae584' conserved in wheat A, B and D subgenomes as probe (Supplemental Fig. 14). Consequently, to avoid any

unnecessary misleading, we revised the manuscript to retain the genome symbols 'Y' for the *Dasypyrum*-related subgenome in *R. kamoji* and 'J' for the *Aegilops*-related subgenome in *Th. intermedium*, respectively.

We thank the reviewer again for emphasizing this important methodological consideration and welcome further clarification if specific taxa/genomes of concern remain unaddressed.

In Figure 2C, the V01-07 mapping signals of V mapping to *R. kamoji* are not very strong, while the mapping signals of other genomes are strong. How did the author obtain the conclusion that *R. kamoji* contain the V genome?

Response: As the reviewer noted, the mapping signals of the diploid V genome (V01-07 chromosomes) to *R. kamoji* were indeed weaker than those of the other genomes. The reconstructed *Triticeae* phylogenetic tree (Fig. 2a) revealed that this Y subgenome diverged relatively earlier from the V clade (node age: ~6.0 MYA for Y) than the other V-related genomes. This suggests that the extant diploid V genome of *D. villosum* was not the direct progenitor of this subgenome. Considering the phylogenetic relationship, genomic mapping patterns and the cytogenetic results collectively, we have maintained the 'Y' designation to characterize this *Dasypyrum*-affiliated subgenome in *R. kamoji* throughout our revised manuscript, as above mentioned.

In conclusion, the study reports a genome constitution of StStVVDD for *Th. intermedium* and StStVVHH of *Roegneria kamoji*. The genome reports of these two species are new and different from previous reports. It is suggested that the authors need more experiments and data to support them.

Response: Following the reviewer's valuable suggestions, we have substantially expanded our experimental evidence by incorporating both genomic analyses and cytogenetic validation to reinforce the phylogenetic conclusions. Specifically, we have performed comprehensive phylogenetic reconstruction using sequencing data from most of the diploid *Triticeae* species to establish more clearly evolutionary relationships within the tribe; then we conducted GISH and FISH experiments to validate the genomic constitution of the two hexaploid species. These additions have significantly improved and strengthened the manuscript's foundational arguments regarding polyploid formation in these two species. We deeply appreciate the reviewer's insightful suggestions that prompted us to enhance these pivotal aspects of our study, ultimately elevating the scientific rigor and impact of this work.

Reviewer #3 (Remarks to the Author):

In this study, Sun and colleagues present high-quality assemblies of two Triticeae species, *Thinopyrum intermedium* and *Roegneria kamoji*. Both species have been used in wide crosses with wheat to introduce novel traits. The two assemblies have been generated with the latest sequencing technologies (PacBio HiFi and Hi-C), matching the expected standards in the field. Phylogenetic analyses allowed sub-genome designations for the two hexaploid species. In addition, the authors identify and validate novel alleles of the *Fhb7 Fusarium* resistance gene that was initially described in *Thinopyrum elongatum*. The manuscript is well written, and the results are presented in a logical order. Below I list some comments that the authors should consider:

Major comments:

• Selection of accessions: How were the two accessions selected? As with any wild plant material, a major issue is to validate that the species designations in the genebanks were correct? It is widely accepted that about 10% of wild genebank material has been wrongly classified on the species level. This is particularly relevant for this study because the authors report some 'surprises' in the subgenome compositions of *Th. intermedium* and *R. kamoji*. How did the authors make sure that the two selected accessions represent 'typical' representatives of their species? Critical information about accession selection and species validation is lacking in the methods section. I could not find any information on *R. kamoji* accession Pr 87-88 353 in Genesys. From which genebank does this accession originate? From the methods section, it appears that the sequenced accessions were not obtained from the respective genebanks directly, but through intermediary labs. This always involves the risks of seed mix-up. How can seeds of the two sequenced accessions be obtained by the international research community?

Response: We sincerely appreciate the reviewer's insightful comments and rigorous evaluation of our manuscript. We acknowledge the critical importance of proper accession selection and species validation in phylogenetic studies, particularly when working with wild plant materials. Below we provide a detailed response to address these concerns:

The *Th. intermedium* accession PI440031 we sequenced was originally obtained from the USDA National Plant Germplasm System (NPGS), collected from Stavropol Russia (45°03'N / 41°58'E) in 1977, in the major region where *Th. intermedium* distributed. It shows advances in grain weight, seed set rating and yield during a population evaluation as perennial grain crop (doi: 10.21203/rs.3.rs-3399539/v1). To minimize risks of seed mix-ups, this accession underwent multiple generations of self-pollination to ensure genetic homogeneity. In addition, we also performed GISH and FISH experiments to confirm the chromosome constitution of the sequenced accession and rule out misclassification or hybridization artifacts (Supplementary Fig.

14). Notably, this accession shows typical chromosomal constitution of *Th. intermedium* that aligns with earlier study by Mahelka et al. (*BMC Evolutionary Biology* 2011, 11:127), who reported the two subgenomes of *Th. intermedium* most probably been contributed by *Pseudoroegneria* (St) and *Dasypyrum* (V) and the third subgenome potentially related with *Aegilops*. In addition, re-sequencing analyses on six additional *Th. intermedium* accessions all exhibited consistent subgenome compositions, confirming the reliability of our sequenced accession.

Roegneria is a large genus of perennial polyploid grasses, mainly distributed across the temperate regions of Central and East Asia. Many species, particularly hexaploids, still require well genetic and cytogenetic validation of their genomic constitution. Actually, we failed to obtain a confirmed representative accession of *R. kamoji* from international germplasm repositories such as the USDA National Plant Germplasm System. To ensure accuracy, we alternatively obtained the *R. kamoji* accession Pr 87-88 353 from the *Triticeae* Research Institute of Sichuan Agricultural University (Ya'an, Sichuan, China), an authoritative center for taxonomical studies of *Triticeae*. This accession was originally collected and taxonomically validated by Prof. Chi Yen and Prof. Junliang Yang, pioneering scholars in *Triticeae* systematics. Their seminal work (*Biosystematics of Triticeae*, Vol.1-5) established the contemporary systematic framework for *Triticeae* species, ensuring the nomenclatural accuracy of this designation. Below figures are the detailed information for this accession.

Figures. The recorded morphological and cytogenetic information for accession Pr 87-88 353 from the *Triticeae* Research Institute of Sichuan Agricultural University

This accession Pr 87-88 353 exhibits standard diagnostic characteristics of hexaploid *R. kamoji* at morphological, cytogenetic, and molecular levels, unequivocally differentiating it from related genera. These characteristics include: (1) A single spikelet per rachis node (Supplementary Fig. 2; In contrast to *Elymus* species typically bear 2–3 spikelets per node;); (2) A hexaploid genome with the chromosome constitution StStYYHH (2n= 6x=42), corroborated by whole-genome sequencing, GISH and FISH analyses (Fig. 2a, Supplementary Fig. 15); (3) Chloroplast phylogenetic reconstruction demonstrated tetraploid *Roegneria* species (StStYY, 2n = 4x = 28) as sister taxa to the hexaploid *Roegneria* clade, suggesting maternal origin from tetraploid *Roegneria* (Supplementary Fig. 17).

Collectively, the integrated evidence conclusively demonstrates that the accession Pr 87-88 353 represents an authentic hexaploid *R. kamoji* with no evidence of seed mix-up. This accession also exhibits high resistance to *Fusarium* head blight (FHB), a persistent challenge in wheat breeding programs, as demonstrated in a previous evaluation of *R. kamoji* populations (Luo et al., Evaluation of *Fusarium* head blight resistance and identification of resistant germplasms in different populations of *Roegneria kamoji*. *Journal of Plant Genetic Resources*. 2021, 22(3): 725-733). Therefore, we selected this accession for genome assembling.

Seeds of these accessions can be obtained from the corresponding author to facilitate independent verification. This availability is explicitly stated in the Data Availability section. To enhance clarity, we have revised the manuscript to include detailed descriptions of accession origins and taxonomic validation in the Methods section.

• Heterozygosity rate: Thinopyrum intermedium is described as a self-sterile, outcrossing species. The heterozygosity rate of 1.35% (Figure S1) seems to be low for an obligate outcrossing species. Do the authors have any explanation for this?

Response: Thanks for the comment. The calculated heterozygosity rate (1.35%) presented in our study was generated through Jellyfish software based on k-mer spectrum analysis (k=17). To establish genetic uniformity in our study, we implemented a rigorous single-plant selection protocol across multiple generations of controlled self-pollination under bagging isolation. While *Th. intermedium* is indeed predominantly outcrossing in natural conditions, this enforced inbreeding regime progressively reduces the heterozygosity level.

• Fhb7: There is a major discrepancy regarding the origin/distribution of the Fhb7 gene with previous literature. In their 2020 paper (Wang et al. (2020), Science 368, eaba5435), the authors concluded ‘this FP-HGT event apparently occurred after the divergence of the E genome from Triticum sp. but before the formation of the decaploid Th. ponticum’. This conclusion was based on the reported absence of Fhb7 orthologs across Triticum species. Here, the authors state ‘These findings support the hypothesis that the horizontal gene transfer of Fhb7 occurred prior to the divergence of the Triticeae family, resulting in the integration of a small genomic fragment into

the plant genome.' This conclusion is very different from the 2020 paper, implying that Fhb7 was introduced much earlier and occurs across a wide range of species. This discrepancy needs to be addressed. Why did the authors miss the Fhb7 orthologs in their 2020 study? Have the authors validated that Fhb7 orthologs are absent from Pooideae tribes other than Triticeae?

Response: We sincerely appreciate the reviewer's critical evaluation of this important evolutionary question. In the previous study published on Science, comprehensive BLAST analyses against the NCBI Refseq database, all publicly available *Triticeae* genome assemblies (including *Triticum aestivum*, *T. dicoccoides*, *Aegilops tauschii*, *Secale cereal* and *Hordeum vulgare* etc.) and transcriptomes of annual diploids (Ns, F, V and Q genomes) failed to detect *Fhb7* orthologs, leading us to hypothesize the horizontal transfer might have occurred after the divergence of the E genome of *Th. elongatum* from *Triticum* sp. but before the formation of the decaploid *Th. ponticum*.

However, recently emerging genomic resources from the *Triticeae* tribe including *Elymus*, *Roegneria*, *Agropyron*, and *Hordeum* etc. have revealed conserved *Fhb7* orthologs in multiple lineages, demonstrating its pan-*Triticeae* distribution. To ensure phylogenetic rigor, we extended comprehensive searches to encompass more assembled *Pooideae* species. Notably, the presence of *Fhb7*'s homologs was also detected in several *Pooideae* species outside *Triticeae* (including *Alopecurus aequalis*, *Bromus sterilis*, *Bromus tectorum* and *Hierochloe odorata*), with absence in other *Poaceae* subfamilies. This result suggests *Fhb7*'s HGT occurred after *Pooideae* diversification and prior to *Triticeae* radiation. However, the occurrence time of this HGT may be refined in the future with expanded sampling of *Pooideae* genomes. We acknowledge the reviewer's important comment and have thoroughly revised the manuscript with the updated conclusions regarding *Fhb7*'s evolutionary trajectory.

• Figure 3B: The Fielder spike looks resistant to me, while the T3_Fhb7RkaH looks susceptible.

Response: Thanks for the comment. As described in the manuscript, the wheat cultivar 'Fielder' (CK) exhibits expected susceptibility, while the T3_Fhb7RkaH transgenic line demonstrates enhanced resistance. After carefully double-checking the original photos, we confirm this is an inadvertent error in the panel labeling of Figure 3B during figure assembly. We have revised the figure labeling that now accurately reflects the phenotypic comparison (Fig. 3b). We sincerely regret this oversight and appreciate the reviewer's comment in ensuring the integrity of our data presentation.

• Data availability: The two accessions numbers provided in the manuscript are not active yet, so I was unable to validate the deposited data. The data availability statement indicates that only assembly and annotation data have been deposited. What about the raw PacBio and Illumina reads? These also need to be deposited and made available. I also suggest to deposit the data in NCBI or ENA.

Response: Following the reviewer's suggestion, we have deposited the raw PacBio HiFi, Hi-C, RNA-seq, resequencing reads and genome assemblies of *Th. intermedium* and *R. kamoji* into NCBI database with accession number PRJNA1235488 and PRJNA1235489. The genome assemblies and annotations of these two species were also deposited to Zenodo (https://doi.org/10.5281/zenodo.15412758). We also updated this information in the Data availability section in the manuscript.

Minor comments:

• Line 57: The term 'elite gene' sounds strange to me. Maybe 'superior allele' might be more appropriate.

Response: Following the comment, we corrected this term in manuscript (line 58).

• Figure 3, legend: invocated or inoculated?

Response: Thanks for the comment, we confirm the term "inoculated" is correct.

Reviewer #4 (Remarks to the Author):

The manuscript by Sun et al. presents the whole-genome sequence of two allohexaploid grass species from the tribe Triticeae, including an outline of the possible origin of these species. They also discuss the analysis of the FHB resistance of both species, which is controlled by the Fhb7 gene.
Overall, my assessment of the manuscript is positive, as it contains interesting and important information backed by a large amount of work and data. The data are plausible and their interpretation is, in my opinion, mostly correct. The main output of the work, the whole genome sequences of these two allohexaploid species, will be a significant step forward in wheat breeding.

Response: We thank the reviewer for the positive assessment of our manuscript, and appreciate the thorough review and extensive effort invested in this work. As the reviewer highlighted, we hope the genomic resources generated in this study will advance wheat breeding programs, particularly in addressing FHB resistance-a major challenge in global cereal production.

What I have a little problem with is the selection of genotypes for sequencing and their characterization. It would be good to characterize the reference genotypes in more detail. Where the plants were collected, what the climatic conditions are there, whether they have been used in previous work and whether they have been characterized in any way, either morphologically or genetically. Both species have large ranges and are certainly not genetically uniform, so it would be good to know which populations the reference genomes actually represent. I also wonder about the selection of genotypes for resequencing. Why were genotypes used from the USA (5 genotypes) and Australia (1), where the species is not native? I could still understand it if (1) these were genotypes used in wheat breeding and (2) the primary goal of the work was the whole genome sequence of genotypes used in wheat breeding. Please explain.

Response: Thank you for raising these critical questions regarding genotype selection and characterization. We appreciate the opportunity to clarify these points.

The representative *Thinopyrum intermedium* accession PI440031 was obtained from the USDA National Plant Germplasm System (NPGS) and subjected to multiple generations of self-pollination to ensure genetic homogeneity, thereby reducing risks of seed mix-ups. This germplasm was originally collected in 1977 from Stavropol, Russia, within the primary distribution range of *Th. intermedium*. It demonstrates significant improvements in grain weight, seed set rating and yield during population-based evaluation as a perennial grain crop (doi: 10.21203/rs.3.rs-3399539/v1). In the revised manuscript, we have incorporated GISH and FISH experiments to confirm its chromosomal constitution, with phylogenetic analysis providing additional validation through sequencing data representing 21 of the 24 recognized diploid genera (Fig. 2a).

The six available *Th. intermedium* accessions, stored in the Chinese Crop Germplasm Resources Information System (https://www.cgris.net/home), document that seeds were acquired from the USA and Australia, though detailed origination information remains unavailable. Here, these resequenced accessions were employed to validate the chromosomal constitution of *Th. intermedium* genome assemble, rather than discussing the population distribution of the *Th. intermedium*. Therefore, we revised the manuscript to make it clearer.

Collectively, multiple levels of evidences from morphology, cytogenetic (Supplementary Fig. 14), phylogenetic (Fig.2a), and genetic resequencing (Supplementary Fig. 13) analyses demonstrate conclusively that the sequenced germplasm represents a standard hexaploid *Th. intermedium*. Following the comment, comprehensive germplasm descriptions have been added to the Methods section of the revised manuscript.

To assure the accuracy of the sequenced germplasm as a typical hexaploid *Roegneria kamoji*, we obtained the *R. kamoji* accession Pr 87-88 353 from the *Triticeae* Research Institute of Sichuan Agricultural University (Ya'an, Sichuan, China), an authoritative institution for *Triticeae* systematics. This accession was originally collected and taxonomically validated by Prof. Chi Yen and Prof. Junliang Yang, pioneering scholars in *Triticeae* systematics. Their seminal work (*Biosystematics of Triticeae*, Vol.1-5) established the contemporary systematic framework for *Triticeae* species, ensuring the nomenclatural accuracy of this designation.

This accession Pr 87-88 353 exhibits standard diagnostic characteristics of hexaploid *R. kamoji* at morphological, cytogenetic, and molecular levels, unequivocally differentiating it from related genera. These characteristics include: (1) A single spikelet per rachis node (Supplementary Fig. 2; In contrast to *Elymus* species typically bear 2–3 spikelets per node;); (2) A hexaploid genome with the chromosome constitution StStYYHH (2n= 6x=42), corroborated by whole-genome sequencing, GISH and FISH analyses (Supplementary Fig. 15); (3) Chloroplast phylogenetic reconstruction demonstrated tetraploid *Roegneria* species (StStYY, 2n = 4x = 28) as sister taxa to the hexaploid *Roegneria* clade, suggesting maternal origin from tetraploid *Roegneria* (Supplementary Fig. 17).

Collectively, the integrated evidence conclusively demonstrates that the accession Pr 87-88 353 represents a typical hexaploid *R. kamoji* with no evidence of seed mix-up. This accession also exhibits high resistance to *Fusarium* head blight (FHB), a persistent challenge in wheat breeding programs, as demonstrated in a previous evaluation of *R. kamoji* populations (Luo et al., Evaluation of *Fusarium* head blight resistance and identification of resistant germplasms in different populations of *Roegneria kamoji*. *Journal of Plant Genetic Resources*. 2021, 22(3): 725-733). Therefore, we selected this accession for genome assembly.

Detailed accession origins, validation protocols, and taxonomic references have been added to the Methods section. Seed availability is explicitly stated in the Data Availability section to facilitate independent verification.

We hope these clarifications underscore the rigor of our genotype selection and characterization. Thank you again for your constructive comments.

Another problem is the interpretation of the genomic composition of the two hexaploids. Phylogenetic analysis depends on the input data, especially the representation of diploid Triticeae species. This was far from complete in this case. I understand that in this case it was necessary to use only those species for which whole-genome sequences are available, because otherwise it would probably have been impossible to obtain sequence data for 129 genes. But then you have to take into account the incomplete sampling and be very cautious in interpreting the data.

Response: Thank you for raising this critical concern regarding the interpretation of the genomic composition of these two hexaploid species. We fully acknowledge the limitations of phylogenetic inference when sampling of diploid *Triticeae* taxa is incomplete. To address this, we have rigorously expanded the dataset and refined our analytical framework as follows:

(1) Comprehensive taxon sampling:

Following Löve's taxonomic system (1984), we incorporated genomic, transcriptomic, or exome data from 21 of the 24 recognized diploid/autopolyploid cytogenetic genera (A–Y) in *Triticeae* (Supplementary Table. 11). For genera lacking whole-genome assemblies, we utilized transcriptomic or exome datasets from public repositories (NCBI), covering key lineages include: *Aegilops* spp. (C, M, T and U), *Pseudoroegneria* spp. (St), *Eremopyrum triticeum* (F), *Crithopsis delileana* (K), *Psathyrostachys juncea* (Ns), *Henrardia persica* (O), *Agropyron cristatum* (P), *Heteranthelium piliferum* (Q), *Taeniatherum caput-medusae* (Ta), *Australopyrum retrofractum* (W), *Peridictyon sanctum* (Xp), and multiple *Hordeum* species (H, I and Xu genomes). These datasets align with prior phylogenetic studies (Mason-Gamer & White, 2024), ensuring consistency with established *Triticeae* evolutionary frameworks. Two exceptions include the G genome in *Triticum timopheevii* (AG) and the N genome of *Aegilops uniaristata*, which were reported to exhibit close phylogenetic affinities with the wheat B-linage and the M genome of *Ae. comosa* respectively, were not included in the phylogenetic analyses. The third exception involves the *Festucopsis* (L) genus, for which no genomic datasets were publicly accessible.

(2) Phylogenetic Robustness:

The reconstructed phylogenetic tree based on genome-wide shared SNP loci among these *Triticeae* species, demonstrates congruence with prior single-copy ortholog genes based phylogenetic reconstructions regarding the genomic composition of *Th. intermedium* and *R. kamoji* (Fig. 2a). The topology strongly supports the *Aegilops*-related ancestry of the J sub-genome in *Th. intermedium* and the *Dasypyrum*-related ancestry of the Y subgenome in *R. kamoji*.

(3) Cautious Interpretation and Symbolization:

Even integrated with data from 21 diploid genera, there are proposed gaps in sampling for phylogenetic analysis (*Festucopsis* L genome and other uncertain

lineages) and emphasize that our interpretations are constrained by data availability. To avoid overstatement, we conservatively labeled unresolved subgenomes as "Y" (for the *Dasypyrum*-related subgenome in *R. kamoji*) and "J" (for the *Aegilops*-related subgenome in *Th. intermedium*), rather than assigning new definitive genome symbols.

Thank you again for your insightful comment emphasizing this important methodological consideration, which has significantly improved the rigor and clarity of our work.

Furthermore, I wonder whether the authors also performed a chloroplast genome assembly. If so, identification of the maternal genome would provide an important clue to the origin of the two hexaploids.

Response: Thank you for raising this insightful question regarding the maternal origin of the two allohexaploid species. Following the comments, we have now performed *de novo* chloroplast genome assembling for both species using PacBio HiFi long-read sequencing data. These assemblies generated high-quality chloroplast sequences (deposited in NCBI GenBank under accession numbers PV435186.1 and PV435187.1), which were phylogenetically analyzed alongside publicly available chloroplast sequences from related *Triticeae* species (Supplementary Table 14).

Our analysis robustly identified *Pseudoroegneria* (St) as the maternal donor for both hexaploids (Supplementary Fig. 17). This finding further supports the evolutionary trajectory proposed in our nuclear genome-based phylogeny. Moreover, the chloroplast phylogeny revealed tetraploid *Roegneria* species (StStYY) as the closest relative of hexaploid *Roegneria* lineage. Combined with nuclear genome evidence, these findings suggest that hexaploidization of *R. kamoji* may have arisen through hybridization between a tetraploid *Roegneria* ancestor and a diploid *Hordeum*-related paternal donor.

We have included detailed descriptions of the chloroplast assembly in the Methods section (subsection "Chloroplast genome assembly and phylogeny analysis") and the Results section in the revised manuscript.

Another comment is more technical, namely that the figures (especially the legends and axis labels) are harder to read throughout the manuscript and the legends are not self-explanatory.

Response: Thank you for highlighting the issues regarding figure readability and clarity. We sincerely apologize for the inconvenience caused by the poor presentation of figures and legends. Following the comment, we have thoroughly revised all figures and legends in the manuscript to ensure improved legibility and self-explanatory design.

Further specific comments are:
l. 48. In species names, authority should be given at first mention in the manuscript.

Please check this throughout the manuscript (l. 60, 61 etc.).

Response: Following the comment, we have revised the manuscript and added the species authority.

l. 73. 'Genomic in situ hybridization analyses...' This sentence is a bit misleading because if GISH is the only marker used (is that the case with ref. 14?), it does not give independent results (the results depend on the probes used). It would be more accurate to write that the study, based solely on in situ hybridization, have designated the genomic formula... In fact, GISH has also been used in other papers to confirm hypotheses generated from sequencing data, and has shown different results to those mentioned by the authors in this sentence.
Btw, journal title is missing in ref. 14 (l. 796).

Response: Following the comment, we have rewritten this sentence (line 80). The journal title for ref. 14 was also added.

l. 85. The sentence 'While Pseudoroegneria is a well-known diploid ancestor...' does not describe the situation. How about 'While the involvement of Pseudoroegneria in the origin of Th. intermedium has been demonstrated, the origin of the other subgenomes has not been reliably identified'?

Response: Following the comment, we have rewritten this sentence (line 97).

l. 90. If species distribution is mentioned for Roegneria, then it would be good to mention it for Thinopyrum as well.

Response: Following the comment, we have added species distribution information for *Th. intermedium* in the manuscript (line 72).

l. 92-93. It might be good to mention that Roegneria is exclusively a polyploid species (as opposed to Thinopyrum where diploids occur).

Response: Following the comment, we have rewritten this sentence (line 103).

l. 92. In references 20, 23, 25 and 26 the representatives of the genus Elymus are mentioned. Is Elymus an older name for current Roegneria? It would be useful to briefly explain what the relationship between the genera Roegneria and Elymus is (in terms of taxonomy), and which designation is actual.

Response: Historically, *Roegneria* was established by Koch (1848) as a distinct genus with *R. caucasica* as the type species, and was further supported by Nevski (1934) and subsequent taxonomists (Keng et al., 1963; Yang et al., 2008; Yen et al., 2011). This genus is endemic to Eurasia, with the majority of species distributed in Central

and East Asia (Yang et al., 2008).

However, some North American botanists (Löve,1984; Tzvelev, 1976; Mason-Gamer, 2001) had synonymized *Roegneria* within super-family of *Elymus* sensu lato due to some overlapping morphological traits and the limited characterization of the Y genome, which is predominant in Asian-distributed *Roegneria* species. Based on their genomic constitution, it has now been widely accepted for classifying the tetraploid species of *Elymus* (StH) and *Roegneria* K.Koch (StY) as separate genera. However, the hexaploid *Roegneria* (StYH), also named as *Campeiostachys* Drobow, still require further investigation for phylogenetic relationships and taxonomic treatment (Yen et al., 2011; Yang et al., 2019).

In this study, the *R. kamoji* Pr 87-88 353 accession exhibits standard diagnostic characteristics at morphological, cytogenetic and genomic levels, distinguishing it from related genera as described above. To avoid any potential confusion, we here have carefully revised the manuscript with more details describing the relationship between the genera *Roegneria* and *Elymus*. This clarification has been added to the manuscript to contextualize the nomenclature used (line 108-116).

Koch C. 1848. Beitrage zueiner Florades Orientes. *Linnaea* 21:289-443

Nevskis A. 1934. Roegneria C. Koch. In: Komarov VL, Rozhevits RY, Shishkin BK eds. *Flora of the USSR*. Jerusalem: Israel Program for Scientific Translations. 2: 599–627.

Keng YL, Chen SL. 1963. A revision of the genus Roegneria C. Koch of China. *Journal of Nanjing University* 3(1): 1–92.

Yang JL, Baum BR, Yen C. 2008. A revision of the genus Roegneria K. Koch (Poaceae: Triticeae). Journal of Sichuan Agricultural University 4: 311–381

Yen C, Yang JL, Baum BR. 2011. *Biosystematics of Triticeae*. Beijing: China Agricultural Press. 4: 337–582

Löve Á. 1984. Conspectus of the *Triticeae*. *Feddes Repert* 95: 425–521.

Tvelev, NN 1976. Grasses of the Soviet Union. *Zlaki* SSSR. Nauka, Leningrad.

Mason‐Gamer RJ. 2001. Origin of North American species of *Elymus* (Poaceae: *Triticeae*) allotetraploids based on granule‐bound starch synthase gene sequences. Systematic Botany 26: 757–768

Yang CR, Baum BR, Johnson DA, Zhang HQ, Zhou YH. 2019. Molecular diversity of the 5S nuclear ribosomal DNA in Campeiostachys with StHY haplome constitution. *Journal of Systematics and Evolution* 58: 69–76.

l. 93-94. How many tetraploids and hexaploids are there?

Response: The precise number of tetraploid and hexaploid cytotypes within *Roegneria* remains unresolved due to persistent taxonomic ambiguity in the genus. While *Flora* of China recognizes approximately about 120 species under *Roegneria* (https://www.iplant.cn/info/Roegneria?t=z), the lack of consensus on species delimitation and evolving polyploidy classifications precludes definitive quantification of ploidy-level distributions.

l. 112-119. This part starting with ‘Our findings...’ can be omitted here as it contains results. It should be replaced by hypotheses and aims of the study.

Response: Thanks for the comment, we have rewritten this part as suggested (line 135).

l. 122. Is there a particular reason for choosing these genotypes? Both species are characterized by a large geographic range and, we can assume, corresponding genetic diversity. In Th. intermedium, six accessions were used for resequencing, but none in R. kamoji. The question arises as to what parameters were used to select the reference genomes for sequencing and to what extent they represent the genetic diversity of the species. Why were additional genotypes not used for resequencing in Roegneria?

Response: Thank you for the insightful comment regarding the genotype selection. Although many perennial genera in the *Triticeae* tribe have been extensively employed as forage breeding materials or genetic resources for wheat improvement, the chromosome constitutions for many genera remain poorly characterized, let alone achieving precise species-level delineation. Elucidating multiple chromosomal sets and their reticulate evolutionary patterns in these genera remains a foundational research objective at this stage. In this study, the selected hexaploid species *Th. intermedium* and *R. kamoji* both contain the St subgenome, independently harboring V, J, Y or H subgenome. One of our objectives in assembling these genomes is to elucidate their subgenome constitutions and evolutionary trajectories during polyploidization.

As described above, the representative *Th. intermedium* accession PI440031 was originally collected from Stavropol, Russia, within the major distribution region of *Th. intermedium*, which shows advances in agronomic traits (doi: 10.21203/rs.3.rs-3399539/v1). As an outcrossing species, genetic contamination through pollen flow may potentially influence germplasm purity. Thus, this study not only implemented multiple generations of self-pollination to ensure genetic homogeneity, but also conducted comparative analyses using six re-sequenced germplasm samples in order to confirm the validity of the sub-genome constitutions.

The *R. kamoji* accession Pr 87-88 353 was collected from its primary distribution range in Sichuan Province, China, representing a native population. More importantly, this accession shows high resistance to FHB, a persistent challenge in wheat breeding programs (*Journal of ePlant Genetic Resources*. 2021, 22(3): 725-733). The pollen of *R. kamoji* is relatively small and predominantly selfing, as reflected by its heterozygosity rate of 0.17% in genomic analyses, with little risk of seeds mix-up.Assembling the hexaploid *Roegneria* genome aimed to trace the origin of the “Y” subgenome. As this accession represents the typical constitution of this species, so we didn’t choose to sequence the other accessions.

l. 123. ‘genome size’ is not the correct term here. DNA content per 1C should be given instead. In addition, the value given for Th. intermedium is quite low. The

2C-DNA database gives higher values (from 13 to 13.5 pg, which is equivalent to between 12.7 and 13.2 GB). How was the genome size estimated? No details are provided in M&M section.

Response: In this study, we estimated the genome size of *Th. intermedium* and *R. kamoji* by using Jellyfish software based on sequencing read depth and k-mer frequency distribution, which is also widely used to calculate genome size in other species. This result is similar with our flow cytometry based analysis for *Th. intermedium* (11.84 Gb) and *R. kamoji* (11.15 Gb), when using the 14.51 Gb T2T genome assembly of wheat cultivar 'Chinese Spring' as a reference (Supplementary Fig. 3). The relative higher DNA per 1C value of *Th. intermedium* in the public database is probably due to different sampled subspecies.

To make clarity, we have now added the methods and results based on flow cytometry in the revised manuscript (Supplementary Fig. 3; line 145).

l. 124. Figures S1 and S2 refer to the morphology of the species, but morphological details are not visible in them. I suggest mentioning the key morphological differences in the text of the Introduction.

Response: Following the comment, we have added more photos for these two species (Supplementary Fig. 1 and 2), followed by detailed description in the Introduction (line7 3-76, line 105-108).

l. 131. 97.7 and 99.4 % of sequence data were assambled into pseudomolecules. It would be useful to mention what amount of data (in MB or GB) was left as unplaced contigs.

Response: Thanks for the comment. About 239 and 66 Mb sequences unanchored in *Th. intermedium* and *R. kamoji* assemblies, respectively. We have added this information in the manuscript (line 157).

l. 135 (744). Table S4. What do the letters C, S, D etc. stand for? Please explain in table heading.

Response: C: complete; S: single; D: duplicate; F: fragment; M: miss. We have added this information in the manuscript.

l. 145 (664). The framed legend is illegible. It should be enlarged.

Response: Following the comment, we have enlarged the figure legend.

l. 150 (666 and 673), Figure S6 and S7. Legends (colour scales) are missing.

Response: Following the comment, we have added the colour scales.

l. 151 (751). Table S8. Pse. Libanotica in table heading should be corrected to Pse. libanotica. Same in Ae. Tauschii and Pse. Libanotica in the table (first row).

Response: We have corrected this typo.

l. 155 (745). The header should explain what LAI means and how it is calculated.

Response: LAI stands for LTR Assembly Index (LAI), a value produced by LAI software to evaluate the LTR completeness in the genome assembly. Following the comment, we have explained the meaning of LAI in the manuscript (line 162).

l. 157-159. It's not accurate to talk about potential progenitors here. This designation is based on the phylogenetic analysis, which depended on the reference genomes used (Fig. 2A). Only a few species were represented in each genus. For example, Hordeum is represented by only two species, while the entire genus contains 31 diploids (with 4 distinct genomes). Therefore, it would be more accurate to say that the TE annotation was also performed on selected representatives of genera whose members are likely to have contributed to the origin of the hexaploids studied. The result of the comparative analysis should also be briefly commented on (not just refer to table S8).

Response: Thanks for the insightful comment, and we have corrected this part in the revised manuscript. A brief description of TE annotation results was also added (line 188).

The other thing is that since the possible progenitors of the polyploids Thinopyrum and Roegneria are mentioned here, but the identification of the individual subgenomes is not described until the next chapter, it might be more logical to place the Molecular phylogeny chapters (both) earlier, after line 142. A paragraph beginning with the sentence 'Transposable element...' would follow the Molecular phylogeny chapters.

Response: Thanks for the helpful comment. This section "Genome assembly and annotation" was mainly to introduce the basic information of genome assembly and annotation results of these two hexaploids, so we think it would be more suitable to put the TE annotation results in this section.

l. 167 (754). Table S10. 'Ratio' on the third row could be replaced with '[%]' or 'Ratio [%]'

Response: Thanks for the comment. We have corrected this typo.

l. 179-182. Mapping of St reads on Thinopyrum chromosomes shows gaps at the right end of the chromosomes. While reads from other diploids map end-to-end on the corresponding chromosomes, this is not the case for St reads on the St genome. Is

there an explanation for this pattern?

Response: The observed mapping gaps at the terminal regions of *Thinopyrum* St subgenome chromosomes were induced by genome length differences between the species. The X-axis of Fig. 2C represents the pseudo-chromosome length of each subgenome. The St subgenome pseudo-chromosomes of *R. kamoji* exhibit significantly greater lengths compared to their homologous counterparts in *Th. intermedium*, with total genome sizes of 3.98 Gb versus 3.04 Gb, respectively. This phenomenon is not unique to the St subgenome, as similar length discrepancies and mapping irregularities are observed in the V subgenome comparisons, as the V subgenome pseudo-chromosomes lengths in *R. kamoji* were shorter than that in *Th. intermedium* (3.21 Gb versus 3.8 Gb).

l. 182 (696 and 700). Fig S10. replace Ae. Tauschii with Ae. tauschii (twice).

Response: We have corrected this typo.

l. 182. not Figure 2B but Figure 2C includes that analysis.

Response: We have corrected this typo.

l. 182-184. 'Notably, the ThinD subgenome was clustered between wheat A lineages and other Aegilops-clade species, diverging approximately 3.4 million years ago.' In phylogenetics, it is not appropriate to say that a species is clustered between some clades. To be correct, the ThinD subgenome is sister to the D- + S-genome Aegilops species (according to what I quickly found in Feldman's classification for Ae. searsii and other species. Feldman, M., Levy, A.A. 2023. Aegilops L. In: Wheat Evolution and Domestication. Springer, Cham. https://doi.org/10.1007/978-3-031-30175-9_9.). To be more accurate in identifying the ThinD subgenome, a more comprehensive representation of Aegilops species would be needed. Therefore, the designation of this subgenome as ThinD should be considered tentative. The text should be corrected.

Response: Following the comment, the statement for phylogenetic relationship has been revised in the manuscript (line 217). Moreover, we recognized that the designation of this sub-genome as "ThinD" is not accurate and remains provisional pending, since probably insufficient samples of *Aegilops* species were included in the phylogeny. To avoid any misleading in taxonomy, we prefer to keep the symbols "J" for the *Aegilops*-affiliated subgenome in *Th. intermedium* throughout the manuscript.

l. 184-189. Since the closest relatives of the ancestors of the hexaploids are not known, the comparison of the TE content of the hexaploids with a particular diploid (shown in Fig. S9) makes little sense. I would omit this analysis except for Thinopyrum vs. Roegneria comparison.

Response: We appreciate the reviewer's critical assessment and have carefully reconsidered the rationale for comparing TE content between hexaploids and diploids. While the exact diploid progenitors of the hexaploids remain unresolved, our phylogenetic analyses (Fig.2a) positioned the ThinJ subgenome as closely related to *Ae. tauschii* (D-genome). This phylogenetic proximity provides a meaningful basis for comparing TE profiles, as transposable element dynamics often reflect lineage-specific evolutionary histories. The observed TE similarities between ThinJ and *Ae. tauschii* further support the hypothesis of a shared ancestral origin within the *Aegilops* clade.

l. 190-192. Figure 2B. How was the synteny analysis done? There are no details provided in methods section.

Response: Thanks for the comment. We have added the synteny analysis methods in the method section (line 678).

l. 196. ...'previously proposed as an ancestral donor.' A reference(s) should be added.

Response: Following the comment, the related reference has been cited (Ref. 16).

l. 200 (756). Table S11. correct Th. Intermedium to Th. intermedium. Which gene bank did these accessions come from?

Response: We have corrected this typo. All these accessions were collected from Chinese Crop Germplasm Resources Information System (https://www.cgris.net/home), which has been added in the methods section (line 496).

l. 213-221. Based on the analyses performed, it appears that the ancestor is indeed Dasypyrum. However, it should be noted that the phylogenetic analysis is far from including all possible representatives of Triticeae. It is therefore theoretically possible that the ancestor is another species related to Dasypyrum that was not included in the analyses. More detailed analyses in this direction would be useful in the future, and it would also be fair to discuss this possibility in the text.

Response: Thanks for the valuable comment. We have reconstructed the phylogenetic tree including 21 of the 24 recognized diploid/autopolyploid cytogenetic genera (A–Y) in *Triticeae* (Fig. 2a), which shows consistence with our previous result for the V close-related subgenome in *Th. intermedium*. Following the comment, we have carefully discussed other possibilities if specific taxa/genomes of concern updated in the future (line 375).

l. 220 (710). Fig S12. not blue but green lines represent the translocation blocks.

Response: We have corrected this typo.

l. 232-234. It is also possible that each polyploid arose from a different (not sampled here) St ancestor.

Response: Thanks for the insightful comment. Based on the reconstructed phylogenetic tree, the RkaSt was clustered with *Pse. strigosa* and ThinSt clustered with another St diploid species of *Pse. stipifolia*, which supports that these subgenomes should have underwent independent tetraploidization events. This result has been described in the revised manuscript (line 291).

l. 235. ... 'a specific inversion in the V03 subgenome'...why V03? Wouldn't be enough V subgenome?

Response: The Y subgenome in *R. kamoji* is sintered with *D. villosum* (V) in phylogeny, displayed exceptionally long branch lengths, suggesting lack of donor species in this phylogenetic tree (Fig.2a). To avoid any unnecessary misleading, we kept the genome symbol "Y" to represent the *Dasypyrum* (V)-close related subgenome in *R. kamoji*. There is a specific inversion in the Y03 subgenome of *R. kamoji* compared with V03 of other species, further indicating its independent evolutionary event.

l. 237. Is R. ciliaris the only known tetraploid, or are there more candidates?

Response: Thanks for the comment. There are more tetraploids in this genus, and R. ciliaris may not the direct ancestor of hexaploids, we have further estimated the polyploidy events by comparing chloroplast genomes from related species, and have now revised this part (line 300).

l. 240. The involvement of genomic in situ hybridization would confirm the identity of the subgenomes and the presence/absence of translocations, and greatly increase the credibility of the results.

Response: We appreciate the reviewer's comment regarding the importance of genomic *in situ* hybridization (GISH) in validating subgenome identity and translocation events. In the revision, we conducted sequential GISH and fluorescence *in situ* hybridization (FISH) analyses to confirm the genomic constitutions of these two hexaploids.

For *R. kamoji*, the 14 chromosomes belonging to the St subgenome were clearly distinguished by GISH, using *Pse. spicata* (St) genomic DNA as probes (Supplementary Fig. 15a). The chromosomes of H subgenomes could also be identified by FISH experiment (Supplementary Fig. 15b,d). However, no clear signals were obtained for the chromosomes of Y subgenomes, when using the proposed *D.*

villosum (V) genomic DNA or specific repeats as probes (Supplementary Fig. 15a,c), which further supports that V genome should be not the direct donor for Y subgenome in *R. kamoji*.

For *Th. intermedium*, GISH with *Pse. spicata* (St) DNA as probe clearly distinguished the St set of chromosomes (Supplementary Fig. 14c). The V related chromosomes could also be identified by FISH using the *D. villosum* (V) specific repeats of Oligo-pDb12H. Considering the close relationship within the *Aegilops* genus, we designed the repeat/LTR probe Oligo-Ae584, specific and conserved in A/B/D genomes (Supplementary Fig. 14a), which clearly identified the *Aegilops* related chromosomes (J) in *Th. intermedium* (Supplementary Fig. 14b,d).

Collectively, the integration of GISH and FISH analyses unequivocally validated the subgenome origins and absence of translocations in both species, ensuring the credibility of our genomic conclusions. We have added this analysis in the revised manuscript.

l. 244-246. ...'Because St genome...shares a closer evolutionary relationship with Dasypyrum than with Aegilops, it is likely that the ThinSt and ThinV subgenomes first combined to form a tetraploid species.' I don't understand what is meant by the 'closer evolutionary relationship with Dasypyrum than with Aegilops'.

Response: We apologize for the lack of clarity in our original statement. We have revised this part in the manuscript (line 304).

l. 260 (714 and 717). Fig S13. Instead of the term comparison, the term alignment would be more appropriate.

Response: Following the comment, we have replaced this term.

l. 260. I think that synteny is not the right term for an analysis that deals with a structure of a single gene. It is just a comparison of the gene region sequences.

Response: Following the comment, we have replaced this term.

l. 262 (394). Fig 3A. 'Infected' instead of 'invocated'? Next, in panels 3B and C, it would be good to indicate which spikes are supposed to be sensitive and which insensitive (transgenic).

Response: Thanks for the comment. We used inoculated to replace the original word "invocated". Upon re-examining the original figures, we identified an inadvertent error in the panel labeling of Figure 3B during figure assembly. The corrected version now accurately reflects the phenotypic comparison: the Fielder control exhibits clear susceptibility, while the T3_Fhb7^{RkaH} transgenic line demonstrates improved FHB resistance. We deeply regret this oversight and appreciate the reviewer's diligence in ensuring the integrity of our visual data presentation.

l. 266. Transcriptome with lowercase t.

Response: We have corrected this typo.

l. 268 (720). Fig S14. Please explain what is on y axis.

Response: We have added the explanation for y axis of Fig S14. The y-axis represents the expression value calculated based on the transcriptome data. FPKM: Fragments Per Kilobase of exon model per Million mapped fragments.

l. 271 (723). Fig S15. The figure as it stands doesn't say much and could be omitted.

Response: Following the comment, we have deleted this figure.

l. 279. Could you briefly describe what deoxynivalenol-glutathione is good for? What is the principle of the resistance?

Response: Deoxynivalenol-glutathione (DON-GSH) is a detoxified conjugate formed when glutathione (GSH) binds covalently to deoxynivalenol (DON), a trichothecene mycotoxin produced by *Fusarium* pathogens. This conjugation reaction, catalyzed by glutathione-S-transferase (GST) enzymes such as $Fhb7^{RkaSt}$-GST and $Fhb7^{RkaH}$-GST in this study, neutralizes DON's toxicity. The detailed mechanism of this gene was illustrated in our previous paper (Wang et al., Horizontal gene transfer of Fhb7 from fungus underlies *Fusarium* head blight resistance in wheat. *Science* 2020, 368.).

l. 284. please correct Fhb7RkaSt (missing t)

Response: We have corrected this typo.

l. 293 (393). Fig 3E. The figure legend should be self-explanatory. It should be explained what is on x axis etc.

Response: Following the comment, we have added the explanation for x axis of Fig. 3E in the manuscript.

l. 305. The expression...'with contributions from the present-day genera'... may give the false impression that the hexaploid was formed recently with the contribution of present-day species. It would probably be better to write 'with contributions from progenitors corresponding to present-day genera...'

Response: Thanks for the helpful comment, we have rewritten this sentence as suggested (line 368).

l. 307. ...suggesting that DIPLOID Thinopyrum species...

Response: We have corrected it following the comment.

l. 309. The contribution from Aegilops have been suggested in other papers. It would be good to add reference(s).

Response: We have added the related references following the comment (Ref. 17).

l. 325. How does genomic stability relate to the hybridization ability with wheat? Many genomes are stable but don't cross with wheat.

Response: We appreciate the reviewer's critical observation. Genomic stability mentioned here was not suitable, we have deleted this description.

l. 332-333. ...'but not in annual species'... Have the annual species been tested? If not, then this statement is unsupported.

Response: Thank you for highlighting this important point. In our study, the statement regarding the absence of *Fhb7* in annual species was based on published genomic datasets of annual taxa. However, we acknowledge that our own experimental work did not explicitly test all annual species in *Triticeae*, and the phrasing could imply overgeneralization. We have revised the text to reflect the scope of current evidence more precisely (line 410).

Materials and Methods
Overall, it seems to me that the methodology could be more detailed.

l. 409. Please add a reference for CTAB method.

Response: Following the comment, the related reference was added (Ref. 48).

l. 433. subjected instead of subjection

Response: We have corrected this typo.

l. 468-469. Accession numbers for the diploids should be provided.

Response: We have added the accession numbers for each genome assembly.

l. 476. What substitution model was used to calculate divergence K?

Response: We used the Kimura substitution correction algorithm to calculate the K value. We have added this information in the manuscript (line 590).

l. 480. Please add a reference paper to the TE-specific mutation rate used.

Response: Following the comment, the related reference was added (Ref. 59).

l. 523. A sentence explaining why the phylogenetic analysis was performed would be useful. The same should be stated for the other chapters.

Response: Following the comment, we have added the analysis purpose at the header of each paragraph.

l. 529. Please add accession number for Ps. stipifolia. What tissues were used for RNA isolation? Please add details on how the RNA seq was done.

Response: We have collected the replacement dataset for *Pse. stipifolia* in the NCBI database (Supplementary Table 11), so we deleted this dataset in the manuscript.

l. 532. In total, 129 single-copy genes were identified... I wonder if only 129 common genes were found for the selected taxa, or if there were more and 129 represents only a subset (in that case I would be interested in the selection criteria).

Response: It's actually 129 single-copy genes were identified among the publicly available assembled species and no selection was performed.

l. 537 and 553. As noted above, a sentence explaining the purpose of the analyses should be added to each chapter.

Response: We have added the analysis purpose at the header of these paragraphs.

l. 565. Please provide the primers' sequences.

Response: All the primers used in this study have been listed in Supplementary Table. 15.

l. 570-571. What cultivar of wheat was used? How many plants of each line were used for the experiment?

Response: The cultivar 'Fielder' was used as transgenic receptor. At least 10 plants were used for FHB resistance for each line. We have added this information in the manuscript (line 750).

l. 584. Which genotypes of Thinopyrum and Roegneria were used? Same as for the sequencing? Regarding wheat lines, were both Fusarium susceptible and transgenic lines used? Please add details in the text.

Response: The genotypes of *Th. intermedium* and *R. kamoji* mentioned here were same as the sequencing ones. We have added the details in the manuscript (line 764).

l. 586. I don't understand the term one-heart stage.

Response: Here is the one-leaf stage. We have corrected this typo.

l. 593. Add a sentence describing the purpose of the analysis.

Response: We have added the analysis purpose at the header of this paragraph (line 774).

REVIEWER COMMENTS

Reviewer #2 (Remarks to the Author):

The author carefully revised the reviewer's questions in the revised manuscript, and I have no other questions. Please supplement and improve a few precautions:
1. The origin time of Thinopyrum intermedium and Roegneria kamoji should be predicted based on the results of whole genome sequencing in this paper.

Response: We thank the reviewer for the helpful comments. Based on the reconstructed phylogenetic tree (Fig. 2a) and inferred speciation mechanism of *R. kamoji* (Fig. 2d), we estimated its origin time at approximately 2.8 Mya. Regarding *Th. intermedium*, as reviewer #4 suggested and due to insufficient evidence, we have removed the description of its speciation and did not infer an origin time for this species. These updates have been incorporated into the revised manuscript (line 302).

2. In supplementary Fig. 10, Where does the data for Pse. stipifolia?

Response: Thanks for the comment. The Fig. 2a presents a relatively comprehensive phylogenetic relationship encompassing diploid/polyploid genera within *Triticeae*, constructed using genome-wide shared SNP loci derived from available genome assemblies, transcriptomes, or exome datasets. Supplementary Fig. 10 includes only species with whole-genome assemblies and was built using single-copy genes. As we lack the genome assembly of *Pse. stipifolia*, thus only its transcriptome data was employed in Fig. 2a. We have deleted this species from Supplementary Fig. 10.

3. In Genome in situ hybridization experiments, if a species genome is used as a probe, the probe name should be V-genome instead of Oligo-pDb12H. Suggest making modifications in the text and images.

Response: Following the comment, we have corrected this mistake throughout the manuscript. The V-genome-specific probe Oligo-pDb12H was used to hybridize to the V subgenome in *Th. intermedium*. Total genomic DNA from the V-genome was used to hybridize to the Y subgenome in *R. kamoji*.

4. How the Oligo-Ae584 probe was obtained should be explained in the methods section.

Response: The Oligo-Ae584 probe sequence was designed according to the genomes of *Aegilops* species using Tandem Repeat Finder software (v 4.09) with default parameters. We have incorporated this methodological detail in the Methods section (line 711).

5. Triticeae does not italicize in the text.

Response: We have corrected this item thoroughly.

6. In supplementary Fig. 14, “Ps. Spicata” should be “Pse. spicata”

Response: We have corrected this typo.

7. In Line 49, “Triticum aestivum Lam.” should be “Triticum aestivum L.”

Response: We have corrected this typo.

8. Genomic symbols, if not bolded, the entire text will not be bolded. Please check.

Response: We have corrected the genomic symbols thoroughly.

Reviewer #3 (Remarks to the Author):

The authors have adequately addressed my comments. I particularly appreciate the additional information on the origin, management, and validation of the sequenced accessions.

Response: We are deeply appreciating the reviewer for the constructive feedback throughout the review process, and we are grateful for the time and expertise you dedicated to evaluating this work.

Reviewer #4 (Remarks to the Author):

The revision of the manuscript increased the credibility of the results, especially the identification of the ancestors of the studied hexaploids by phylogenetic analyses. I welcome the use of cytogenetic methods to verify genomic composition, but I have reservations about this section, both in terms of the methodology used and the interpretation. Specific comments are mentioned below.
l. 46. The authority for Triticum aestivum is L. Please check the International Plant Names Index (IPNI) for species names and their authorities.

Response: Thank you for this helpful comment. The authority for *Triticum aestivum* has been corrected to "L." All other species authorities in the manuscript have also been verified.

l. 62. According to IPNI, Roegneria kamoji (Ohwi) Ohwi ex Keng is the correct name for R. kamoji.

Response: We have corrected this error (line 62).

l. 72-76. What source is the species' distribution from? Th. intermedium is also native to a major part of Europe (ca from Spain to the Ural mts.). Please check and add reference(s). Add a reference for the morphological description, too.

Response: Thanks for the comment. We have revised the distribution description of *Th. intermedium* according to previous study (Ref.14) and Plants of the World Online database (POWO, https://powo.science.kew.org/). Additionally, references for the morphological descriptions of both *Th. intermedium* and *R. kamoji* have also been included in the revised manuscript (line 76 and 107).

l. 92. Dasypyrum breviaristatum (H. Lindb.) Fred.

Response: We have corrected this error.

l. 144-148. Keep in mind that these are 1C values. It should be specified, e.g., on l. 147 as follows:...'flow cytometry measurements (11.84 Gb and 11.16 Gb per 1C, respectively)'...

Response: Following the comment, we have added 'per 1C' in the revised manuscript (line 146).

l. 169-170. The sentence sounds weird. I suggest rephrasing, e.g., Transposable element (TE) annotation revealed that TEs occupy 77.15% and 78.49% of the genome sequences in Th. intermedium and R. kamoji, respectively.

Response: We have rephrased this sentence as suggested (line 168).

In this context, I don't understand the category 'Transposons' in supplementary Figs. S6 and S7 (first line and first column in the tables). It has no values and could be omitted in both tables.

Response: Following the comment, we have removed this term from supplementary Tables. S6 and S7.

l. 189. replace 'transposable elements' with 'TEs'

Response: Following the comment, we have replaced this term.

l. 212-213. This interpretation is incorrect. The Thinopyrum sequence does not cluster with the D-clade of Aegilops. Instead, it is a sister species to a clade comprising species ranging from Ae. sharonensis to Ae. tauschii. As the clade can rotate at the node, ThinJ is as closely related to Ae. tauschii as to Ae. sharonensis. The analysis shows that ThinJ is not closely related to any of the Aegilops species present today (as

expected), but it is likely to come from one of the species included in the aforementioned sister group comprising various Aegilops species (node XI). If we looked deeper into the three, we would see that the ThinJ sequence falls within the paraphyletic clade comprising the Aegilops and Triticum genomes (node IX). In conclusion, I suggest rephrasing the sentence as follows (or something similar):
...'yielded concordant topologies. They revealed that the subgenomes of Th. intermedium clustered with the St-clade of Pseudoroegneria, the V-clade of Dasypyrum, and a heterogeneous clade of Aegilops species, which we designate as the ThinSt, ThinV, and ThinJ subgenomes, respectively.

Response: We sincerely thank the reviewer for highlighting these important points. We have rephrased this sentence as suggested (line 212).

Finally, it would be good to include more information in the legend of Fig. 2a. For example, it would be useful to specify what type of phylogenetic analysis the tree shows. It would also be desirable to add bootstrap values to the tree. The same applies to Supplementary Fig. 10.

Response: This phylogenetic tree was constructed using whole-genome shared SNP loci among *Triticeae* species by maximum likelihood method. Branch support values were based on 1000 standard bootstrap replicates. Following the comment, we have added this methodological information and bootstrap values to the figure legend of Fig. 2a and Supplementary Fig. 10.

l. 240–251. I welcome the inclusion of cytogenetic verification. However, I am not convinced by it. Specifically, I am missing information about how the probes were chosen and how their specificity was verified.

Response: We thank the reviewer for highlighting the importance of probe selection and specificity in cytogenetic verification. Some Oligo-FISH probes used in this study were obtained from previous cytogenetic studies (Supplementary Table 13). The specificity of these probes has been relatively well-established and they are widely used within the molecular cytogenetic community, particularly for chromosome engineering studies involving wheat-alien hybrids. To further verify probe specificity, we aligned the sequences of these probes to available *Triticeae* genome assemblies using BLASTN (Table below).

Oligo-HvCSR: This probe was designed based on *Hordeum*-specific centromeric repeats, and was used as an ND-FISH probe to characterize the H genome chromosomes [1]. Our sequence alignment confirmed its high specificity, with exclusive amplification in the H subgenome of *R. kamoji* and *H. vulgare*. This probe enabled detection of RkaH chromosomes in *R. kamoji* (Supplementary Fig. 15b).

Oligo-pTa535: This probe was originally found to be variously distributed in wheat

A/B/D subgenomes and other *Aegilops* species [2-5]. Subsequently, it was found to be highly amplified in wild *Hordeum* and used to identify wild *Hordeum*-related chromosomes in wheat-*Hordeum* hybrids [1, 6]. This probe shows low copy numbers in St or Y-subgenomes of *R. kamoji*, and we thus used Oligo-pTa535 together with the H-genome specific probe Oligo-HvCSR to help detect the RkaH chromosomes in *R. kamoji* (Supplementary Fig. 15b).

Oligo-pDb12H: This probe was designed based on specific tandem repeats (TRs) with the highest accumulation levels in *D. villosum* [7]. It has subsequently been used to analyze wheat–*Dasypyrum* hybrids and the V-genome in wild relatives [8-9]. Sequence alignment further confirmed its strong preferential amplification in the V-subgenome of *Th. intermedium* and *D. villosum*. Therefore, this probe was employed to detect the ThinV chromosomes in *Th. intermedium* (Supplementary Fig. 14b,d).

Oligo-Ae584: This probe was *de novo* designed in this study and was primarily amplified in *Aegilops* species including wheat A/B/D subgenomes. Our FISH result also revealed that Oligo-Ae584 could exhibit clear signal for the *Aegilops*-related ThinJ subgenome in *Th. intermedium* (Supplementary Fig. 14b,d).

Table. Copy numbers of Oligo probes in each assembled *Triticeae* species

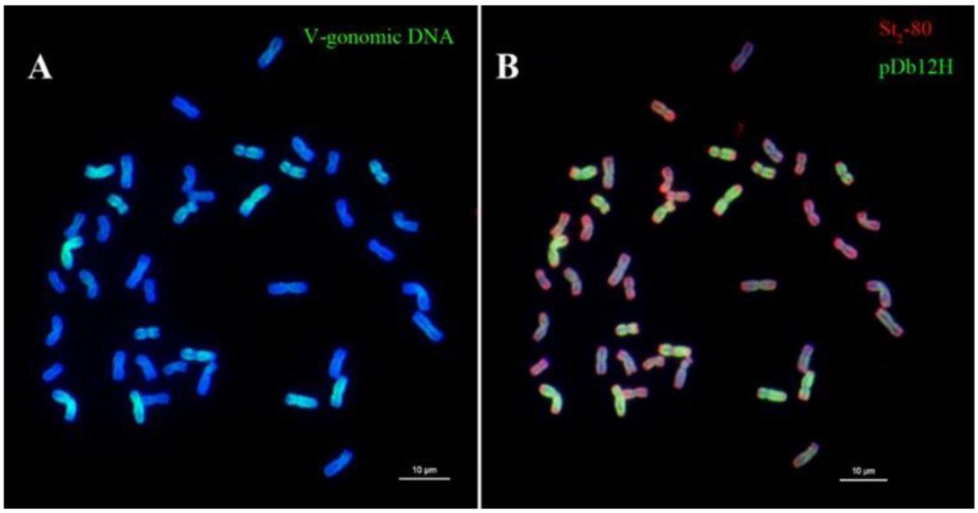
	Oligo-HvCSR	Oligo-pTa535	Oligo-pDb12H	Oligo-Ae584
Ae. sharonensis	438	1,099	182	23,287
Ae. longissima	568	1,132	193	25,900
Ae. bicornis	301	1,569	184	29,707
Ae. searsii	230	1,004	166	20,229
T. aestivum subD	91	4,146	245	31,454
Ae. tauschii	133	8,819	331	33,046
Th. intermedium subJ	159	16,133	403	21,341
T. urartu	302	4,804	236	28,069
T. aestivum subA	241	4,880	143	25,527
Ae. speltoides	337	861	55	11,942
T. aestivum subB	187	385	56	30,022
R. kamoji subSt	189	2,255	2	138
Pse. libanotica	181	1,937	12	126
Th. intermedium subSt	156	945	200	197
Th. intermedium subV	2,141	1,893	19,861	469
D. villosum	695	4,553	12,394	182
R. kamoji subY	114	857	0	1,029
H. marinum	340	20,380	32	1
R. kamoji subH	20,647	11,830	10	26
H. vulgare	53,705	1,082	95	0

1. Jiang C, Liu X, Yang Z, Li G. Chromosome Rearrangement in Elymus dahuricus Revealed by ND-FISH and Oligo-FISH Painting. *Plants*. 2023; 12(18):3268.
2. Tang Z, Yang Z, Fu S (2014) Oligonucleotides replacing the roles of repetitive sequences pAs1, pSc119.2, pTa535, pTa71, CCS1, and pAWRC.1 for FISH analysis. *J Appl Genet* 55:313–318
3. Li, Z.; Ren, Z.L.; Tan, F.Q.; Tang, Z.X.; Fu, S.L.; Yan, B.J.; Ren, T.H. Molecular Cytogenetic Characterization of New Wheat-Rye 1R (1B) Substitution and Translocation Lines from a Chinese *Secale cereal* L. Aigan with Resistance to Stripe Rust. *PLoS ONE* 2016, 11, e0163642
4. Liu, C.; Li, G.R.; Gong, W.P.; Li, G.Y.; Han, R.; Li, H.S.; Song, J.M.; Liu, A.F.; Cao, X.Y.; Chu, X.S.; et al. Molecular and cytogenetic characterization of a powdery mildew-resistant wheat-*aegilops mutica* partial amphiploid and addition line. *Cytogenet. Genome Res*. 2016, 147, 186–194.
5. Ren, T.; He, M.; Sun, Z.; Tan, F.; Luo, P.; Tang, Z.; Fu, S.; Yan, B.; Ren, Z.; Li, Z. The Polymorphisms of Oligonucleotide Probes in Wheat Cultivars Determined by ND-FISH. *Molecules* 2019, 24, 1126.
6. Delgado, A.; Carvalho, A.; Martín, A.C.; Martín, A.; Lima-Brito, J. Use of the synthetic oligo-pTa535 and oligo-pAs1 probes for identification of *Hordeum chilense*-origin chromosomes in hexaploid tritordeum. *Genet. Resour. Crop Evol.* 2016, 63, 945–951.
7. Yu Z, Wang H, Xu Y, Li Y, Lang T, Yang Z, Li G. Characterization of Chromosomal Rearrangement in New Wheat—*Thinopyrum intermedium* Addition Lines Carrying Thinopyrum—Specific Grain Hardness Genes. *Agronomy*. 2019; 9(1):18
8. Jiang, C.; Jiang, W.; Liu, M.; Wang, H.; Yang, E.; Yang, Z.; Li, G. Characterization of a Wheat-*Dasypyrum breviaristatum* Chromosome Addition and Its Derived Progenies Carrying Novel *Dasypyrum*-Specific Gliadin Genes. *Agronomy* 2022, 12, 1673.
9. Qi, F.; Liang, S.; Xing, P.; Bao, Y.; Wang, R.R.-C.; Li, X. Genome Analysis of *Thinopyrum intermedium* and Its Potential Progenitor Species Using Oligo-FISH. Plants 2023, 12, 3705.

Th. intermedium: I am no expert in cytogenetics, but the assignment of individual chromosomes to particular progenitors seems somewhat arbitrary. Have you tried hybridizing genomic DNA probes from Aegilops and Dasypyrum as well? Mapping of reads (Fig. 2c) suggests that GISH with these probes could work (since both mapping of reads and GISH show repeatome homology).

Response: Thanks for the valuable comment. In addition to St-subgenome validation by GISH experiments, we have also attempted direct hybridization of total genomic DNA from *Aegilops tauschi* to *Th. intermedium* chromosomes, but no distinct signals corresponding to ThinJ chromosomes were observed (In our previous point-by-point response letter line85). This is a common phenomenon in GISH experiments even when genomic sequences show close homology to the proposed subgenomes, since many repetitive sequences can cross-hybridize with non-target subgenomes. Successful GISH signal detection requires not only optimal hybridization stringency but also depends on the efficiency of DNA blocking. Therefore, we developed the *Aegilops*-species-specific probe Oligo-Ae584 for FISH analysis in this study. This probe successfully hybridized to the ThinJ chromosomes in *Th. intermedium*.

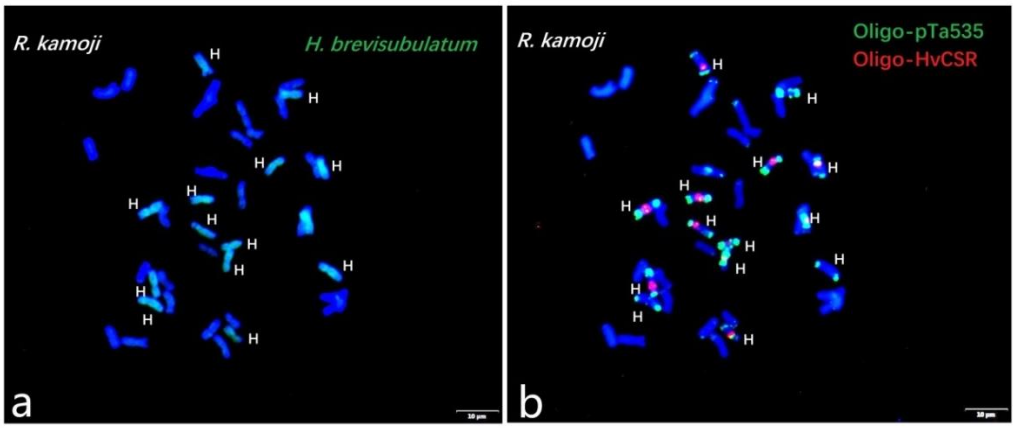
Regarding the ThinV subgenome, our colleagues had previously confirmed its identity through GISH using V-genomic DNA and FISH using Oligo-pDb12H as probes (Figure below, Qi, et al., 2023). We therefore only utilized the V-genome specific probe Oligo-pDb12H in this study, which clearly identified 14 chromosomes associated with ThinV in *Th. intermedium* (Supplementary Fig. 14b).



Chromosomes of *Th. intermedium* probed with (A) genomic DNA of *D. villosum* and (B) oligos pDb12H (green) and St2-80 (red) (Qi et al., 2023)

R. kamoji: GISH (Fig. 15a) – Why not hybridize a total genomic DNA probe from Hordeum? Mapping of Hordeum reads shows the greatest homology (suggesting Hordeum is indeed a progenitor), so this experiment would help to identify the Hordeum chromosomes first. The remaining chromosomes could then be analyzed using oligo-specific (or genomic) probes.

Response: Thanks for the valuable comment. We have now obtained the genomic DNA of wild *Hordeum* (*H. brevisubulatum*) and conducted the GISH experiment (Figure below)**.** The results clearly distinguished the RkaH chromosomes in *R. kamoji*, consistent with our FISH approach using both the *Hordeum*-specific centromeric repeat probe Oligo-HvCSR and the high-copy tandem repeat probe Oligo-pTa535.



Chromosomes of *R. kamoji* probed with (a) genomic DNA of *H. brevisubulatum* and (b) Oligo-pTa535 (green) and Oligo-HvCSR (red)

Fig. 15b and d. Are you sure that the oligo-pTa535 is Hordeum-specific as stated in the legend? Its name suggests that it is Triticum aestivum-specific. If it was Hordeum-specific, both probes should hybridize on the same chromosomes.

Response: We apologize for this misleading statement. The probe Oligo-pTa535 was originally characterized in the wheat D subgenome. Since it was also found highly amplified in wild *Hordeum* but not in RkaSt or RkaY subgenome, we therefore used Oligo-pTa535 to complement detection of RkaH chromosomes in *R. kamoji.* We have revised the main text and figure legend for Supplementary Fig. 15.

Fig. 15c. The figure is confusing. Why is there a signal on H chromosomes when the probes are St- and V-specific??

Response: We appreciate the comment and apologize for the confusion. Our BLASTN analysis shows that probe Oligo-pDb12H is specific to the V-lineage, while probe Oligo-B11 lacked sufficient specificity and has now been deleted (Supplementary Fig. 15c). Actually, there was no clear signal distinguishing the RkaY or RkaH subgenome using Oligo-pDb12H. The unexpected signal likely resulted from residual staining due to incomplete washing after the GISH experiment. To avoid confusion, we have removed Supplementary Fig. 15c,d.

I suggest that the authors clarify which probes were used, what their specificity was, and what sequence they target. To make the use of probes reliable, a reciprocal cross-hybridization test should be performed on all diploids (St, V, J).

Response: We sincerely thank the reviewer for emphasizing the need for clarity regarding probe characteristics and validation. For GISH analyses, total genomic DNA of *Pse. spicata*, *D. villosum*, *H. brevisubulatum* served as probes for the ThinSt/RkaSt, ThinV, and RkaH subgenome detection, respectively. The specificity of the probes Oligo-HvCSR, Oligo-pTa535 and Oligo-pDb12H has been established in

prior cytogenetic studies and have been widely used in numerous chromosome engineering studies within the community. The Oligo-Ae584 probe was *de novo* designed in this study to target *Aegilops*-specific tandem repeats. To validate specificity, we aligned all Oligo probe sequences against high-quality genome assemblies of multiple *Triticeae* species using BLASTN, which revealed the expected amplification patterns and preferential high copy numbers in their target genomes/subgenomes. Consequently, we revised our cytogenetic analyses (Supplementary Fig. 14–15), omitting the non-specific Oligo-B11 probe in FISH.

For *Th. intermedium*: The ThinSt subgenome was clearly identified by GISH using *Pse. Spicata* genomic DNA. The ThinV subgenome, previously detected by GISH and FISH experiments, was further confirmed using the V-specific probe Oligo-pDb12H. As GISH with *Ae. tauschii* genomic DNA failed to detect the ThinJ subgenome, we therefore developed the *Aegilops*-species-specific probe Oligo-Ae584 for FISH analysis, which successfully hybridized to the ThinJ subgenome.

For *R. kamoji*: The St and H subgenomes were clearly identified by GISH using *Pse. spicata* and *H. brevisubulatum* genomic DNA, respectively. The identity of RkaH subgenome was further confirmed by the *Hordeum* centromeric specific probe Oligo-HvCSR and the high-copy tandem repeat probe Oligo-pTa535. However, no distinct signals were obtained for the RkaY subgenome using either *D. villosum* genomic DNA (GISH) or the V-specific Oligo-pDb12H (FISH), likely due to the distant evolutionary relationship between V and Y subgenomes as revealed by genomic analyses.

Collectively, these cytogenetic results consistently support the genome constitution of *Th. intermedium* and *R. kamoji* as revealed in our genomic and phylogenetic analysis. It is noted that both GISH and FISH produce signals depending on DNA hybridyzation density in certain chromosomal regions, and both face technical difficulties and inherent limitations. Sometimes, clear-cut GISH signals could not be obtained as expected since even distant species can share many similar repeat sequences. Even for the same species or genus in *Triticeae*, repeat sequences often exhibit complex evolutionary patterns within species or during polyploidization. Given lack of genome assembles for most *Triticeae* species, we currently can only design or validate probe specificity in specific species or contexts. We hope that in the future, high-quality assemblies will become available for most species, enabling systematically validated specific probe design and verification across the tribe.

l. 257. Please add the bootstrap values to the tree in Fig. 2a. If clade VII is supported, the V-genome ancestry conclusion is correct.

Response: Following the reviewer's suggestion, we have added the bootstrap values to Fig. 2a and Supplementary Fig. 10. The bootstrap values of clade VII and VIII were all 100, which support the placement of ThinV and RkaY within the V-clade of *Dasypyrum*.

l. 297 (Suppl. Fig. 17). Please add all necessary details to the phylogenetic analysis so

that the legend is self-explanatory.

Response: We have added the detailed information to the legend of Supplementary Fig. 17 as suggested.

l. 304-305. No, there is no such close relationship. Pseudoroegneria is clearly the maternal lineage. Omit this part from l. 304 to 311 and do not complicate the story.

Response: We have deleted this part as suggested.

l. 396-400. In line with the previous comment, I suggest deleting this part (starting with 'whereas...' on l. 395).

Response: We have deleted this part as suggested.

l. 515. petri dish with uppercase P

Response: We have corrected this typo.

l. 650. Aegilops caudata

Response: We have corrected this typo.

l. 655. Pse. cognata

Response: We have corrected this typo.

l. 663. Secale cereale

Response: We have corrected this typo.

l. 686. high-molecular with lowercase h

Response: We have corrected this typo.

l. 911. predicted instead of predicated

Response: We have corrected this typo.