

# Chromatin accessibility landscape and its association with heterosis in maize hybrids

Corresponding Author: Professor Zhaobin Dong

**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)

Huang and colleagues utilized MH-seq and RNA-seq data from maize B73, Mo17, and F1 hybrids to explore the potential contribution of chromatin accessibility to heterosis in maize. The aim of this work is commendable, and the authors employed a pan-genome approach for analyzing F1 data, effectively minimizing the mapping bias commonly encountered in heterosis studies. The writing is concise and clear, with a logical flow, and the figures appear well-prepared. I have the following questions for the authors to address.

Major:

1. The authors identified ACRs in B73 and Mo17 using MH-seq. It would be beneficial for the authors to analyze the consistency between this method and previously published ATAC-seq data for B73. The authors discovered that among the syntenic ACRs between B73 and Mo17, approximately 20% are inbred-specific. Further data is needed to clarify whether these specificities are due to data saturation issues or represent genuine genotype-specific ACRs.
2. The authors analyzed MH-seq and RNA-seq data from B73, Mo17, and the F1 hybrid (BM). Theoretically, changes of ACRs in the F1 should be tightly linked to the inheritance patterns of gene expression. However, the authors identified widespread transcriptional downregulation without a corresponding downregulation in chromatin accessibility. How can this be explained?
3. In Lines 205-218, studying the heterosis-associated ACRs from the perspective of Andropogoneae evolution could provide insights into the evolutionary history of heterosis. However, I did not discern the main conclusion from this part of the analysis? Furthermore, to my knowledge, the Andropogoneae CNS are all based on the B73 reference genome. How were the various types of ACRs handled under the Mo17 reference genome?
4. The authors utilized high-resolution ChIA-PET and HiChIP data to identify different categories of heterosis ACRs and their interacting targets. A major concern here is that these chromatin interaction data are all from B73, which likely prevents a complete view of ACR interaction partners in the F1 state. The authors should, at a minimum, explicitly state the genotype of these data in the manuscript and discuss the limitations arising from the lack of 3D omics data for the F1 hybrid.
5. Which target genes are regulated by the transgressive ACRs in the F1, and do these target genes also exhibit transgressive expression?
6. The use of the term "Super-Enhancer like" requires extra caution in plants. I would recommend the authors use "CRE-cluster" or "ACR-cluster" to avoid unnecessary debate.
7. In Line 367, the authors analyzed the contribution of transgressive MHS to GWAS loci. I am particularly curious because the authors' ACR data is limited to seedlings, while these developmental traits (e.g., maturity, flowering time) are primarily assessed at later stages. Does this imply that these heterosis-associated ACRs are very stable throughout development? Considering this question, MH-seq or ATAC-seq data from tissues other than seedlings would seem highly valuable.
8. In Lines 397-414, the authors identified sm2 as a candidate gene and associated ACRs related to heterosis using the node number trait. Transgenic F1 plants indeed showed increased plant height. My question is: is this increase in plant height a consequence of increased node number, or is it due to increased internode length? Can node number itself exhibit heterosis in maize F1 hybrids?

Minor:

1. Line 77: There is a typo for "Dnase-seq" (should be "DNase-seq").
2. Lines 290-292 seem to discuss the same issue as Lines 187-188? Also, I would not recommend using "greater heritability" to describe "higher conservation."
3. Line 299: "genomic reach" might need rephrasing; Line 321: "To address this, we grouped genes...."

4. Avoid repeatedly using vague terms like "regulatory mechanism" (e.g., Line 337).  
5. Line 401: Conclusions about interactions in the F1 cannot be drawn using HiChIP data from B73.

Reviewer #2

(Remarks to the Author)

In the manuscript "Chromatin Accessibility Changes in Maize Hybrids Contribute to Heterosis", Huang and colleagues tried to reveal the molecular mechanism of maize heterosis from the variation of accessible chromatin region between hybrid and its parents. They characterized chromatin accessibility in inbred lines B73 and Mo17 maize and their hybrids by using MH-seq method, and identified the inheritance mode and cis-/trans-regulatory basis of chromatin accessibility. Finally, the authors proposed a framework for the inheritance and regulatory mechanisms of chromatin accessibility in maize hybrids, and suggested the role of chromatin accessibility in heterosis. This is the first MH-seq analysis on hybrid maize and its parents. The sequencing data seem properly analyzed, and the manuscript is well structured, explains the experimental design and results clearly. However, I have some major concerns about the manuscript.

(1) One main issue here is that whether heterosis is existed in the hybrid of this cross combination. It is well known that not any two maize varieties will produce heterosis when crossed. This research conducts a cross between two maize inbred lines, B73 and Mo17, and then analyzes the molecular mechanism of heterosis in the F1 hybrid. Lines 145-146 showed "Given the extensive variation in ACRs between B73 and Mo17, we wondered how these regions were inherited in their F1 hybrid (BM), which exhibits hybrid vigor for various phenotypes compared to the parents". However, this study did not show that the hybrid offspring of these two varieties had heterosis. If there was heterosis, it should also be stated at which developmental stages and in which organs (tissues) it was occurred. The material used in the research was the leaves (shoots) at the seedling stage, and it should also be stated whether there was heterosis in the seedling leaves and what the phenotypic or physiological characteristics of the heterosis were.

(2) The experiment design in this study is not integrity and comprehensive. The experiments on hybrids and their parents only involved chromatin accessibility and transcriptome analysis, while the data used for histone modification and GWAS analysis were obtained from other literatures. The materials used in those studies were not the same as the hybrid and its parent combinations in this study. This indicates that the conclusions are not accurate and comprehensive enough. The authors should conduct studies on 3D genome, WGBS, histone modifications, etc. to explain the mechanism of chromatin accessibility changes and their regulation on gene expression and phenotype during the hybridization process.

(3) In section "Functional consequence of open chromatin with different inheritance modes" and other sections, some important results and conclusions were cited from references. Since this is a section in Results part, the authors should indicate the results and conclusions supported by their own experiments.

Reviewer #3

(Remarks to the Author)

This is an excellent paper that presents complex and detailed analyses of chromatin accessibility in hybrids. The work addresses an important knowledge gap. While prior studies have extensively examined gene expression, small RNAs, and some chromatin features in hybrids, chromatin accessibility in hybrid combinations has received comparatively less attention in maize. This study provides a robust and comprehensive foundation for understanding accessibility patterns in hybrids. However, I recommend that the title and conclusions be moderated to better align with what the data can definitively support.

The analyses are thorough and provide compelling evidence that chromatin accessibility at some loci exhibits non-additive and transgressive patterns in hybrids. The depth and rigor of the analyses are good, though I note that the density of some sections may require multiple readings to fully digest, which could present challenges for a broader audience. Nonetheless, the results are valuable and contribute meaningfully to the field.

Major:

My primary reservation is that the current data, while correlative and intriguing, do not conclusively prove that chromatin accessibility changes contribute to or are required for heterosis. Establishing causality would require more extensive functional validation beyond the correlative patterns presented. Specifically, I would recommend: (1) analysis across multiple hybrid combinations, including reciprocal crosses, with increased biological replication (the current study uses a single non-reciprocal(?) hybrid with only 2 replicates); and (2) more detailed functional validation with appropriate controls. The preliminary functional work on the sm2 MHS locus is a promising start but requires additional detail and controls to be fully convincing.

It should also noted that a recent elegant study analyzing chromatin accessibility and transcription factor occupancy in hybrids using MOA-seq was published earlier this year (<https://doi.org/10.1038/s41588-025-02246-7>). This reduces the novelty of the current work, although the two studies can be viewed as complementary. I recommend that the discussion explicitly address the similarities and differences between these two approaches.

Despite these concerns, this remains a significant contribution that warrants consideration for publication, provided the claims in the title and conclusions are appropriately tempered to reflect what the data can support.

Other specific comments:

Line 119, please specify how the correlation was calculated, including eg Spearman or Pearson and which reads – eg all reads binned or reads in peaks?

Line 120 it is mentioned that there are around 60k MHS sites in each genome, does this only include peaks after IDR filtering? Please mention how many peaks there were before and after IDR; what was the proportion that were reproducible?

Line 126 and Line 693, it is stated that “we extracted only the best uniquely aligned reads from the bam files”, considering there is often confusion in the community about what a uniquely aligned read is (or the infamous “multimapping” read), please specify the bioinformatic approach to identify the “best uniquely aligned read” and “multiple aligned reads”, including the specific approach to filtering the bam, software and arguments. Considering this is central to the main findings of the paper, this is essential to explain in sufficient detail and code.

Line 129, Figure 1A is a bit confusing, the light blue box is too big for 1% and its not clear why there is a row each for B73 and Mo17 are the percentages the same for each (seems like on B73 is labelled?). I do really like Supp Figure 1B, I would suggest making this the main panel figure to replace Fig 1A and adding the percentages directly to Supp Figure 1b. In this section in the text, it would be helpful to explicitly say that differentially accessible regions that are also syntenic are only 1% and state how many regions (~600?).

Line 148 and 619 – were both replicates of the hybrids B x M, or was one the reciprocal M x B?

Line 155 – it mentioned that a substantial number of novel open chromatin regions are detected in the hybrid are not detected in either parent, but it is unclear where this is shown in the figure because each panel only compare the hybrid on one parent (not both combined). Please mention in the text how many novel open chromatin regions are detected in the hybrid are not detected in either parent.

Line 168 – please state the total number of MHS in this part of the analysis.

Line 188 The authors also identify some interesting difference between ACRs and gene expression in hybrids for example noting accessibility has more conservative heritability than transcription, but it should be discussed if this is in part due to technical differences between MH-seq and RNA-seq sensitivity and dynamic range.

Line 209 and Figure 2D – it is unclear why this is a box plot and not a single value for the fraction? Please explain in more detail why this is the case and why multiple comparisons of the fraction values were conducted by the least significant difference (LSD) method.

Line 210 – Text should be Figure 2D (not 2C). The figure call outs for this section appear to be out for E and F as well.

Line 237 – this section compares chromatin accessibility in the hybrid to chromatin interaction data but presumably the chromatin interaction data is all from B73 (?). If so, the significance of comparing hybrid MHS to B73 interaction data should be made clear at the beginning of the section. The source of the data should also be mentioned in the text and/or Figure legend.

Line 277 – the authors speculate about transgressive syntenic super enhancer MHS hubs. It is an evocative idea, but to support this speculation, it would be good to mention how many MHS loci fit this description. Is it more than just this example? What fraction?

Line 361 – it is mentioned there is an enrichment for developmental traits, but which of the traits tested were not developmental, seems like most are developmental; if so, can it really be said that there is an enrichment?

Line 398 – please detail the re-analysis of HiChIP data. Why was reanalysis required and how was it performed differently?

Line 401 – was HiChIP performed in the hybrid?

Line 397 – it is difficult to understand which is the transgressive up regulate locus in this example (Fig 5C). B73 has a value of 23, BM is  $17 + 12 = 29$  and Mo is 34? It would be useful to explain the calculation using an example, to make this clear.

Line 410 and Fig 5E – was the height of the inbreds (LH244, LH244-OE and T13) also collected? It would be useful to include the inbreds to demonstrate hybrid vigor was observed and the effect of OE in inbred background. Please also include the number of plants/plots tested and trial details. For multiple of the tests (eg years and locations) some of the OE lines were not significantly taller than the control hybrid, this should be discussed. A reciprocal cross would also be idea, but at least including all the control is necessary.

Line 445 – it is unclear what is mean by “poised” state?

Line 471 – it is unclear why the more stable and ‘robust’ feature would be more likely to provide an explanation to heterosis. It seems more likely that the feature that is more transgressive/variable in hybrids would have more potential to offer an explanation? This assertion should needs more reasoning.

Line 508 – it is stated that chromatin accessibility changes in hybrids have lower magnitude change of expression. This is

significant and should be explained with more detail in both results and discussion, including some numbers to define the magnitude (or refer to a relevant Figure if it is already created).

#### Reviewer #4

##### (Remarks to the Author)

The study presented by Huang et al. addresses the question if chromatin accessibility changes in maize hybrids can be linked them to heterosis. Overall, the study provides a bioinformatic pipeline to assess differential chromatin accessibility, measured by MNase Hypersensitivity sequencing. The method was employed to compare accessible chromatin (ACRs) and transcriptomes from two maize accessions, Mo4 and B73 and their hybrids using 10-day old seedlings or leaves (the information is not quite consistent). Interesting insights include the observation that chromatin accessibility is less variant than transcript levels and that 20% of ACR regions in syntenic regions show variation that is most likely explained by sequence divergence. A further focus of the manuscript were non-additive/transgressive ACRs. Transgressive upregulated ACRs were found to be more conserved in sequence and linked to potential super-enhancers based on published Hi-C data. They were linked to genes upregulated in hybrids. Overexpression of a candidate gene showing linkage to an upregulated ACR led to increase plant height, which the authors interpret as potential contribution to heterosis. The dataset is rich, and the analysis of cis-/trans-regulatory patterns and transgressive chromatin accessibility is conceptually strong and relevant. However, the manuscript is not written for a more general audience and a could be improved by more explicit introduction of the analyses performed. Furthermore, the manuscript contains numerous errors in figure citations, labeling, and legends. The frequency of such mistakes is distracting and make reviewing unnecessarily time-consuming, distracting from the otherwise solid scientific content.

##### Major comments

1. Figure 1A. It is not immediately clear how the numbers add up in the graph. 81% of the ACRs were syntenic, 19 non-syntenic. The block within classifies 80% of syntenic as non-differential, but at first glance the figure seems to indicate that only 1% of the syntenic sites were non-consistent open chromatin.
2. Figure 1B. It is not quite clear why the curves for SNP density for “syntenic accessible” are not the same for both genotypes – should they not be identical? The figure makes a point about differences between variant and non-variant ACRs and this is quite clear, but the different baseline is confusing.
3. In Figure 2A, chromatin accessibility and gene expression inheritance are classified into four modes: Conserved, Additive, Dominant, and Transgressive. First, transgressive should be defined as transgressive up and transgressive down (beyond and below parental values, respectively)
4. Additivity should better be Additive to align grammatically with dominant, transgressive etc.
5. According to the description, Additive means that hybrid (BM) chromatin accessibility equals the mid-parent value (lines 165-166),  $BM = (B73 + Mo17)/2$ , falls between both parents ( $B73 > BM > Mo17$  or  $B73 < BM < Mo17$ ). Transgressive is defined as  $BM > B73$  &  $BM > Mo17$  or  $BM < B73$  &  $BM < Mo17$ . Could the authors clarify how cases are categorized when  $B73 > BM > Mo17$  or  $B73 < BM < Mo17$  but  $BM \neq (B73 + Mo17)/2$ ? For example, in Figure 5C, sm2-MHS is classified as a “transgressive up-regulated” MHS (line 397). However,  $BM (17.82 + 12.05 = 29.87)$  does not exceed Mo17 (34.50).  $Mo17 > BM > B73$ . Actually, the mean of B73 and Mo17 (29.12) is similar to BM. Could you explain why this case was classified as transgressive up rather than additive?
6. Similarly, for Zm00001eb290700 (line 277) and Zm00001eb080410 (line 417) gene expression, which are described as down- and up-regulated respectively:
  - Zm00001eb290700 in BM ( $2.93 + 1.69 = 4.69$ ) is lower than B73 (13.39) but higher than Mo17 (1.76) ( $B73 > BM > Mo17$ ;  $BM \neq (B73 + Mo17)/2$ ) (Figure 3E).
  - Zm00001eb080410 in BM ( $18.18 + 1.50 = 19.68$ ) is greater than B73 (0.60) but similar to Mo17 (21.23) ( $Mo17 > BM > B73$ ;  $BM \neq (B73 + Mo17)/2$ ) (Supplementary Figure 8C). Please clarify how up- and down-regulation were defined for these cases.
7. Line 168: Fix grammar
8. Line 237 and following: The comparisons were based on published chromatin data that must certainly show a bias towards the maize variety that was used to generate them. This should be assessed and discussed. For example, the bias for H3K27me3 regions towards Mo17 dominant lower MNH sites could be explained by H3K27me3 data being generated in Mo17.
9. Figure 3E and 5B are not intuitive and need a better legend (at least the first time). It is not intuitive to plot one accession above, the other below the x-axis, unless they are plotted on the same x-axis. The number labels are so small that these cannot be recognized without zooming
10. Novel open chromatin sites in hybrids (lines 154–155). The statement “BM harbored a substantial number of novel open chromatin sites not detected in either parent” (lines 154–155), the current comparison (BM vs. B73 and BM vs. Mo17 separately in Supplementary Figure 2CD) is indirect. It would be more convincing if BM were compared to an integrated B73–Mo17 pangenome rather than to B73 and Mo17 separately (as in Supplementary Figure 2C–D).
11. 299ff The analysis of distance decline of cis-variation is not clear from the description.
12. Line 356: What are broad TF types?
13. In Figure 5C, the Mo17 Chr2 side lacks the gray bar and red dashed box used to indicate sm2-MHS and its abundance. Was this omitted intentionally (e.g., due to lack of synteny) or inadvertently?
14. sm2-overexpression lines (OE1–OE3). In Figure 5, the height of the sm2 overexpression (OE) line OE3 is more frequently and significantly higher than that of the WT control compared with OE1 or OE2. Is the sm2 expression level in OE3 correspondingly higher than in OE1 and OE2? Please provide the sm2 expression levels in all OE lines and the control to determine whether plant height correlates with sm2 expression.
15. Figure citation corrections  
Please revise the following for consistency with figure content:
  - Lines 188–189: “Figure 2A” → “Figure 2B”.

- Lines 191–192: Remove “Figure 2B”.
- Lines 210–218: “Figure 2C” → “Figure 2D”; “Figure 2D” → “Figure 2E”.
- Line 227: “Figure 2E” → “Figure 2F”.
- Lines 303–305: “Figure 4B” → “Figure 4C”.
- Line 415: “Supplementary Figure 8B” → “Supplementary Figure 8C”.
- Regarding Supplementary Figure 3 (Lines 195-199), the suggested fix is:
- Switch the order of panels C and D in Supplementary Figure 3.
- Then, change the text citation in Line 199 from “Supplementary Figure 3C” to “Supplementary Figure 3D” accordingly.
- 16. Abbreviation inconsistency (lines 366–367 and Supplementary Figure 7D). “Chromatin accessibility genome-wide association” is abbreviated as CAWAS (line 366), but CaGWAS is used in Supplementary Figure 7D, and CWAS in the legend. If these refer to the same analysis, please standardize the abbreviation; if not, define them clearly.
- 17. Line 265: The sentence is unclear. Did you bulk all inheritance modes and that overlapped with 5% of the superenhancers? Probably not but the sentence sounds as if.
- 18. You refer to Zm00001eb324900 as being “described in Figure 3E,” but it does not appear there. Please correct the figure citation to the accurate panel or explicitly label the gene in Figure 3E.
- 19. Figure legends and abbreviations. Nature Communications requires that abbreviations be defined at first use in figure legends. Please define terms such as CNS, SV, and GERP (Figure 2), and SS and NSS (Supplementary Figure 7). While these are standard in maize research, definitions will aid broader readership.
- 20. Supplemental 1E: colors are different in graph and legend
- 21. Supplemental 1F: the legend states that the alignment is per peak summit, not to each TSS as indicated in the x axes
- 22. In the legend of Figure 4C, the text states “0–50 kb (Top panel) and 0–400 kb (Bottom panel),” but the actual plots display 0–50 kb (bottom) and 0–500 kb (top). Please correct this discrepancy.
- 23. In the legend of Supplementary Figure 1H and I, terms Syn-Incon and Syn-Con should be clarified—are they equivalent to Syn (Acc) and Syn (Inacc) used previously? If so, ensure consistent terminology.
- 24. Please review Supplementary Figure 7 and confirm that the legends for panels D and E should be switched to align with the visual presentation of the data.
- 25. A similar study has recently been published by Engelhorn et. al. in Nature Genetics, the authors should discuss their data in context of this study.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

The authors have made a great job to revise their manuscript, the manuscript is now significantly improved. I suggest the authors to put the Response Figures regarding BM HiChIP data into SuppFigures to better clarify the chromatin interaction patterns of F1 seedlings. I have no further questions.

Reviewer #2

(Remarks to the Author)

The authors have revised the manuscript according to the suggestions.

Reviewer #3

(Remarks to the Author)

Excellent revision, all my comments/concerns were addressed.

One reply in the response to reviewers was a little unclear – comment 16, reviewer 3 – it is mentioned that this is a non-syntenic region (Fig 5c, right panel), and if so, then the analysis makes sense. But if that is the case, why is there a peak at virtually the same position in the Mo17 track (which region does the correspond to in B73?) and why is it labelled with a count, if the count is not relevant? I have probably misunderstood, but this might be clarified.

Other than that, I no further concerns.

Reviewer #4

(Remarks to the Author)

The authors have addressed my points to my satisfaction. Also based on the serious reply to all reviewer's comments, I find the paper much improved.

Only some minor points remain:

- 1.) Title: should better use the plural of hybrid
- 2.) Line 117: merging peaks detected in both replicates
- 3.) Line 382: Supplemental Table X

Version 2:

Reviewer comments:

Reviewer #3

(Remarks to the Author)

Thanks for the clarification, no further concerns.

**Open Access** This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

## **Point-by-point responses to the comments:**

### **Reviewer 1:**

#### **General Comments:**

Huang and colleagues utilized MH-seq and RNA-seq data from maize B73, Mo17, and F1 hybrids to explore the potential contribution of chromatin accessibility to heterosis in maize. The aim of this work is commendable, and the authors employed a pan-genome approach for analyzing F1 data, effectively minimizing the mapping bias commonly encountered in heterosis studies. The writing is concise and clear, with a logical flow, and the figures appear well-prepared. I have the following questions for the authors to address.

#### **Reply:**

We thank the reviewer for the positive and constructive assessment. We appreciate the recognition of our pan-genome approach and the clarity of the presentation. We have carefully addressed the reviewer's specific questions and comments, corrected inconsistencies, and revised the manuscript and figures where needed.

### **Major:**

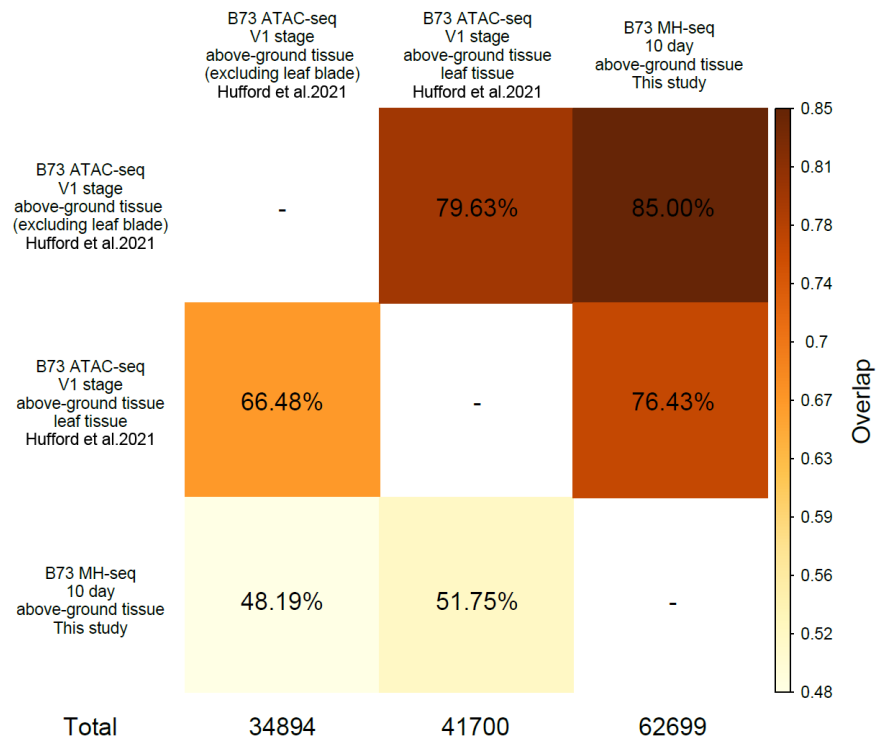
#### **1. Comment:**

The authors identified ACRs in B73 and Mo17 using MH-seq. It would be beneficial for the authors to analyze the consistency between this method and previously published ATAC-seq data for B73. The authors discovered that among the syntenic ACRs between B73 and Mo17, approximately 20% are inbred-specific. Further data is needed to clarify whether these specificities are due to data saturation issues or represent genuine genotype-specific ACRs.

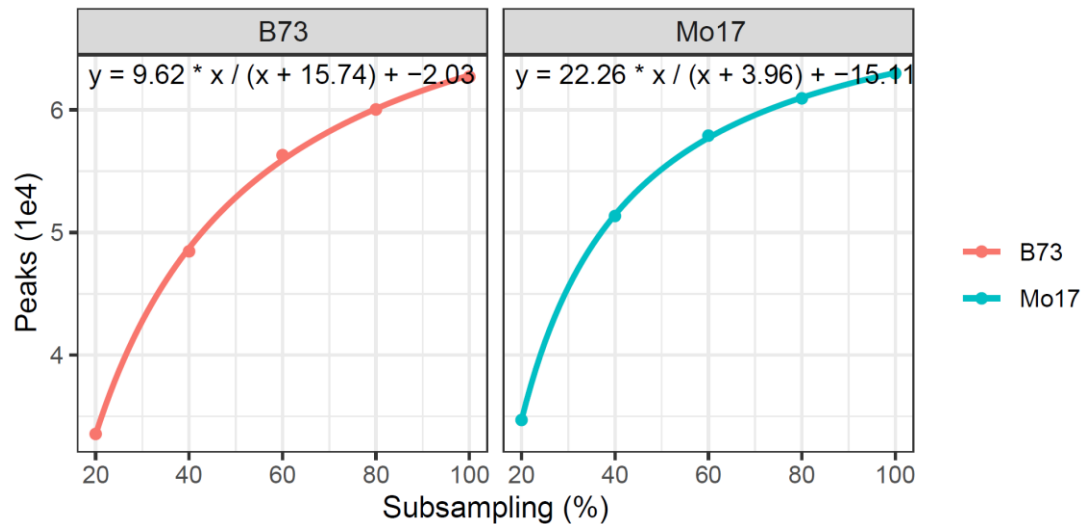
#### **1. Reply:**

Thank you for your feedback. We have analyzed the consistency between our MH-seq and previously published ATAC-seq data from similar seedling above-ground B73 tissues (Hufford et al. 2021). To ensure a fair comparison, we used a consistent analysis pipeline with MH-seq data (see Methods). Finally, we identified 34894, 41700, and 62699 peaks for two ATAC-seq experiments and our MH-seq dataset, respectively. Approximately 76.43–85.00% of the peaks identified by ATAC-seq were covered by MH-seq, ~48.19-51.75% MH-seq peaks are covered by ATAC-seq (Response-Fig. 1). These results were largely consistent with previous reports in Arabidopsis (Zhao et al. 2020), indicating that MH-seq can provide a valuable tool for identifying and cataloging more comprehensive classes of open chromatin in maize.

To clarify whether the inbred-specific ACRs are due to data saturation issues, we performed subsampling on our real MH-seq data, ranging from 20% to 100%, to test for saturation. After fitting the model  $y = a_1 \times x / (x + a_2) + a_3$  to estimate the growth curve, we can calculate the saturation ratio with:  $Saturation\ Ratio = y_{observed} / (a_1 + a_3)$ , where  $y_{observed}$  are the actual peak counts identified with real MH-seq data. We found ~83% and 88% of the peaks have been captured by current data, indicating our MH-seq data are nearing saturation (Response-Fig. 2). So, our analysis between B73 and Mo17 represent genuine genotype-specific ACRs.



**Response-Fig. 1:** Overlaps among MH-seq and ATAC-seq peaks. The percentage in each grid represents the overlapping rate between the features in rows and the features in columns. The percentage in each cell at the intersection of a row and column represents the overlapping rate between the source feature (row) and the linked feature (column). For example, the cell at the intersection of the MH-seq row and the ATAC-seq (leaf tissue) column indicates that 51.75% of MH-seq peaks overlapped with ATAC-seq peaks.



**Response-Fig. 2:** The saturation curve of B73 and Mo17 MH-seq data. The real MH-seq data were subsampled, ranging from 20% to 100%. The growth curve was fitting with the model  $y = a_1 \times x / (x + a_2) + a_3$ .

## 2. Comment:

The authors analyzed MH-seq and RNA-seq data from B73, Mo17, and the F1 hybrid (BM). Theoretically, changes of ACRs in the F1 should be tightly linked to the inheritance patterns of gene expression. However, the authors identified widespread transcriptional downregulation without a corresponding downregulation in chromatin accessibility. How can this be explained?

## 2. Reply:

Thank you for your comments. The apparent decoupling can be explained by multilayered controls for gene expression beyond chromatin accessibility, including histone modifications, transcription factor binding, post-transcriptional processing, translation rates, and mRNA degradation. While chromatin accessibility contribute largely to gene expression, trans-acting factors or others accounted for the variation in gene expression. We have added the above clarification to the Discussion.

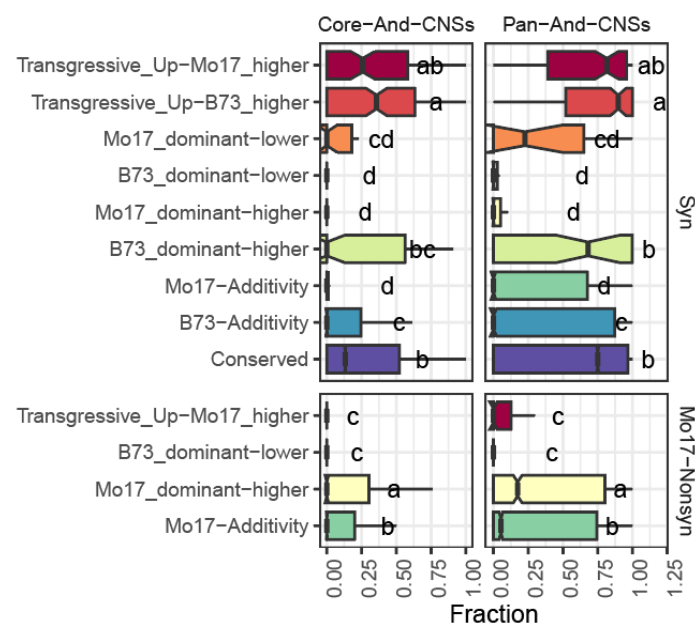
## 3. Comment:

In Lines 205-218, studying the heterosis-associated ACRs from the perspective of Andropogoneae evolution could provide insights into the evolutionary history of heterosis. However, I did not discern the main conclusion from this part of the analysis? Furthermore, to my knowledge, the Andropogoneae CNS are all based on the B73 reference genome. How were the various types of ACRs handled under the Mo17 reference genome?

### 3. Reply:

Thank you. The main conclusion from this part is that the transgressive up-regulated MHSs are the most evolutionary conserved (overlap the largest with *Andropogoneae* CNS) among various types of MHSs, which pinpoint the evolutionary importance of transgressive up-regulated MHSs. We revise this part to reclarify this conclusion.

The published *Andropogoneae* CNS dataset is indeed based on the B73 reference. In the previous version of our manuscript, we did not perform the analysis using the Mo17 reference because we expected the enrichment patterns to be largely consistent across different references. To respond to your concerns, we have now added the analysis based on the Mo17 reference and found that the results are indeed consistent: the transgressive up-regulated MHSs in the syntenic region exhibited the highest fraction of overlap with conserved non-coding sequences (Response-Fig. 3).



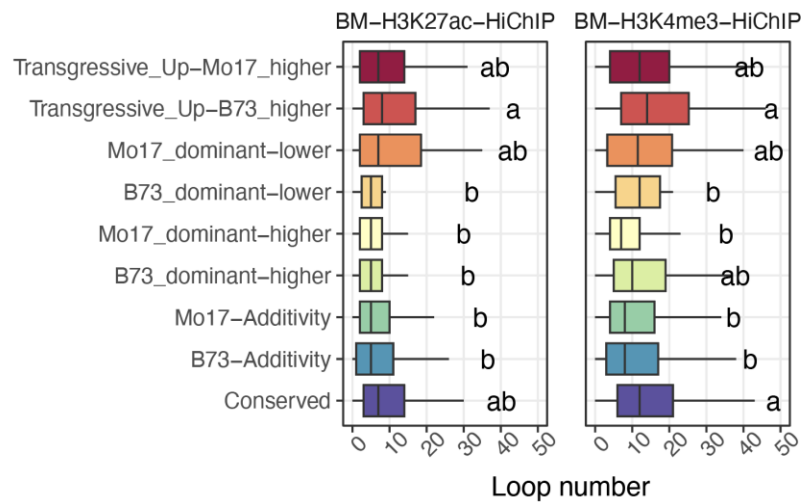
**Response-Fig. 3:** Fraction of MHSs of each inheritance mode overlapping with CNS (based on the Mo17 reference). Multiple comparisons of the fraction values were conducted by the least significant difference (LSD) method.

### 4. Comment:

The authors utilized high-resolution ChIA-PET and HiChIP data to identify different categories of heterosis ACRs and their interacting targets. A major concern here is that these chromatin interaction data are all from B73, which likely prevents a complete view of ACR interaction partners in the F1 state. The authors should, at a minimum, explicitly state the genotype of these data in the manuscript and discuss the limitations arising from the lack of 3D omics data for the F1 hybrid.

#### 4. Reply:

Thanks for your valuable suggestions. We have collected recently published H3K4me3-HiChIP and H3K27ac-HiChIP data from BM F1 hybrid seedling leaf tissues and conducted the analysis (H. Liu et al. 2025). The results indicate that in the F1 genome, transgressive MHS are indeed more likely to involve 3D interactions (Response-Fig. 4), which is consistent with our findings in B73.



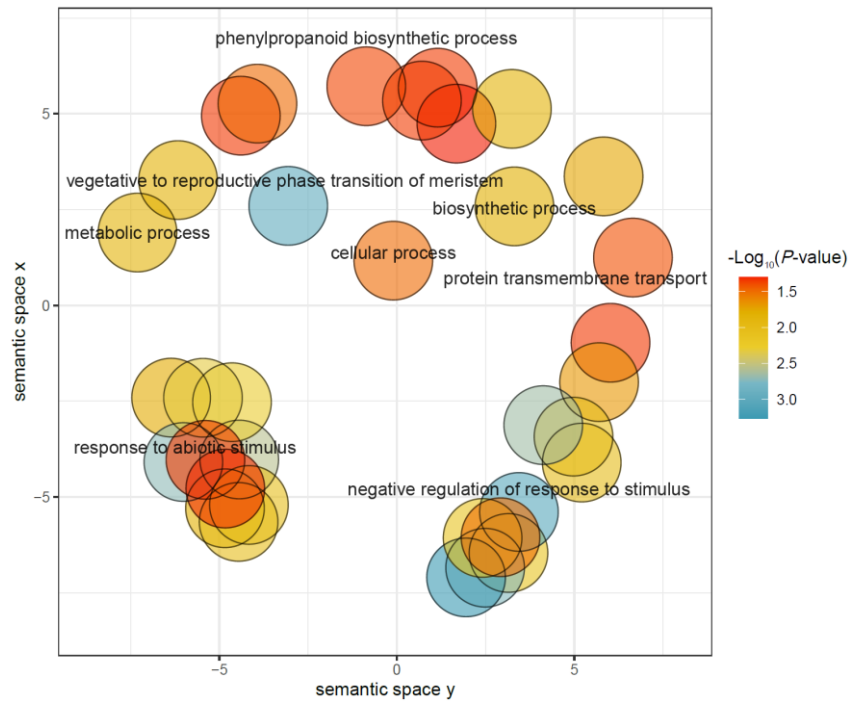
**Response-Fig. 4:** The number of chromatin loops in the BM F1 hybrid in which different inheritance modes of syntenic MHSs are involved.

#### 5. Comment:

Which target genes are regulated by the transgressive ACRs in the F1, and do these target genes also exhibit transgressive expression?

#### 5. Reply:

Thank you for your valuable feedback. We defined the closest genes to the MHSs as the putative targeted genes and identified a total of 251 expressed genes for transgressive MHSs. GO enrichment analysis revealed that these genes were enriched in broad pathways, particularly in the metabolic and biosynthetic processes, stimulus responses, etc (Response-Fig. 5). Of these 251 targeted genes, 16 genes exhibited non-additive expression (4 genes and 12 genes were transgressively and dominantly expressed, respectively). More details for these genes are listed in the Supplementary Table 7.



**Response-Fig. 5:** Enriched gene ontology terms for putative targeted genes regulated by the transgressive MHSs. Semantically similar GO terms should remain close together in the plot. Circle color indicates the  $-\log_{10}$ -transformed  $P$ -value.

#### 6. Comment:

The use of the term "Super-Enhancer like" requires extra caution in plants. I would recommend the authors use "CRE-cluster" or "ACR-cluster" to avoid unnecessary debate.

#### 6. Reply:

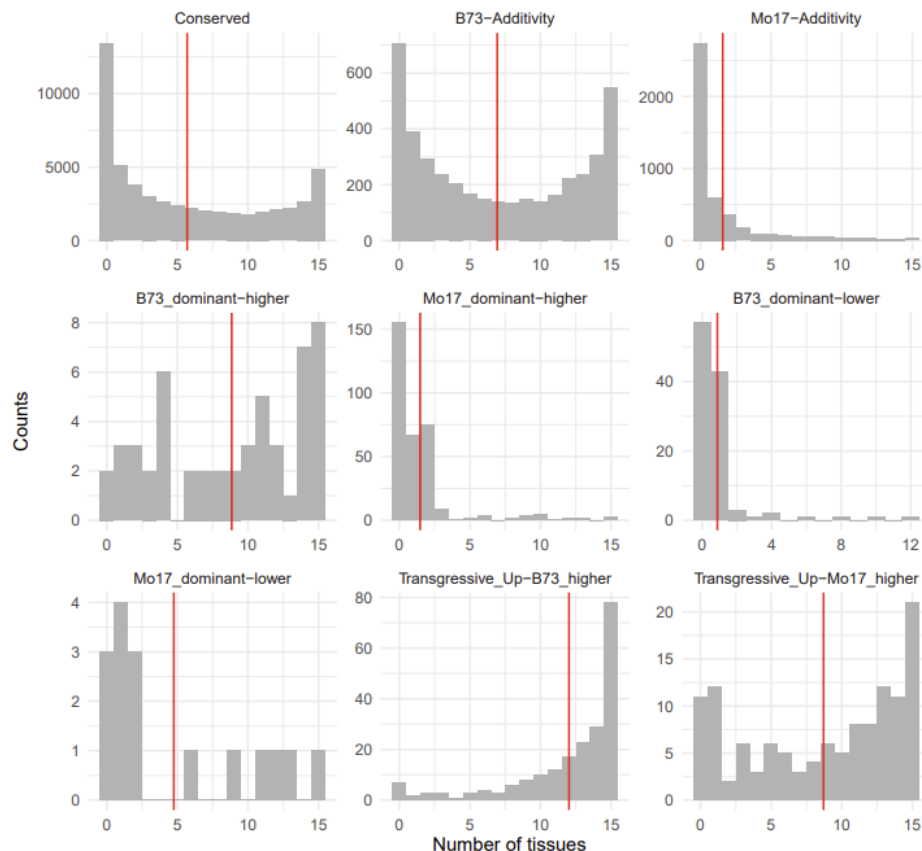
Thank you for your suggestion. We used "CRE-cluster" instead.

#### 7. Comment:

In Line 367, the authors analyzed the contribution of transgressive MHS to GWAS loci. I am particularly curious because the authors' ACR data is limited to seedlings, while these developmental traits (e.g., maturity, flowering time) are primarily assessed at later stages. Does this imply that these heterosis-associated ACRs are very stable throughout development? Considering this question, MH-seq or ATAC-seq data from tissues other than seedlings would seem highly valuable.

#### 7. Reply:

Thank you for your suggestion. We collected previously published ATAC-seq data from 15 major maize tissues (Y. Liu et al. 2025) to test the tissue specificity for our MHS from different inheritance modes. We found that transgressive MHS indeed showed more constitutive and stable accessible features than the other types (accessible in more different tissues) (Response-Fig. 6), which indicated that these heterosis-associated ACRs are tend to be stable throughout development.



**Response-Fig. 6:** Histogram showing the MHSs with different inheritance modes that conserved or specifically existed in ATAC-seq peaks across 15 major maize tissues. The mean number of tissues was indicated by the red line.

## 8. Comment:

In Lines 397-414, the authors identified sm2 as a candidate gene and associated ACRs related to heterosis using the node number trait. Transgenic F1 plants indeed showed increased plant height. My question is: is this increase in plant height a consequence of increased node number, or is it due to increased internode length? Can node number itself exhibit heterosis in maize F1 hybrids?

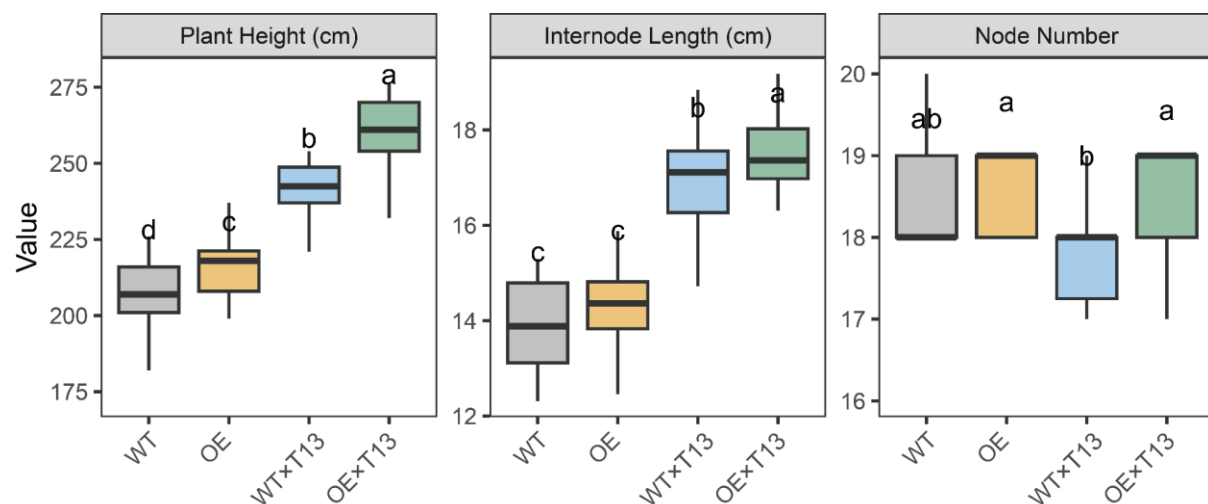
## 8. Reply:

We thank the reviewer for this insightful question regarding the components of plant height. To clarify the contribution of node number versus internode length to the observed increase

in plant height, we collected and analyzed these traits across the wild-type LH244 (WT), overexpression (OE) line, and their respective F1 hybrids (WT×T13 and OE×T13) at Hainan in 2026 (Response-Fig. 7).

Our results indicate that the increase in plant height in the transgenic F1 plants is a result of both major increased internode length and minor increase in node number. The OE F1 plants exhibit the highest internode length, significantly exceeding both the inbreds and the control hybrid (Response-Fig. 7). While the node number does not always show strong positive correlation with plant height (noting that the control hybrid actually showed a slight decrease compared to WT), our data shows that the over-expression of *sm2* significantly increases node number in the hybrid background (Response-Fig. 7).

Overall, the increased plant height in the transgenic F1 is a combined effect: *sm2* promotes both internode elongation and increases the total node number relative to the hybrid control. This confirms that *sm2* acts on multiple components of plant architecture to drive the observed height heterosis.



**Response-Fig. 7:** Phenotypic analysis for plant height, internode length and node number across the wild-type LH244 (WT), *sm2*-overexpression (OE) line, and their respective F1 hybrids (WT×T13 and OE×T13) at Hainan in 2026 ( $n = 25, 20, 22$  and  $29$  for WT, OE, WT×T13 and OE×T13, respectively). Multiple comparisons were conducted by the least significant difference (LSD) method.

**Minor:**

### 1. Comment:

Line 77: There is a typo for "Dnase-seq" (should be "DNase-seq").

### 1. Reply:

Fixed.

### 2. Comment:

Lines 290-292 seem to discuss the same issue as Lines 187-188? Also, I would not recommend using "greater heritability" to describe "higher conservation."

### 2. Reply:

Thank you for your comments. These are two issues for the conservation level of inheritance mode (Lines 187-188) and *cis/trans* regulatory mechanisms (Lines 293-294). We changed the "greater heritability" to "higher conservation".

### 3. Comment:

Line 299: "genomic reach" might need rephrasing; Line 321: "To address this, we grouped genes...."

### 3. Reply:

We rephrased these sentences as:

"To quantify the genomic extent of the effects of *cis*-acting variants, we analyzed the distance-dependent decay of shared *cis*-signals between MHS pairs, following an established strategy." (Lines 301-302).

and:

"To address this, we grouped genes by their expression changes upon hybridization and assessed the corresponding log2-fold change of adjacent MHSs." (Lines 325-327).

### 4. Comment:

Avoid repeatedly using vague terms like "regulatory mechanism" (e.g., Line 337).

### 4. Reply:

We changed it to "regulatory categories".

**5. Comment:**

Line 401: Conclusions about interactions in the F1 cannot be drawn using HiChIP data from B73.

**5. Reply:**

Thank you. We confirmed this interaction was observed both in B73 and the F1 hybrid (Supplementary Figure 8B).

**Reviewer 2:****General Comments:**

In the manuscript “Chromatin Accessibility Changes in Maize Hybrids Contribute to Heterosis”, Huang and colleagues tried to reveal the molecular mechanism of maize heterosis from the variation of accessible chromatin region between hybrid and its parents. They characterized chromatin accessibility in inbred lines B73 and Mo17 maize and their hybrids by using MH-seq method, and identified the inheritance mode and cis-/trans-regulatory basis of chromatin accessibility. Finally, the authors proposed a framework for the inheritance and regulatory mechanisms of chromatin accessibility in maize hybrids, and suggested the role of chromatin accessibility in heterosis. This is the first MH-seq analysis on hybrid maize and its parents. The sequencing data seem properly analyzed, and the manuscript is well structured, explains the experimental design and results clearly. However, I have some major concerns about the manuscript.

**Reply:**

We sincerely thank the reviewer for the positive and constructive evaluation of our manuscript. We have revised the manuscript to address all the concerns raised, and provide detailed responses below.

**Major:****1. Comment:**

One main issue here is that whether heterosis is existed in the hybrid of this cross combination. It is well known that not any two maize varieties will produce heterosis when crossed. This research conducts a cross between two maize inbred lines, B73 and Mo17, and then analyzes the molecular mechanism of heterosis in the F1 hybrid. Lines 145-146 showed “Given the extensive variation in ACRs between B73 and Mo17, we wondered how these regions were inherited in their F1 hybrid (BM), which exhibits hybrid vigor for various phenotypes compared to the parents”. However, this study did not show that the hybrid offspring of these two varieties had heterosis. If there was heterosis, it should also be stated at which developmental stages and in which organs (tissues) it was occurred. The material used in the research was the leaves (shoots) at the seedling stage, and it should also be stated whether there was heterosis in the seedling leaves and what the phenotypic or physiological characteristics of the heterosis were.

**1. Reply:**

Thanks for your suggestions. The trio combination of parent lines from different heterotic groups is considered an excellent system for the study and application of heterosis (Springer and Stupar 2007). The Reid group, represented by the B73 inbred line, and the Lancaster group, represented by the Mo17 inbred line, constitute two of the most classical maize

heterotic groups. This combination was widely grown commercially in the United States and other nations, many maize breeding programs still make extensive use of the resulting materials.

B73-Mo17 hybrids exhibit significant growth vigor at the early seedling stage, including biomass, plant height, and fresh weight, establishing this trio as a longstanding, most widely used and representative material set for heterosis studies (Springer and Stupar 2007). Thus, many studies have collected extensive molecular data based on the B73-Mo17 trio combination at the seedling stage to study heterosis (Crisp et al. 2020; Luo et al. 2021; Li et al. 2020; He et al. 2013).

We stated that our sampling of 10-day-old seedling leaves followed the heterosis phenotypic characteristics of biomass, plant height, and fresh weight (Lines 670-671).

## **2. Comment:**

The experiment design in this study is not integrity and comprehensive. The experiments on hybrids and their parents only involved chromatin accessibility and transcriptome analysis, while the data used for histone modification and GWAS analysis were obtained from other literatures. The materials used in those studies were not the same as the hybrid and its parent combinations in this study. This indicates that the conclusions are not accurate and comprehensive enough. The authors should conduct studies on 3D genome, WGBS, histone modifications, etc. to explain the mechanism of chromatin accessibility changes and their regulation on gene expression and phenotype during the hybridization process.

## **2. Reply:**

Thanks for your valuable comment. We acknowledge that GWAS and some other data may not perfectly match the heterozygous condition. Nevertheless, given the rationale based on the fact that a high frequency of loci maintains functional conservation between inbreds and hybrids, annotating open chromatin function using GWAS hits from an inbred population could be seen as an insightful reference for the hybrid, indeed this analytical paradigm follows established practices in similar studies (Y. Liu et al. 2025; Zhu et al. 2024). Although the background materials in studies may differ from those used in GWAS, GWAS loci reflect the potential functional phenotypic outcomes of these regions.

Additional supporting evidence for understanding the functions of different types of MHS, particularly at the genetic (GWAS, selected sweeps) and epigenetic (TFBS) levels, also provided meaningful annotation. Therefore, we think these analyses can serve as indirect evidence for understanding the function of MHS. Notably, we specifically focused on the transgressive up-regulated MHS type, which shows significant colocalization with GWAS-associated loci related to photoperiod and metabolite levels. This finding aligns with key regulatory networks proposed in previous heterosis studies (Chen 2013) and provides supporting evidence for the contribution of transgressive MHS to heterosis.

We appreciate the reviewer's suggestion to incorporate additional omics layers such as 3D genome organization, WGBS, and histone modifications. We fully recognize that integrating such data generated from the exact same hybrid–parent combinations would ideally strengthen the mechanistic interpretation. However, we would like to highlight that the existing multi-omics datasets referenced in this study, though derived from different studies, still offer valuable and broadly applicable insights into the epigenetic regulatory principles governing chromatin accessibility, given the well-documented functional conservation. Meanwhile, generating a complete suite of new omics data for the specific trio combinations used in this study would represent a substantial investment of both time and resources, extending well beyond the intended scope and timeline of the current work. We therefore believe that the current analytical framework, centered on chromatin accessibility with supporting evidence from orthogonal public datasets, provides a coherent and well-grounded basis for our conclusions, while leaving room for more comprehensive multi-omics investigations in future studies.

### **3. Comment:**

In section “Functional consequence of open chromatin with different inheritance modes” and other sections, some important results and conclusions were cited from references. Since this is a section in Results part, the authors should indicate the results and conclusions supported by their own experiments.

### **3. Reply:**

We appreciate the reviewer’s suggestion. The primary objective of this part of the results, is to provide multi-level evidence for understanding the functions of different types of MHS, particularly at the genetic (GWAS, selected sweeps) and epigenetic (TFBS, CAWAS) levels.

To clarify, all of the conclusions in the “Functional consequence...” section are directly supported by analyses performed in this study. Specifically, by integrating the MHSs identified in this study with public ChIP-seq from >100 maize TFs, GWAS and CAWAS hits, and breeding/improvement region, we provide multiple interpretative perspectives for the MHSs dataset. We show that (1) inheritance modes differ in TF-binding preferences (additive MHSs have the fewest TFBSs; transgressive and conserved MHSs show distinct TF enrichments), (2) transgressive syntenic MHSs are preferentially enriched for developmental and metabolic GWAS traits (Fig 5A) and for CAWAS signals (Supplementary Figure 7D), and (3) transgressive and dominant sites are preferentially found in breeding/improvement regions whereas conserved sites are depleted (Supplementary Figure 7E).

In the revised manuscript, we have rewritten the Results text to (i) use direct, data-centric language (“we found”, “our analysis shows”) when reporting these findings and (ii) reserve prior citations for background/context rather than as evidence for these results. All results are linked to the corresponding figures and supplementary panels.

**Reviewer 3:****General Comments:**

This is an excellent paper that presents complex and detailed analyses of chromatin accessibility in hybrids. The work addresses an important knowledge gap. While prior studies have extensively examined gene expression, small RNAs, and some chromatin features in hybrids, chromatin accessibility in hybrid combinations has received comparatively less attention in maize. This study provides a robust and comprehensive foundation for understanding accessibility patterns in hybrids. However, I recommend that the title and conclusions be moderated to better align with what the data can definitively support.

The analyses are thorough and provide compelling evidence that chromatin accessibility at some loci exhibits non-additive and transgressive patterns in hybrids. The depth and rigor of the analyses are good, though I note that the density of some sections may require multiple readings to fully digest, which could present challenges for a broader audience. Nonetheless, the results are valuable and contribute meaningfully to the field.

**Reply:**

Thank you very much for your positive and constructive comments on our manuscript. We appreciate your recognition of the study's contribution to understanding chromatin accessibility in maize hybrids. We have revised the title and toned down causal language throughout the conclusions and Discussion. Rather than claiming that chromatin accessibility changes "contribute to" heterosis, we now frame our findings as revealing regulatory patterns "associated with" heterosis and providing a mechanistic framework for future functional dissection. We address the specific concerns point by point below.

**Major:****1. Comment:**

My primary reservation is that the current data, while correlative and intriguing, do not conclusively prove that chromatin accessibility changes contribute to or are required for heterosis. Establishing causality would require more extensive functional validation beyond the correlative patterns presented. Specifically, I would recommend: (1) analysis across multiple hybrid combinations, including reciprocal crosses, with increased biological replication (the current study uses a single non-reciprocal(?) hybrid with only 2 replicates); and (2) more detailed functional validation with appropriate controls. The preliminary functional work on the *sm2* MHS locus is a promising start but requires additional detail and controls to be fully convincing.

It should also be noted that a recent elegant study analyzing chromatin accessibility and transcription factor occupancy in hybrids using MOA-seq was published earlier this year (<https://doi.org/10.1038/s41588-025-02246-7>). This reduces the novelty of the current work,

although the two studies can be viewed as complementary. I recommend that the discussion explicitly address the similarities and differences between these two approaches.

Despite these concerns, this remains a significant contribution that warrants consideration for publication, provided the claims in the title and conclusions are appropriately tempered to reflect what the data can support.

## 1. Reply:

We thank the reviewer for this thoughtful and constructive comment. We fully agree that causal claims require careful qualification and ideally direct perturbation; at the same time, we respectfully note that conclusive functional validation often requires substantial further experimental work (additional crosses, longer growth/transform timelines) that may not be feasible within the standard revision window. To address the reviewer's concerns, we have taken a multi-pronged approach using the existing experimental data and the literature. Specifically, we have addressed these concerns as follows:

**Addressing causality and scope:** We acknowledge the limitation of using a single hybrid combination (B73×Mo17). While adding reciprocal crosses or additional hybrids is beyond the time scope of this revision, we have explicitly discussed this as a limitation and shifted our conclusions to emphasize that these patterns represent a comprehensive blueprint for future functional studies.

**Enhanced functional validation:** We have added more detail and internal controls to the *sm2*-MHS validation (see revised Supplementary Figure 8CD and Methods) to strengthen the evidence for its regulatory role, which including more phenotype data (node number, internode length), more controls (LH244, LH244-OE, T13), these new data confirmed that *sm2* indeed exhibited height heterosis in F1 hybrid background, and demonstrate that *sm2* enhances plant height by coordinately modulating both internode length and node number. We also included more detailed description about experiments (sample/replicate numbers, trait details and expression level of OE lines), and confirmed the positive dose-dependent relationship between *sm2* transcript abundance and plant height.

**Comparison with MOA-seq study:** This landmark paper provides a high-resolution pan-cistrome of the maize leaf and demonstrates how genetic variation at TF binding sites explains phenotypic heritability (Engelhorn et al. 2025). We have added a dedicated section in the Discussion to discuss our findings in the context of this significant work. We highlight that while both explore hybrid regulation, our work focuses on inheritance modes and links to heterosis, providing a complementary perspective on the maize regulatory genome:

“A recent study utilized MOA-seq to construct a high-resolution pan-cistrome across ~25 F1 hybrids and used footprint-level allele bias to map binding QTLs and link TF-occupancy variation to trait heritability (Engelhorn et al. 2025), which provides an auspicious approach to make functional *cis*-variation accessible at scale. Our study started with a trio-combination approach, enabling us to capture the dynamic patterns of chromatin accessibility during maize hybridization and their potential association with heterosis. These two approaches together offer a comprehensive multi-scale framework for deciphering the molecular engines of complex trait in maize”

## Minor:

### 1. Comment:

Line 119, please specify how the correlation was calculated, including eg Spearman or Pearson and which reads – eg all reads binned or reads in peaks?

### 1. Reply:

To evaluate the reproducibility of MH-seq data, we divided the maize genome into 20-kb non-overlapping windows. Read counts were calculated for each bin across all biological replicates. The Pearson correlation coefficient was then computed based on these genomic bins to assess global reproducibility (Supplementary Figure 1).

### 2. Comment:

Line 120 it is mentioned that there are around 60k MHS sites in each genome, does this only include peaks after IDR filtering? Please mention how many peaks there were before and after IDR; what was the proportion that were reproducible?

### 2. Reply:

Yes, the 60k only includes peaks after IDR filtering (Lines 116-117). The reproducibility is shown as follows (Supplementary Table 8):

|               | Rep1   | Rep2  | IDR filtered (proportion %) |
|---------------|--------|-------|-----------------------------|
| B73           | 129464 | 79963 | 62699 (48.43-78.42%)        |
| Mo17          | 77123  | 87473 | 62984 (72.00-81.67%)        |
| BM (B73 ref)  | 90285  | 83404 | 60316 (66.81-72.32%)        |
| BM (Mo17 ref) | 87969  | 84008 | 60588 (68.87-72.12%)        |

### 3. Comment:

Line 126 and Line 693, it is stated that “we extracted only the best uniquely aligned reads from the bam files”, considering there is often confusion in the community about what a uniquely aligned read is (or the infamous “multimapping” read), please specify the bioinformatic approach to identify the “best uniquely aligned read” and “multiple aligned reads”, including the specific approach to filtering the bam, software and arguments.

Considering this is central to the main findings of the paper, this is essential to explain in sufficient detail and code.

### 3. Reply:

Thank you for your suggestions. We provided detailed workflow and code for read filtering (see Methods, Supplementary Figure 9B).

We mapped the F1 reads against the B73, Mo17 and the B73-Mo17 synthetic genome using bowtie2 with parameters "--very-sensitive", respectively:

```
bowtie2 -p 72 --very-sensitive -x ${hybridref} -1 ${wkdir}/${i}-${rep}_1P.fq.gz -2 ${wkdir}/${i}-${rep}_2P.fq.gz |samtools sort -@ 36 -T ${i}-${rep}.hybrid -o ./${outdir}/${i}-${rep}.hybrid.bam
```

We retained only the best uniquely aligned reads ( $\text{MAPQ} \geq 2$ ) from the BAM files mapped to the synthetic genome, which were designated as "BM unique reads". We extracted only multiple aligned reads ( $\text{MAPQ} = 1$ ) from BAM files of B73 and Mo17 genome, respectively, marked as "B73 multiple reads" and "Mo17 multiple reads":

```
# extract BM "unique reads"
samtools view -@ 72 -h -f3 -q 2 BM-1.hybrid.bam | samtools view -@ 72 -b > BM-1.hybrid.q2.bam
samtools view BM-1.hybrid.q2.bam | cut -f1 | LC_ALL=C sort -S 80% | uniq > BM-1.hybrid.q2.bam.ID

# extract BM "multiple reads"
samtools view -@ 80 BM-1.B73.MD.bam |awk '{if($5==1){print}}'|cut -f1|uniq|LC_ALL=C sort -S 80%|uniq > BM-1.B73.multi
samtools view -@ 80 BM-1.Mo17.MD.bam |awk '{if($5==1){print}}'|cut -f1|uniq|LC_ALL=C sort -S 80%|uniq > BM-1.Mo17.multi
```

In Bowtie2 outputs, multimapping reads consistently receive MAPQ scores below 2 (<https://biofinysics.blogspot.com/2014/05/how-does-bowtie2-assign-mapq-scores.html>).

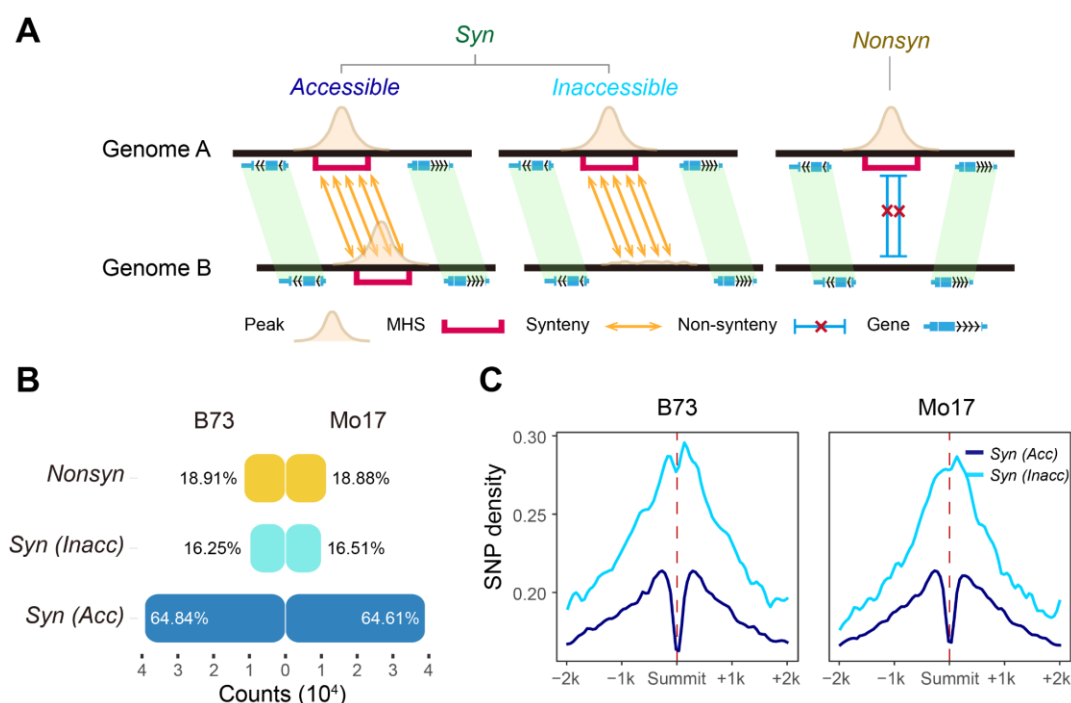
### 4. Comment:

Line 129, Figure 1A is a bit confusing, the light blue box is too big for 1% and its not clear why there is a row each for B73 and Mo17 are the percentages the same for each (seems like on B73 is labelled?). I do really like Supp Figure 1B, I would suggest making this the main panel figure to replace Fig 1A and adding the percentages directly to Supp Figure 1b. In this section in the text, it would be helpful to explicitly say that differentially accessible regions that are also syntenic are only 1% and state how many regions (~600?).

#### 4. Reply:

We apologize for any confusion in these descriptions. What we actually wanted to express here is: based on sequence alignment, approximately 81% of MHSs (we termed as “Syn”) exhibited syntenic relationships between the two genomes, whereas 19% were non-syntenic (we termed as “Nonsyn”). **Among the syntenic MHSs**, about 80% of them showed conserved chromatin accessibility between two genomes [we termed as “Syn (Acc)”], while the remaining 20% exhibited accessibility in only one genome [we termed as “Syn (Inacc)”]. Both “Syn (Acc)” and “Syn (Inacc)” belonged to the “Syn” type.

We now combined the previous Fig 1A and Supplementary Figure 1B:



**Response-Fig. 8:** Open chromatin landscape between B73 and Mo17. **A:** Schematic illustration of open chromatin comparison in a gene collinearity-based pan-genome analysis. **B:** Comparison of MHSs between B73 and Mo17. **C:** B73-Mo17 SNP density around the summits of different types of MHSs. “*Nonsyn*” means non-syntenic MHS; “*Syn*” means sequence-syntenic MHS; “*Accessible*” or “*Syn (Acc)*” means sequence-syntenic MHS that showed consistent accessible chromatin; “*Inaccessible*” or “*Syn (Inacc)*” means sequence-syntenic MHS that showed inconsistent accessible chromatin.

**5. Comment:**

Line 148 and 619 – were both replicates of the hybrids B x M, or was one the reciprocal M x B?

**5. Reply:**

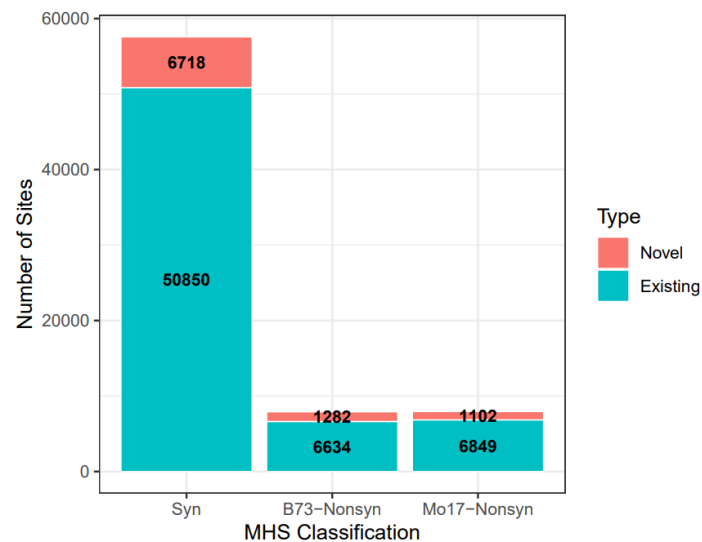
We apologize for the ambiguity in this sentence. We used 2 replicates of the hybrids B73×Mo17 here.

**6. Comment:**

Line 155 – it mentioned that a substantial number of novel open chromatin regions are detected in the hybrid are not detected in either parent, but it is unclear where this is shown in the figure because each panel only compare the hybrid on one parent (not both combined). Please mention in the text how many novel open chromatin regions are detected in the hybrid are not detected in either parent.

**6. Reply:**

We thank the reviewer for pointing out this lack of clarity. We agree that a direct comparison against a combined parental background is essential to define "novel" sites accurately. To address this, we have now performed a pangenome-scale comparison by merging the open chromatin repertoires of both B73 and Mo17. We now have explicitly stated in the revised text that a total of 9,102 novel MHSs were identified in the BM hybrid. Among these, 6,718 are located in syntenic regions, while 1,282 and 1,102 are situated in B73- and Mo17-non-syntenic regions, respectively (Response-Fig. 9). We have revised the Supplementary Figure 2B, and the corresponding description has been added to the main text (Lines 152-153).



**Response-Fig. 9:** Distribution of total and novel MHSs in the BM hybrid. Stacked bar chart showing the composition of total MHSs identified in the BM hybrid across three genomic contexts: *Syn* (Syntenic regions), *B73-Nonsyn* (B73-specific non-syntenic regions), and *Mo17-Nonsyn* (Mo17-specific non-syntenic regions).

#### 7. Comment:

Line 168 – please state the total number of MHS in this part of the analysis.

#### 7. Reply:

A total of 86965 MHSs were used in the analysis of the inheritance mode.

#### 8. Comment:

Line 188 The authors also identify some interesting difference between ACRs and gene expression in hybrids for example noting accessibility has more conservative heritability than transcription, but it should be discussed if this is in part due to technical differences between MH-seq and RNA-seq sensitivity and dynamic range.

#### 8. Reply:

We thank the reviewer for this important point. We agree that intrinsic differences between MH-seq and RNA-seq (molecular readout and dynamic range) could partially contribute to the observed difference in apparent inheritance patterns. We have added a paragraph to the Discussion explicitly clarifying that, despite this potential partial effect, multiple independent analyses in our study (inheritance-mode classification, cis/trans decomposition, and cis-

regulatory enrichment analyses) support the biological interpretation that accessibility tends to be a more stable regulatory substrate than transcript abundance:

“We observe that ACRs generally exhibit more conservative inheritance patterns than gene expression, this may be partially interpreted in light of technical feature differences between MH-seq and RNA-seq. MH-seq assays probe chromatin accessibility at the DNA level and report local enrichment with a comparatively constrained dynamic range, whereas RNA-seq measures steady-state transcript abundance, which integrates transcriptional output, RNA processing and transcript stability and typically exhibits a much wider dynamic range and higher dispersion across biological conditions. These distinctions imply that RNA-seq is intrinsically more sensitive to small perturbations and to post-transcriptional modulation, which may inflate apparent variability. Nevertheless, multiple orthogonal analyses in our study (inheritance mode classification, cis/trans enrichment) support a biological basis for the relative stability of chromatin, so assay-level differences are likely to explain only part of the observed disparity.”

## **9. Comment:**

Line 209 and Figure 2D – it is unclear why this is a box plot and not a single value for the fraction? Please explain in more detail why this is the case and why multiple comparisons of the fraction values were conducted by the least significant difference (LSD) method.

## **9. Reply:**

We calculated the overlapped fraction with conserved non-coding sequences (CNSs) set for each MHS, so it is a boxplot and not a single value, where the sample size is the total number of MHSs in each category.

Multiple comparisons were performed using Fisher's Least Significant Difference (LSD) test within each predefined biological group (e.g., inheritance category), rather than across all samples globally. For each group, we first fitted a one-way ANOVA to test whether fraction values differed among comparison classes. Only after a significant ANOVA, we conducted pairwise comparisons using the LSD framework, which leverages the pooled variance from the ANOVA and provides higher power for a limited number of planned contrasts.

Importantly, to control for multiple testing within each group, we applied Benjamini-Hochberg (BH) false discovery rate correction to the pairwise p-values. This testing strategy balances statistical power and error control, and is appropriate for comparing fraction distributions derived from biological replicates and resampling. We have clarified this procedure in the Methods section.

## **10. Comment:**

Line 210 – Text should be Figure 2D (not 2C). The figure call outs for this section appear to be out for E and F as well.

## 10. Reply:

Fixed.

## 11. Comment:

Line 237 – this section compares chromatin accessibility in the hybrid to chromatin interaction data but presumably the chromatin interaction data is all from B73 (?). If so, the significance of comparing hybrid MHS to B73 interaction data should be made clear at the beginning of the section. The source of the data should also be mentioned in the text and/or Figure legend.

## 11. Reply:

We apologize for the confusion caused by our previous descriptions. The conclusion was also drawn from the analysis of BM hybrid chromatin interaction data. We have collected H3K4me3-HiChIP and H3K27ac-HiChIP data from BM F1 hybrid seedling leaf tissues and conducted the analysis (H. Liu et al. 2025). The results indicate that in the F1 genome, transgressive MHS are indeed more likely to involve 3D interactions (Response-Fig. 4), which is consistent with our findings in B73.

## 12. Comment:

Line 277 – the authors speculate about transgressive syntenic super enhancer MHS hubs. It is an evocative idea, but to support this speculation, it would be good to mention how many MHS loci fit this description. Is it more than just this example? What fraction?

## 12. Reply:

Thank you for your suggestions. We analyzed genes that interacting with each MHS, we found 99 of 332 (~29.8%) transgressive up-regulated MHSs interacted with non-additive genes, while 15158 of 64115 (~23.6%) other types of MHSs interacted with non-additive genes, indicating that transgressive MHSs are more likely to interact with non-additive genes than other types ( $P$ -value = 0.009985, Chi-square test).

## 13. Comment:

Line 361 – it is mentioned there is an enrichment for developmental traits, but which of the traits tested were not developmental, seems like most are developmental; if so, can it really be said that there is an enrichment?

### 13. Reply:

We apologize for any confusion in the description. These traits were indeed all developmental traits. We now rephrased the wording to avoid misleading:

“Transgressive syntenic MHSs, in particular, were specifically enriched for some of the agronomic traits, including photoperiodic traits, internode length, cob diameter, ear height, and northern leaf blight (Figure 5A).”

### 14. Comment:

Line 398 – please detail the re-analysis of HiChIP data. Why was reanalysis required and how was it performed differently?

### 14. Reply:

Thanks for your comments. Since the published HiChIP results were based on the earlier B73 AGPv4 reference genome, we thus need to liftover these B73v4-based HiChIP interaction intervals to B73 version 5 to fit our workflow and results. The lifts were conducted using CrossMap (<https://crossmap.sourceforge.net/>) from chain files ([https://download.maizegdb.org/Zm-B73-REFERENCE-NAM-5.0/chain\\_files/](https://download.maizegdb.org/Zm-B73-REFERENCE-NAM-5.0/chain_files/)).

### 15. Comment:

Line 401 – was HiChIP performed in the hybrid?

### 15. Reply:

Yes, we also re-analyzed the HiChIP data from the BM F1 hybrid (H. Liu et al. 2025).

### 16. Comment:

Line 397 – it is difficult to understand which is the transgressive up regulate locus in this example (Fig 5C). B73 has a value of 23, BM is  $17 + 12 = 29$  and Mo is 34? It would be useful to explain the calculation using an example, to make this clear.

### 16. Reply:

We apologize for any confusion in the description. This MHS was classified as B73 non-syntenic (*Nonsyn*) MHS, because the sequence structure variation between the B73 and

Mo17. For the B73 *Nonsyn* MHS, we only focused on the abundance of the B73 haplotype in the parent and the hybrid, and treated the Mo17 haplotypes as null alleles (absence). So in this instance, the B73 parent has an RPKM (Reads Per Kilobase of transcript per Million mapped read) value of 23.74, the mid-parent value is  $(23.74+0)/2=11.87$ , the B73 allele in the BM hybrid has a RPKM value of 17.82 while Mo17 allele contribute 0, we can get:  $RPKM_{MPV} = 11.87 \ll RPKM_{BM} = 17.82$ . Combined with the adjusted *P*-value and log2 of fold change calculated by DESeq2, this excess is statistically significant, and we therefore classify this locus as transgressive up-regulated.

#### 17. Comment:

Line 410 and Fig 5E – was the height of the inbreds (LH244, LH244-OE and T13) also collected? It would be useful to include the inbreds to demonstrate hybrid vigor was observed and the effect of OE in inbred background. Please also include the number of plants/plots tested and trial details. For multiple of the tests (eg years and locations) some of the OE lines were not significantly taller than the control hybrid, this should be discussed. A reciprocal cross would also be idea, but at least including all the control is necessary.

#### 17. Reply:

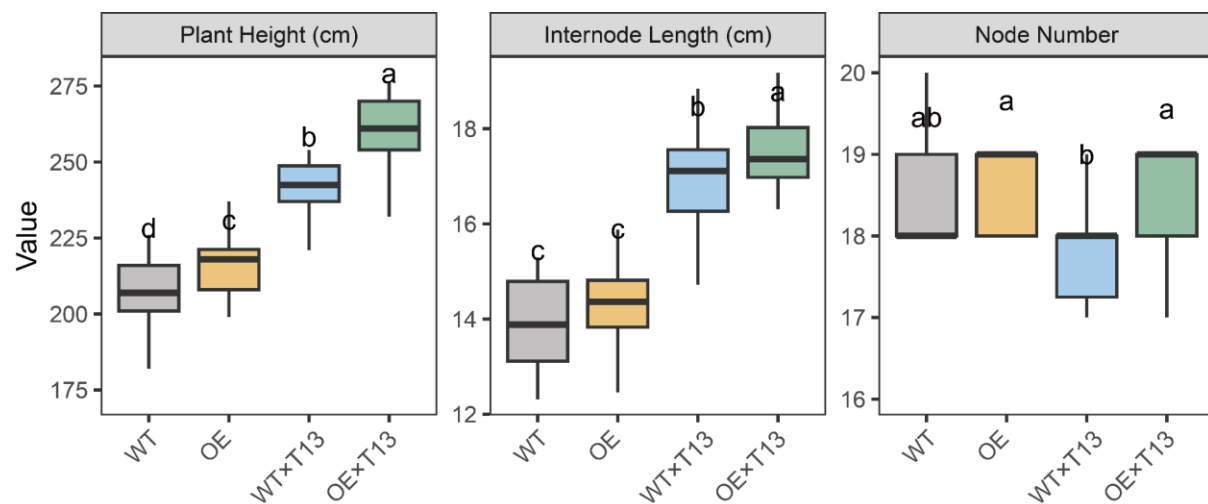
We agree with the reviewer that including the inbred parents provides essential context for demonstrating hybrid vigor and the specific effect of the overexpression (OE) construct in different genetic backgrounds. We have updated our analysis to include these controls as shown in the provided figure (Response-Fig. 10).

The data confirms significant hybrid vigor for plant height and internode length. For the effect of OE in inbred and hybrid backgrounds, the overexpression of *sm2* in the inbred background (OE) results in a statistically significant increase in plant height compared to the wild-type (WT). Interestingly, the effect of OE is even more pronounced in the hybrid background. The OE hybrid is significantly taller and shows more node number and higher internode length than the control hybrid. This indicates that the OE construct not only improves the performance of the inbred line but also further enhances the existing heterosis in the F1 generation.

While the OE F1 line consistently outperformed the control hybrid across most trials, we acknowledge that in specific years or locations, the height difference was not statistically significant. This observation could be attributed to several factors including environmental sensitivity and genotype-by-environment interactions, we discussed it in the main text:

“We found the *sm2*-OE hybrid outperformed the control hybrid across majority of multi-year, multi-location trials, demonstrating the robustness and reproducibility of the *sm2*-OE effect, despite of that the advantage did not reach statistical significance in certain specific trails. This is not unexpected, as heterosis is a quantitative, polygenic trait, and the contribution of a single gene such as *sm2* is likely subject to modulation by genotype-by-environment interactions. The overall pattern across trials nonetheless supports the functional validity of *sm2* in promoting plant height and enhancing hybrid performance, and future evaluations

should further characterize the *sm2*-mediated effects across diverse environmental conditions.”



**Response-Fig. 10:** Phenotypic analysis for plant height, internode length and node number across the wild-type LH244 (WT), *sm2*-overexpression (OE) line, and their respective F1 hybrids (WT×T13 and OE×T13) at Hainan in 2026 ( $n = 25, 20, 22$  and  $29$  for WT, OE, WT×T13 and OE×T13, respectively). Multiple comparisons were conducted by the least significant difference (LSD) method.

## 18. Comment:

Line 445 – it is unclear what is mean by “poised” state?

## 18. Reply:

Poised CRMs occupy an intermediate state between repression and activation and are therefore primed to become either fully active or fully inactive (Rada-Iglesias et al. 2011; Koenecke et al. 2017). These elements are chromatin-accessible, bound by few TFs, and show enrichment for H3K27me3 alongside low levels of activating histone modifications (e.g., HAc) (Lu et al. 2019; Ricci et al. 2019). Our identified *Syn* (*Inacc*) type MHSs were consistent with these special features.

We added a one-sentence explanation here (Lines 459-463):

“Interestingly, regions with identical sequences but inconsistent chromatin states between the parents were mostly in a poised state of open chromatin which occupy an intermediate state between repression and activation, and show enrichment for H3K27me3 alongside low levels of activating histone modifications (Rada-Iglesias et al. 2011; Koenecke et al. 2017; Lu et al. 2019; Ricci et al. 2019).”

#### 19. Comment:

Line 471 – it is unclear why the more stable and ‘robust’ feature would be more likely to provide an explanation to heterosis. It seems more likely that the feature that is more transgressive/variable in hybrids would have more potential to offer an explanation? This assertion should need more reasoning.

#### 19. Reply:

We apologize for the confusion here. We have clarified the sentence to state that by “more stable and robust” we mean open chromatin provides a more reliable substrate for dissecting regulatory mechanisms (because its accessibility is recurrent and easier to interpret across backgrounds); we also explicitly note that variable/transgressive chromatin events and transcriptomic changes can provide complementary, high-impact contributions to heterosis:

“These results suggest that open chromatin, being relatively more stable and robust than more dynamic genomic layers such as the transcriptome, provides a better substrate for dissecting the regulatory mechanisms underlying heterosis.”

#### 20. Comment:

Line 508 – it is stated that chromatin accessibility changes in hybrids have lower magnitude change of expression. This is significant and should be explained with more detail in both results and discussion, including some numbers to define the magnitude (or refer to a relevant Figure if it is already created).

#### 20. Reply:

We have discussed this discrepancy in detail (included the relevant Supplementary Figure 6A-B), the primary reasons include that chromatin accessibility represents only one source of *cis*-regulatory acting, while other various regulatory processes, including *trans*-acting and other *cis*-acting factors, may also affect gene expression; Accurately matching *cis*-regulatory elements with target genes remains a challenge, as we only considered proximal regulatory elements in our analysis, potentially overlooking distal regulatory elements. Even in the proximal region, it is common for multiple MHSs to correspond to a single gene. Determining which MHSs are more important for gene expression is challenging, including MHSs unrelated to gene expression could lower the overall association level.

#### **Reviewer 4:**

##### **General Comments:**

The study presented by Huang et al. addresses the question if chromatin accessibility changes in maize hybrids can be linked them to heterosis. Overall, the study provides a bioinformatic pipeline to assess differential chromatin accessibility, measured by MNase Hypersensitivity sequencing. The method was employed to compare accessible chromatin (ACRs) and transcriptomes from two maize accessions, Mo17 and B73 and their hybrids using 10-day old seedlings or leaves (the information is not quite consistent). Interesting insights include the observation that chromatin accessibility is less variant than transcript levels and that 20% of ACR regions in syntenic regions show variation that is most likely explained by sequence divergence. A further focus of the manuscript were non-additive/transgressive ACRs. Transgressive upregulated ACRs were found to be more conserved in sequence and linked to potential super-enhancers based on published Hi-C data. They were linked to genes upregulated in hybrids. Overexpression of a candidate gene showing linkage to an upregulated ACR led to increase plant height, which the authors interpret as potential contribution to heterosis.

The dataset is rich, and the analysis of cis-/trans-regulatory patterns and transgressive chromatin accessibility is conceptually strong and relevant. However, the manuscript is not written for a more general audience and a could be improved by more explicit introduction of the analyses performed. Furthermore, the manuscript contains numerous errors in figure citations, labeling, and legends. The frequency of such mistakes is distracting and make reviewing unnecessarily time-consuming, distracting from the otherwise solid scientific content.

##### **Reply:**

We thank the reviewer for the careful reading and for the constructive comments. We apologize for the confusion and for the errors in figure citations, labeling and legends, and we appreciate the time the reviewer spent identifying these issues. We have thoroughly revised the manuscript to improve clarity, correct the mistakes, and make the analyses and experimental design clearer.

##### **Major:**

###### **1. Comment:**

Figure 1A. It is not immediately clear how the numbers add up in the graph. 81% of the ACRs were syntenic, 19 non-syntenic. The block within classifies 80% of syntenic as non-differential, but at first glance the figure seems to indicate that only 1% of the syntenic sites were non-consistent open chromatin.

###### **1. Reply:**

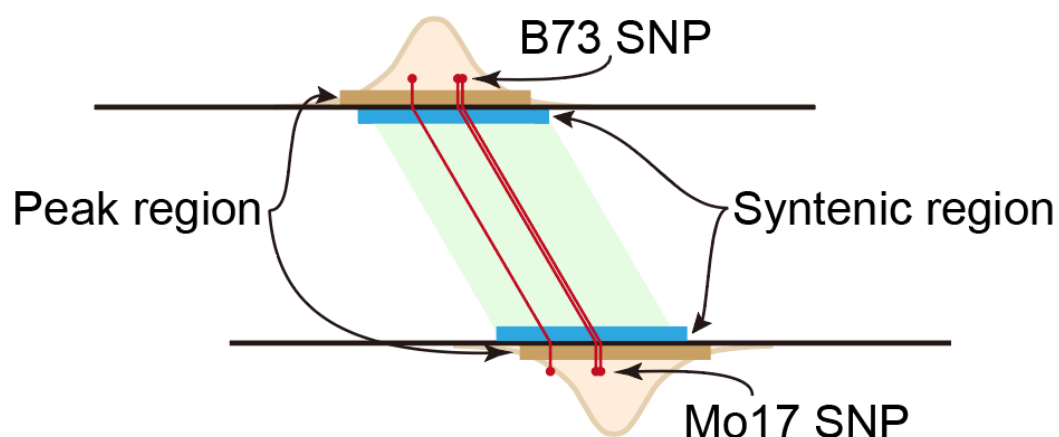
We apologize for any confusion in these descriptions. What we actually wanted to express here is: Using sequence alignments between the two genomes, we classified MHSs into syntenic and non-syntenic groups: approximately 81% were syntenic (denoted “*Syn*”) and 19% non-syntenic (“*Nonsyn*”). We further subdivided “*Syn*” type according to conservation of chromatin accessibility: about 80% of *Syn* elements retained accessibility in both genomes (denoted “*Syn (Acc)*”), while ~20% were accessible in only one genome (denoted “*Syn (Inacc)*”). Note that both “*Syn (Acc)*” and “*Syn (Inacc)*” are subsets of the *Syn* category. The analyses and plots previously shown in Fig. 1A and Supplementary Fig. S1B have been merged into the updated Fig. 1A.

## 2. Comment:

Figure 1B. It is not quite clear why the curves for SNP density for “syntenic accessible” are not the same for both genotypes – should they not be identical? The figure makes a point about differences between variant and non-variant ACRs and this is quite clear, but the different baseline is confusing.

## 2. Reply:

Thanks for your comments. It is nearly impossible that the curves for both genotypes are fully identical. We present a schematic illustration here (Response-Fig. 11). Although the syntenic region between the two genomes is perfectly matched, the actual peak region in B73 and Mo17 may have a discrepancy, which made the SNP location in the B73 peak and the Mo17 peak inconsistent. Moreover, the small InDels inside the syntenic region can also influence the SNP locations.



**Response-Fig. 11:** Schematic illustration of peak region and SNP location discrepancies between genotypes. Although the underlying syntenic regions (blue/green block) are precisely matched between the two genomes, the actual enriched chromatin “peak regions” (beige blocks) may shift or vary in width. Such discrepancies in peak boundaries result in inconsistent relative locations for associated SNPs (B73 SNP vs. Mo17 SNP).

### 3. Comment:

In Figure 2A, chromatin accessibility and gene expression inheritance are classified into four modes: Conserved, Additive, Dominant, and Transgressive. First, transgressive should be defined as transgressive up and transgressive down (beyond and below parental values, respectively). Additivity should better be Additive to align grammatically with dominant, transgressive etc.

### 3. Reply:

Thanks for your suggestions. Yes, we first classified the inheritance modes into four major types: Conserved, Additive, Dominant, and Transgressive, then we in-depth divided them into more detailed types as follows, which included transgressive up and transgressive down. We agreed that “Additivity” should be “Additive” and fixed it.

### 5. Comment:

According to the description, Additive means that hybrid (BM) chromatin accessibility equals the mid-parent value (lines 165-166),  $BM = (B73 + Mo17)/2$ , falls between both parents ( $B73 > BM > Mo17$  or  $B73 < BM < Mo17$ ). Transgressive is defined as  $BM > B73$  &  $BM > Mo17$  or  $BM < B73$  &  $BM < Mo17$ . Could the authors clarify how cases are categorized when  $B73 > BM > Mo17$  or  $B73 < BM < Mo17$  but  $BM \neq (B73 + Mo17)/2$ ? For example, in Figure 5C, sm2-MHS is classified as a “transgressive up-regulated” MHS (line 397). However,  $BM (17.82 + 12.05 = 29.87)$  does not exceed Mo17 (34.50).  $Mo17 > BM > B73$ . Actually, the mean of B73 and Mo17 (29.12) is similar to BM. Could you explain why this case was classified as transgressive up rather than additive?

### 5. Reply:

We appreciate the reviewer’s careful inspection of Figure 5C. We apologize for the lack of clarity in our initial description regarding the classification of *Nonsyn* (non-syntenic) MHS. The classification logic differs between *Syn* (syntenic) and *Nonsyn* MHS due to the presence or absence of the underlying genomic sequence.

For *Syn*-MHS: We used the standard formula for Mid-Parent Value (MPV) is:  $MPV = (P1 + P2)/2$ . If the BM value is significantly higher than the higher parent or lower than the lower parent, it is classified as transgressive.

For *Nonsyn*-MHS (like sm2-MHS): These MHS sequences are presence/absence variations (PAVs) or highly divergent between two parents. In the case of a B73-specific *Nonsyn*-MHS, the Mo17 allele (absent parental allele) is treated as a null allele (value = 0) when computing

the mid-parent value, because the corresponding regulatory element does not exist in the Mo17 genome. This convention follows published studies of PAV/expression complementation in plants, and reflects the biological absence of the sequence in one parent (Paschold et al. 2012; Li et al. 2021). Therefore, the MPV is calculated as:  $MPV = (B73 + 0)/2$ . In the example of sm2-MHS:  $B73 = 23.74$ ,  $Mo17 = 0$ , thus  $MPV = 11.87$ . The observed value in the BM hybrid:  $BM_{B73\_haplotype} = 17.82$ . Since  $BM_{B73\_haplotype}$  (17.82) is significantly greater than the MPV (11.87) (confirmed by adjusted *P*-value and log2 of fold change that was calculated by DESeq2), we categorized it as transgressive up-regulation.

We have now updated the Methods and Results sections (Lines 406-407, 805-813) to explicitly define this calculation for non-syntenic MHSs to avoid further confusion.

## 6. Comment:

Similarly, for Zm00001eb290700 (line 277) and Zm00001eb080410 (line 417) gene expression, which are described as down- and up-regulated respectively:

- Zm00001eb290700 in BM ( $2.93 + 1.69 = 4.69$ ) is lower than B73 (13.39) but higher than Mo17 (1.76) ( $B73 > BM > Mo17$ ;  $BM \neq (B73 + Mo17)/2$ ) (Figure 3E).
- Zm00001eb080410 in BM ( $18.18 + 1.50 = 19.68$ ) is greater than B73 (0.60) but similar to Mo17 (21.23) ( $Mo17 > BM > B73$ ;  $BM \neq (B73 + Mo17)/2$ ) (Supplementary Figure 8C). Please clarify how up- and down-regulation were defined for these cases.

## 6. Reply:

We appreciate the reviewer's request for clarification on our terminology. In the context of these two cases, "up-regulated" and "down-regulated" in hybrids are defined relative to the Mid-Parent Value (MPV), representing a significant deviation from additive inheritance (non-additivity).

While the hybrid value ( $BM$ ) may fall between the two parental values ( $P1$  and  $P2$ ), it can still be statistically higher or lower than the expected MPV  $[(P1 + P2)/2]$ . We used DESeq2 to test the significance of this difference ( $BM$  vs.  $MPV$ ). Specifically:

For Zm00001eb290700 (Figure 3E):

$B73 = 13.39$ ,  $Mo17 = 1.76$ ,  $MPV = (13.39 + 1.76)/2 = 7.575$ . The hybrid  $BM = 4.69$ , which is significantly lower than  $MPV$ . Thus, we classified it as "non-additive down-regulated" relative to the additive expectation.

For Zm00001eb080410 (Supp. Figure 8C):

$B73 = 0.60$  ,  $Mo17 = 21.23$  ,  $MPV = (0.60 + 21.23)/2 = 10.915$  . The hybrid  $BM = 19.68$  , which is significantly higher than the  $MPV$  . Thus, we classified it as "non-additive up-regulated" relative to the additive expectation.

To avoid confusion, we have clarified this definition in the revised manuscript (Lines 805-813) and ensured that "up/down-regulated" in the hybrid context is consistently referred to as "relative to the MPV" or "non-additive" throughout the text.

## 7. Comment:

Line 168: Fix grammar

## 7. Reply:

Fixed.

## 8. Comment:

Line 237 and following: The comparisons were based on published chromatin data that must certainly show a bias towards the maize variety that was used to generate them. This should be assessed and discussed. For example, the bias for H3K27me3 regions towards Mo17 dominant lower MNH sites could be explained by H3K27me3 data being generated in Mo17.

## 8. Reply:

We thank the reviewer for raising this point. We agree that the published epigenomic datasets used in our analysis were mostly generated from B73 genotype, and therefore genotype-specific bias cannot be fully excluded. In particular, the enrichment of H3K27me3-associated interactions in Mo17-dominant-lower MHSs could partially reflect the genotype origin of the available datasets. Unfortunately, there are no currently publicly available H3K27me3-related data for the Mo17 genotype. Nevertheless, analysis of recently published HiChIP data from the B73–Mo17 F1 hybrid supports our main conclusion that transgressive MHSs are more likely to participate in 3D chromatin interactions (Response-Fig. 4), consistent with the pattern observed in B73. We have now explicitly acknowledged this limitation in the revised manuscript (lines 538–541).

“Although this pattern should be interpreted with caution given that some available interaction datasets were generated from a single genotype, our conclusion about transgressive MHSs still holds given the consistent observation in F1 hybrids (Supplementary Figure 4C).”

## 9. Comment:

Figure 3E and 5B are not intuitive and need a better legend (at least the first time). It is not intuitive to plot one accession above, the other below the x-axis, unless they are plotted on the same x-axis. The number labels are so small that these cannot be recognized without zooming.

## 9. Reply:

We sincerely apologize for the visual complexity and the legibility issues in Figures 3E and 5B.

The decision to plot B73 tracks upward and Mo17 tracks downward was made to facilitate the direct comparison of syntenic regions across two different reference genomes. As these two inbred lines have distinct genomic coordinates, placing them on a single horizontal axis would obscure the structural variations and syntenic relationships. By using a mirrored layout centered around the "Synteny" ribbons (light blue), we aim to show how the chromatin environment of each parental allele is maintained or altered in the hybrid (BM) relative to its respective progenitor.

We have improved the legends to provide a clearer "roadmap" for the reader. The revised legend explicitly defines the orientation (top vs. bottom), the specific tracks (histone marks, MH-seq, and allele-specific RNA-seq), and the significance of the central synteny and looping arcs. We have increased the font size of all numerical labels and axis titles to ensure they are clearly legible.

## 10. Comment:

Novel open chromatin sites in hybrids (lines 154–155). The statement "BM harbored a substantial number of novel open chromatin sites not detected in either parent" (lines 154–155), the current comparison (BM vs. B73 and BM vs. Mo17 separately in Supplementary Figure 2CD) is indirect. It would be more convincing if BM were compared to an integrated B73–Mo17 pangenome rather than to B73 and Mo17 separately (as in Supplementary Figure 2C–D).

## 10. Reply:

We sincerely thank the reviewer for this constructive suggestion. We agree that comparing the hybrid (BM) directly to an integrated parental repertoire is a more rigorous approach to identifying truly novel open chromatin sites. We now make the comparison at an B73-Mo17 pangenome scale. Compared to both parents, we identified a total of 9102 novel MHSs in BM hybrid, 6718 MHSs are located in syntenic genomic region (*Syn* MHSs), while 1282 and 1102 MHSs are located in B73 and Mo17 non-syntenic region (B73-Nonsyn and Mo17-Nonsyn MHSs), respectively (Response-Fig. 9). We have updated Supplementary Figure 2B

to include this direct comparison and revised the corresponding text in the main text (Lines 152-153).

## 11. Comment:

299ff The analysis of distance decline of cis-variation is not clear from the description.

## 11. Reply:

We appreciate the reviewer's feedback regarding the clarity of our distance-decay analysis. We have now expanded the description in the Results (Lines 301-311) and the Methods section to provide a step-by-step explanation of the statistical approach.

Specifically, we followed the strategy described by Wong et al. to infer the regulatory reach of cis-acting variants (Wong et al. 2017). We have revised the text and the Methods section to include the details and make it clear.

Main text:

To quantify the genomic extent of the effects of cis-acting variants, we analyzed the distance-dependent decay of shared cis-signals between MHS pairs, following an established strategy (see Methods) (Wong et al. 2017). We calculated the Spearman's correlation  $\rho$  of allelic proportions across increasing genomic intervals to specifically isolate cis-regulatory effects from trans-acting noise. Our results revealed a logarithmic decline in correlation as distance increased, whereas a random genomic null background showed no such correspondence (Figure 4B). Using the elbow method to identify the inflection point, we determined the effective genomic scope of cis-acting influence to be approximately 8 kb (Figure 4B). This scope is consistent with observations in mice, which has a genome size comparable to maize (Wong et al. 2017), suggesting that cis variation effects on chromatin accessibility are predominantly local to influence proximal regions, with weaker effects also detectable in the distal region.

Methods:

Estimation of the effective reach of cis-regulatory variation.

To infer the genomic span of cis-acting regulatory influence in maize, we adopted a spatial correlation strategy as previously described. The analytical framework was implemented as follows: (1) Pairwise correlation of allelic proportions. For a given genomic interval  $i$ , each cis-mediated MHS was defined as an "anchor." We then identified all MHSs located at distance  $i$  from these anchors to form MHS-anchor pairs. To specifically isolate cis-acting signals, we calculated the Spearman's correlation coefficient  $\rho$  of the allelic proportions (B73 / [B73 + Mo17]) between all paired MHSs and their respective anchors. (2) Distance decay modeling. The analysis was performed iteratively, with the interval size  $i$  starting at 1 kb and increasing by 1 kb increments. We utilized a linear regression model where log-transformed physical distances served as predictors and Spearman's  $\rho$  estimates served as outcome

variables to quantify the decay of the correlated cis-regulatory signals. (3) Null distribution and threshold determination. A genomic null distribution was generated by pairing anchor MHSs with an equivalent number of MHSs randomly sampled from across the genome. The effective genomic scope of cis-acting variants on chromatin accessibility was determined using the "elbow" method, which identifies the inflection point where the decay curve stabilizes and the correlation no longer significantly exceeds the null background.

## 12. Comment:

Line 356: What are broad TF types?

## 12. Reply:

We apologize for the imprecise terminology. We have revised the text to: "..., were significantly enriched in some specific TFBSs families, particularly those of types HB, HMG, WRKY, and ZIM....".

## 13. Comment:

In Figure 5C, the Mo17 Chr2 side lacks the gray bar and red dashed box used to indicate sm2-MHS and its abundance. Was this omitted intentionally (e.g., due to lack of synteny) or inadvertently?

## 13. Reply:

The omission of the gray bar and red dashed box on the Mo17 side in Figure 5C was intentional since sm2-MHS is classified as a *Nonsyn* (non-syntenic) type MHS. This means that the regulatory element is present in the B73 genome but is physically absent in the corresponding genomic region of Mo17. Consequently, we did not include the gray bar (which indicates syntenic correspondence) or the red dashed box (which highlights signal abundance) on the Mo17 side, as there is no biological feature to demarcate.

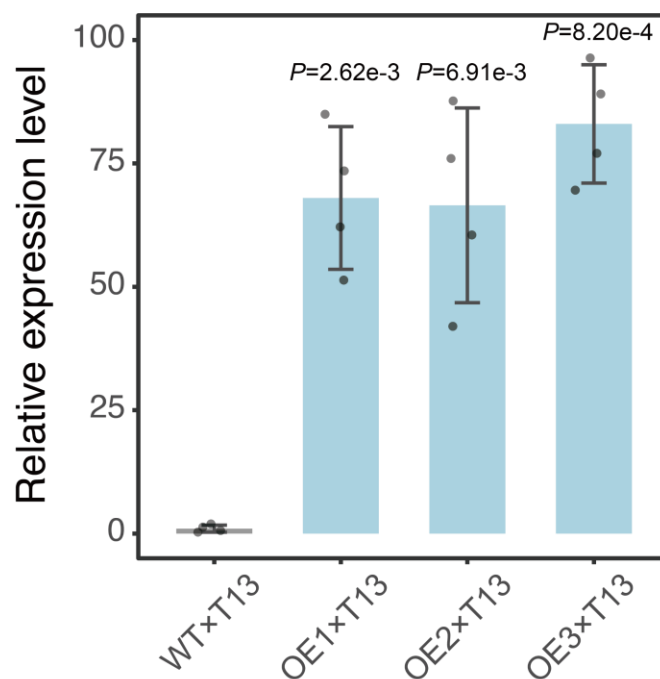
We have clarified this in the revised figure legend to ensure it is clear to the readers that the absence of these markers denotes a lack of synteny.

## 14. Comment:

sm2-overexpression lines (OE1–OE3). In Figure 5, the height of the sm2 overexpression (OE) line OE3 is more frequently and significantly higher than that of the WT control compared with OE1 or OE2. Is the sm2 expression level in OE3 correspondingly higher than in OE1 and OE2? Please provide the sm2 expression levels in all OE lines and the control to determine whether plant height correlates with sm2 expression.

#### 14. Reply:

We thank the reviewer for this insightful suggestion. We measured *sm2* expression levels in OE1–OE3 and WT by qRT-PCR. All three OE lines showed significantly elevated *sm2* transcript levels compared to WT, confirming successful overexpression across all lines. Among the OE lines, OE3 exhibited the highest *sm2* expression level, which is consistent with its more pronounced and frequent increase in plant height. This pattern may suggest a dose-dependent relationship between *sm2* transcript abundance and plant height, whereby the relatively higher expression in OE3 may have surpassed a threshold sufficient to produce a more robust phenotypic response compared to OE1 and OE2. However, we note that these measurements only reflect mRNA-level differences, and we cannot exclude the possibility that post-transcriptional or translational regulatory processes contribute to differential protein accumulation among the lines, which could also partially account for the variation in phenotypic expressivity observed. These data and discussion have been incorporated into the revised manuscript



**Response-Fig. 12:** Relative expression level of *sm2* in *sm2*-OE and WT lines ( $n = 4$  for all lines). Pairwise t-tests were performed between WT and each *sm2*-OE lines.

#### 15. Comment:

Figure citation corrections

Please revise the following for consistency with figure content:

- Lines 188–189: “Figure 2A” → “Figure 2B”.
- Lines 191–192: Remove “Figure 2B”.
- Lines 210–218: “Figure 2C” → “Figure 2D”; “Figure 2D” → “Figure 2E”.
- Line 227: “Figure 2E” → “Figure 2F”.
- Lines 303–305: “Figure 4B” → “Figure 4C”.
- Line 415: “Supplementary Figure 8B” → “Supplementary Figure 8C”.
- Regarding Supplementary Figure 3 (Lines 195-199), the suggested fix is:
- Switch the order of panels C and D in Supplementary Figure 3.
- Then, change the text citation in Line 199 from "Supplementary Figure 3C" to "Supplementary Figure 3D" accordingly.

#### **15. Reply:**

Fixed.

#### **16. Comment:**

Abbreviation inconsistency (lines 366–367 and Supplementary Figure 7D). “Chromatin accessibility genome-wide association” is abbreviated as CAWAS (line 366), but CaGWAS is used in Supplementary Figure 7D, and CWAS in the legend. If these refer to the same analysis, please standardize the abbreviation; if not, define them clearly.

#### **16. Reply:**

We apologize for the inconsistent abbreviations. We confirm that CAWAS, CaGWAS, and CWAS all refer to the same analysis. To ensure clarity and consistency, we have standardized the abbreviation to CAWAS throughout the revised manuscript.

#### **17. Comment:**

Line 265: The sentence is unclear. Did you bulk all inheritance modes and that overlapped with 5% of the superenhancers? Probably not but the sentence sounds as if.

#### **17. Reply:**

We apologize for the ambiguity in this sentence. The reviewer is correct; we did not mean that 5% of all super-enhancers (SEs) overlapped with other inheritance modes. Instead, we intended to state that within the pool of all other inheritance modes (excluding transgressive up-regulation), only 5% of the sites were associated with SEs. To clarify this, we have rephrased the sentence here:

“Analysis of inheritance modes revealed that SEs were significantly more prevalent among transgressive up-regulated sites (14%) compared to sites exhibiting other inheritance modes (5%) (Figure 3D,  $P$ -value =  $4.11\text{e-}08$ , Chi-square test).”

**18. Comment:**

You refer to Zm00001eb324900 as being “described in Figure 3E,” but it does not appear there. Please correct the figure citation to the accurate panel or explicitly label the gene in Figure 3E.

**18. Reply:**

We sincerely apologize for this oversight. You are correct that Zm00001eb324900 was incorrectly cited as being in Figure 3E. We have deleted this description.

**19. Comment:**

Figure legends and abbreviations. Nature Communications requires that abbreviations be defined at first use in figure legends. Please define terms such as CNS, SV, and GERP (Figure 2), and SS and NSS (Supplementary Figure 7). While these are standard in maize research, definitions will aid broader readership.

**19. Reply:**

We fully agree with your suggestion. We have updated all figure legends to ensure that every abbreviation is defined at its first occurrence.

**20. Comment:**

Supplemental 1E: colors are different in graph and legend

**20. Reply:**

Fixed.

**21. Comment:**

Supplemental 1F: the legend states that the alignment is per peak summit, not to each TSS as indicated in the x axes

**21. Reply:**

Fixed. It should be the peak summit here.

**22. Comment:**

In the legend of Figure 4C, the text states “0–50 kb (Top panel) and 0–400 kb (Bottom panel),” but the actual plots display 0–50 kb (bottom) and 0–500 kb (top). Please correct this discrepancy.

**22. Reply:**

Fixed. It should be 0-500 kb here.

**23. Comment:**

In the legend of Supplementary Figure 1H and I, terms Syn-Incon and Syn-Con should be clarified—are they equivalent to Syn (Acc) and Syn (Inacc) used previously? If so, ensure consistent terminology.

**23. Reply:**

Thank you. We have revised these terms to ensure consistency.

**24. Comment:**

Please review Supplementary Figure 7 and confirm that the legends for panels D and E should be switched to align with the visual presentation of the data.

**24. Reply:**

We sincerely apologize for this oversight. We have switched them.

## **25. Comment:**

A similar study has recently been published by Engelhorn et. al. in Nature Genetics, the authors should discuss their data in context of this study.

## **25. Reply:**

We thank the reviewer for pointing to the recent MOA-seq study (Engelhorn et al. 2025). This landmark paper provides an elegant, high-resolution map of haplotype-specific TF footprints and associated bQTL across many F1s, demonstrates how genetic variation at TF binding sites explains phenotypic heritability. This work and ours are complementary: Engelhorn et al. mainly focus on TF footprinting and bQTL that capture trait heritability, whereas our MH-seq study focuses on chromatin accessibility inheritance modes, cis/trans regulatory origins, enhancer-hub (super-enhancer-like) behaviour and direct functional links to heterosis (including transgressive MHSs and experimental validation). We have expanded the Discussion to explicitly compare these two works:

“A recent study utilized MOA-seq to construct a high-resolution pan-cistrome across ~25 F1 hybrids and used footprint-level allele bias to map binding QTLs and link TF-occupancy variation to trait heritability (Engelhorn et al. 2025), which provides an auspicious approach to make functional cis-variation accessible at scale. Our study started with a trio-combination approach, enabling us to capture the dynamic patterns of chromatin accessibility during maize hybridization and their potential association with heterosis. These two approaches together offer a comprehensive multi-scale framework for deciphering the molecular engines of complex trait in maize”

## References

- Crisp, Peter A., Reza Hammond, Peng Zhou, et al. 2020. "Variation and Inheritance of Small RNAs in Maize Inbreds and F1 Hybrids." *Plant Physiology* 182 (1): 318–331. <https://doi.org/10.1104/pp.19.00817>.
- Engelhorn, Julia, Samantha J. Snodgrass, Amelie Kok, et al. 2025. "Genetic Variation at Transcription Factor Binding Sites Largely Explains Phenotypic Heritability in Maize." *Nature Genetics* 57 (9): 2313–2322. <https://doi.org/10.1038/s41588-025-02246-7>.
- He, Guangming, Beibei Chen, Xuncheng Wang, et al. 2013. "Conservation and Divergence of Transcriptomic and Epigenomic Variation in Maize Hybrids." In *Genome Biology*, vol. 14, no. 6. Preprint. <https://doi.org/10.1186/gb-2013-14-6-r57>.
- Hufford, Matthew B., Arun S. Seetharam, Margaret R. Woodhouse, et al. 2021. "De Novo Assembly, Annotation, and Comparative Analysis of 26 Diverse Maize Genomes." *Science* 373 (6555): 655–662. <https://doi.org/10.1126/science.abg5289>.
- Koenecke, Nina, Jeff Johnston, Qiye He, Samuel Meier, and Julia Zeitlinger. 2017. "Drosophila Poised Enhancers Are Generated during Tissue Patterning with the Help of Repression." *Genome Research* 27 (1): 64–74. <https://doi.org/10.1101/gr.209486.116>.
- Liu, Han, Xuxu Ma, Siqi Jiang, et al. 2025. "Conservation and Variability of Long-Range Interactions in Structurally Diverse Maize Genomes." *Nature Communications* 16 (1): 10105. <https://doi.org/10.1038/s41467-025-65065-8>.
- Liu, Yuting, Xiang Gao, Hongjun Liu, et al. 2025. "Constraint of Accessible Chromatins Maps Regulatory Loci Involved in Maize Speciation and Domestication." *Nature Communications* 16 (1): 2477. <https://doi.org/10.1038/s41467-025-57932-1>.
- Li, Zhi, Peng Zhou, Rafael Della Coletta, et al. 2021. "Single-Parent Expression Drives Dynamic Gene Expression Complementation in Maize Hybrids." *The Plant Journal: For Cell and Molecular Biology* 105 (1): 93–107. <https://doi.org/10.1111/tpj.15042>.
- Li, Zhi, Andan Zhu, Qingxin Song, Helen Y. Chen, Frank G. Harmon, and Z. Jeffrey Chen. 2020. "Temporal Regulation of the Metabolome and Proteome in Photosynthetic and Photorespiratory Pathways Contributes to Maize Heterosis." *The Plant Cell* 32 (12): 3706–3722. <https://doi.org/10.1105/tpc.20.00320>.
- Luo, Jin-Hong, Min Wang, Gui-Fang Jia, and Yan He. 2021. "Transcriptome-Wide Analysis of Epitranscriptome and Translational Efficiency Associated with Heterosis in Maize." *Journal of Experimental Botany* 72 (8): 2933–2946. <https://doi.org/10.1093/jxb/erab074>.
- Lu, Zefu, Alexandre P. Marand, William A. Ricci, Christina L. Ethridge, Xiaoyu Zhang, and Robert J. Schmitz. 2019. "The Prevalence, Evolution and Chromatin Signatures of Plant Regulatory Elements." *Nature Plants* 5 (12): 1250–1259. <https://doi.org/10.1038/s41477-019-0548-z>.
- Paschold, Anja, Yi Jia, Caroline Marcon, et al. 2012. "Complementation Contributes to Transcriptome Complexity in Maize (*Zea Mays* L.) Hybrids Relative to Their Inbred Parents." *Genome Research* 22 (12): 2445–2454. <https://doi.org/10.1101/gr.138461.112>.
- Rada-Iglesias, Alvaro, Ruchi Bajpai, Tomek Swigut, Samantha A. Brugmann, Ryan A. Flynn, and Joanna Wysocka. 2011. "A Unique Chromatin Signature Uncovers Early Developmental Enhancers in Humans." *Nature* 470 (7333): 279–283.

<https://doi.org/10.1038/nature09692>.

Ricci, William A., Zefu Lu, Lexiang Ji, et al. 2019. "Widespread Long-Range Cis-Regulatory Elements in the Maize Genome." *Nature Plants* 5 (12): 1237–1249.

<https://doi.org/10.1038/s41477-019-0547-0>.

Springer, Nathan M., and Robert M. Stupar. 2007. "Allelic Variation and Heterosis in Maize: How Do Two Halves Make More than a Whole?" *Genome Research* 17 (3): 264–275.

<https://doi.org/10.1101/gr.5347007>.

Wong, Emily S., Bianca M. Schmitt, Anastasiya Kazachenka, et al. 2017. "Interplay of Cis and Trans Mechanisms Driving Transcription Factor Binding and Gene Expression Evolution." *Nature Communications* 8 (1): 1092. <https://doi.org/10.1038/s41467-017-01037-x>.

Zhao, Hainan, Wenli Zhang, Tao Zhang, et al. 2020. "Genome-Wide MNase Hypersensitivity Assay Unveils Distinct Classes of Open Chromatin Associated with H3K27me3 and DNA Methylation in *Arabidopsis Thaliana*." *Genome Biology* 21 (1): 24.

<https://doi.org/10.1186/s13059-020-1927-5>.

Zhu, Tao, Chunjiao Xia, Ranran Yu, et al. 2024. "Comprehensive Mapping and Modelling of the Rice Regulome Landscape Unveils the Regulatory Architecture Underlying Complex Traits." *Nature Communications* 15 (1): 6562. <https://doi.org/10.1038/s41467-024-50787-y>.

## **Point-by-point responses to the comments:**

### **Reviewer 1:**

#### **Comments:**

The authors have made a great job to revise their manuscript, the manuscript is now significantly improved. I suggest the authors to put the Response Figures regarding BM HiChIP data into SuppFigures to better clarify the chromatin interaction patterns of F1 seedlings. I have no further questions.

#### **Reply:**

We thank the reviewer for the positive feedback and the constructive suggestion. Following the reviewer's advice, we have put the response figures regarding the BM HiChIP data to the Supplementary Figure 4C to better clarify the chromatin interaction patterns of F1 seedlings.

### **Reviewer 2:**

#### **Comments:**

The authors have revised the manuscript according to the suggestions.

#### **Reply:**

We thank the reviewer for the positive evaluation of our revised manuscript.

### **Reviewer 3:**

#### **Comments:**

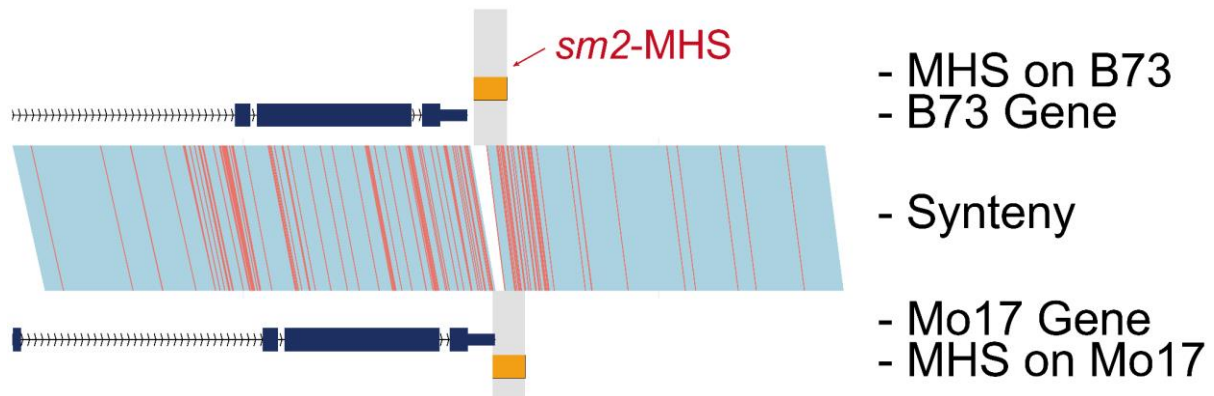
Excellent revision, all my comments/concerns were addressed.

One reply in the response to reviewers was a little unclear – comment 16, reviewer 3 – it is mentioned that this is a non-syntenic region (Fig 5c, right panel), and if so, then the analysis makes sense. But if that is the case, why is there a peak at virtually the same position in the Mo17 track (which region does the correspond to in B73?) and why is it labelled with a count, if the count is not relevant? I have probably misunderstood, but this might be clarified.

Other than that, I no further concerns.

#### **Reply:**

We sincerely thank the reviewer for the positive evaluation and recognition of our revisions. Although the peaks in Figure 5c (right panel) appear visually aligned, a closer inspection reveals that this region involves a structural variation (SV) between the two genomes (Response-Fig. 1). This SV results in significant sequence divergence at this locus, which led to its classification as a non-syntenic region. To avoid any potential misunderstanding regarding the direct comparability of these signals, we have removed the count label from the Mo17 track in the revised Figure 5c and updated the figure legend accordingly.



**Response-Fig. 1:** Detailed syntenicity of the *sm2-MHS* region. A mirrored layout centered around the "Synteny" ribbons (light blue) was displayed. For syntenic blocks, red lines indicate SNPs/InDels between B73 and Mo17; the blank area within the syntenicity ribbons represents SV, indicating regions where the B73 and Mo17 sequences are too divergent to align. This zoom-in visualization, centered around the structural comparison between the B73 (top) and Mo17 (bottom) genomes, reveals significant local sequence divergence.

#### Reviewer 4:

#### Comments:

The authors have addressed my points to my satisfaction. Also based on the serious reply to all reviewer's comments, I find the paper much improved.

Only some minor points remain:

- 1.) Title: should better use the plural of hybrid
- 2.) Line 117: merging peaks detected in both replicates
- 3.) Line 382: Supplemental Table X

#### Reply:

We sincerely thank the reviewer for the positive assessment and the final constructive suggestions. We have carefully addressed each of the remaining minor points as follows:

Title: We have changed "hybrid" to its plural form "hybrids".

Line 117: We have revised this sentence to "merging peaks detected in both replicates".

Line 382: The placeholder has been updated: "Supplementary Table 7".

## Response to Reviewers

Reviewer #3:

Thanks for the clarification, no further concerns.

Response:

We thank Reviewer for the positive feedback and for the time and effort throughout the review process.