

Conserved and divergent gene regulatory networks for crop drought resistance

Corresponding Author: Professor Liangsheng Shi

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #3

(Remarks to the Author)

My review was transferred from another Nature journal to this one. The authors have fully and successfully addressed all my remarks. I do not have any further comments and am fully satisfied with their responses.

(Remarks on code availability)

The GitHub depository is well-structured.

Reviewer #4

(Remarks to the Author)

Understanding the mechanisms of crop drought resistance is crucial for sustaining grain production, especially in the face of an ever-changing climate. Deng et al. conducted large-scale gene regulatory network analysis across several C3 and C4 crops, which behave differently in response to drought stress. They proposed it is the topological structure of GRNs, rather than a few genes, that shapes the drought resistance across different crops. Their findings are expected to help us understand the evolutionary dynamics of the GRNs controlling plant complex traits, such as drought resistance. I have read the manuscript and the rebuttal carefully, and found that although the authors have improved their manuscript to some extent according to the reviewers' suggestions and comments, the readabilities of figures and result sections still need further improvement.

I agree with reviewer 1's comments, that substantial description about the metadata, workflow, figures and results should be included in the Result section, rather than embedded in the Methods, Supplementary Notes and Supplementary figures. However, in the revised manuscript, the readability of the Result section is still relatively poor. It is suggested that:

1) A brief introduction about the new algorithm the authors developed, or a workflow should be included in the main text or main figures, to help improve the readability of this work. Fig. S17 should be moved as a main figure before the results section. In addition, "a novel water-encoding layer", this novelty should be integrated in the workflow. "Performance evaluations across transcriptional, protein, and metabolic layers". Integrate the transcriptional, protein, and metabolic layers in the workflow as well. A brief description was also needed for what are the transcriptional, protein and metabolic information from, and how they were processed into the network establishment.

2) Considering that this work collected a number of samples under normal and drought conditions, as shown in the Supplementary table, it is suggested to give a short summary about how these data were treated, grouped and used, before the results. For example, before the authors mentioned DEGs the first time (Line 152), a short description should be given how these DEGs were obtained. By collectively analyzing all the drought vs normal samples? Or choosing the union or intersection of all drought-responsive genes across all samples for each species?

3) Short descriptions are needed for all the technical terms in the first place, especially for these computational ones. Such as CH index, attention scores.

4) The supplementary table "data meta" sheet, is kind of a raw file with webpage crawling. It is suggested to filter this table for useful information for the samples, and display it in a readable sheet. The authors should double check the information in the

table to ensure they are informative and correct. For example, what do “A”, “Illumina HiSeq 2000” and “PAIRED” in the column “Sample Name” mean?

5) I agree with Reviewer 1 and Reviewer 2’s comments that only results from two species each in C3 and C4 might not be sufficient to draw a general and sound conclusion for the drought-responsive pathways in C3 and C4 crops. As an intermediate species between rice and maize/sorghum and a C3 crop, wheat display intermediate drought tolerance. Thus, the question is, how the authors distinguish these two scenarios: differences in drought responses between rice/wheat and maize/sorghum reflect differences between C3 and C4 crops, or between rice/wheat lineage and maize/sorghum lineage, simply due to their evolutionary differences? To reinforce these findings, it would be beneficial to include C3 and C4 species from plant families other than grasses.

6) The Fig. 8 is not informative to illustrate the differences in drought responses between rice/wheat and maize/sorghum lineages.

In the rebuttal, the authors said that “our study is designed as a meta-analysis (or integrative analysis) of publicly available transcriptomic datasets, not as a report on a single new experiment.” A brief explanation about “meta-analysis” should be given to help the readers understand this strategy.

In the rebuttal, the authors said that “We constructed drought-stress-specific networks and compared multiple networks across different species, ensuring an equal number of genes and focusing on those exhibiting the greatest changes in degree centrality.” However, in the revised manuscript, it went “Generating multiple networks with varying gene numbers for different species and then partitioning them into supernodes,” There is discrepancy between these two statements. “ensuring an equal number of genes” vs “with varying gene numbers”

Reviewer 1 asked “Of course this will be smaller than the total number of gene families, but does this deviate statistically from background?” However, the authors didn’t answer this question, which is important to know whether wheat (with totally different background gene numbers due to the hexaploid genome). The authors only interpreted it as “Wheat displayed the greatest number of differentially expressed genes, which is likely attributable to its hexaploid genome (Fig. 1b).”

The quality of the figures and writing are also suggested to be improved further:

Lines 212-213, “Performance evaluations across transcriptional, protein, and metabolic layers”, what did the authors or the algorithms do for each layer? Construct GRNs? What are the benchmark datasets came from for each layer? So the authors can calculate the AUC and AUPR.

For Fig. 1,

- 1) The numbers of genes in a were inconsistent with the numbers in b;
- 2) Why not order the species in a according to C3 or C4 groups? The species ordering in a and b should be consistent as well.
- 3) The x-axis label “Gene Num” in b was partially covered by d.
- 4) The legend for “Gene Count” in c-d only contains one situation: gene count = 1. It is not informative to infer the gene counts for other circle sizes.
- 5) Line 145, should be “group-specific”, rather than “species-specific”
- 6) Line 146-147: “Figure (f) shows the expression clustering of conserved DEGs of top significant GO annotations across species”. It is not clear what these genes list in Fig. 1f are and how they were selected. For a top significant GO annotation (may list a few example) across species, whether all the orthologous genes from five species investigated were selected? Or whether all the genes listed in Fig. 1f are from one significant GO annotation?
- 7) Line 147-148: “with samples grouped into six clusters based on their average expression values”. The description is not clear. According to the figure, there should be two (green and pink color) or 12 clusters of the samples, rather than 6 clusters. Further more, “their” is not clear for what.
- 8) Is it reasonable to mixed the up- and down-regulated DEGs in a single figure (Fig. 1a)? DEGs with up- or down-regulation upon drought stress should conduct different functions.

Line 158-159: “Figure 1a shows the number of orthologous and species-specific genes.” However, in the Fig. 1 legend, they are “orthologous DEGs”, rather than “orthologous genes”. They are totally different.

Line 160-163: 34 and 141 orthologous genes exhibited differential expression, specifically, in C3 and C4 crops, respectively. Should emphasize the specificity.

Line 168-171: may specify what GO terms listed in Fig. 1c correspond to these three pathways, i.e., as water channel activity, cell wall biogenesis, and water transport.

Line 175-178: GO terms such as “phosphoribulokinase activity”, “oxidoreductase activity”, “proline catabolic process”, and “fatty acid biosynthetic process” were shown in Fig. 1d. however, according to the Fig. 1 legend, Fig. 1d was results for C3-specific GOs. This is inconsistent with the statement “C4 crops exhibited a higher proportion of top-significant terms directly linked to drought tolerance, including phosphoribulokinase activity, oxidoreductase activity, proline catabolic process, and fatty acid biosynthetic process (p value < 0.05), compared to their C3 counterparts”.

Line 182: “with red indicating activation and blue indicating repression.” This sentence should be in Figure legend.

Line 83: “EES02191, EES05444, EER95621, and EER90645” which species do these genes belong to?

Lines 179-186: this paragraph only mentioned the relatively conserved expression profiles between four “EE” genes and three indica genes. But have not explained the expression profiles of other genes in Fig. 1f. What did these results tell us?

For Figure 2,

1) What do the abbreviations such as “AD”, “APL” mean?

2) What do the red numbers in Fig. 2d mean? Should specify in the legend.

3) What do the red and blue colors mean in Fig. 2e mean? Activation and repression, respectively? Should be specified in the legend and in the figure.

4) Lines 195-196: “the percentage distribution of changes in average degree in the co-regulatory network across species”. What do the changes in average degree mean? Changes in average degree when comparing what and what?

5) What do the data points in Fig. 2d mean? A single comparison between drought and normal conditions? Should be specified.

6) Line 201: TF activity in what aspects? Expression or what? Should be specified as well. It is suggested that to move the sentence “transcription factor activity, which corresponds to edge strength in these networks 46,” in the Discussion section to here when the TF activity was first mentioned.

Line 203-210: Fig. S10 shows the venn analysis of gene families across different species, rather than the diagram for the authors’ graph network algorithm, as would be expected after reading these sentences. In addition, the name of the authors’ tool should be spelled out here and in Fig. S4.

Line 216: what do attention scores mean? Should explain all these computational terms to readers with more broader background, especially those in biological field.

Fig. S5: what the differences between “normal” subfigure and “drought” subfigure in Fig. S5a we are supposed to observe?

Line 220: the authors wanted to cite Fig. S5, rather than Fig. S4, right? The same for line 223 as well.

Line 226-231: what do the higher AD, CC, APL, SIM, MIM, and FFL mean? Should be explained somehow. Such as: higher AD and CC indicate higher tightness of the network?

Line 233: “For each species, 80 networks of different sizes were generated.” What are these 80 networks? What is the motivation to generate networks of different sizes? What are these different sizes?

Line 245-246: what are the “average degree”?

The top-left sub-figure in Fig. 3a was partially covered by the top-middle one.

Lines 262-263: Figure (d) illustrates degree distributions of modules related to hormone (left panel?) and photosynthesis (right panel?) response.

Line 285-286: According to Fig. 3b, only the GO term “response to ethanol” was shared between indica-A and maize-L modules.

Line 288: there was only one GO term “oxidoreductase activity” shared between maize-A and sorghum-L modules.

Line 290-292: there were no shared GO terms between Japonica-A and other modules.

Line 297: what does CH index stands for? And what does it mean?

For Fig. 4a, there were two rice species. How did the author deal with the numbers in two rices in the bar plot? When comparing two groups, there would be only three categories: A unique, B unique and overlap between A and B. So what are the $1 - (A \cup B)$ for rice-wheat and maize-sorghum comparisons?

Lines 325-328: Figure (e) provides GO annotations for drought-responsive genes within the TCP (Teosinte branched1/Cycloidea/Proliferating cell factor)-PP2C (right panel?) and ERF (Ethylene-responsive transcription factor)-2OGD (left panel?) regulatory gene families, respectively. The ordering of the two sub-figures should be consistent with the ordering in the legend.

Line 325: provides, instead of “provide”

Line 339: “while maize and sorghum were more conserved” how this conclusion was inferred from Fig. 4a?

A detailed legend for Fig. 4d is needed. Datapoints: species or samples? Under normal or drought conditions? Values: TPM?

Line 350: "the overall expression patterns remain consistent," how was this inferred from Fig. 4c and Fig. S11?

Line 367-368: "Figure (b) illustrates the collinear positions of the 2OGD and PP2C gene families across species, revealing their genomic relationships." Which sub-figure was for which gene family? Should be specified.

Line 368-370: "Figure (c) shows the distribution of gene contraction and expansion counts for the 2OGD and PP2C families based on clustering results." Which sub-figure was for which number? Should be clear.

Line 371: What is the definition of paralogs in this study? Reciprocal best matched gene pairs? Should be specified.

Line 370-372: "Figure (d) highlights the distribution of Ka/Ks values both" missing a word "for" or "of" before "both". Again, which sub-figure was for which gene family? Should be specified.

Fig. 5f What do the data points stand for? Genes of 2OGD and PP2C families from all the five species? Or what?

Lines 401-402: "between recently duplicated paralogs within a species and orthologs between closely related species" this sentence should be mentioned when the authors talked about the Fig. 5d.

Lines 436-439: "Wheat, diverging earliest (~56.7 MYA), shows the largest expansion (+11,250) and moderate contraction (-5,189), followed by maize, sorghum, and rice. These genomic adjustments reflect the dynamic adaptations, functional innovations, and evolutionary pressures that a species undergoes." First, the authors didn't solved the comment raised by reviewer 1. Second, "wheat, diverging earliest" is not correct.

Lines 454-455: the figure should be Fig. S15, rather than Fig. S14

Line 464: there was no Fig. 6f in the figure.

Lines 464-466: "This provides further evidence that C3 crops combine growth regulation to escape drought, while C4 crops are directly adapted to drought tolerance in the representative species we studied". The authors should give more interpretation about why the above GO enrichment results would lead to this conclusion. Sentences of Lines 596-599 should be moved here to support this conclusion.

The comments for Fig. 7 legends are the same as for Fig. 5.

Line 513-514: "The whole-genome duplication event in wheat was estimated to have occurred between 0.36 MYA ($K_s = 0.01$) and 177.8 MYA ($K_s = 4.98$)" What does this sentence serve for in this paragraph?

Lines 849-851: "Finally, the contribution of a regulatory pattern to drought response was determined by summing the scores of the two gene families it regulates." What do "the two gene families" mean? And what does "it" mean in this sentence?

(Remarks on code availability)

Version 1:

Reviewer comments:

Reviewer #5

(Remarks to the Author)

This manuscript aims to elements of drought-specific gene regulatory network (GRN) structure which are conserved and divergent between species with differing drought response. The authors apply a novel gene regulatory network (GRN) construction method to produce GRNs across a variety of C3 and C4 crops under drought conditions. By comparing these networks, the authors were able to identify network topological measures which differ between the drought-specific GRNs of more and less drought-tolerant crops, drought GRN module traits and properties, and conserved and differential connections between genes in drought conditions.

The authors' responses to Reviewer 4's comments appear satisfactory to me. They have introduced some details of their novel algorithm earlier in the text, provided summaries of data and terminology were applicable, improved the supplementary materials, included relevant details on the evaluation of their algorithm, and made appropriate corrections and clarifications where requested throughout the text. The additional analysis of soybean, potato, foxtail millet, and teff does a great deal to strengthen the argument that the observed changes in network structure are general to C3 and C4 plants and not a specific combination of crops.

I believe that this work is exciting and has scientific merit, but the manuscript would benefit from minor improvements to more clearly convey this.

I also agree with prior reviewers that the code access for this manuscript is well-organized.

Primary Comments

Introduction: A brief mention in this section of some of the key findings of the study would greatly enhance the manuscript. For example, brief discussions of high-level outcomes such as those at line 724-728. This is to some extent a matter of stylistic preference but this addition would seem to be in line with other articles in this journal.

The figures and figure captions in this manuscript still require improvement. The authors have done well to incorporate the suggestions made by Reviewer 4 in this regard, but there are still numerous opportunities to improve the clarity of the results by using labels to reduce the reliance on figure captions and reducing the number of subfigures where possible. I have noted some individual improvements in my minor comments below, but I also encourage the authors to identify additional opportunities to simplify and clarify the figures.

Line 592-596: It would be clearer to state the p value and test when stating that the differences are significant. It is not clear to me how the Kruskal-Wallis (KW) test (which evaluates whether an arbitrary number of distributions have the same mean) would provide insight into the cause of that difference, as seems to be stated on line 594. It seems reasonable to note the difference and discuss possible causes for it, but the delineation between that discussion of possible causes and the statement of the results of the statistical test needs to be clearer. I also think the conclusion that wheat and maize have the highest expansion is clear for DREB, but for MADS it is less so: although at least 1 of the 5 species does not have the same median as shown by the KW test, Indica rice appears to have a similar median to Maize. Follow-up pairwise testing could be used to support these more specific claims. It would be interesting to see if these tests reveal differences in gene family expansion between DREB and MADS.

The discussion section repeats a lot of the details from the results section. This is necessary to some extent, but I think reducing the specificity in the discussion to more clearly communicate a broad, high-level overview of the results would make this section more useful.

Line 841-872: These two sections would benefit from being more closely tied together. In particular, the section on baseline methods would be improved by a clearer statement of what training and testing data are used and how they were divided into training and test sets for the supervised methods.

Minor Comments:

Line 119-122: It would strengthen the merit of this novel computational approach to explain the limitations in existing approaches that require the development of a new approach to address this analytical task.

Line 164: The phrase "is likely attributable" seems too strong for the level of support provided. Additionally, a citation here to support the discussion of the impact of ploidy on regulatory interactions would be beneficial.

Line 234-239: This first discussion of metrics such as AUC and AUPR should include some context as to the ground truth labels used to construct them. This full context is eventually provided in methods, but a brief mention of where the labels for these metrics come from would be appropriate here.

Fig 3a: The Y axis on this plot is somewhat confusing. From the text, it seems that these network metrics for each plant have been scaled so that they sum to 1. Presenting these values as a stacked plot with the Y axis labeled as percentage would seem to imply that these are proportions of an overall whole rather than distinct values for each network. This is a very compact representation, but I think sacrifices some clarity. Particularly as the discussion at 273-286 compares these measures between species, stacking the bars makes the comparison difficult. It may also be worth considering which of these metrics are necessary to support the conclusions that are being drawn later in the text. Plotting fewer metrics would allow for a more interpretable figure.

Figure 3c: I think that Fig 3d alone sufficiently supports the claims made using 3c, since 3c is a more qualitative view of the same phenomenon as 3d provides. I would consider moving 3c to supplementary to provide more space for the other elements of figure 3.

Fig 3c, 3e: these elements are not well labeled. The column labels for species and row labels for transcription factor family should be in the figure element rather than requiring reference to the figure caption.

Line 271: what "original database" is being referred to here?

Line 289-296: The discussion of network properties here would benefit from some interpretation as to what higher centrality under drought conditions might reflect. What does a higher average degree suggest about the system?

Figure 4a: the individual networks should be labeled by species in the figure, not just in the caption.

Figure 4d: the individual panels should be labeled as hormone or photosynthesis in the figure.

Line 336-338: This claim would be strengthened by a quantitative measure for scale free properties such as power law regression R^2 or that proposed by Li et al: <https://doi.org/10.1080/15427951.2005.10129111>

Line 340: What kind of differences might this steeper decline in C4 vs C4 crops reflect? Is there a broader connection to the regulatory processes governing drought response in these groups that this observation can connect to? I'm not sure if there

is basis to make that kind of statement, but if there is it would be beneficial.

Fig 5a: I'm not sure this subfigure is legible enough to be useful. For the first 2 elements, UpSet plots might be more readable than the current Venn diagrams, and these elements should be labeled as gene families and regulatory patterns in the figure as well as in the caption. The visualization for chi squared tests seems unnecessary in a main body figure and might be better moved to the supplements.

Fig 5b: The color coding for this subfigure does not have sufficient contrast between categories, and the labeling scheme needs to be revised so it can be more easily interpreted without repeated reference back to the caption. For example, since "4-1" represents indica-japonica-wheat-maize, the grouping might be better labeled as LJWM. This would also allow the figure caption to be substantially simplified. It is also difficult to see the connections.

Line 513: If the word "significantly" here refers to a statistical test, it would be appropriate to parenthetically mention the test and threshold used. If it does not, another word should be selected.

Fig 8b, 8b, 8d, 8e, and 8f: panels should be labeled in the figure, not only in the caption.

625-637: This is a very interesting result, and the kind of discussion I think could be beneficial to introduce in an abbreviated form at lines 289-296 and 340 (see comments there) to contextualize the results and help convey their importance to the overall conclusions of the paper.

(Remarks on code availability)

The code is well-documented and organized. It is possible to reproduce the results of the paper with the code provided, and the procedure to create a working environment in which to run the code is clear. The authors have provided an exceptionally transparent codebase for this work.

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/>

Dear Editors and Reviewers:

We are grateful to receive comments and suggestions on our manuscript of "NCOMMS-25-43275-T" entitled "Conserved and divergent gene regulatory networks for crop drought resistance" from you. These comments and suggestions are precious, and we have revised the manuscript based on them.

All of the modifications in the revised manuscript are pointed in responses. The responses to the reviewer's comments are appended for your convenience.

We hope the revised manuscript is more suitable for the potential publication in *Nature Communications*.

Yours sincerely,
Liangsheng Shi on behalf of all co-authors
Wuhan University

Response to Reviewer 3's comments

My review was transferred from another Nature journal to this one. The authors have fully and successfully addressed all my remarks. I do not have any further comments and am fully satisfied with their responses.

Remarks on code availability:

The GitHub depository is well-structured.

Response :

We are delighted by your positive feedback and your confirmation that our revisions have fully addressed all your previous remarks on all fronts, specifically your satisfaction with the manuscript and the well-compiled code. We sincerely appreciate your thoroughness and valuable feedback throughout the review process, which has significantly improved the quality of our work.

Response to Reviewer 4's comments

Understanding the mechanisms of crop drought resistance is crucial for sustaining grain production, especially in the face of an ever-changing climate. Deng et al. conducted large-scale gene regulatory network analysis across several C_3 and C_4 crops, which behave differently in response to drought stress. They proposed it is the topological structure of GRNs, rather than a few genes, that shapes the drought resistance across different crops. Their findings are expected to help us understand the evolutionary dynamics of the GRNs controlling plant complex traits, such as drought resistance. I have read the manuscript and the rebuttal carefully, and found that although the authors have improved their manuscript to some extent according to the reviewers' suggestions and comments, the readability of figures and result sections still need further improvement.

Response :

Thank you for carefully reviewing our manuscript and for your constructive comments. We would like to revise the figures and result sections according to your suggestions to improve the readability.

I agree with reviewer 1's comments, that substantial description about the metadata, workflow, figures and results should be included in the Result section, rather than embedded in the Methods, Supplementary Notes and Supplementary figures. However, in the revised manuscript, the readability of the Result section is still relatively poor. It is suggested that:

Response :

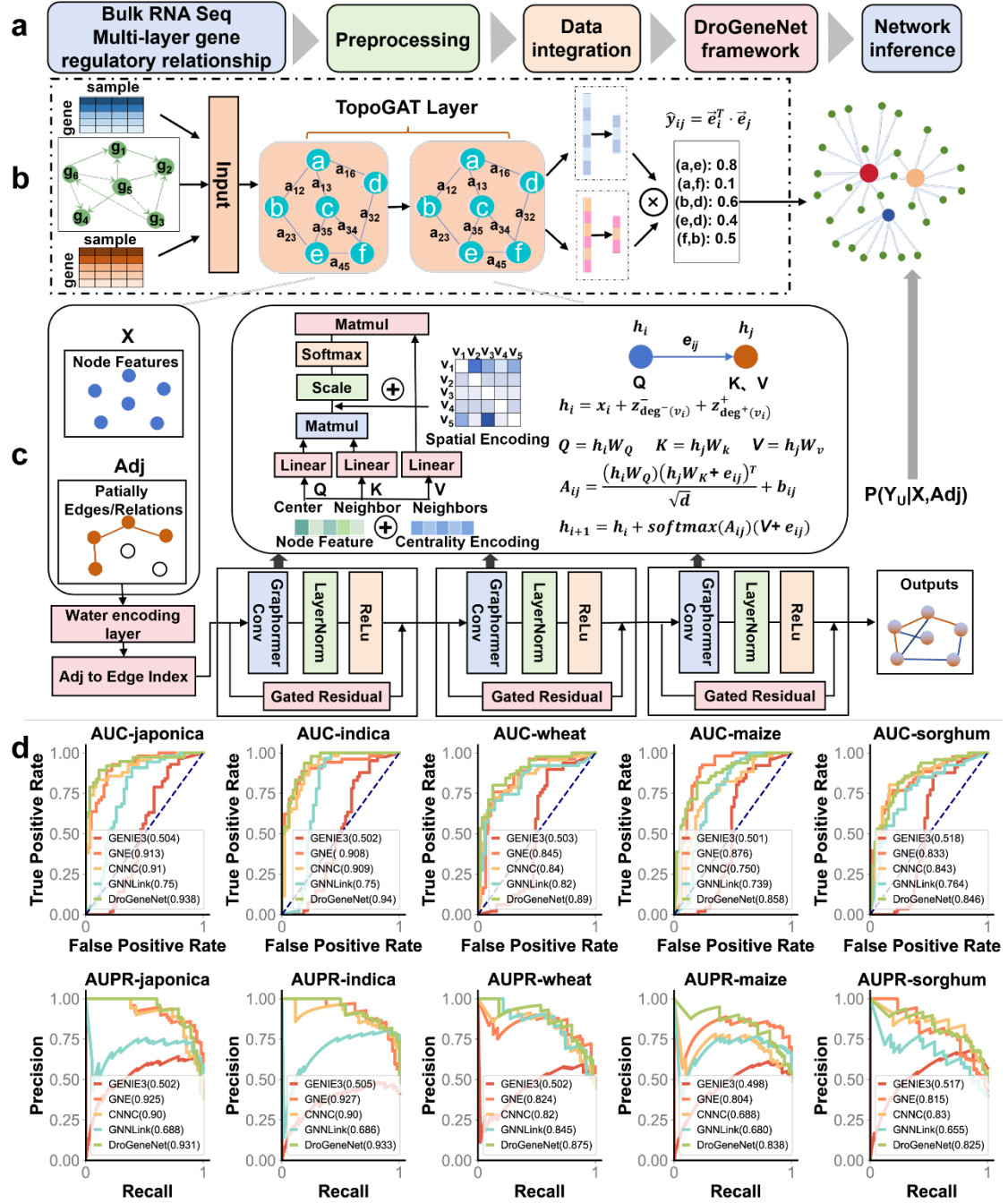
Thanks for your valuable comments. We would like to revise our results section based on your suggestions. Our point-to-point response is given below.

1) A brief introduction about the new algorithm the authors developed, or a workflow should be included in the main text or main figures, to help improve the readability of this work. Fig. S17 should be moved as a main figure before the results section. In addition, “a novel water-encoding layer”, this novelty should be integrated in the workflow. “Performance evaluations across transcriptional, protein, and metabolic layers”. Integrate the transcriptional, protein, and metabolic layers in the workflow as well. A brief description was also needed for what are the transcriptional, protein and metabolic information from, and how they were processed into the network establishment.

Response:

Thank you for your constructive comments. We have reorganized the algorithmic framework into a main figure included in the Results section. In the workflow, we have indicated the position of the water-encoding layer and highlighted the utilization of the transcriptional, protein, and metabolic layers. We also provided a brief introduction to the acquisition of multi-layer regulatory relationships and their integration into the network, while detailed implementation steps are described in the Methods section.

Line 203-248:



“Fig. 2 Overview of the DroGeneNet model framework and comparison of accuracy for each model in transcriptional regulatory network. DroGeneNet integrates a graph neural network with an attention mechanism. A prior network Adj is first built using regulatory relationships from chromatin accessibility data. Gene expression inputs are mapped into a uniform dimensional space via a water encoding layer, where each node represents a gene’s expression vector. This encoding allows data from different conditions, such as normal and drought, to be processed within a single model. During inference, batch-corrected expression data from normal or drought conditions can be input separately to reconstruct respective gene networks. To handle memory limitations with over 10,000 genes, the adjacency matrix is converted into an edge index format. Figure (a) shows the data processing pipeline, including bulk RNA-seq data collection,

preprocessing, data integration, the DroGeneNet framework, and subsequent network inference. Figure (b) presents the DroGeneNet architecture featuring the TopoGAT layer, which captures gene interactions across multiple samples to construct regulatory networks. Figure (c) provides a detailed view of the TopoGAT layer, highlighting node features (X), adjacency relations (Adj), spatial and centrality encoding, and gated residual connections used to model complex regulatory dependencies. Figure (d) provides comparison of accuracy for each model in all five species. Notably, the GENELink model was untrainable due to the large number of genes in the dataset.

Aiming to unravel the key determinants of GRN topology in drought resistance, we developed a novel graph network algorithm DroGeneNet for constructing highly accurate GRNs under drought-responsive conditions (Fig. 2). Our innovative approach pioneers the construction of GRNs under varying water conditions by incorporating a water-encoding layer. This layer allows the embedding of gene expression matrix data from different water conditions into the GRN inference process, enabling the reconstruction of gene regulatory networks that reflect gene expression patterns under specific experimental conditions (Fig. 2, Methods S1, Methods S2, Methods S3). Furthermore, our model overcomes the memory limitations associated with large adjacency matrices in traditional methods (Notes S1, Notes S2). Within a consistent model framework, gene expression data, along with transcriptional, protein, and metabolic regulatory relationships derived from database integration (Methods), were input to construct the regulatory network for each respective layer. Performance evaluations across transcriptional, protein, and metabolic layers demonstrate that our model outperforms existing tools, achieving improvements in AUC (a measure of classification accuracy) ranging from 0.02 to 0.19 and in AUPR (a measure of classification accuracy) ranging from 0.01 to 0.24, establishing a reliable platform for dissecting drought adaptation mechanisms (Table S1, Table S2, Fig. 2d). We further validated the reliability of our model by visualizing the attention scores from the intermediate layers, which represent the relative regulatory influence of each gene interaction and reflect the context-specific control logic learned by the model. We observed that a single transcription factor regulates multiple target genes (Fig. S4). Furthermore, we discovered that genes with higher attention scores are more extensively connected, underscoring their central roles within the network (Table S3, Fig. S4). Core genes identified in GRNs, primarily transcription factors, exhibit the highest centrality. Notably, we found that the ERF (ethylene response factor) family consistently plays a significant role in transcriptional regulation across species under both conditions (Fig. S4).”

2)Considering that this work collected a number of samples under normal and drought conditions, as shown in the Supplementary table, it is suggested to give a short summary about how these data were treated, grouped and used, before the results. For example, before the authors mentioned DEGs the first time (Line 152), a short description should be given how these DEGs were obtained. By collectively analyzing all the drought vs normal samples? Or choosing the union or intersection of all drought-responsive genes across all samples for each species?

Response :

Thank you for your valuable comments. We added a short summary before the results to describe how the transcriptome data were processed, grouped, and analyzed. A large set of transcriptome data under normal and drought conditions was collected and subjected to quality filtering. Samples from normal and drought conditions were grouped, and batch effects were corrected. This yielded gene expression data from multiple samples under both conditions, and a differential expression analysis was conducted across all samples to identify DEGs.

Furthermore, we have rechecked the entire manuscript and appended a short summary to the data processing section before the results.

Line 153-158:

“We collected a large amount of transcriptomic data under both normal and drought conditions. After filtering, grouping normal and drought samples, and correcting for batch effects, we obtained gene expression data for multiple samples under both conditions. To unravel the genomically evolutionary underpinnings of drought resistance, we first performed differential expression analysis using all available samples to identify drought-responsive genes across different species.”

Line 232-235:

“Within a consistent model framework, gene expression data, along with transcriptional, protein, and metabolic regulatory relationships derived from database integration (Methods), were input to construct the regulatory network for each respective layer.”

Line 270-273:

“We first explored the initial topological features of the transcriptional regulatory network, using regulatory relationships from the original database, and then examined how these features dynamically changed under drought conditions based on networks constructed with the DroGeneNet framework for both normal and drought conditions.”

Line 323-327:

“We further identified gene modules using the Louvain algorithm to understand drought-resistant gene regulatory networks across species. To minimize the potential bias introduced by different gene numbers, we constructed, for each species, a series of networks containing the same number of genes for cross-species comparisons, and then partitioned these networks into supernodes.”

Line 354-356:

“We delved deeper into the gene expression patterns and the key response genes within modules. We analyzed the gene expression data from all responsive genes within the hormone-responsive gene modules.”

Line 401-405:

“Advancing our research, we mapped genes in the drought-responsive GRNs into gene families, constructing a supergene regulatory network guided by these gene families, which facilitates the comparison of interaction mechanisms across species. Further insights into conserved mechanisms were discovered through the interaction relationships of gene families.”

3) Short descriptions are needed for all the technical terms in the first place, especially for these computational ones. Such as CH index, attention scores.

Response:

Thank you for your constructive comments. We have re-examined the entire manuscript and added brief descriptions whenever computational terms appear for the first time, with detailed explanations provided in the Methods section and Appendix.

Line 358-362:

“However, under drought stress, we observed a pronounced convergence of gene expression patterns across species. In the UMAP-embedded expression space, the Calinski-Harabasz (CH) index increased substantially (Fig. 4e), indicating fewer and more compact clusters and thus a more coherent hormone-related transcriptional response to drought across these species.”

Line 239-242:

“We further validated the reliability of our model by visualizing the attention scores from the intermediate layers, which represent the relative regulatory influence of each gene interaction and reflect the context-specific control logic learned by the model.”

Line 158-161:

“We observed significant changes in gene expression across different crop species, characterized by both upregulated and downregulated DEGs (1.3-fold change, q value < 0.05; fold change reflects expression change magnitude, while q value denotes FDR-adjusted statistical significance) (Fig. S2).”

Line 178-182:

“The conserved differentially expressed genes across all species were enriched in pathways such as water channel activity (GO:0015250), cell wall biogenesis (GO:0009834), and water transport (GO:0006833) (p value < 0.05, denoting statistical significance), indicating that all species adjust water transport processes in response to drought stress (Fig. 1e, Fig.S18d).”

Line 235-239:

“Performance evaluations across transcriptional, protein, and metabolic layers demonstrate that our model outperforms existing tools, achieving improvements in AUC (a measure of classification accuracy) ranging from 0.02 to 0.19 and in AUPR (a measure of classification accuracy) ranging from 0.01 to 0.24, establishing a reliable platform for dissecting drought adaptation mechanisms (Table S1, Table S2, Fig. 2d).”

Line 273-284:

“The initial topological characteristics revealed that rice exhibits the highest average degree (AD, where larger values indicate denser overall regulatory connectivity) among species, while wheat has the lowest. The clustering coefficient (CC, with higher values reflecting tighter local gene network connectivity) distribution aligns with average degree trends across species. In contrast, average path length (APL, for which higher values correspond to longer average shortest regulatory distances between genes) and single-input motifs (SIM, where an increased value denotes a greater prevalence of simple TF-target regulation) display consistent distributions with minimal variability. Wheat has the highest proportion of multi-input motifs (MIM, where larger values represent more complex multi-TF regulatory convergence), while rice demonstrates the largest proportion of feedback loop motifs (FFL, with higher values indicating more abundant regulatory feedback circuits).”

4)The supplementary table “data meta” sheet, is kind of a raw file with webpage crawling. It is suggested to filter this table for useful information for the samples, and display it in a readable sheet. The authors should double check the information in the table to ensure they are informative and correct. For example, what do “A”, “Illumina HiSeq 2000” and “PAIRED” in the column “Sample Name” mean?

Response:

Thank you for your valuable comments. We originally used the “data meta” sheet in the supplementary table to present the raw data information collected from online sources. The processed and utilized data are shown in the “data info” sheet. The “data meta” sheet was indeed somewhat messy in format, so we reorganized and filtered the original data information to ensure accuracy. The updated table is now cleaner and easier to read. In the “Sample Name” column, the letter “A” was meaningless due to formatting issues. “Illumina HiSeq 2000” should be in the “Instrument” column, and “PAIRED” should be in the “LibraryLayout” column. We have corrected all of these errors one by one.

Supplementary table “data meta” (simplified version)

Run	Assay Type	AvgSpotLen	Bases	BioProject	BioSample	...	version	tissue	treatment
SRR14322728	RNA-Seq	145	6957954858	PRJNA725109	SAMN18869952	...	1	leaf	drought
SRR14322729	RNA-Seq	145	6665095850	PRJNA725109	SAMN18869951	...	1	leaf	drought
SRR14322730	RNA-Seq	144	6940510295	PRJNA725109	SAMN18869950	...	1	leaf	control

SRR14322731	RNA-Seq	145	6375745548	PRJNA725109	SAMN18869949	...	1	leaf	control
SRR14322732	RNA-Seq	145	7039418949	PRJNA725109	SAMN18869948	...	1	leaf	control
...	...	...	...	...	...	...	...	...	...
SRR8743251	RNA-Seq	101	1868680689	PRJNA527782	SAMN11159519	...	1	Roots	Control
SRR8743252	RNA-Seq	101	1867060447	PRJNA527782	SAMN11159518	...	1	Roots	Control
SRR8743253	RNA-Seq	101	2498685359	PRJNA527782	SAMN11159517	...	1	Roots	Control
SRR8743254	RNA-Seq	101	1904925246	PRJNA527782	SAMN11159516	...	1	Roots	Control
SRR8743255	RNA-Seq	101	2628923546	PRJNA527782	SAMN11159515	...	1	Roots	Control

5)I agree with Reviewer 1 and Reviewer 2's comments that only results from two species each in C₃ and C₄ might not be sufficient to draw a general and sound conclusion for the drought-responsive pathways in C₃ and C₄ crops. As an intermediate species between rice and maize/sorghum and a C₃ crop, wheat display intermediate drought tolerance. Thus, the question is, how the authors distinguish these two scenarios: differences in drought responses between rice/wheat and maize/sorghum reflect differences between C₃ and C₄ crops, or between rice/wheat lineage and maize/sorghum lineage, simply due to their evolutionary differences? To reinforce these findings, it would be beneficial to include C₃ and C₄ species from plant families other than grasses.

Response:

Thank you for your constructive comments. We fully agree that, ideally, C₃ and C₄ species from different families and genera should be included to obtain more generalizable conclusions.

To address this limitation and enhance the universality of our study, we actively explored the incorporation of data from non-Poaceae C₃ and C₄ crops. For C₃ crops, we additionally collected data from soybean (*Glycine max*) in Fabaceae and potato (*Solanum tuberosum*) in Solanaceae. For C₄ crops, we added the Poaceae species foxtail millet (*Setaria italica*) and teff (*Eragrostis tef*). For these species, we collected transcriptomic data and regulatory relationship data, constructed networks under both control and drought conditions, and reanalyzed the main results to verify their robustness. All additional bioinformatics data have also been included in the supplementary tables. The corresponding results have been provided as supplementary appendix material, as follows:

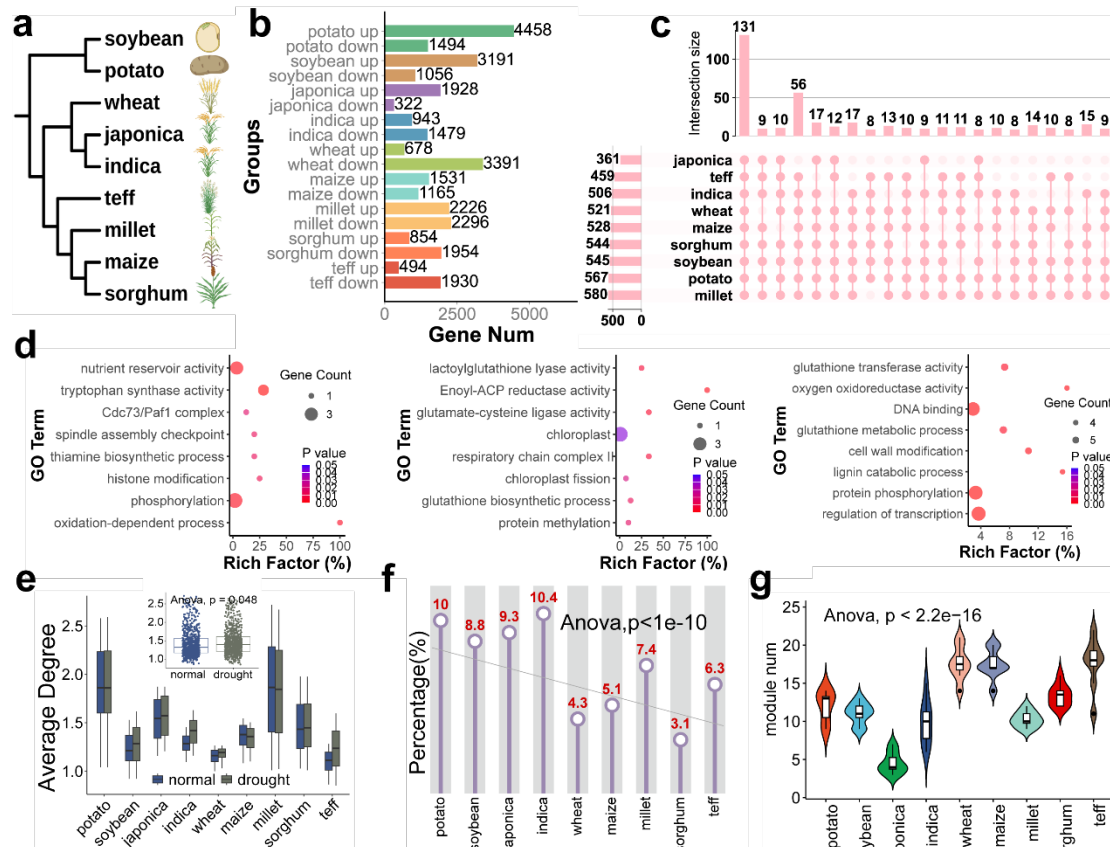


Fig. S18 Comparative analysis of the C₃ crops (potato, soybean, japonica, indica, wheat) and C₄ crops (maize, millet, sorghum, teff) in terms of differentially expressed genes, functional annotation, gene regulatory network topology, and gene modules. Figure (a) illustrates the species tree of the 9 species studied. Figure (b) illustrates the number of upregulated and downregulated DEGs in each species. Figure (c) presents an UpSet plot of groups of orthologous DEGs shared by various species under drought stress. Figure (d) detail top significant Gene Ontology (GO) annotations of group-specific DEGs. The left panel focuses on C₃ crops (soybean, potato, japonica rice, indica rice, and wheat), the center panel on C₄ crops (millet, teff, maize, and sorghum), and the right panel shows GO terms of common orthologous DEGs shared by all species under drought stress. Figure (e) compares changes in average network degree between normal and drought conditions. Figure (f) shows the percentage distribution of absolute changes in average degree in the co-regulatory network across species between normal and drought conditions. Red numbers indicate the maximum absolute percentage change in average degree between normal and drought conditions for each species. Figure (g) shows the number of gene modules identified under drought conditions.

Figure S18c shows the number of groups of orthologous and species-specific DEGs. A total of 131 orthologous genes were identified as differentially expressed across all species. Gene Ontology (GO) annotations were employed for functional analysis of these conserved and divergent genes. The conserved differentially expressed genes across all species were enriched in pathways such as cell wall modification (GO:0042545), glutathione metabolic process (GO:0006749), and oxygen oxidoreductase activity (GO:0052716) (p value < 0.05 , denoting statistical significance), indicating that all species adjust water transport processes in response to

drought stress (Fig. S18d). Among the compared network centrality measures, the drought group generally exhibited significantly higher values than the normal group based on the mean values, with the difference being statistically significant ($p = 0.048$) (Fig. S18e). Further analysis of the average degree changes under drought conditions across different species' networks revealed significant differences among species (ANOVA $p < 1e-10$). Notably, potato, soybean and rice exhibited a greater magnitude of change compared to the other species, while maize and sorghum tended to show smaller fluctuations (Fig. S18f). We found that among the representative species we studied, crops with greater drought tolerance generally possess smaller but more numerous gene modules, while less drought-tolerant crops tend to exhibit fewer, larger modules (Anova, $p < 2.2e-16$) (Fig. S18g).

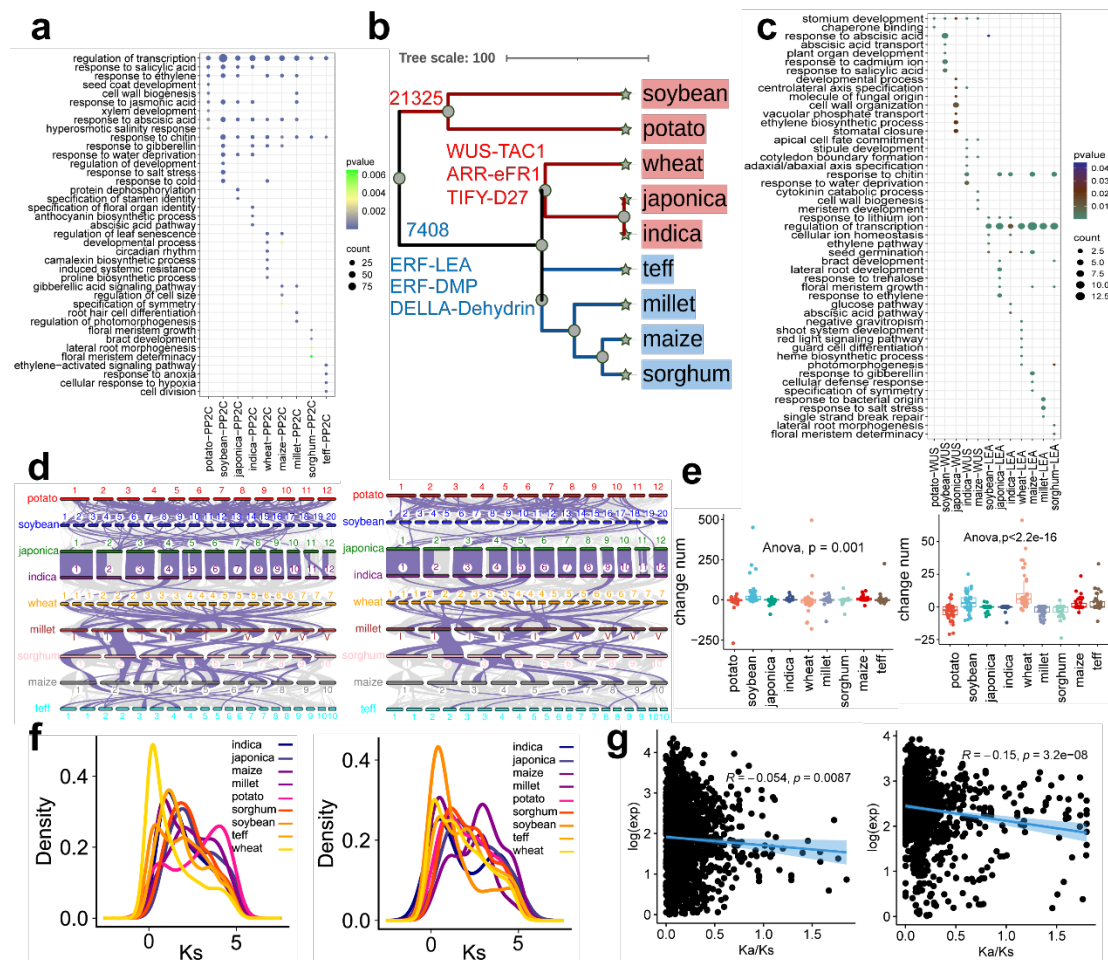


Fig. S19 Functional annotation and phylogenetic analysis of conserved and divergent gene regulatory networks under drought stress. Figure (a) provides GO annotations for drought-responsive genes within the PP2C (Protein Phosphatase 2C) regulatory conserved gene family. Figure (b) features a drought-regulatory gene tree across species, with branch numbers indicating the count of regulatory relationships unique to each branch compared to adjacent ones. Red branches represent C3 crops, and blue branches represent C4 crops. The highlighted regulatory patterns focus on top-ranked drought-response relationships. Figure (c) provide GO annotations for genes in differentiated drought-regulatory families, including WUS (WUSCHEL-like) and LEA (Late embryogenesis abundant protein, SMP subgroup domain), highlighting their functional

roles under drought conditions. Figure (d) illustrates the collinear positions of the ERF (AP2/ERF transcription factor, ERF subfamily) (left panel) and PP2C (right panel) gene families across species, revealing their genomic relationships. Figure (e) shows the distribution of gene contraction and expansion counts for the ERF (left panel) and PP2C (right panel) families based on clustering results. Figure (f) displays the distribution of synonymous substitution rates (Ks) values for ERF and PP2C families within homolog clusters, providing insights into evolutionary timescales and species-specific differences. Figure (g) explores the correlation between Ka/Ks values and gene expression levels (log(FPKM)) for drought-responsive ERF and PP2C family genes across all species, indicating evolutionary pressure and functional significance.

GO annotations were used to reveal the functions of genes from the PP2C, WUS and LEA families. We found that the PP2C families show highly conserved drought-regulatory pathways, such as transcriptional regulation (regulation of transcription) and hormone response (abscisic acid and chitin responses) (Fig. S19a). In contrast, pathways associated with the WUS and LEA families are more variable between species (Fig. S19c). Further collinearity analysis shows conserved regions for the ERF and PP2C gene families, particularly between japonica and indica, which retain similar chromosomal positions, indicating high conservation (Fig. S19d). Obvious expansion and contraction of the ERF and PP2C gene families were observed among species. Both families show greater expansion in wheat, soybean ($p = 0.001$ for ERF; $p < 2.2 \times 10^{-16}$ for PP2C) (Fig. S19e). Due to allopolyploid nature and recent genome duplication, wheat exhibited lower mean Ks values within paralogs (Fig. S19f). We found that correlation analysis between Ka/Ks ratios and gene expression levels (log(exp)) reveals a negative relationship for both families, with $R = -0.054$ ($p = 0.0087$) for ERF and $R = -0.15$ ($p = 3.2 \times 10^{-8}$) for PP2C. This suggests that genes under stronger purifying selection tend to have higher expression levels, emphasizing their critical roles in drought resistance (Fig. S19g).

After including more species, we observed a similar trend even in non-Poaceae crops, which further strengthens our conclusions.

We have revised the manuscript accordingly, the conclusions in our manuscript mainly focus on Poaceae, with the other non-Poaceae crops serving as additional validation:

- 1) We have updated the wording of the abstract accordingly.

Line 31-32:

“Other crops in Fabaceae and Solanaceae were also included for further validation.”

- 2) Incorporate the validation results of the newly added species into the description of the Results section of the manuscript.

Line 178-182:

“The conserved differentially expressed genes across all species were enriched in pathways such as water channel activity (GO:0015250), cell wall biogenesis (GO:0009834), and water transport (GO:0006833) (p value < 0.05 , denoting statistical significance), indicating that all species adjust water transport processes in response to drought stress (Fig. 1e, Fig.S18d).”

Line 289-292:

“Among the compared network centrality measures, the drought group generally exhibited significantly higher values than the normal group based on the mean values (T-stat = -2.174), with the difference being statistically significant ($p = 0.014$) (Fig. 3b, Fig.S18e).”

Line 294-296:

Notably, rice, potato and soybean exhibited a greater magnitude of change compared to the other species, while maize and sorghum tended to show smaller fluctuations (Fig. 3c, Fig. 3d, Fig. S5, Fig.S18f).

Line 328-331:

“We found that among the representative species we studied, crops with greater drought tolerance generally possess smaller but more numerous gene modules, while less drought-tolerant crops tend to exhibit fewer, larger modules (Anova, $p = 2.1e-10$) (Fig. 4a, Fig. S18g).”

Line 433-437:

“We found that the PP2C and 2OGD families show highly conserved drought-regulatory pathways, such as transcriptional regulation (DNA binding, transcription factor activity, nuclear localization) and physiological processes (ethylene and chitin responses for PP2C; redox and oxidoreductase activity for 2OGD) (Fig. 5e, Fig. S19a).”

Line 465-467:

“Further collinearity analysis shows conserved regions for the 2OGD and PP2C gene families, particularly between japonica and indica, which retain similar chromosomal positions, indicating high conservation (Fig. 6b, Fig. S19d).”

Line 480-484:

“When comparing Ks values between recently duplicated paralogs within a species and orthologs between closely related species, we observed that Ks values were generally higher within paralogs (Fig. 6e, Fig. S13). Due to allopolyploid nature and recent genome duplication, wheat exhibited lower mean Ks values within paralogs (Fig. 6e, Fig. S19f).”

Line 490-492:

“This suggests that genes under stronger purifying selection tend to have higher expression levels, emphasizing their critical roles in drought resistance (Fig. 6f, Fig. S19g).”

Line 552-555:

“These results provide further evidence that representative C₃ crops combine growth regulation to escape drought, while representative C₄ crops are directly adapted to drought tolerance in the representative species we studied (Fig. 7e, Fig. S19c).”

Line 623-625:

“We observed that changes in co-expression and co-regulation network average degree indicate that species with weaker drought resistance undergo more extensive rewiring under drought stress (Fig. 3c, Fig. 3d, Fig. S5, Fig. S18f).”

Line 643-645:

“In contrast, wheat, maize, and sorghum display more gradual curves, suggesting that their drought response systems are supported by multiple independent modules, enhancing network robustness (Fig. 4a, Fig. S18g).”

Line 685-687:

“This selective pressure supports the efficiency of the drought response, especially for genes with high expression levels, which tend to have lower Ka/Ks ratios (Fig. 6f, Fig. S19g) ⁶⁶.”

Line 767-770:

“To ensure the reliability of the results, we further obtained transcriptome data under control and drought conditions for soybean (*Glycine max*) in Fabaceae, potato (*Solanum tuberosum*) in Solanaceae, as well as foxtail millet (*Setaria italica*) and teff (*Eragrostis tef*).”

- 3) Tighten the wording in the Results and Discussion by revising previous generalized statements about “C₃ crops” and “C₄ crops” into comparisons between “these representative C₃ crops and representative C₄ crops” or between “the two lineages of rice/wheat and maize/sorghum,” and avoid simply extrapolating the conclusions to all C₃/C₄ crops, so as to make the conclusions more rigorous.

Line 35-37:

“Divergent mechanisms were identified, including SPL–PELP in rice/wheat lineage and ERF–Psb28 in maize/sorghum lineage from large-scale comparative network analysis.”

Line 184-190:

“Through functional annotation of orthologous differentially expressed genes across species, we uncovered a distinct functional divergence between representative C₃ and representative C₄ crops. Notably, representative C₄ crops exhibited a higher proportion of top-significant terms directly linked to drought tolerance, including phosphoribulokinase activity, oxidoreductase activity, proline catabolic process, and fatty acid biosynthetic process (p value < 0.05), compared to their C₃ counterparts (Fig. 1c, Fig.1d, Fig.S3, Fig.S18d).”

Line 338-341:

“The degree distribution curves show steeper declines in representative C₃ crops compared to representative C₄ crops (Spearman correlation coefficient test, $p < 2.2 \times 10^{-16}$), reflecting differences in network connectivity and regulation (Fig. 4d).”

Line 411-412:

“The analysis revealed significant differences between rice/wheat lineage and maize/sorghum lineage ($\chi^2 = 341.98$, $p = 2.36 \times 10^{-76}$).”

Line 505-507:

“Figure (d) illustrates gene expression distributions for selected drought-differentiated families: ASN (Asparagine synthase, glutamine-hydrolyzing) and Psb28 for rice/wheat lineage and maize/sorghum lineage.”

Line 524-527:

“Rubisco/PEPC (Phosphoenolpyruvate carboxylase) patterns are maize/sorghum lineage-specific, while SPL-PELP (Pistil-specific extensin-like protein) / Dof (Dof zinc finger protein)-PELP are rice/wheat lineage-specific. Within rice, DREB is prominent in indica, and GAGA (GAGA-binding transcriptional activator) in japonica (Fig. 7a, Fig. 7b).”

Line 529-531:

“Comparative gene expression analysis (Fig. 7d) showed that ASN and Psb28 levels were significantly higher in maize/sorghum lineage compared to rice/wheat lineage (t-test, $p < 2.22 \times 10^{-16}$ and $p = 5.6 \times 10^{-8}$).”

Line 552-555:

“These results provide further evidence that representative C₃ crops combine growth regulation to escape drought, while representative C₄ crops are directly adapted to drought tolerance in the representative species we studied (Fig. 7e, Fig. S19c).”

Line 732-734:

“(3) Divergence in drought-resistance mechanisms is also evident among species. Rice/wheat lineage rely on distinct SPL-PELP regulation, while maize/sorghum lineage use ERF-Psb28 and TCP-ASN pathways.”

- 4) Add separate sentences in the Discussion to explain that the effects of photosynthetic type and those of lineage cannot be fully disentangled, and explicitly identify this as one of the limitations of the present study.

Line 743-747:

“The observed differences between C₃ and C₄ crops are inherently intertwined with their distinct evolutionary lineages. In this study, the effects of photosynthetic type and those of phylogenetic history cannot be fully separated, which represents a limitation of our comparative framework.”

- 5) In the future directions section, emphasize that incorporating C₃ and C₄ species beyond Poaceae and systematically comparing their drought responses across different families will be crucial for further validating and extending the findings of this study.

Line 747-750:

“Future work that incorporates more C₃ and C₄ species from additional plant families beyond Poaceae, and systematically compares their drought responses across lineages, will be essential to decouple these factors and validate the generalizability of our findings.”

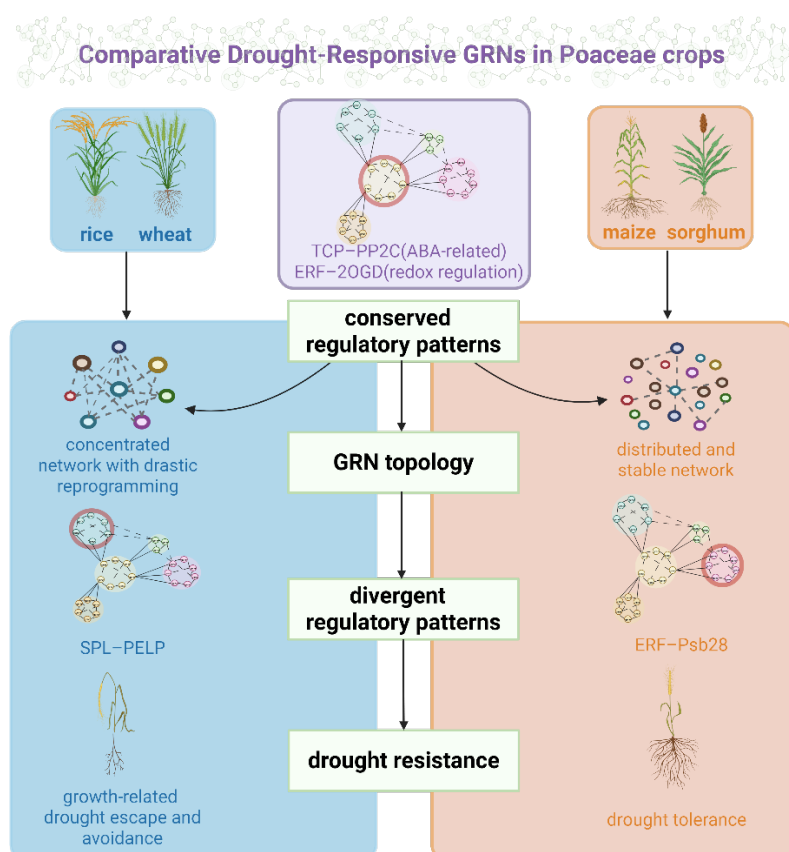
We believe that this result is highly representative. If the reviewers consider it necessary to further expand the species comparison, we are willing to explore feasible approaches

for additional improvement and investigation, although this would entail substantial additional experimental work.

6)The Fig. 8 is not informative to illustrate the differences in drought responses between rice/wheat and maize/sorghum lineages.

Response :

Thank you for your valuable comment on Figure 8. We agree that the current presentation may not clearly convey the differences in drought responses between the rice/wheat and maize/sorghum lineages. In the revised manuscript, we have improved the figure by applying more distinct visual separation between the two crop groups (e.g., split panels and color coding) and adding additional schematic information (such as illustrations of gene regulatory networks and drought response patterns) to better support the textual descriptions. We believe these modifications make the differences between the two lineages more intuitive and information-rich. Furthermore, as an additional main figure has been included in the Results section, to avoid redundancy, and after considering the relative importance and function of each figure, we have moved this figure to the Appendix to maintain overall brevity and readability of the manuscript.



“Fig. S17 A regulatory model diagram of conserved and divergent gene regulatory networks in Poaceae crops (rice, wheat, maize, and sorghum) under drought conditions (Created with BioRender).”

In the rebuttal, the authors said that “our study is designed as a meta-analysis (or integrative analysis) of publicly available transcriptomic datasets, not as a report on a single new experiment.” A brief explanation about “meta-analysis” should be given to help the readers understand this strategy.

Response:

Thank you for your constructive comment. We have added a brief explanation about “meta-analysis” in the Data Collection Methods section to help readers better understand.

Line 770-781:

“Our study is designed as a meta-analysis of publicly available transcriptomic datasets. This approach allows us to synthesize findings from multiple independent experiments, thereby increasing statistical power, identifying robust and conserved patterns, and drawing more generalizable conclusions than might be possible from any single study alone. By integrating data from a wide range of experiments, we aim to capture a more comprehensive picture of drought responses across different species and conditions. The rigorous selection and harmonization of these datasets are crucial for the validity of a meta-analysis, ensuring that the combined results accurately reflect the underlying biological phenomena. We performed a strict initial screening of the data. When a study includes different genotypes, tissues, or varying degrees of drought conditions, it is treated as separate datasets for inclusion in a meta-analysis. For data containing different time points, we only include data from one time point.”

In the rebuttal, the authors said that “We constructed drought-stress-specific networks and compared multiple networks across different species, ensuring an equal number of genes and focusing on those exhibiting the greatest changes in degree centrality.” However, in the revised manuscript, it went “Generating multiple networks with varying gene numbers for different species and then partitioning them into supernodes,” There is discrepancy between these two statements. “ensuring an equal number of genes” vs “with varying gene numbers”

Response:

Thank you for your constructive comments. After re-examining our previous response and the manuscript, as well as carefully checking our code, we found that our earlier wording was ambiguous. What we intended to convey is that, for each species, we used the same number of genes when making comparisons. To ensure the reliability of the results, we also generated multiple networks with varying numbers of genes. Specifically, for each of the five species, we constructed a series of networks containing 1,000, 1,500, 2,000, 2,500, 3,000, and 3,500 genes, respectively. Networks with the same number of genes were compared across species, while networks of different sizes

were used to validate the conclusions. We have revised the corresponding sentence in the manuscript to make this point clearer.

Line 324-328:

“To minimize the potential bias introduced by different gene numbers, we constructed, for each species, a series of networks containing the same number of genes for cross-species comparisons, and then partitioned these networks into supernodes. Networks of different sizes were also generated to verify the robustness of the conclusions.”

Reviewer 1 asked “Of course this will be smaller than the total number of gene families, but does this deviate statistically from background?” However, the authors didn’t answer this question, which is important to know whether wheat (with totally different background gene numbers due to the hexaploid genome). The authors only interpreted it as “Wheat displayed the greatest number of differentially expressed genes, which is likely attributable to its hexaploid genome (Fig. 1b).”

Response:

We sincerely appreciate your valuable comments. We have re-checked our previous response and manuscript and found that our earlier description was insufficient. Reviewer 1, in commenting on the initially submitted version of the manuscript, referred to the sentence “Our analysis of drought-responsive gene families, which constituted a smaller subset compared to the total gene families, showed reduced overlap between species, indicating species-specific adaptations to drought stress (Fig. 4a)” (this sentence is no longer in the current version), and raised the question: “Of course this will be smaller than the total number of gene families, but does this deviate statistically from background?”

We agree that it is essential to determine whether the observed number and overlap of drought-responsive gene families deviate statistically from what would be expected by chance.

In the revised analysis, we explicitly tested this. We first identified differentially expressed genes (DEGs) in each species using identical fold-change and p-value thresholds (Supplementary Fig. 10). Orthologs were inferred with OrthoFinder and annotated by InterPro, yielding a set of 1,572 conserved gene families present in all five species. All subsequent analyses of drought-responsive families were restricted to this common set, so that the gene-family background was identical across species.

We mapped the DEGs to this conserved gene family set, ensuring that the background gene family number used for comparison was consistent across species. As such, mapping DEGs to gene families does not inherently cause wheat, which is hexaploid and has more DEGs, to also display the highest number of responsive gene families (Supplementary table sheet “Fig 4a gene family venn”). As a hexaploid plant, wheat’s

polyploidy has led to substantial gene family expansion, increasing both total gene numbers and the size of individual families.

To make this point clearer, we have revised the wording in the manuscript accordingly.

Line 163-166:

“Wheat displayed the greatest number of differentially expressed genes, a phenomenon is likely attributable to its hexaploid nature, which provides three homologous copies of many genes and thus a larger pool of transcripts available for regulation under stress (Fig. 1a).”

Line 407-411:

“Biologically relevant families were retained across species using InterPro annotations (chi-squared test, $p < 1e-6$) (Fig. S9). We first performed a Venn analysis of drought-responsive gene families and their regulatory patterns. Further analysis was conducted on the 1572 gene families present in all five species (Fig. S9). Drought-responsive gene families are derived from differentially expressed genes.”

The quality of the figures and writing are also suggested to be improved further:

Lines 212-213, “Performance evaluations across transcriptional, protein, and metabolic layers”, what did the authors or the algorithms do for each layer? Construct GRNs? What are the benchmark datasets came from for each layer? So the authors can calculate the AUC and AUPR.

Response:

Thank you for your valuable suggestion. Our algorithm adopts the same framework, where the input to the model consists of gene expression matrices under normal and drought conditions. For the transcriptional, protein, and metabolic layers, we perform modeling separately; in each layer, the corresponding regulatory relationships are combined with the expression matrix and fed into the model for training. Specifically, the transcriptional layer corresponds to the transcriptional regulatory network, the protein layer to the protein-protein interaction network, and the metabolic layer to the metabolic regulatory network. In other words, we construct gene regulatory networks for each individual layer.

The gene expression matrices are derived from extensive processing of transcriptomic datasets, whereas the regulatory relationships are collected from different databases, which integrate information from experimental studies and literature sources. During model training, the complete gene expression matrix is used as input, but the regulatory relationships are partitioned into training, validation, and test sets. This means the model is provided with only a subset of regulatory relationships and must infer the remaining ones based on the gene expression data, an approach commonly used in gene

regulatory network modeling, as detailed in the referenced literature and the baseline algorithm frameworks used for comparison.

To improve clarity, we have included further methodological details in the revised manuscript.

Line 232-239:

“Within a consistent model framework, gene expression data, along with transcriptional, protein, and metabolic regulatory relationships derived from database integration (Methods), were input to construct the regulatory network for each respective layer. Performance evaluations across transcriptional, protein, and metabolic layers demonstrate that our model outperforms existing tools, achieving improvements in AUC (a measure of classification accuracy) ranging from 0.02 to 0.19 and in AUPR (a measure of classification accuracy) ranging from 0.01 to 0.24, establishing a reliable platform for dissecting drought adaptation mechanisms (Table S1, Table S2, Fig. 2d).”

Line 858-872:

To evaluate the effectiveness of our model in predicting GRNs, we compared it against five state-of-the-art methods commonly used for GRN inference, as detailed below:

GENIE3⁹⁷: A machine learning method based on random forests that constructs GRNs using regression weight coefficients. It achieved first place in the DREAM5 In Silico Network Challenge in 2010.

GNE⁹⁸: A Multi-Layer Perceptron (MLP) approach to encode gene expression profiles and network topology for predicting gene dependencies.

CNNC⁹⁹: Converts gene expression data into joint probability distribution images and employs a deep convolutional neural network to infer GRNs.

GENELink⁴⁷: A graph attention network method for inferring potential GRNs.

GNNLink¹⁰⁰: A graph convolutional network approach for GRN inference.

These models were selected based on their relevance to our research problem, their status as state-of-the-art baseline methods, the diversity of their approaches, and the availability of reproducible code. To ensure consistency and comparability, we used the default parameters provided in the original implementations of all baseline methods.

For Fig. 1,

1)The numbers of genes in a were inconsistent with the numbers in b;

Response:

Thank you for your valuable comment. We have re-examined our data. Figure 1a shows the number of orthologous gene groups among the differentially expressed genes, as presented in the supplementary table “Fig 1a deg_ortho_upset”. Figure 1b illustrates the numbers of all upregulated and downregulated differentially expressed genes for each species. To improve the accuracy of our description, we have revised the corresponding sentence in the manuscript. For better readability and a smoother flow of the manuscript, we have changed the order of Figures 1a and 1b.

Line 140-142:

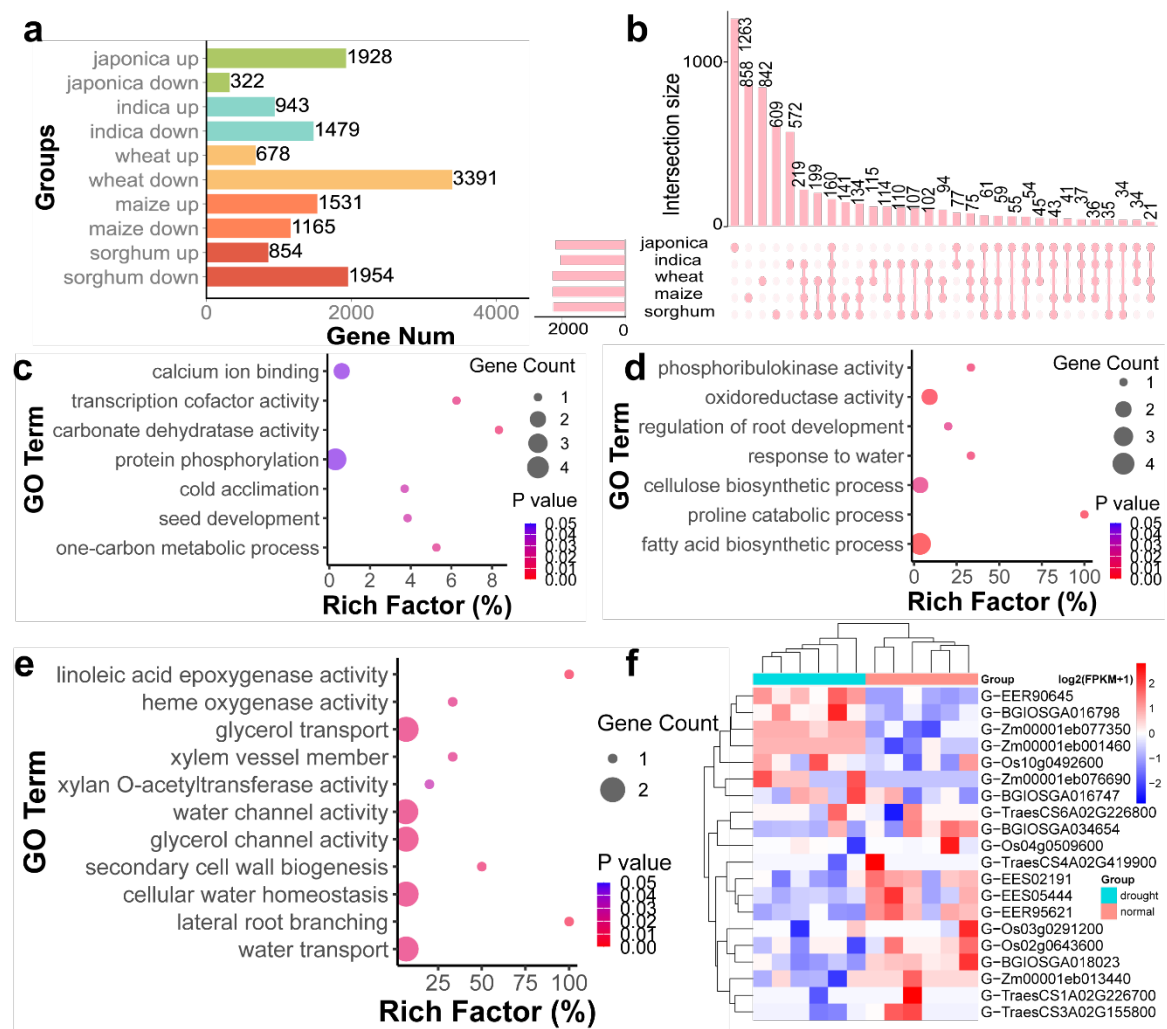
“Figure (a) illustrates the number of upregulated and downregulated DEGs in each species. Figure (b) presents an UpSet plot of groups of orthologous DEGs shared by various species under drought stress.”

2)Why not order the species in a according to C₃ or C₄ groups? The species ordering in a and b should be consistent as well.

Response:

Thank you for your valuable comment. We have redrawn the images in the order japonica-indica-wheat-maize-sorghum and ensured that the species order in figures a and b is consistent.

Line 138-152:



“Fig. 1 Upregulated and downregulated genes in different species under drought stress and functional annotation of orthologous genes. Figure (a) illustrates the number of upregulated and downregulated DEGs in each species. Figure (b) presents an UpSet plot of groups of orthologous DEGs shared by various species under drought stress.

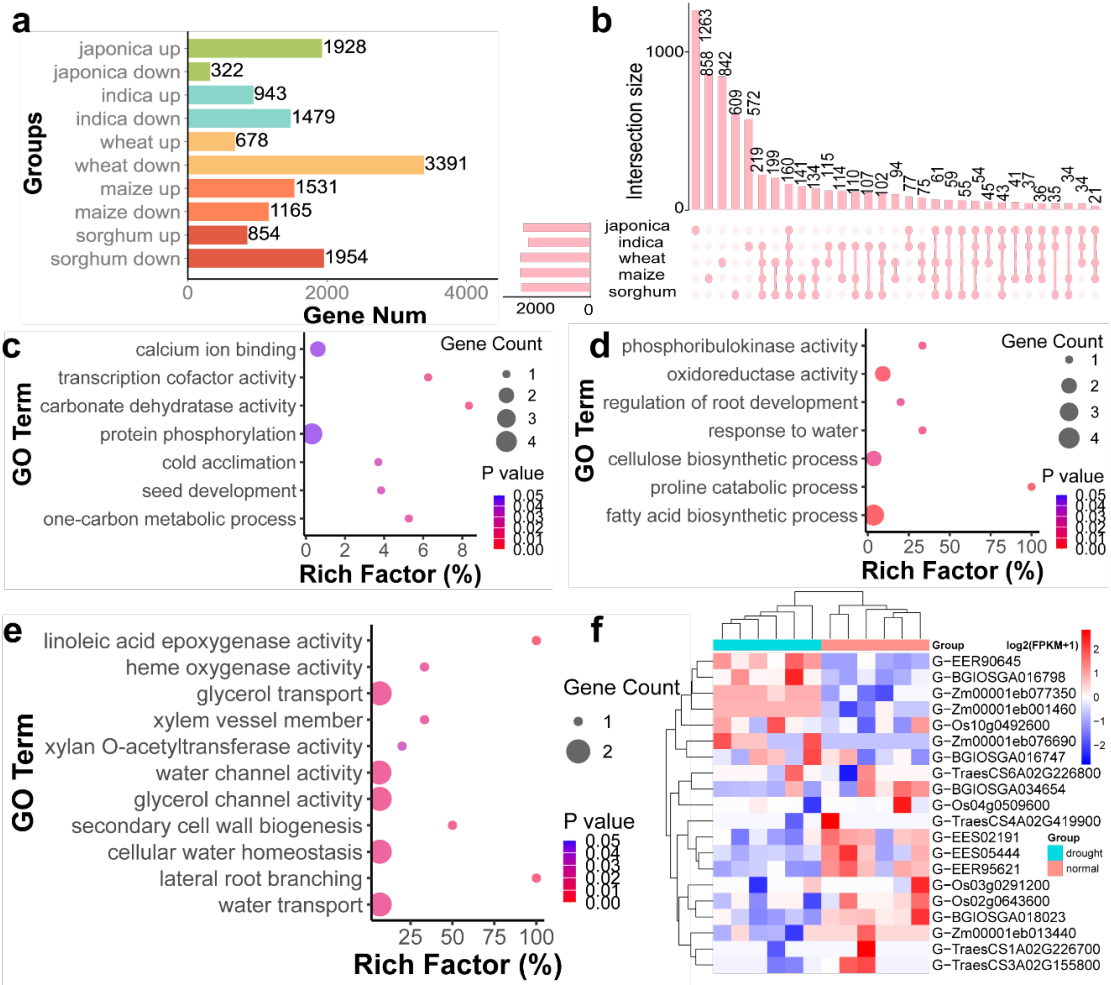
Figures (c) and (d) detail top significant Gene Ontology (GO) annotations of group-specific DEGs, with (c) focusing on representative C₃ crops (japonica rice, indica rice, and wheat) and (d) on representative C₄ crops (maize and sorghum). Figure (e) provides top significant GO annotations of common orthologous DEGs shared by all species under drought stress. Figure (f) shows the expression clustering of conserved DEGs of top significant GO annotations across species. In the heatmap, red indicates activation and blue indicates repression. Samples are divided into normal and drought groups, each further grouped into six clusters based on the average expression values of the conserved DEGs within each cluster. The detailed list is in the appendix table. G-xx represents a group of orthologous genes using gene xx as the representative.”

3)The x-axis label “Gene Num” in b was partially covered by d.

Response:

Thank you for your valuable comment. We have repositioned figure b. In the updated version, the x-axis label 'Gene Num' is now visible.

Figure 1

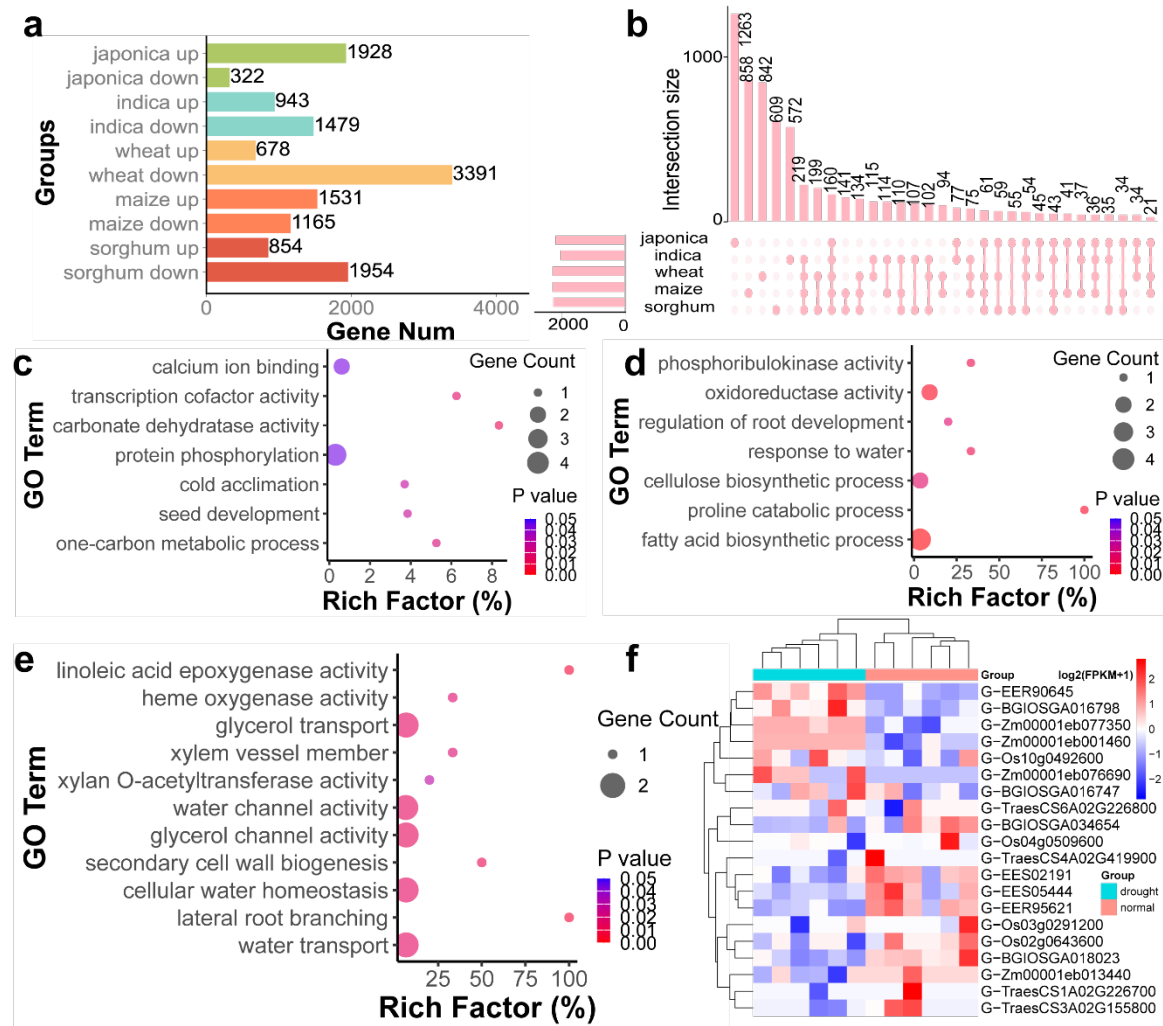


4)The legend for “Gene Count” in c-d only contains one situation: gene count = 1. It is not informative to infer the gene counts for other circle sizes.

Response:

Thank you for your constructive comment. We have readjusted the legend so that gene counts for other circle sizes can also be inferred.

Line 138-152:



“Fig. 1 Upregulated and downregulated genes in different species under drought stress and functional annotation of orthologous genes. Figure (a) illustrates the number of upregulated and downregulated DEGs in each species. Figure (b) presents an UpSet plot of groups of orthologous DEGs shared by various species under drought stress. Figures (c) and (d) detail top significant Gene Ontology (GO) annotations of group-specific DEGs, with (c) focusing on representative C₃ crops (japonica rice, indica rice, and wheat) and (d) on representative C₄ crops (maize and sorghum). Figure (e) provides top significant GO annotations of common orthologous DEGs shared by all species under drought stress. Figure (f) shows the expression clustering of conserved DEGs of top significant GO annotations across species. In the heatmap, red indicates activation

and blue indicates repression. Samples are divided into normal and drought groups, each further grouped into six clusters based on the average expression values of the conserved DEGs within each cluster. The detailed list is in the appendix table. G-xx represents a group of orthologous genes using gene xx as the representative.”

5)Line 145, should be “group-specific”, rather than “species-specific”

Response:

Thank you for your valuable feedback. We have accordingly changed 'species-specific' to 'group-specific'.

Line 142-145:

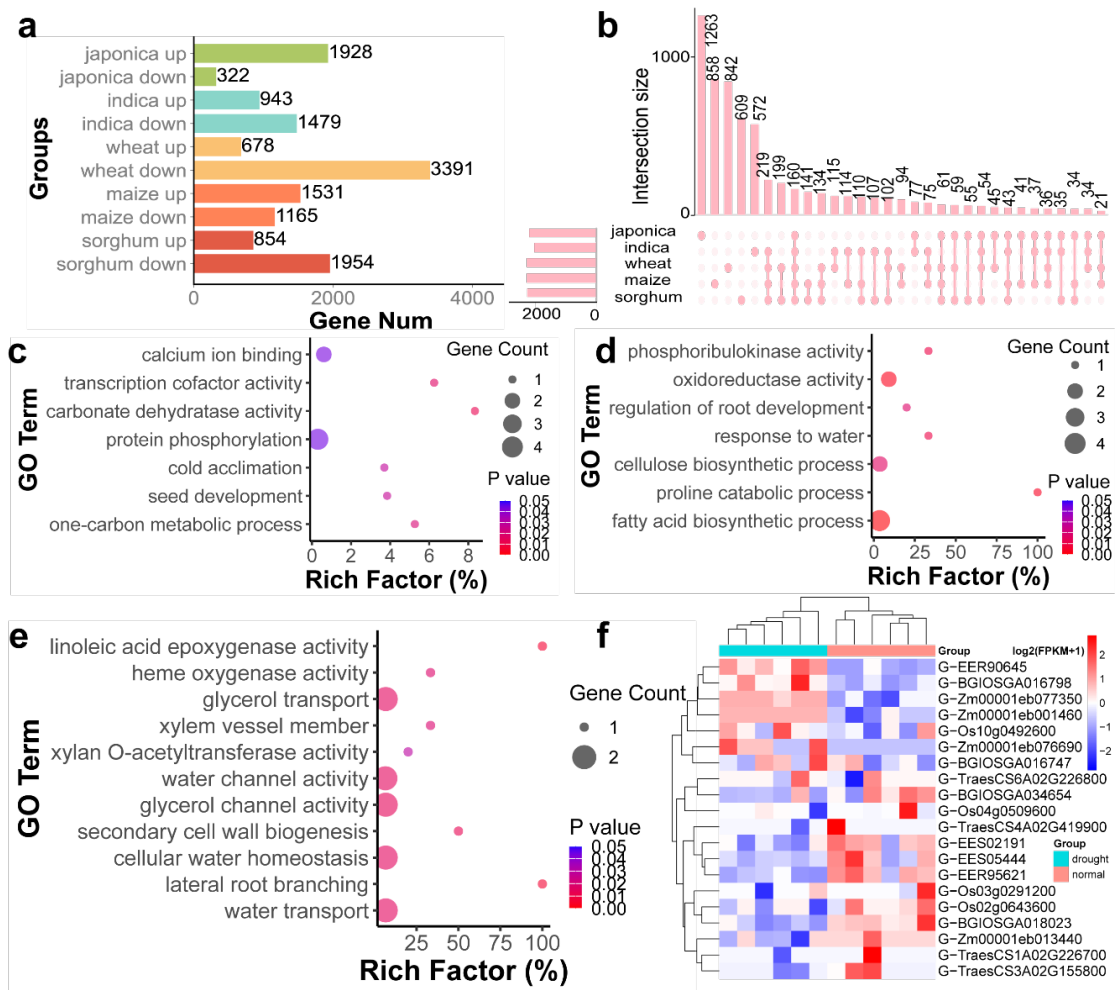
“Figures (c) and (d) detail top significant Gene Ontology (GO) annotations of group-specific DEGs, with (c) focusing on representative C₃ crops (japonica rice, indica rice, and wheat) and (d) on representative C₄ crops (maize and sorghum).”

6)Line 146-147: “Figure (f) shows the expression clustering of conserved DEGs of top significant GO annotations across species”. It is not clear what these genes list in Fig. 1f are and how they were selected. For a top significant GO annotation (may list a few example) across species, whether all the orthologous genes from five species investigated were selected? Or whether all the genes listed in Fig. 1f are from one significant GO annotation?

Response:

Thank you for your constructive comment. We first identified orthologs among the differentially expressed genes (DEGs) of the five species, and obtained 160 orthologous groups in which all five species contain DEGs. We then performed GO annotation on these 160 orthologous DEG groups, as shown in Fig. 1e. From the significantly enriched GO terms (water transport, GO:0006833; cellular water homeostasis, GO:0009992; glycerol channel activity, GO:0015254; water channel activity, GO:0015250; glycerol transport, GO:0015793; lateral root branching, GO:0080181; secondary cell wall biogenesis, GO:2000652; xylan O-acetyltransferase activity, GO:1990538 — see Supplementary Table “Fig 1cde deg_GO”), we extracted the expression values of the genes from the five species (Supplementary Table “Fig 1f ortho_deg_exp”) and used them for plotting. When a species contained multiple orthologous genes within a group, we used the mean expression value of these genes to represent that group. In other words, the gene name shown in the figure actually represents the expression of an orthologous gene group. The corresponding GO term information, gene lists, and averaged expression values have all been added to Supplementary Table “Fig 1f ortho_deg_exp”. To make the figure clearer, we have revised the gene labels in the plot and added corresponding descriptions of the data processing and figure labeling in the manuscript.

Figure 1



Supplementary table “Fig 1f ortho_deg_exp” (Simplified version)

all species top significant GO term	GO term ID
water transport;cellular water homeostasis;glycerol channel activity;water channel activity;glycerol transport	GO:0006833,GO:0009992,GO:0015254,GO:0015250,GO:0015793
lateral root branching	GO:0080181
secondary cell wall biogenesis	GO:2000652
xylan O-acetyltransferase activity	GO:1990538
GO term	ortho DEGs
GO:0006833,GO:0009992,GO:0015254,GO:0015250,GO:0015793	Os10g0492600,BG10SGA016798,"TraesCS3A02G155800, TraesCS6B02G275100",Zm00001eb076690,"EES02191, EES12617, KXG31086"
GO:0080181	Os04g0509600,BG10SGA016747,"TraesCS6A02G226800, TraesCS6B02G254800","Zm00001eb077350, Zm00001eb247430","EES05444, KXG26692"
GO:2000652	"Os02g0643600, Os06g0101800",BG10SGA018023,"TraesCS4A02G419900, TraesCS4A02G420100, TraesCS4D02G345300,

	TraesCS5A02G519300", "Zm00001eb001460, Zm00001eb260850", EER95621	
GO:1990538	Os03g0291200, BGIOGA034654, "TraesCS1A02G226700, TraesCS1B02G240700, TraesCS1D02G228100, TraesCS3A02G141500, TraesCS3A02G235900, TraesCS3B02G264000, TraesCS3D02G235200, TraesCS4A02G109700, TraesCS4B02G194100, TraesCS4B02G194500, TraesCS4D02G195200, TraesCS5A02G152100, TraesCS5B02G150700, TraesCS5D02G157300, TraesCS6B02G353800", "Zm00001eb013440, Zm00001eb130530, Zm00001eb347490", "EER90645, EER90646, EER95036, EES01702, EES07854, EES12472, EES19366"	
groups	gene	species
G-Os10g0492600	Os10g0492600	japonica
G-Os04g0509600	Os04g0509600	japonica
G-Os02g0643600	Os02g0643600, Os06g0101800	japonica
...		...
G-EES05444	EES05444, KXG26692	sorghum
G-EER95621	EER95621	sorghum
G-EER90645	EER90645, EER90646, EER95036, EES01702, EES07854, EES12472, EES19366	sorghum

Line 148-152:

“Samples are divided into normal and drought groups, each further grouped into six clusters based on the average expression values of the conserved DEGs within each cluster. The detailed list is in the appendix table. G-xx represents a group of orthologous genes using gene xx as the representative.”

7)Line 147-148: “with samples grouped into six clusters based on their average expression values”. The description is not clear. According to the figure, there should be two (green and pink color) or 12 clusters of the samples, rather than 6 clusters. Further more, “their” is not clear for what.

Response:

Thank you for your constructive comments. Indeed, our previous description was not entirely clear. We have two groups, 'normal' and 'drought,' with each group containing six clusters, resulting in a total of twelve clusters. We have revised our description accordingly. Previously, 'their' referred to all the transcriptomic samples mentioned. To avoid confusion, we have rephrased our description.

Line 148-152:

“Samples are divided into normal and drought groups, each further grouped into six clusters based on the average expression values of the conserved DEGs within each cluster. The detailed list is in the appendix table. G-xx represents a group of orthologous genes using gene xx as the representative.”

8) Is it reasonable to mix the up- and down-regulated DEGs in a single figure (Fig. 1a)? DEGs with up- or down-regulation upon drought stress should conduct different functions.

Response:

Thank you for your valuable comments. We fully agree that up- and down-regulated genes generally perform different biological functions. The goal of our Fig. 1a (now revised as Fig. 1b), however, was specifically to identify the core set of orthologous genes that are consistently responsive to drought stress across species, irrespective of their direction of regulation, as this conserved gene set itself is of great biological interest. The expression regulatory patterns (up-regulation or down-regulation) of such orthologous genes can diverge due to differences in species' evolutionary trajectories, ecological niches, or physiological strategies. Our core objective is to identify a “core set of conserved genes” that show significant responses to drought stress across all species, regardless of the direction of regulation. Similarly, we only explore whether a certain biological functional pathway will respond under drought conditions, regardless of whether it is activated or inhibited. Therefore, the Upset plot in Fig. 1a (now revised as Fig. 1b) was used to visualize the intersection of these responsive DEGs as a whole.

Line 158-159: “Figure 1a shows the number of orthologous and species-specific genes.” However, in the Fig. 1 legend, they are “orthologous DEGs”, rather than “orthologous genes”. They are totally different.

Response:

Thank you for your constructive comments. This figure was generated by identifying orthologous genes from the differentially expressed genes of all species. Our previous description was inaccurate, and we have now revised the wording accordingly.

Line 168-169:

“Figure 1b shows the number of groups of orthologous and species-specific DEGs.”

Line 160-163: 34 and 141 orthologous genes exhibited differential expression, specifically, in C₃ and C₄ crops, respectively. Should emphasize the specificity.

Response:

Thank you for your constructive comments. We have revised this sentence and strengthened the corresponding description.

Line 170-173:

“In total, 34 orthologous genes were specifically differentially expressed only in the rice/wheat lineage, and 141 orthologous genes were specifically differentially expressed only in the maize/sorghum lineage, indicating lineage-specific transcriptional responses to drought.”

Line 168-171: may specify what GO terms listed in Fig. 1c correspond to these three pathways, i.e., as water channel activity, cell wall biogenesis, and water transport.

Response:

Thank you for your constructive comments. We have added the corresponding GO term IDs.

Line 178-182:

“The conserved differentially expressed genes across all species were enriched in pathways such as water channel activity (GO:0015250), cell wall biogenesis (GO:0009834), and water transport (GO:0006833) (p value < 0.05, denoting statistical significance), indicating that all species adjust water transport processes in response to drought stress (Fig. 1e, Fig.S18d).”

Line 175-178: GO terms such as “phosphoribulokinase activity”, “oxidoreductase activity”, “proline catabolic process”, and “fatty acid biosynthetic process” were shown in Fig. 1d. however, according to the Fig. 1 legend, Fig. 1d was results for C₃-specific GOs. This is inconsistent with the statement “C₄ crops exhibited a higher proportion of top-significant terms directly linked to drought tolerance, including phosphoribulokinase activity, oxidoreductase activity, proline catabolic process, and fatty acid biosynthetic process (p value < 0.05), compared to their C₃ counterparts”.

Response:

Thank you for your valuable comments. In the previous version of the manuscript, there was an issue with the correspondence between the figure legend and the figure itself. We have now revised this.

Line 142-146:

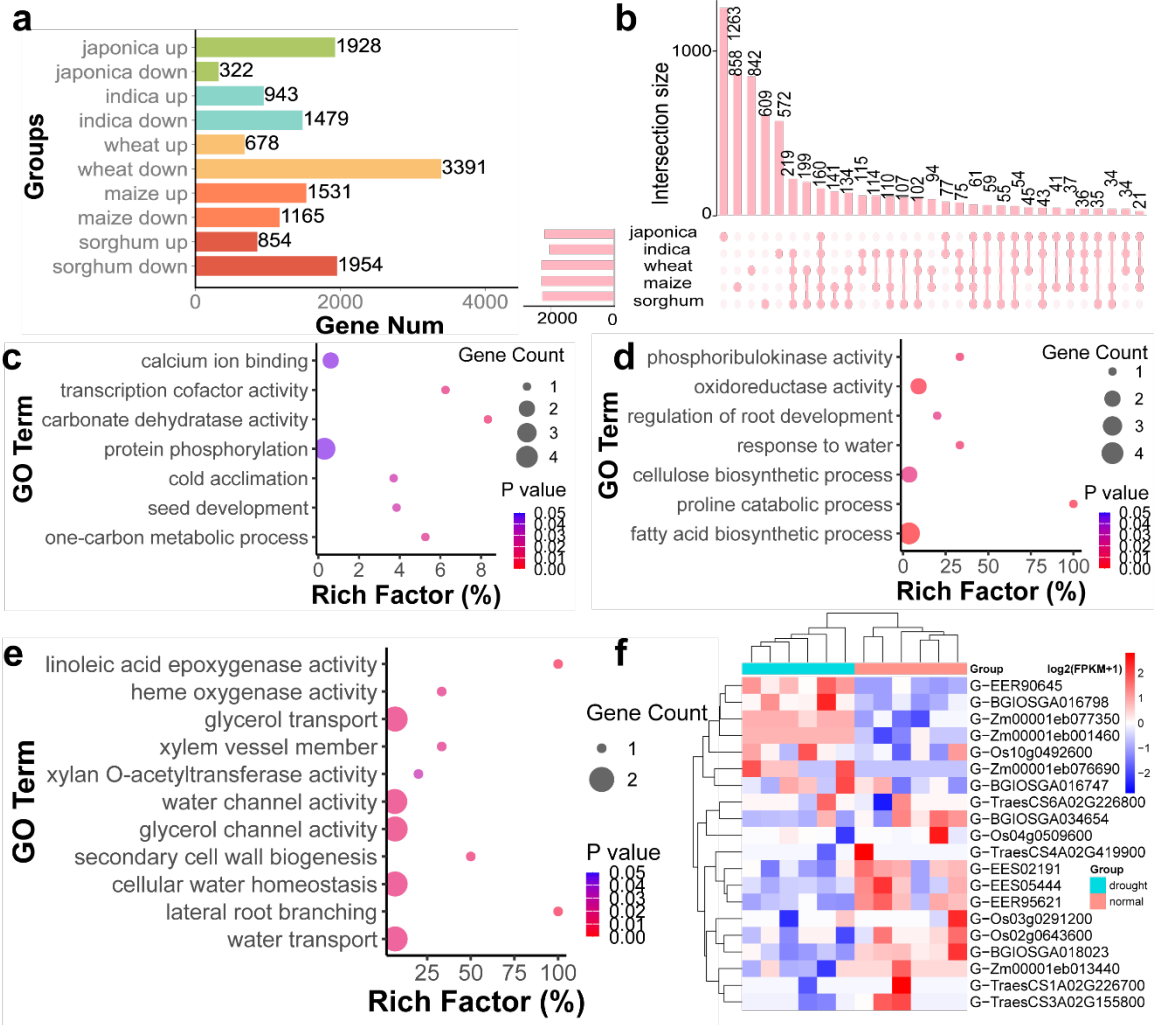
“Figures (c) and (d) detail top significant Gene Ontology (GO) annotations of group-specific DEGs, with (c) focusing on representative C₃ crops (japonica rice, indica rice, and wheat) and (d) on representative C₄ crops (maize and sorghum). Figure (e) provides top significant GO annotations of common orthologous DEGs shared by all species under drought stress.”

Line 186-190:

“Notably, representative C₄ crops exhibited a higher proportion of top-significant terms directly linked to drought tolerance, including phosphoribulokinase activity,

oxidoreductase activity, proline catabolic process, and fatty acid biosynthetic process (p value < 0.05), compared to their C₃ counterparts (Fig. 1c, Fig.1d, Fig.S3, Fig.S18d).”

Line 138-152:



“Fig. 1 Upregulated and downregulated genes in different species under drought stress and functional annotation of orthologous genes. Figure (a) illustrates the number of upregulated and downregulated DEGs in each species. Figure (b) presents an UpSet plot of groups of orthologous DEGs shared by various species under drought stress. Figures (c) and (d) detail top significant Gene Ontology (GO) annotations of group-specific DEGs, with (c) focusing on representative C₃ crops (japonica rice, indica rice, and wheat) and (d) on representative C₄ crops (maize and sorghum). Figure (e) provides top significant GO annotations of common orthologous DEGs shared by all species under drought stress. Figure (f) shows the expression clustering of conserved DEGs of top significant GO annotations across species. In the heatmap, red indicates activation and blue indicates repression. Samples are divided into normal and drought groups, each further grouped into six clusters based on the average expression values of the conserved DEGs within each cluster. The detailed list is in the appendix table. G-xx represents a group of orthologous genes using gene xx as the representative.”

Line 182: “with red indicating activation and blue indicating repression.” This sentence should be in Figure legend.

Response :

Thank you for your valuable comments. We have made the corresponding revisions.

Line 148:

“In the heatmap, red indicates activation and blue indicates repression.”

Line 193-194:

“Drought stress induces notable upregulation and downregulation.”

Line 183: “EES02191, EES05444, EER95621, and EER90645” which species do these genes belong to?

Response :

Thank you for your valuable comments. These genes belong to sorghum, and to make this clearer, we have revised the wording in the manuscript. In addition, we have added species labels for each gene in the supplementary table “Fig 1f ortho_deg_exp.”

Line 194-201:

“We identified two distinct groups of genes with similar expression patterns. One group shows activation of gene expression under drought stress, including G-EER90645 in sorghum, G-BGIOGA016798 in indica rice, and G-Zm00001eb077350 and G-Zm00001eb001460 in maize. The other group shows repression of gene expression under drought stress, including G-EES02191, G-EES05444, and G-EER95621 in sorghum; G-Os03g0291200 and G-Os02g0643600 in japonica rice; G-BGIOGA018023 in indica rice; G-Zm00001eb013440 in maize; and G-TraesCS1A02G226700 and G-TraesCS3A02G155800 in wheat (Fig. 1f).”

Supplementary table “Fig 1f ortho_deg_exp” (simplified version)

groups	gene	species
G-Os10g0492600	Os10g0492600	japonica
G-Os04g0509600	Os04g0509600	japonica
G-Os02g0643600	Os02g0643600, Os06g0101800	japonica
...	...	
G-EES05444	EES05444, KXG26692	sorghum
G-EER95621	EER95621	sorghum
G-EER90645	EER90645, EER90646, EER95036, EES01702, EES07854, EES12472, EES19366	sorghum

Lines 179-186: this paragraph only mentioned the relatively conserved expression

profiles between four “EE” genes and three indica genes. But have not explained the expression profiles of other genes in Fig. 1f. What did these results tell us?

Response:

Thank you for your valuable comments. We have revised the wording of this paragraph and have also provided explanations for the other genomic expression patterns. This result is mainly intended to show that among the orthologous differentially expressed genes in the top GO pathways with the highest significance, there are shared gene expression patterns.

Line 194-201:

“We identified two distinct groups of genes with similar expression patterns. One group shows activation of gene expression under drought stress, including G-EER90645 in sorghum, G-BGIOSGA016798 in indica rice, and G-Zm00001eb077350 and G-Zm00001eb001460 in maize. The other group shows repression of gene expression under drought stress, including G-EES02191, G-EES05444, and G-EER95621 in sorghum; G-Os03g0291200 and G-Os02g0643600 in japonica rice; G-BGIOSGA018023 in indica rice; G-Zm00001eb013440 in maize; and G-TraesCS1A02G226700 and G-TraesCS3A02G155800 in wheat (Fig. 1f).”

For Figure 2,

1)What do the abbreviations such as “AD”, “APL” mean?

Response:

Thank you for your valuable comments. These abbreviations are computational metrics used to describe the topological structure of gene regulatory networks. We have added brief explanations in the manuscript, and the detailed calculation formulas are provided in the supporting information in Methods S4.

Line 273-284:

“The initial topological characteristics revealed that rice exhibits the highest average degree (AD, where larger values indicate denser overall regulatory connectivity) among species, while wheat has the lowest. The clustering coefficient (CC, with higher values reflecting tighter local gene network connectivity) distribution aligns with average degree trends across species. In contrast, average path length (APL, for which higher values correspond to longer average shortest regulatory distances between genes) and single-input motifs (SIM, where an increased value denotes a greater prevalence of simple TF-target regulation) display consistent distributions with minimal variability. Wheat has the highest proportion of multi-input motifs (MIM, where larger values represent more complex multi-TF regulatory convergence), while rice demonstrates the largest proportion of feedback loop motifs (FFL, with higher values indicating more abundant regulatory feedback circuits).”

2)What do the red numbers in Fig. 2d mean? Should specify in the legend.

Response:

Thank you for your valuable comments. The red numbers in Fig. 2d represent, for each species, the maximum absolute change in the average degree of the 80 generated networks between the normal and drought conditions. To improve clarity, we have added the corresponding legend in the figure.

Line 258-259:

“Red numbers indicate the maximum absolute percentage change in average degree between normal and drought conditions for each species.”

3)What do the red and blue colors mean in Fig. 2e mean? Activation and repression, respectively? Should be specified in the legend and in the figure.

Response:

Thank you for your valuable comments. The images we showed before were not clear enough. In these figures, blue represents the normal condition and red represents the drought condition. We have added explanatory text to the legend and included the corresponding legend in the figures.

Line 261-263:

“Figure (e) examines changes in transcription factor family activity between normal and drought conditions, in which blue represents the normal condition and red represents the drought condition.”

Line 249-269:

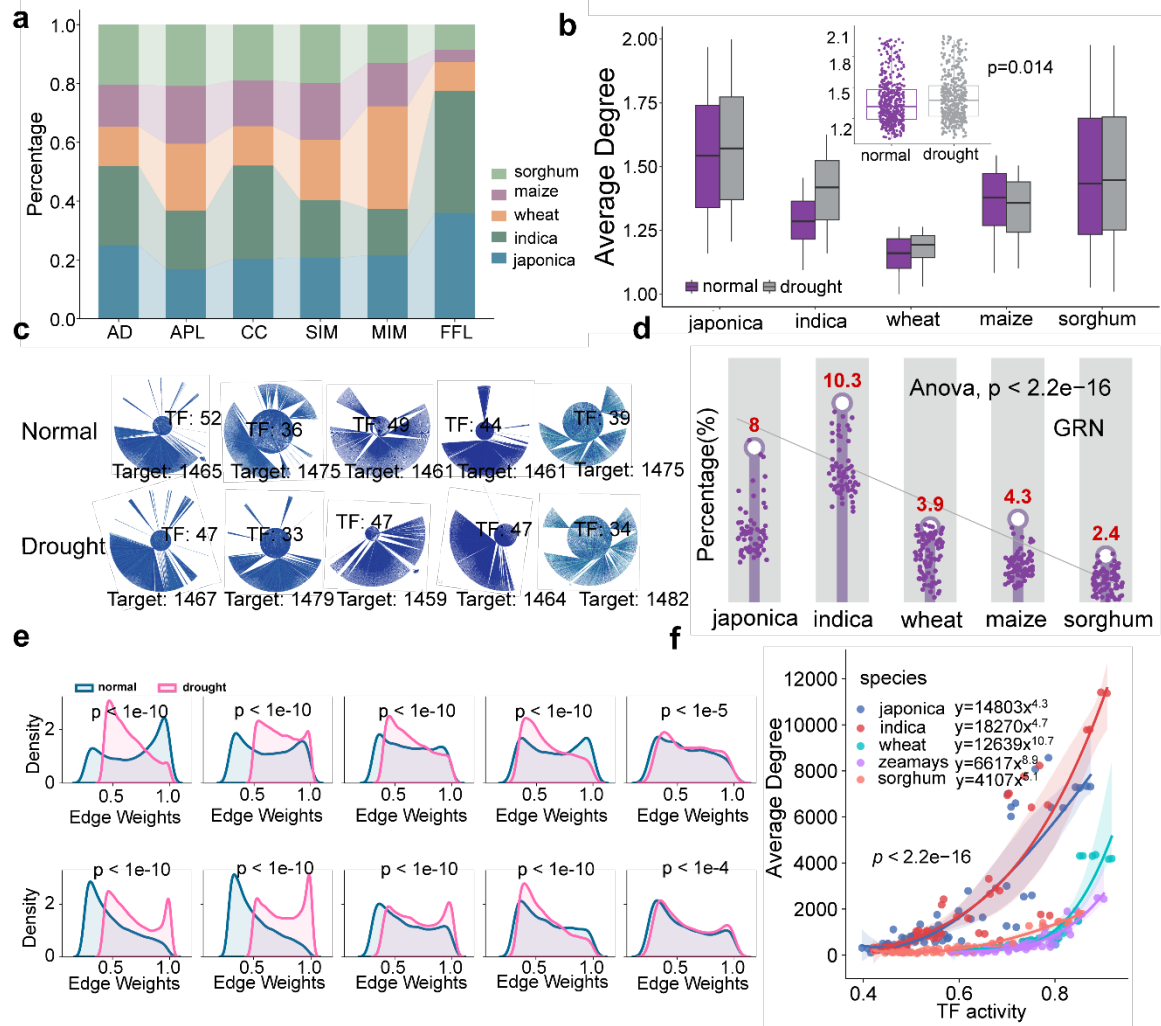


Fig. 3 Topological analysis of gene regulatory networks in different species under drought conditions. Figure (a) shows the proportions of topological properties. Figure (b) compares changes in average network degree between normal and drought conditions. A schematic of a subnetwork comprising the top 1,500 genes with the largest changes in degree centrality under drought stress is presented in Figure (c), where transcription factors occupy the inner circle and target genes occupy the outer circle, with japonica, indica, wheat, maize, and sorghum shown from left to right. Figure (d) shows the percentage distribution of absolute changes in average degree in the co-regulatory network across species between normal and drought conditions. Red numbers indicate the maximum absolute percentage change in average degree between normal and drought conditions for each species. The data points represent the absolute percentage change in the average degree of networks with different numbers of genes under normal and drought conditions. Figure (e) examines changes in transcription factor family activity between normal and drought conditions, in which blue represents the normal condition and red represents the drought condition. The top row represents the bZIP transcription factor family, and the bottom row represents the MYB transcription factor family. From left to right, the species are japonica rice, indica rice, wheat, maize, and sorghum. Figure (f) explores the relationship between the average degree and transcription factor family activity, which corresponds to edge strength in

these networks, across species. This comprehensive analysis links topological properties to drought-induced regulatory changes.

4)Lines 195-196: “the percentage distribution of changes in average degree in the co-regulatory network across species”. What do the changes in average degree mean? Changes in average degree when comparing what and what?

Response:

Thank you for your valuable comments. In a network, the degree is the number of edges (regulatory interactions) connected to each node (gene/transcription factor), and the average degree is calculated as the sum of the degrees of all nodes divided by the total number of nodes. In a gene regulatory network, it reflects how many regulatory relationships each gene participates in on average. A higher average degree means that each gene is, on average, connected to more genes, indicating denser regulatory interactions and a more complex network. A lower average degree means fewer regulatory connections, resulting in a sparser network with a simpler structure.

When comparing normal and drought conditions, an increase in average degree may indicate that more regulatory edges are activated or newly established, making the network under that condition tighter and more complex, with stronger information transfer or coordinated regulation among genes. A decrease in average degree may suggest that the network becomes sparser and more dispersed under that condition, with weakened overall coordinated regulation. In our case, the change in average degree reflects the differences between the networks constructed under normal and drought conditions.

To make the description clearer, we have revised our wording here.

Line 256-258:

“Figure (d) shows the percentage distribution of absolute changes in average degree in the co-regulatory network across species between normal and drought conditions.”

5)What do the data points in Fig. 2d mean? A single comparison between drought and normal conditions? Should be specified.

Response:

Thank you for your valuable comments. The data points in Fig. 2d represent the absolute percentage change in the average degree of networks with different numbers of genes under normal and drought conditions. To make the description clearer, we have added an explanatory note to the figure legend.

Line 259-261:

“The data points represent the absolute percentage change in the average degree of networks with different numbers of genes under normal and drought conditions.”

6)Line 201: TF activity in what aspects? Expression or what? Should be specified as

well. It is suggested that to move the sentence “transcription factor activity, which corresponds to edge strength in these networks 46,” in the Discussion section to here when the TF activity was first mentioned.

Response :

Thank you for your valuable comments. TF activity here is determined by integrating gene expression data and regulatory network information to infer the net regulatory effect intensity of transcription factors on their target genes during normal and drought conditions. For better clarity, we have moved this explanation to the figure caption.

Line 266-268:

“Figure (f) explores the relationship between the average degree and transcription factor family activity, which corresponds to edge strength in these networks, across species.”

Line 203-210: Fig. S10 shows the venn analysis of gene families across different species, rather than the diagram for the authors’ graph network algorithm, as would be expected after reading these sentences. In addition, the name of the authors’ tool should be spelled out here and in Fig. S4.

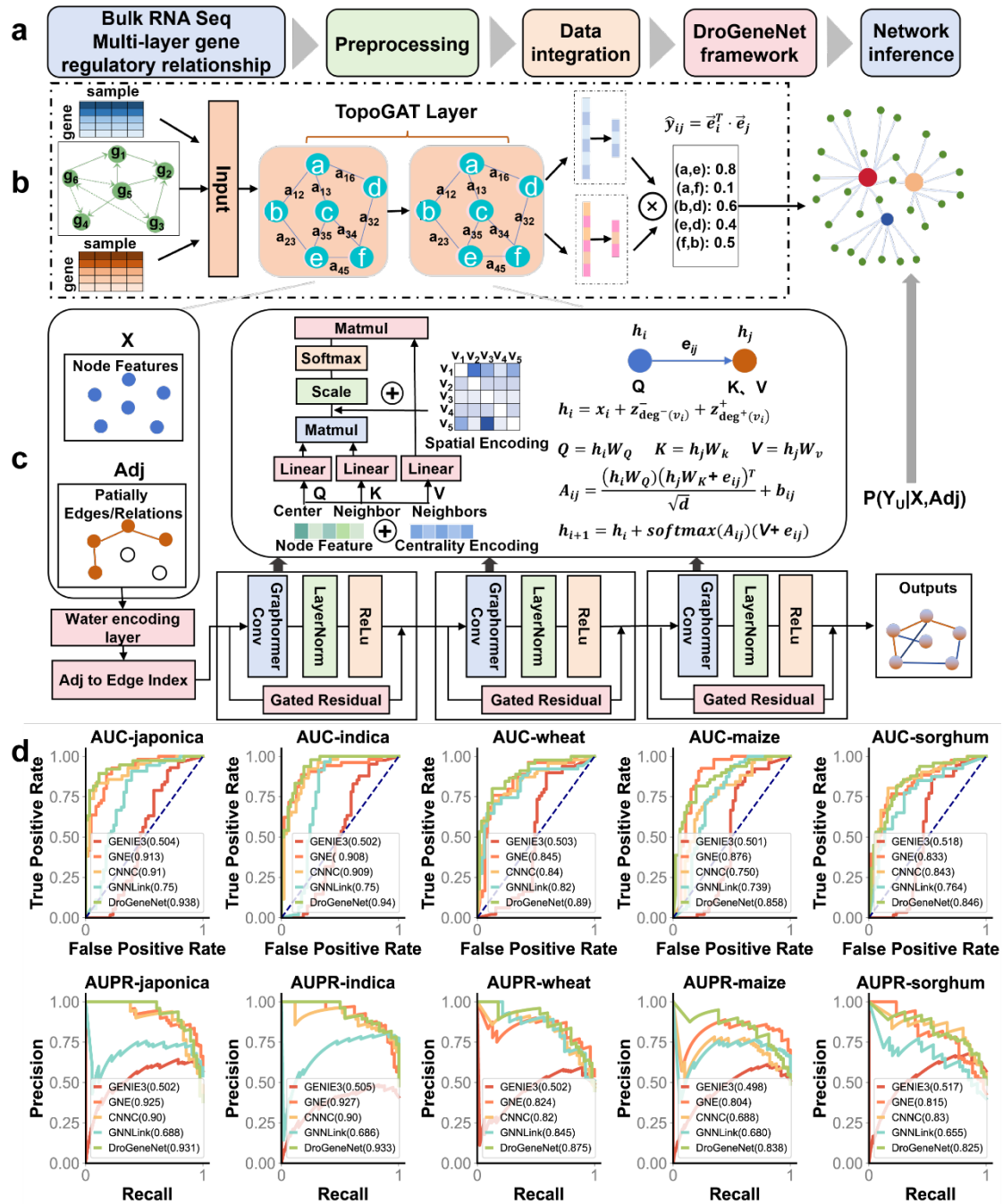
Response :

Thank you for your valuable comments. The figure numbering in our previous version was incorrect, and we have now revised it. In addition, the algorithm framework flowchart previously shown as Fig. S4 has been moved to the main figures as Fig. 2 in the Results section. We have also added the name of our algorithm in both the text and the figure.

Line 223-230:

“Aiming to unravel the key determinants of GRN topology in drought resistance, we developed a novel graph network algorithm DroGeneNet for constructing highly accurate GRNs under drought-responsive conditions (Fig. 2). Our innovative approach pioneers the construction of GRNs under varying water conditions by incorporating a water-encoding layer. This layer allows the embedding of gene expression matrix data from different water conditions into the GRN inference process, enabling the reconstruction of gene regulatory networks that reflect gene expression patterns under specific experimental conditions (Fig. 2, Methods S1, Methods S2, Methods S3).”

Figure 2:



Line 216: what do attention scores mean? Should explain all these computational terms to readers with more broader background, especially those in biological filed.

Response:

Thank you for your valuable comments. In the field of deep learning and artificial intelligence, attention scores are generally treated as the importance weights of the model for other positions relative to the current position. In graph networks, they correspond to edge-level weights learned by the model and are usually interpreted in biology as the regulatory importance or contribution of one gene (or TF) to another gene. To make this clearer, we have added an additional explanation of the attention

scores. In addition, we have checked the other computational metrics in the manuscript and added brief corresponding descriptions.

Line 239-242:

“We further validated the reliability of our model by visualizing the attention scores from the intermediate layers, which represent the relative regulatory influence of each gene interaction and reflect the context-specific control logic learned by the model.”

Line 359-362:

“In the UMAP-embedded expression space, the Calinski-Harabasz (CH) index increased substantially (Fig. 4e), indicating fewer and more compact clusters and thus a more coherent hormone-related transcriptional response to drought across these species.”

Line 273-284:

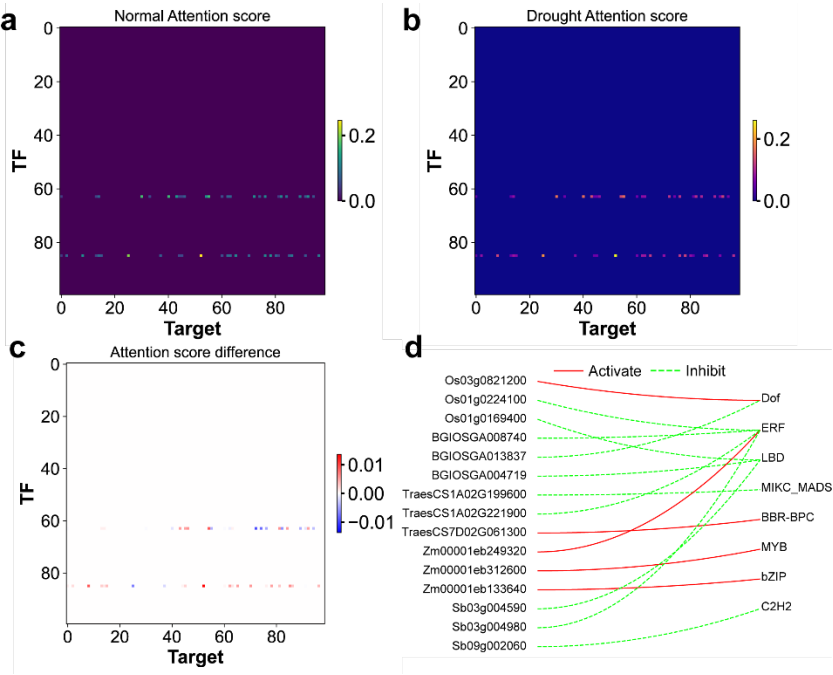
“The initial topological characteristics revealed that rice exhibits the highest average degree (AD, where larger values indicate denser overall regulatory connectivity) among species, while wheat has the lowest. The clustering coefficient (CC, with higher values reflecting tighter local gene network connectivity) distribution aligns with average degree trends across species. In contrast, average path length (APL, for which higher values correspond to longer average shortest regulatory distances between genes) and single-input motifs (SIM, where an increased value denotes a greater prevalence of simple TF-target regulation) display consistent distributions with minimal variability. Wheat has the highest proportion of multi-input motifs (MIM, where larger values represent more complex multi-TF regulatory convergence), while rice demonstrates the largest proportion of feedback loop motifs (FFL, with higher values indicating more abundant regulatory feedback circuits).”

Fig. S5: what the differences between “normal” subfigure and “drought” subfigure in Fig. S5a we are supposed to observe?

Response:

Thank you for your valuable comments. What we aim to demonstrate is the variation in attention scores between normal and drought conditions. The subnetworks we initially selected show only minor differences between these two conditions, and the range of the color scale in the legend was relatively large, so the changes were not very apparent. We therefore re-selected a color bar to make these differences more pronounced and the visualization clearer. We also added a new figure showing the difference in attention scores between drought and normal conditions, which directly reflects their changes. Although the regulatory relationships may not undergo substantial changes, we should be able to observe shifts in the model's attention scores between normal and drought conditions, specifically reflected in the color changes of

the highlighted areas (indicating regulatory relationships). Additionally, we have included this data in the supplementary table, Fig S4 attention score.



“Fig. S4 The core regulatory genes explained by the attention scores of the model. Figure (a) shows the network attention scores under normal conditions. Figure (b) shows the network attention scores under drought conditions. Higher x-axis and y-axis values indicate greater confidence in regulatory interactions. Figure (c) shows the difference in network attention scores between drought and normal conditions. Figure (d) highlights core genes and their families, identified based on topological degree centrality, with colors indicating activation or repression under drought stress.”

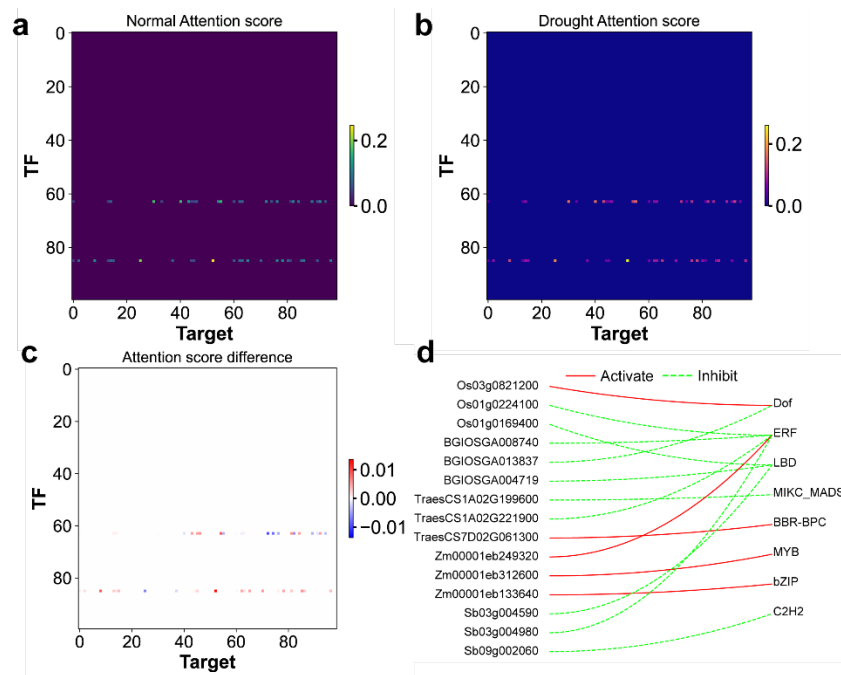
Supplementary table “Fig S4 attention score” (simplified version)

control					
0	1	2	...	98	99
...	...	...	...	...	...
0	0.043843713	0	...	0	0
drought					
0	1	2	...	98	99
...	...	...	...	...	...
0	0.043550413	0	...	...	...
difference					
0	1	2	...	98	99
...	...	...	...	...	...
0.000815958	0	0.00191908	...	0	0
...	...	...	...	...	...

Line 220: the authors wanted to cite Fig. S5, rather than Fig. S4, right? The same for line 223 as well.

Response:

Thank you for your valuable comments. The previous figure numbering was inaccurately labeled, and we have now revised it accordingly. Due to the algorithm framework diagram being presented as a main figure earlier, the numbering of figures in the appendix has been adjusted accordingly.



“Fig. S4 The core regulatory genes explained by the attention scores of the model. Figure (a) shows the network attention scores under normal conditions. Figure (b) shows the network attention scores under drought conditions. Higher x-axis and y-axis values indicate greater confidence in regulatory interactions. Figure (c) shows the difference in network attention scores between drought and normal conditions. Figure (d) highlights core genes and their families, identified based on topological degree centrality, with colors indicating activation or repression under drought stress.”

Line 242-248:

“We observed that a single transcription factor regulates multiple target genes (Fig. S4). Furthermore, we discovered that genes with higher attention scores are more extensively connected, underscoring their central roles within the network (Table S3, Fig. S4). Core genes identified in GRNs, primarily transcription factors, exhibit the highest centrality. Notably, we found that the ERF (ethylene response factor) family consistently plays a significant role in transcriptional regulation across species under both conditions (Fig. S4).”

Line 226-231: what do the higher AD, CC, APL, SIM, MIM, and FFL mean? Should be explained somehow. Such as: higher AD and CC indicate higher tightness of the network?

Response:

Thank you for your valuable comments. We have accordingly added descriptions of the biological implications for these metrics.

Line 273-284:

“The initial topological characteristics revealed that rice exhibits the highest average degree (AD, where larger values indicate denser overall regulatory connectivity) among species, while wheat has the lowest. The clustering coefficient (CC, with higher values reflecting tighter local gene network connectivity) distribution aligns with average degree trends across species. In contrast, average path length (APL, for which higher values correspond to longer average shortest regulatory distances between genes) and single-input motifs (SIM, where an increased value denotes a greater prevalence of simple TF-target regulation) display consistent distributions with minimal variability. Wheat has the highest proportion of multi-input motifs (MIM, where larger values represent more complex multi-TF regulatory convergence), while rice demonstrates the largest proportion of feedback loop motifs (FFL, with higher values indicating more abundant regulatory feedback circuits).”

Line 233: “For each species, 80 networks of different sizes were generated.” What are these 80 networks? What is the motivation to generate networks of different sizes? What are these different sizes?

Response:

Thank you for your valuable comments. These 80 networks were generated to enhance the robustness of our conclusions and to minimize the impact of randomness from any single network. Specifically, we constructed networks of different sizes, using a series of values from 100 to 3000 with an interval of 50. For greater clarity, we added a supplementary table listing all 80 networks for each species. In addition, we revised the Results section to include a more detailed description of the data processing procedures, making the analysis clearer.

Line 286-289:

“For each species and condition, we generated networks of different sizes, ranging from 100 to 3,000 genes in steps of 50, to increase robustness and reduce the influence of random variation in any single network.”

Supplementary table Fig3b 80 networks (simplified network):

Species	TF	Target	Size
indica	BGIOSGA030019	BGIOSGA023571	2230

indica	BGIOGA004955	BGIOGA018782	2230
indica	BGIOGA008740	BGIOGA027553	2230
...	...	...	...
wheat	TraesCS1A02G221900	TraesCS4A02G267600	2289
wheat	TraesCS1A02G221900	TraesCS5D02G405000	2289
wheat	TraesCS1A02G221900	TraesCS6B02G371500	2289

Line 245-246: what are the “average degree”?

Response:

Thank you for your valuable comments. The average degree refers to the mean degree centrality of all genes within each gene family in the network. The detailed raw data are provided in Supplementary Fig. 3f “TFfamily-activity-degree.”

Line 300-302:

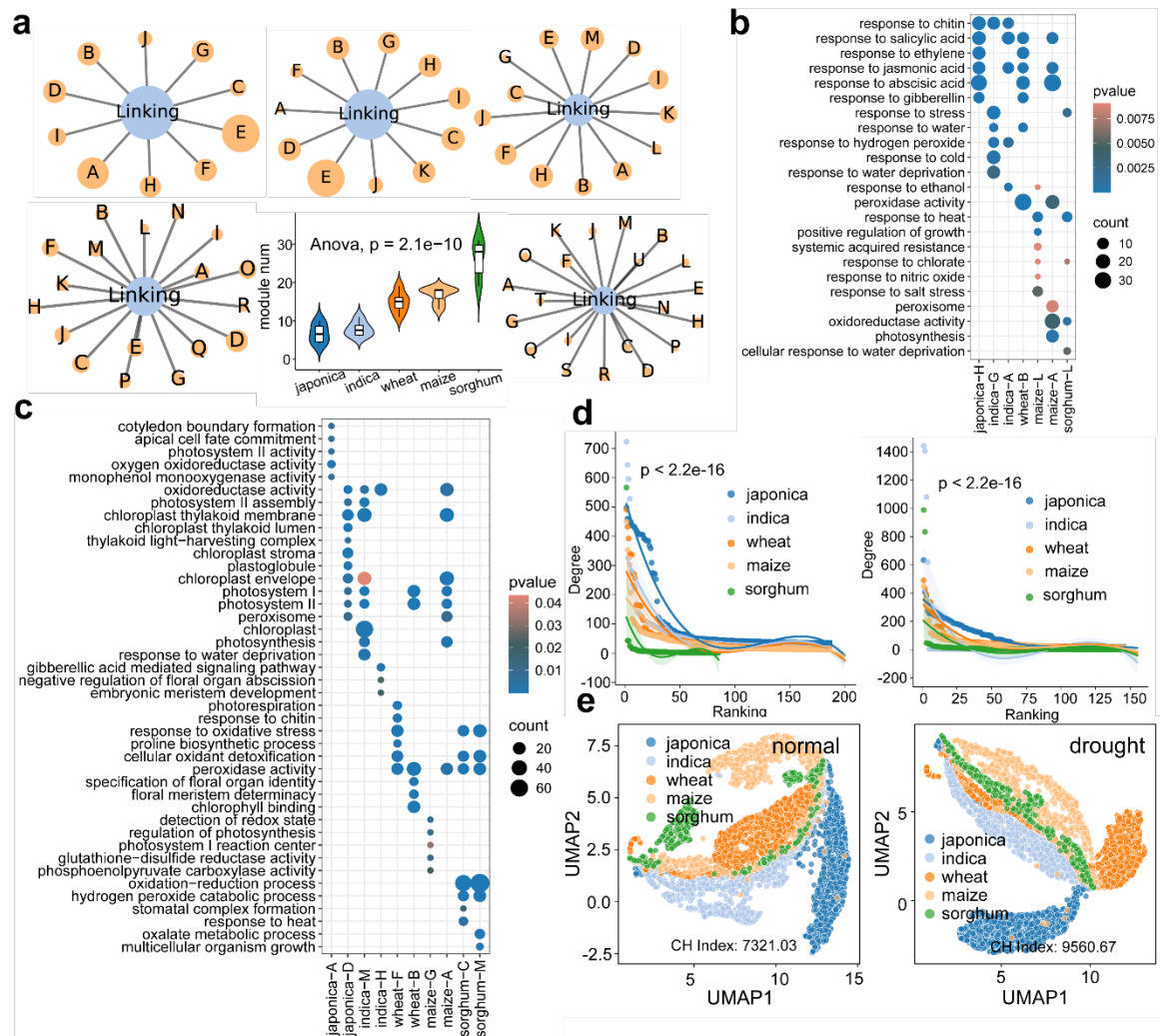
“Additionally, we identified a previously unreported power law relationship between TF activity and average degree across gene families ($p < 2.2e-16$).”

The top-left sub-figure in Fig. 3a was partially covered by the top-middle one.

Response:

Thank you for your valuable comments. We have already moved this sub-figure so that it won’t be blocked.

Line 311-322:



“Fig. 4 The construction and analysis of a super gene regulatory network, gene modules, and process crosstalk within subnetworks revealed key underlying characteristics. The first row of Figure (a) shows super gene regulatory networks for japonica, indica, and wheat, while the second row depicts those for maize and sorghum, along with the number of gene modules identified under drought conditions. The size of the circles reflects the relative quantity of genes. GO annotations for gene modules participating in similar biological processes are shown in Figure (b), while Figure (c) highlights GO annotations for modules exhibiting biological process crosstalk. And Figure (d) illustrates degree distributions of modules related to hormone (left panel) and photosynthesis (right panel) response. Figure (e) provides UMAP plots showing gene expression patterns in hormone-responsive modules under normal and drought conditions.”

Lines 262-263: Figure (d) illustrates degree distributions of modules related to hormone (left panel?) and photosynthesis (right panel?) response.

Response:

Thank you for your valuable comments. On the left are the degree distributions of modules related to hormone response, and on the right are the degree distributions of

modules related to photosynthesis. The corresponding descriptions have already been revised.

Line 319-320:

“And Figure (d) illustrates degree distributions of modules related to hormone (left panel) and photosynthesis (right panel) response.”

Line 285-286: According to Fig. 3b, only the GO term “response to ethanol” was shared between indica-A and maize-L modules.

Response:

Thank you for your valuable comments. Our previous description was inaccurate. We have revised this sentence to make the expression clearer.

Line 345-347:

“Abiotic stress responses, including those to water (module G of indica and module B of wheat), hydrogen peroxide (module G of indica and module A of indica), and heat (module L of maize and module L of sorghum), are co-regulated in modules (Fig. 4b).”

Line 288: there was only one GO term “oxidoreductase activity” shared between maize-A and sorghum-L modules.

Response:

Thank you for your valuable comments. Your observation is correct. However, what we intended to express with this sentence is the crosstalk between different biological processes within the modules. To make this clearer, we have revised the sentence.

Line 349-353:

“For example, module G of indica, module B of wheat, and module L of sorghum exhibit interactions between abiotic stress-related processes and oxidoreductase-related processes (Fig. 4b). Similarly, cross-talk between oxidoreductase-related processes and photosynthesis-related processes is observed within module A of japonica, module M of indica, module F of wheat, and module A of maize (Fig. 4c).”

Line 290-292: there were no shared GO terms between Japonica-A and other modules.

Response:

Thank you for your valuable comments. Your observation is correct. However, what we intended to express with this sentence is the crosstalk between different biological processes within the modules. To make this clearer, we have revised the sentence.

Line 351-353:

“Similarly, cross-talk between oxidoreductase-related processes and photosynthesis-related processes is observed within module A of japonica, module M of indica, module F of wheat, and module A of maize (Fig. 4c).”

Line 297: what does CH index stands for? And what does it mean?

Response:

Thank you for your valuable comments. The Calinski-Harabasz (CH) index reflects the ratio between between-cluster dispersion and within-cluster compactness; a higher value indicates a clearer clustering structure. In the UMAP-embedded expression space, we clustered the samples based on their transcript expression patterns and used the CH index to quantify cluster compactness and separation. Under drought stress, the CH index increased markedly. Together with the UMAP visualization (Fig. 3e), this increase in the CH index supports the visual impression that samples from different species under drought stress are grouped into fewer and more unified clusters. This quantitatively demonstrates that their gene expression patterns (transcriptomic responses) shift from being distinct to becoming more similar, i.e., they exhibit convergence. The definition of this index and its exact formula are provided in Methods S5 of the Supporting Information. To make this clearer, we have revised the sentence accordingly.

Line 358-362:

“However, under drought stress, we observed a pronounced convergence of gene expression patterns across species. In the UMAP-embedded expression space, the Calinski-Harabasz (CH) index increased substantially (Fig. 4e), indicating fewer and more compact clusters and thus a more coherent hormone-related transcriptional response to drought across these species.”

For Fig. 4a, there were two rice species. How did the author deal with the numbers in two rices in the bar plot? When comparing two groups, there would be only three categories: A unique, B unique and overlap between A and B. So what are the $1 - (A \cup B)$ for rice-wheat and maize-sorghum comparisons?

Response:

Thank you for your valuable comments. We rechecked our code and the data in Supplementary Table “Fig 4a gene family venn”. There are two rice subspecies, and we merged them into a single set (taking unique gene families) to compare with wheat. As you pointed out, when comparing two groups there are only three categories: A-unique, B-unique, and the overlap between A and B. In our analysis, we also considered the background set of gene families across all five species; here, “1” denotes this entire background gene family set. “ $1 - (A \cup B)$ ” represents the number of gene families that belong to neither A nor B.

We have added additional explanations to the legend to make this clearer.

Line 372-378:

“There are two rice subspecies, and we merged them into a single set by taking unique gene families. A and B denote gene families that are uniquely drought-responsive in

each of the two species being compared, and $A \cap B$ denotes the gene families that are drought-responsive in both species. “1” represents the complete background set of gene families across all five species, and $1 - (A \cup B)$ corresponds to the background gene families that are not drought-responsive in either of the two species in that comparison.”

Lines 325-328: Figure (e) provides GO annotations for drought-responsive genes within the TCP (Teosinte branched1/Cycloidea/Proliferating cell factor)-PP2C (right panel?) and ERF (Ethylene-responsive transcription factor)-2OGD (left panel?) regulatory gene families, respectively. The ordering of the two sub-figures should be consistent with the ordering in the legend.

Response:

Thank you for your valuable comments. We have rephrased the legend to match the order in the image and enhance clarity.

Line 396-400:

“Figure (e) provides GO annotations for drought-responsive genes within the ERF (Ethylene-responsive transcription factor)-2OGD (left panel) and TCP (Teosinte branched1/Cycloidea/Proliferating cell factor)-PP2C (right panel) regulatory gene families, respectively.”

Line 325: provides, instead of “provide”

Response:

Thank you for your valuable comments. We have implemented the corresponding modifications.

Line 396-400:

“Figure (e) provides GO annotations for drought-responsive genes within the ERF (Ethylene-responsive transcription factor)-2OGD (left panel) and TCP (Teosinte branched1/Cycloidea/Proliferating cell factor)-PP2C (right panel) regulatory gene families, respectively.”

Line 339: “while maize and sorghum were more conserved” how this conclusion was inferred from Fig. 4a?

Response:

Thank you for your valuable comments. In our analysis, Fig. 4a is based on chi-square tests comparing differences in the composition of drought-responsive gene families among different species (or species groups). The results show that the difference between C₃ and C₄ plants is highly significant ($\chi^2 = 341.98$, $p = 2.36 \times 10^{-76}$), and there is also a significant difference between rice and wheat ($\chi^2 = 13.75$, $p = 2.08 \times 10^{-4}$). In contrast, for maize and sorghum the test yields $\chi^2 = 1.70$, $p = 0.19$, indicating no statistically significant difference at commonly used significance levels.

Here, “conserved” refers to the situation where the composition of gene families remains highly similar between species, implying relatively limited evolutionary change. Based on our statistical results, the lack of a statistically significant difference ($p > 0.05$) suggests that these two species show a higher degree of similarity in their drought-responsive gene family composition. Compared with the clear differences observed between C₃ vs. C₄ and rice vs. wheat, maize and sorghum are more similar in this respect, and we therefore described them as “more conserved.” This reflects that, for these key gene families, their composition has likely experienced relatively fewer perturbations or changes during evolution.

In the original manuscript, we used the phrase “maize and sorghum were more conserved,” which could indeed be interpreted as a strong conclusion about “conservation” drawn directly and strictly from Fig. 4a/the chi-square tests, which is not fully rigorous from a statistical perspective. A more accurate statement is that, under the conditions of our analysis, no significant difference was detected in the composition of drought-responsive gene families between maize and sorghum. In the revised manuscript, we have therefore changed the wording to “no statistically significant difference was detected between maize and sorghum ($\chi^2 = 1.70$, $p = 0.19$),” and we have clarified the definitions of A, B, $A \cap B$ and the background set in the figure legend to avoid misunderstanding.

Line 412-414:

“Rice and wheat also showed significant differences ($\chi^2 = 13.75$, $p = 2.08 \times 10^{-4}$), while no statistically significant difference was detected between maize and sorghum ($\chi^2 = 1.70$, $p = 0.19$) (Fig. 5a).”

Line 373-378:

“A and B denote gene families that are uniquely drought-responsive in each of the two species being compared, and $A \cap B$ denotes the gene families that are drought-responsive in both species. “1” represents the complete background set of gene families across all five species, and $1 - (A \cup B)$ corresponds to the background gene families that are not drought-responsive in either of the two species in that comparison.”

A detailed legend for Fig. 4d is needed. Datapoints: species or samples? Under normal or drought conditions? Values: TPM?

Response:

Thank you for your valuable comments. To clarify, the data points presented in Fig. 4d represent gene expression values for specific gene families from different species. For each species included in the analysis, we have data under both normal (control) and drought stress conditions. Therefore, each species contributes expression values from these two distinct conditions. The values plotted are FPKM (Fragments Per Kilobase of transcript per Million mapped reads), which is a normalized measure of gene expression. We have added a detailed legend for Fig. 4d.

Line 393-396:

“Each datapoint represents the expression of a given gene family in one species under a specific condition (control or drought). For each species, two groups of values are shown, corresponding to expression under normal (well-watered) and drought conditions. Gene expression levels are quantified as FPKM.”

Line 350: “the overall expression patterns remain consistent,” how was this inferred from Fig. 4c and Fig. S11?

Response:

Thank you for your valuable comments. UMAP is an unsupervised dimensionality reduction method that provides an intuitive visualization of the relative relationships among samples in a low-dimensional space. Our original intention was that, for the four conserved gene families PP2C, CYP450, Cyclin, and 2OGD, the expression profiles of different species under normal and drought conditions do not show clear separation in the UMAP space but instead display broadly similar clustering patterns; therefore, we infer that their expression patterns are relatively stable across species. In the revised manuscript, we have further clarified in the main text the specific visual evidence underlying this observed “consistency” to make the description more precise.

Line 424-426:

“While drought conditions induce some shifts, the overall expression patterns remain consistent as indicated by the similar clustering distribution, reflecting the conserved nature of these families (Fig. 5c, Fig. S10).”

Line 367-368: “Figure (b) illustrates the collinear positions of the 2OGD and PP2C gene families across species, revealing their genomic relationships.” Which sub-figure was for which gene family? Should be specified.

Response:

Thank you for your valuable comments. We have added corresponding annotations.

Line 443-444:

“Figure (b) illustrates the collinear positions of the 2OGD (left panel) and PP2C (right panel) gene families across species, revealing their genomic relationships.”

Line 368-370: “Figure (c) shows the distribution of gene contraction and expansion counts for the 2OGD and PP2C families based on clustering results.” Which sub-figure was for which number? Should be clear.

Response:

Thank you for your valuable comments. We have added corresponding annotations.

Line 444-446:

“Figure (c) shows the distribution of gene contraction and expansion counts for the 2OGD (left panel) and PP2C (right panel) families based on clustering results.”

Line 371: What is the definition of paralogs in this study? Reciprocal best matched gene pairs? Should be specified.

Response :

Thank you for your valuable comments. In our manuscript, we use the term “paralogous genes (paralogs)” in its standard evolutionary sense, i.e., homologous genes that originate from gene duplication events within the same lineage. Our previous version did not clearly explain how we operationally identified paralogs in this study. Here, our operational definition is that all genes from the same species that are assigned by OrthoFinder to the same orthogroup are considered intraspecific paralogs. In other words, whenever a species has multiple genes within the same orthogroup, we regard these genes as paralogs in that species. We have added a more detailed description of this definition in the revised manuscript.

Line 446-450:

“Figure (d) highlights the distribution of Ka/Ks values for both orthologs (one-to-one genes between species within the same orthogroup) and paralogs (multiple genes from the same species within the same orthogroup) in the 2OGD (left panel) and PP2C (right panel) families, providing insights into their divergence.”

Line 370-372: “Figure (d) highlights the distribution of Ka/Ks values both” missing a word “for” or “of” before “both”. Again, which sub-figure was for which gene family? Should be specified.

Response :

Thank you for your valuable comments. We have revised this sentence and clearly indicated the corresponding part of the figure in the legend.

Line 446-450:

“Figure (d) highlights the distribution of Ka/Ks values for both orthologs (one-to-one genes between species within the same orthogroup) and paralogs (multiple genes from the same species within the same orthogroup) in the 2OGD (left panel) and PP2C (right panel) families, providing insights into their divergence.”

Fig. 5f What do the data points stand for? Genes of 2OGD and PP2C families from all the five species? Or what?

Response :

Thank you for your valuable comments. These data points represent the Ka/Ks values and gene expression levels of drought-responsive differentially expressed genes in the 2OGD and PP2C families across all five species, as shown in the supplementary table “Fig 6f KaKs exp”. To improve clarity, we have provided a more detailed description in the figure legend.

Line 452-455:

“Figure (f) explores the correlation between Ka/Ks values and gene expression levels (log(FPKM)) for drought-responsive 2OGD (left panel) and PP2C (right panel) family genes across the five species, indicating evolutionary pressure and functional significance.”

Lines 401-402: “between recently duplicated paralogs within a species and orthologs between closely related species” this sentence should be mentioned when the authors talked about the Fig. 5d.

Response:

Thank you for your valuable comments. We have already added the relevant description when discussing Fig. 6d.

Line 474-478:

“We assessed selective pressures using Ka/Ks ratios for recently duplicated paralogs within each species and for orthologs between closely related species, and found that purifying selection predominates within and between species, as most ratios are below 1. However, higher Ka/Ks ratios exceeding 1 in PP2C genes of wheat and sorghum suggest instances of positive selection, potentially driving adaptation (Fig. 6d).”

Lines 436-439: “Wheat, diverging earliest (~56.7 MYA), shows the largest expansion (+11,250) and moderate contraction (-5,189), followed by maize, sorghum, and rice. These genomic adjustments reflect the dynamic adaptations, functional innovations, and evolutionary pressures that a species undergoes.” First, the authors didn’t solved the comment raised by reviewer 1. Second, “wheat, diverging earliest” is not correct.

Response:

Thank you for your valuable comments. We agree that the statement “Wheat, diverging earliest (~56.7 MYA)” in the current manuscript is inaccurate and potentially misleading. We have removed this wording and, in the revised version, we describe the phylogenetic relationships among species more cautiously, retaining only the estimated divergence times from databases such as TimeTree as reference information, without using phrases such as “diverging earliest” for broad generalization.

Regarding reviewer 1’s earlier comment that the large number of gene family expansions in wheat is related to polyploidy rather than directly to drought adaptation, we acknowledge that our initial response was not sufficiently clear. We fully agree that, as an allopolyploid species, the extensive gene family expansions and relatively low Ks values observed in wheat are largely driven by polyploidization events themselves and cannot be simply interpreted as evidence of accelerated molecular evolutionary rates or specific drought adaptation. In the revised manuscript, we have therefore weakened and clarified the relevant statements. The original sentence “these genomic adjustments reflect the dynamic adaptations, functional innovations, and evolutionary pressures that

a species undergoes” could imply an overly direct causal relationship. We have now replaced it with a more cautious and descriptive formulation.

Line 516-520:

“Wheat, having undergone allopolyploidization, shows the largest expansion (+11,250) and moderate contraction (-5,189), followed by maize, sorghum, and rice (Fig. 7a). These patterns largely reflect the consequences of allopolyploidy and subsequent genome fractionation in wheat, as well as lineage-specific duplications and losses in the other cereals.”

Lines 454-455: the figure should be Fig. S15, rather than Fig. S14

Response:

Thank you for your valuable comments. Since part of the content from the supplementary figures has been moved into the main figures, we have correspondingly updated and adjusted all supplementary figure numbers.

Line 534-536:

“Our correlation analysis showed a positive relationship between the DREB-RAD and ERF-Psb28 gene families, with ERF-Psb28 correlating more strongly ($R = 0.79$, $p = 0.028$) (Fig. S14).”

Line 464: there was no Fig. 6f in the figure.

Response:

Thank you for your valuable comments. We have made the corresponding revisions.

Line 543-545:

“SPL was enriched in anther development in rice, vascular bundle development in wheat, ethylene response in maize, and methyltransferase activity in sorghum (Fig. 7e).”

Lines 464-466: “This provides further evidence that C_3 crops combine growth regulation to escape drought, while C_4 crops are directly adapted to drought tolerance in the representative species we studied”. The authors should give more interpretation about why the above GO enrichment results would lead to this conclusion. Sentences of Lines 596-599 should be moved here to support this conclusion.

Response:

Thank you for your valuable comments. We have added more detailed interpretations explaining why the above GO enrichment results support this conclusion. In addition, the sentences previously found in Lines 596 – 599 have been integrated into the newly added interpretation.

Line 545-555:

“These GO terms in representative C₃ crops are predominantly associated with developmental and reproductive processes, such as the timing of floral meristem and anther formation and vascular bundle development, which are classic components of drought escape or avoidance strategies through growth adjustment and are mediated in part by SPL-PELP modules. By contrast, the terms enriched in representative C₄ crops are mainly related to photosynthetic processes, metabolic reprogramming and stress signaling, and are implemented through modules such as ERF-Psb28 and TCP-ASN, indicating more direct physiological and biochemical adaptations to drought. These results provide further evidence that representative C₃ crops combine growth regulation to escape drought, while representative C₄ crops are directly adapted to drought tolerance in the representative species we studied (Fig. 7e, Fig. S19c).”

The comments for Fig. 7 legends are the same as for Fig. 5.

Response:

Thank you for your valuable comments. We have already made the corresponding changes to the legend in Fig. 8 (previous Fig. 7).

Line 560-572:

“Figure (b) shows the syntenic positions of the PELP (left panel) and RbcS (right panel) gene families across various species, highlighting their conserved genomic locations. Figure (c) presents the expansion and contraction distributions of the DREB (left panel) and MADS (right panel) gene families based on gene family clustering. Figure (d) illustrates the Ka/Ks distribution for both orthologs (one-to-one genes between species within the same orthogroup) and paralogs (multiple genes from the same species within the same orthogroup) in the DREB (left panel), SPL (right panel) gene families. Figure (e) displays the distribution of synonymous substitution rates (Ks) values for DREB and SPL families within homolog clusters, offering insights into the selective pressures shaping their evolution. Figure (f) examines the relationship between Ka/Ks values and gene expression levels (log(FPKM)) for drought-responsive DREB and SPL family genes across the five species.”

Line 513-514: “The whole-genome duplication event in wheat was estimated to have occurred between 0.36 MYA (Ks = 0.01) and 177.8 MYA (Ks = 4.98)” What does this sentence serve for in this paragraph?

Response:

Thank you for your valuable comments. The original sentence, which presented only the time range of wheat whole-genome duplication (WGD), was indeed disconnected from the main logical flow of the paragraph and did not clearly demonstrate its role in the argument. Our initial intention was to use the timing of wheat WGD to support the inference about gene duplication time mentioned earlier in the paragraph. In the revised manuscript, we will rewrite or integrate this sentence so that it serves as supplementary information to that point.

Line 603-605:

“Similarly, due to allopolyploid nature and genome duplication estimated to have occurred between 0.36 MYA ($K_s = 0.01$) and 177.8 MYA ($K_s = 4.98$), wheat exhibited lower mean K_s values within paralogs.”

Lines 849-851: “Finally, the contribution of a regulatory pattern to drought response was determined by summing the scores of the two gene families it regulates.” What do “the two gene families” mean? And what does “it” mean in this sentence?

Response:

Thank you for your valuable comments. In this sentence, “it” refers to the regulatory pattern mentioned earlier. To make the expression clearer, we have revised this sentence.

Line 958-960:

“Finally, the contribution of a regulatory pattern to drought response was determined by summing the scores of the two gene families that the regulatory pattern contains.”

Dear Editors and Reviewers:

Thank you very much for your constructive comments and suggestions on our manuscript, "Conserved and divergent gene regulatory networks for crop drought resistance" (Manuscript ID: NCOMMS-25-43275A).

We greatly appreciate the time and effort that you and the reviewer have devoted to the evaluation of our work. Following the editorial requests and the remaining reviewer comments, we have carefully revised the manuscript accordingly. All changes made in the revised manuscript are detailed in our point-by-point responses for your convenience.

We are pleased to submit the final revised version of our manuscript and hope that it is now suitable for acceptance and publication in *Nature Communications*.

Yours sincerely,

Liangsheng Shi on behalf of all co-authors

Wuhan University

Response to Reviewer 5's comments

This manuscript aims to elements of drought-specific gene regulatory network (GRN) structure which are conserved and divergent between species with differing drought response. The authors apply a novel gene regulatory network (GRN) construction method to produce GRNs across a variety of C₃ and C₄ crops under drought conditions. By comparing these networks, the authors were able to identify network topological measures which differ between the drought-specific GRNs of more and less drought-tolerant crops, drought GRN module traits and properties, and conserved and differential connections between genes in drought conditions.

The authors' responses to Reviewer 4's comments appear satisfactory to me. They have introduced some details of their novel algorithm earlier in the text, provided summaries of data and terminology were applicable, improved the supplementary materials, included relevant details on the evaluation of their algorithm, and made appropriate corrections and clarifications where requested throughout the text. The additional analysis of soybean, potato, foxtail millet, and teff does a great deal to strengthen the argument that the observed changes in network structure are general to C₃ and C₄ plants and not a specific combination of crops.

I believe that this work is exciting and has scientific merit, but the manuscript would benefit from minor improvements to more clearly convey this.

I also agree with prior reviewers that the code access for this manuscript is well-organized.

Response:

We are very grateful to you for the careful evaluation of our revised manuscript and for the positive and encouraging comments. We sincerely appreciate your recognition of the scientific value of this study, the improvements made in response to Reviewer 4, the added comparative analyses across additional crop species, and the transparency of the code resources. We have carefully considered all of your remaining suggestions and have further revised the manuscript to improve the clarity and presentation of the work. Our detailed responses to each comment are provided below.

Primary Comments

Introduction: A brief mention in this section of some of the key findings of the study would greatly enhance the manuscript. For example, brief discussions of high-level outcomes such as those at line 724-728. This is to some extent a matter of stylistic preference but this addition would seem to be in line with other articles in this journal.

Response:

We thank the reviewer for this helpful suggestion. In response, we have revised the final paragraph of the Introduction to briefly summarize the major results and conclusions of the study, in line with the journal style. Specifically, we now highlight the contrasting drought-responsive network topologies of representative C₃ and C₄ crops, the conserved regulatory modules, and the lineage-specific modules that distinguish the rice/wheat and maize/sorghum lineages.

Line 131-139:

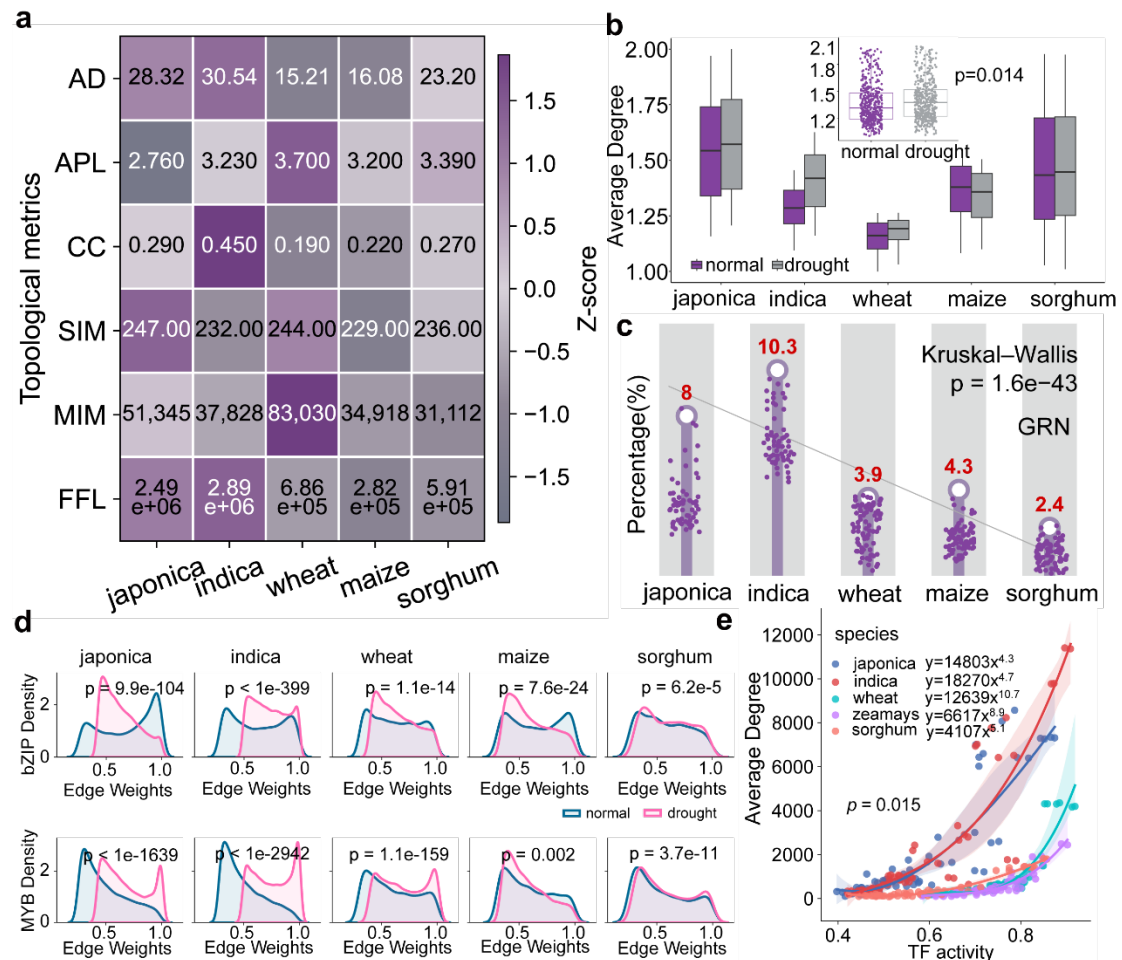
“By exploring drought resistance networks across multiple species, we show that drought adaptation is shaped by both conserved and divergent regulatory architectures. Representative C₃ crops exhibit more centralized and strongly rewired networks under drought, whereas representative C₄ crops maintain more distributed and stable networks. We further identify conserved core modules and lineage-specific regulatory programs that distinguish the rice/wheat and maize/sorghum lineages. Together, these findings indicate that drought resistance is constrained not only by individual genes, but also by the evolutionary topology of gene regulatory networks, with implications for the future improvement of drought-resistant crops.”

The figures and figure captions in this manuscript still require improvement. The authors have done well to incorporate the suggestions made by Reviewer 4 in this regard, but there are still numerous opportunities to improve the clarity of the results by using labels to reduce the reliance on figure captions and reducing the number of subfigures where possible. I have noted some individual improvements in my minor comments below, but I also encourage the authors to identify additional opportunities to simplify and clarify the figures.

Response:

We thank the reviewer for this constructive suggestion. In addition to addressing the specific figure-related comments listed below, we have re-examined all figures and figure legends across the manuscript to identify further opportunities for simplification and clarification. Where appropriate, we improved in-panel labeling to make the figures more self-explanatory, reduced reliance on the captions, and simplified or consolidated subpanels to improve readability while preserving the key information.

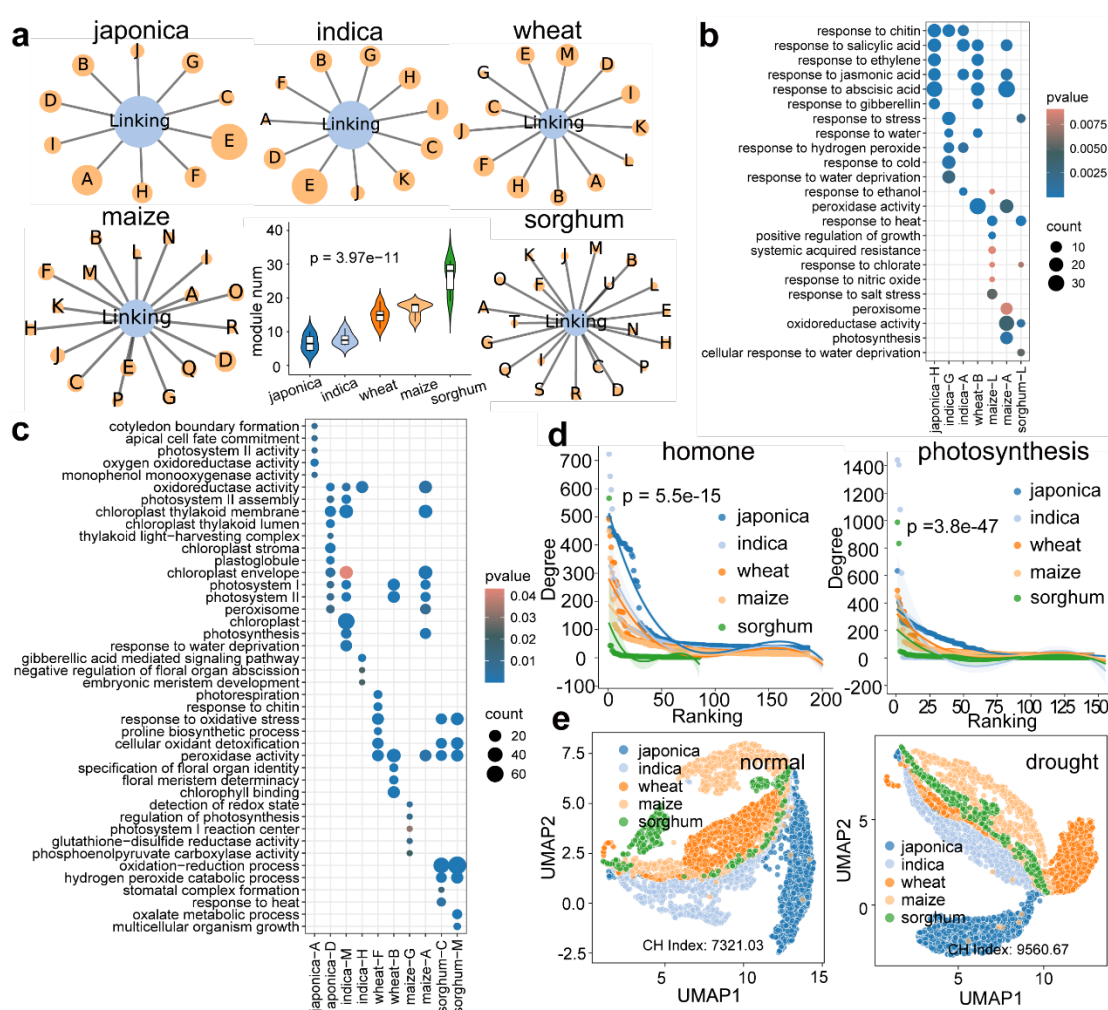
Figure 3:



“Fig. 3 Topological analysis of gene regulatory networks under drought across species. Figure (a) shows heatmap of species-specific initial topological properties. Cell colors indicate Z-scores standardized independently across species for each metric; labels show raw values. Figure (b) compares changes in average network degree between normal and drought conditions. Box plots show the median (centre line), the interquartile range (box bounds), and the whiskers extend to $1.5 \times \text{IQR}$; points represent individual observations. Significance was assessed using a two-sided Wilcoxon rank-sum test. Figure (c) shows the percentage distribution of absolute changes in average degree in the co-regulatory network across species between normal and drought conditions. Red numbers indicate the maximum absolute percentage change for each species. Points represent networks with different gene numbers. Statistical significance was assessed using the Kruskal–Wallis test ($H = 206.36$, exact $P = 1.61 \times 10^{-43}$). Figure (d) examines changes in transcription factor family activity between normal and drought conditions, in which blue represents the normal condition and red represents the drought condition. The top row represents the bZIP transcription factor family, and the bottom row represents the MYB transcription factor family. From left to right, the species are japonica rice, indica rice, wheat, maize, and sorghum. Statistical significance was evaluated using a two-sided Mann-Whitney U test with continuity correction. Source data list U, Z, exact/approximate P values, rank-biserial effect sizes, and BH/Bonferroni-adjusted P values. Figure (e) explores the relationship between the

average degree and transcription factor family activity, which corresponds to edge strength in these networks, across species. This comprehensive analysis links topological properties to drought-induced regulatory changes. Associations were analyzed separately for each species using two-sided Spearman rank correlation tests, with effect size reported as Spearman's ρ and 95% confidence intervals estimated by Fisher's z transformation. Differences among species were evaluated using a Fisher z-based heterogeneity Q test, and pairwise comparisons of correlation coefficients used two-sided Fisher z tests with Benjamini-Hochberg correction for multiple comparisons. Exact test statistics, degrees of freedom, confidence intervals, raw P values and adjusted P values are provided in Source Data."

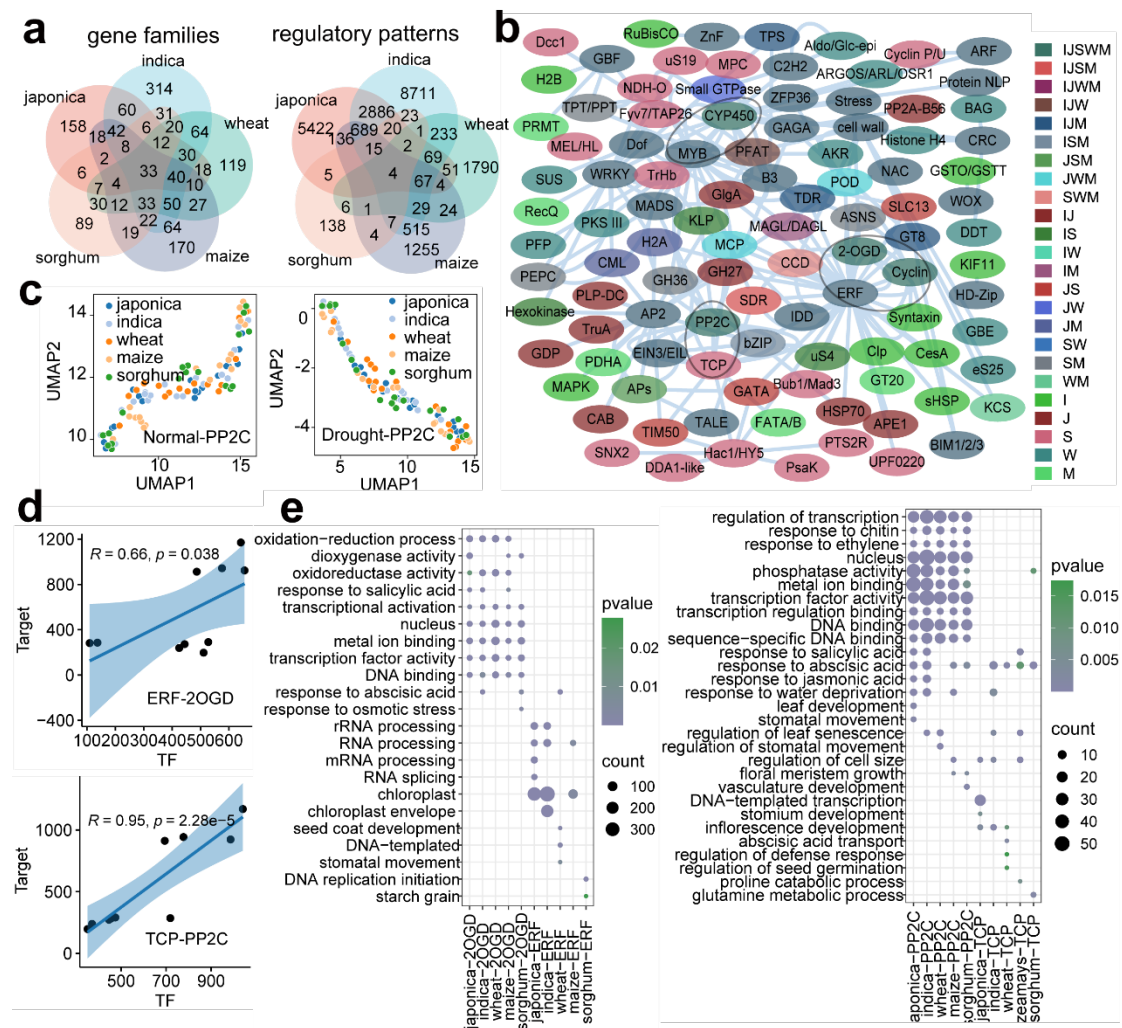
Figure 4:



“Fig. 4 The construction and analysis of a super gene regulatory network, gene modules, and process crosstalk within subnetworks revealed key underlying characteristics. The first row of Figure (a) shows super gene regulatory networks for japonica, indica, and wheat, while the second row depicts those for maize and sorghum, along with the number of gene modules identified under drought conditions. The size of the circles reflects the relative quantity of genes. Violin plots show the data distribution; centre lines indicate the median, boxes indicate the interquartile range (25th to 75th

percentiles), and whiskers extend to $1.5 \times \text{IQR}$. Differences in module number among species were assessed using a two-sided Friedman test with network-size thresholds treated as paired blocks. Pairwise comparisons were performed using two-sided exact Wilcoxon signed-rank tests with Benjamini-Hochberg correction for multiple comparisons. Exact P values and effect sizes are provided in Source Data. GO annotations for gene modules participating in similar biological processes are shown in Figure (b), while Figure (c) highlights GO annotations for modules exhibiting biological process crosstalk. Enrichment significance was assessed using Fisher's exact test, and the adjusted P values shown were corrected for multiple comparisons using the Benjamini-Hochberg method. And Figure (d) illustrates degree distributions of modules related to hormone (left panel) and photosynthesis (right panel) response. Species-specific degree distributions were fitted using a power-law model after log-log transformation. For each species, the exponent b was tested against 0 using a two-sided t-test from the log-log linear regression; effect sizes are reported as the fitted power-law exponent b , together with 95% confidence intervals and R^2 values. Differences among species-specific exponents were assessed using a Cochran Q-like heterogeneity test, followed by two-sided z-tests for pairwise comparisons with Benjamini-Hochberg correction for multiple comparisons. Detailed values are provided in Supplementary Table 4 and the Source Data. Figure (e) provides UMAP plots showing gene expression patterns in hormone-responsive modules under normal and drought conditions."

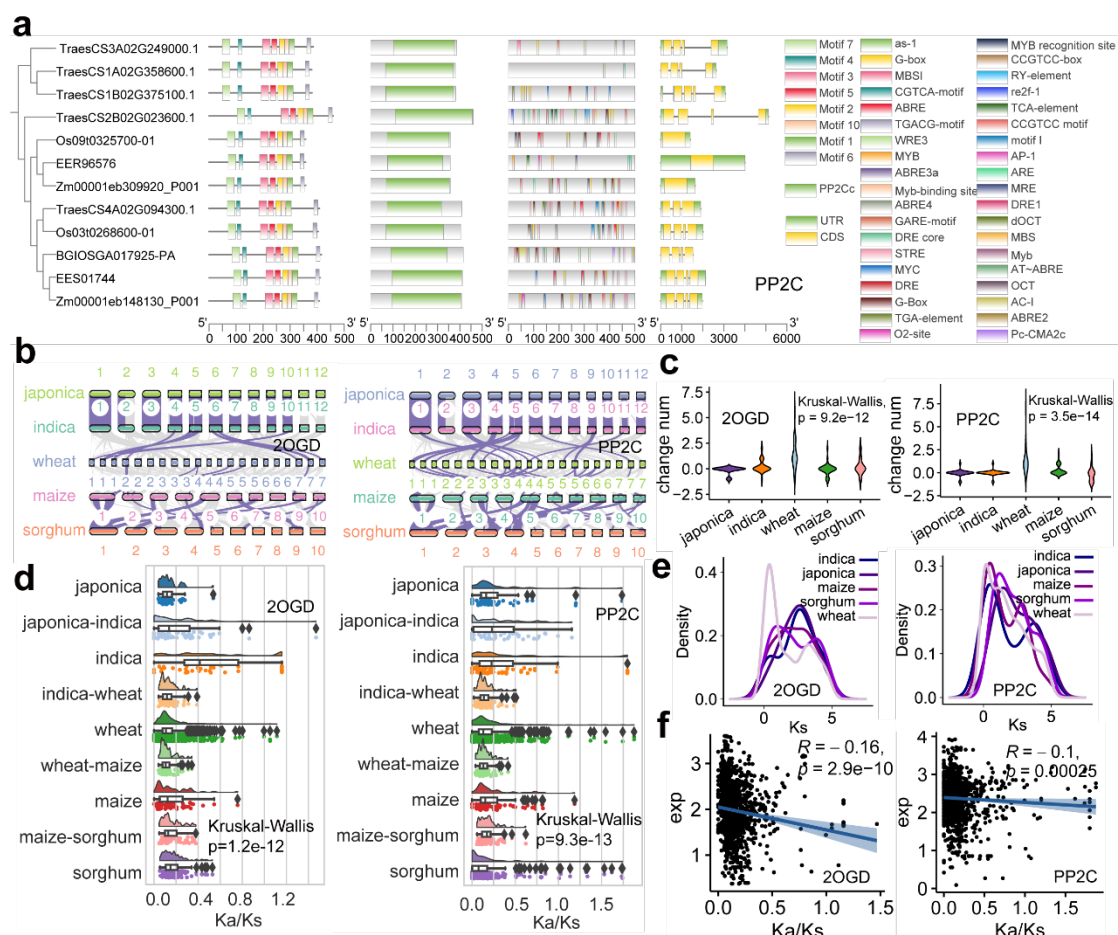
Figure 5:



“Fig. 5 The conserved drought-regulation mechanism in gene regulatory networks across different species. Figure (a) includes two subfigures: the first Venn diagram represents drought-responsive gene families across species, and the second illustrates regulatory patterns among these families. The corresponding chi-squared test results are provided in Supplementary Fig. 11. Figure (b) displays a partial view of the regulatory patterns of drought-responsive gene families, with all gene size families detailed in supplementary data. Different colors indicate drought responses in different species. IJSWM represents indica-japonica-sorghum-wheat-maize, IJWM represents indica-japonica-wheat-maize, IJSM represents indica-japonica-sorghum-maize, IJM represents indica-japonica-maize, ISM represents indica-sorghum-maize, JSM represents japonica-sorghum-maize, JWM represents japonica-wheat-maize, IJW represents indica-japonica-wheat, SWM represents sorghum-wheat-maize, IJ represents indica-japonica, IM represents indica-maize, IW represents indica-wheat, SW represents sorghum-wheat, SM represents sorghum-maize, JM represents japonica-maize, JW represents japonica-wheat, JS represents japonica-sorghum, IS represents indica-sorghum, WM represents wheat-maize, J represents japonica, I represents indica, W represents wheat, M represents maize, and S represents sorghum. Figure (c) shows UMAP plots of gene expression for conserved drought-responsive gene families PP2C (Protein Phosphatase 2C) under normal and drought conditions. Figure (d) presents a

correlation analysis of gene expression in conserved drought-regulatory mechanisms among related gene families. Each datapoint represents the expression of a given gene family in one species under a specific condition (control or drought). For each species, two groups of values are shown, corresponding to expression under normal (well-watered) and drought conditions. Gene expression levels are quantified as FPKM. Lines indicate the fitted linear regression, and shaded bands indicate the 95% confidence interval. Spearman correlation coefficients and two-sided p values are shown. Effect sizes are reported as Spearman's rho (ρ), with 95% confidence intervals estimated by Fisher's z transformation. Exact test statistics, confidence intervals, and raw and adjusted P values are provided in the Source Data. Figure (e) provides GO annotations for drought-responsive genes within the ERF (Ethylene-responsive transcription factor)-2OGD (left panel) and TCP (Teosinte branched1/Cycloidea/Proliferating cell factor)-PP2C (right panel) regulatory gene families, respectively. Enrichment significance was assessed using Fisher's exact test, and the adjusted P values shown were corrected for multiple comparisons using the Benjamini-Hochberg method."

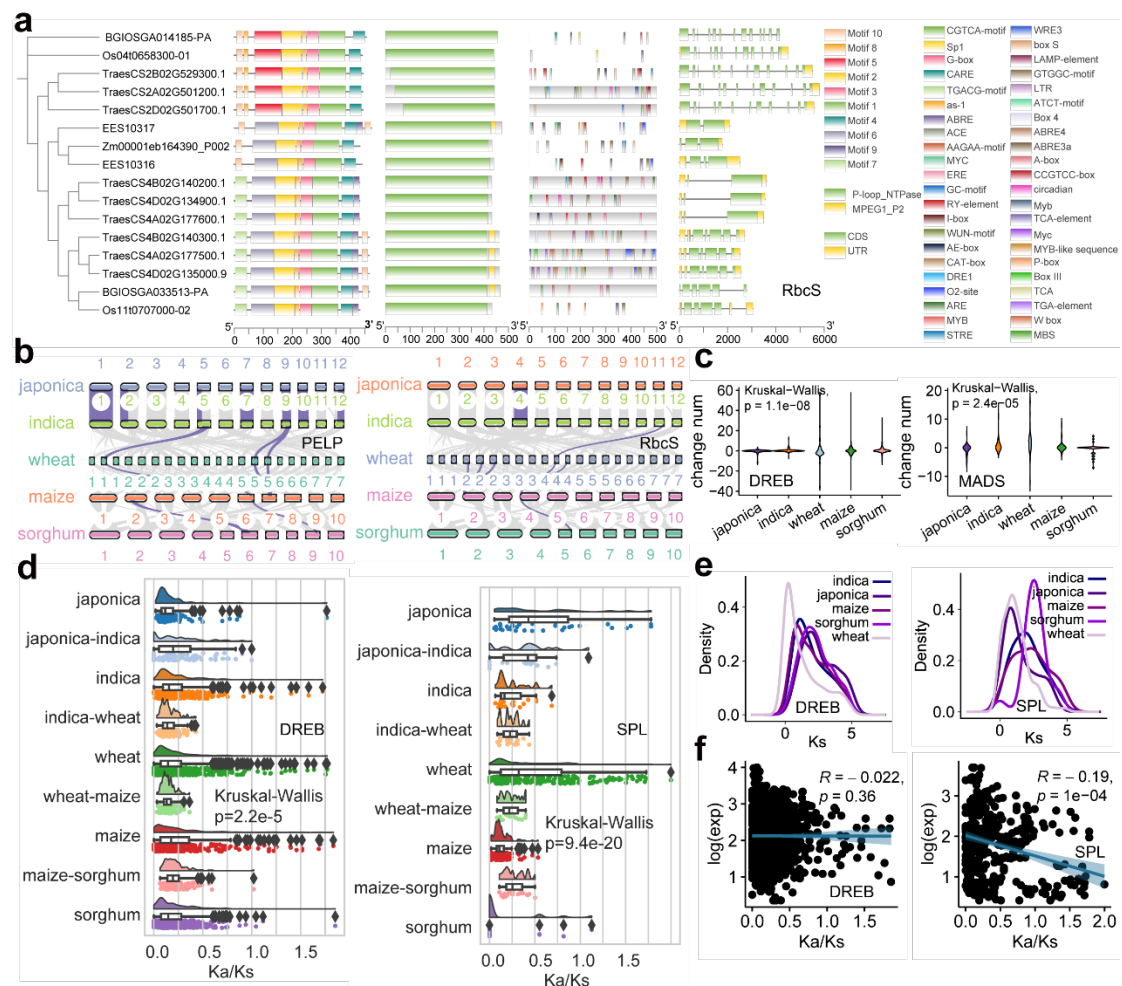
Figure 6:



“Fig. 6 Evolutionary analysis of conserved drought-resistance regulatory gene families. Figure (a) presents a phylogenetic tree for the PP2C gene family within the drought-regulatory network, along with motifs, domains, cis-regulatory elements, and gene

structures. Figure (b) illustrates the collinear positions of the 2OGD (left panel) and PP2C (right panel) gene families across species. Figure (c) shows the distribution of gene contraction and expansion counts for the 2OGD (left panel) and PP2C (right panel) families based on clustering results. Violin plots show distributions, centre dots indicate medians, and vertical lines indicate minima and maxima. Counts were compared among species using two-sided Kruskal-Wallis tests followed by two-sided Mann-Whitney U tests for pairwise comparisons; Benjamini-Hochberg and Bonferroni corrections were applied within each family. Exact raw and adjusted P values are provided in the Source Data. Figure (d) highlights the distribution of Ka/Ks values for both orthologs (one-to-one genes between species within the same orthogroup) and paralogs (multiple genes from the same species within the same orthogroup) in the 2OGD (left panel) and PP2C (right panel) families, providing insights into their divergence. Horizontal box plots show medians (centre line), interquartile ranges (box), $1.5 \times \text{IQR}$ whiskers, and individual data points. Groups were compared using Kruskal-Wallis tests followed by two-sided Mann-Whitney U tests with Benjamini-Hochberg and Bonferroni correction. Exact raw and adjusted P values are provided in the Source Data. Figure (e) displays the distribution of synonymous substitution rates (Ks) values for 2OGD and PP2C families within homolog clusters, providing insights into evolutionary timescales and species-specific differences. Figure (f) explores the correlation between Ka/Ks values and gene expression levels ($\log(\text{FPKM})$) for drought-responsive 2OGD (left panel) and PP2C (right panel) family genes across the five species, indicating evolutionary pressure and functional significance. Lines indicate fitted linear regressions and shaded bands indicate 95% confidence intervals. Associations were tested using two-sided Spearman rank correlation; effect sizes are reported as Spearman's ρ with 95% confidence intervals estimated by Fisher's z transformation. Exact raw and adjusted P values are provided in the Source Data."

Figure 8:



“Fig. 8 The evolutionary analysis of differentiated drought-regulatory gene families. Figure (a) illustrates the phylogenetic tree of the RbcS gene family (Ribulose biphosphate carboxylase/oxygenase activase-like) within the drought-regulatory network, providing detailed information on motifs, domains, cis-acting elements, and gene structures. Figure (b) shows the syntenic positions of the PELP (left panel) and RbcS (right panel) gene families across various species, highlighting their conserved genomic locations. Figure (c) presents the expansion and contraction distributions of the DREB (left panel) and MADS (right panel) gene families based on gene family clustering. Violin plots show distributions, centre dots indicate medians, and vertical lines indicate minima and maxima. Counts were compared among species using two-sided Kruskal-Wallis tests followed by two-sided Mann-Whitney U tests for pairwise comparisons; Benjamini-Hochberg and Bonferroni corrections were applied within each family. Exact raw and adjusted P values are provided in the Source Data. Figure (d) illustrates the Ka/Ks distribution for both orthologs (one-to-one genes between species within the same orthogroup) and paralogs (multiple genes from the same species within the same orthogroup) in the DREB (left panel), SPL (right panel) gene families. Horizontal box plots show medians (centre line), interquartile ranges (box), $1.5 \times$ IQR whiskers, and individual data points. Groups were compared using Kruskal-Wallis tests followed by two-sided Mann-Whitney U tests with Benjamini-Hochberg and Bonferroni correction. Exact raw and adjusted P values are provided in the Source

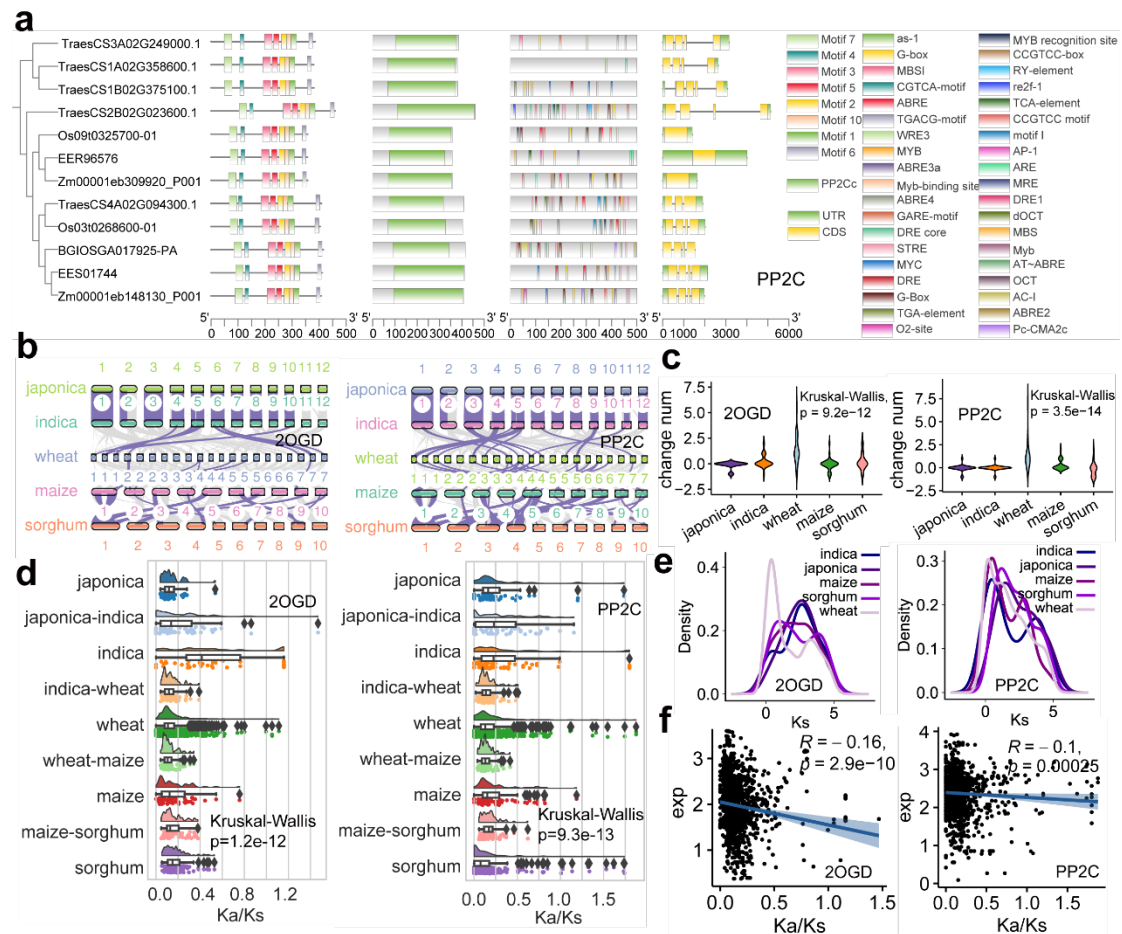
Data. Figure (e) displays the distribution of synonymous substitution rates (K_s) values for DREB and SPL families within homolog clusters, offering insights into the selective pressures shaping their evolution. Figure (f) examines the relationship between K_a/K_s values and gene expression levels ($\log(\text{FPKM})$) for drought-responsive DREB and SPL family genes across the five species. Lines indicate fitted linear regressions and shaded bands indicate 95% confidence intervals. Associations were tested using two-sided Spearman rank correlation; effect sizes are reported as Spearman's ρ with 95% confidence intervals estimated by Fisher's z transformation. Exact raw and adjusted P values are provided in the Source Data."

Line 592-596: It would be clearer to state the p value and test when stating that the differences are significant. It is not clear to me how the Kruskal-Wallis (KW) test (which evaluates whether an arbitrary number of distributions have the same mean) would provide insight into the cause of that difference, as seems to be stated on line 594. It seems reasonable to note the difference and discuss possible causes for it, but the delineation between that discussion of possible causes and the statement of the results of the statistical test needs to be clearer. I also think the conclusion that wheat and maize have the highest expansion is clear for DREB, but for MADS it is less so: although at least 1 of the 5 species does not have the same median as shown by the KW test, Indica rice appears to have a similar median to Maize. Follow-up pairwise testing could be used to support these more specific claims. It would be interesting to see if these tests reveal differences in gene family expansion between DREB and MADS.

Response:

We thank the reviewer for this important comment. We agree that the original wording did not sufficiently distinguish the statistical results from the biological interpretation. In the revised manuscript, we have separated these two aspects more clearly. Specifically, we now report the Kruskal-Wallis results with exact p values in Fig. 8c and added post hoc pairwise comparisons (Mann-Whitney U test with Benjamini-Hochberg correction) to identify the species contributing to the observed differences. We also revised the text to ensure a more rigorous statement. We also revised the corresponding content in Fig. 6c.

Figure 6:



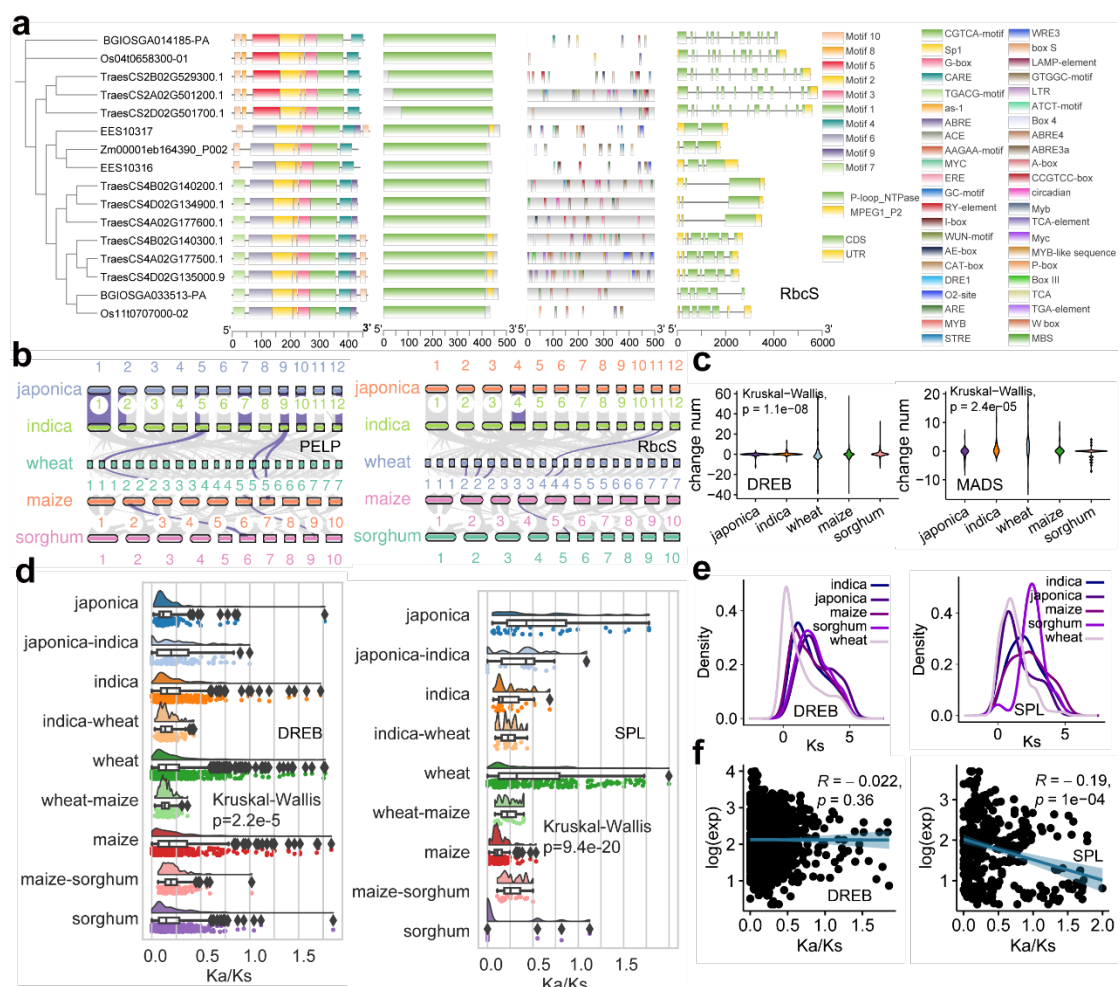
“Fig. 6 Evolutionary analysis of conserved drought-resistance regulatory gene families. Figure (a) presents a phylogenetic tree for the PP2C gene family within the drought-regulatory network, along with motifs, domains, cis-regulatory elements, and gene structures. Figure (b) illustrates the collinear positions of the 2OGD (left panel) and PP2C (right panel) gene families across species. Figure (c) shows the distribution of gene contraction and expansion counts for the 2OGD (left panel) and PP2C (right panel) families based on clustering results. Violin plots show distributions, centre dots indicate medians, and vertical lines indicate minima and maxima. Counts were compared among species using two-sided Kruskal-Wallis tests followed by two-sided Mann-Whitney U tests for pairwise comparisons; Benjamini-Hochberg and Bonferroni corrections were applied within each family. Exact raw and adjusted P values are provided in the Source Data. Figure (d) highlights the distribution of Ka/Ks values for both orthologs (one-to-one genes between species within the same orthogroup) and paralogs (multiple genes from the same species within the same orthogroup) in the 2OGD (left panel) and PP2C (right panel) families, providing insights into their divergence. Horizontal box plots show medians (centre line), interquartile ranges (box), $1.5 \times \text{IQR}$ whiskers, and individual data points. Groups were compared using Kruskal-Wallis tests followed by two-sided Mann-Whitney U tests with Benjamini-Hochberg and Bonferroni correction. Exact raw and adjusted P values are provided in the Source Data. Figure (e) displays the distribution of synonymous substitution rates (Ks) values for 2OGD and PP2C families within homolog clusters, providing insights into evolutionary timescales and

species-specific differences. Figure (f) explores the correlation between Ka/Ks values and gene expression levels (log(FPKM)) for drought-responsive 2OGD (left panel) and PP2C (right panel) family genes across the five species, indicating evolutionary pressure and functional significance. Lines indicate fitted linear regressions and shaded bands indicate 95% confidence intervals. Associations were tested using two-sided Spearman rank correlation; effect sizes are reported as Spearman's ρ with 95% confidence intervals estimated by Fisher's z transformation. Exact raw and adjusted P values are provided in the Source Data."

Line 365-370:

"Obvious expansion and contraction of the 2OGD and PP2C gene families were observed among species. Both families showed interspecies differences in expansion patterns, with wheat exhibiting the strongest expansion in both 2OGD and PP2C, and maize showing relatively greater expansion in PP2C than rice (Supplementary Table 5). Kruskal-Wallis tests confirmed significant interspecies differences in expansion and contraction counts across species ($p = 9.2\text{e-}12$ for 2OGD; $p = 3.5\text{e-}14$ for PP2C) (Fig. 6c)."

Figure 8:



"Fig. 8 The evolutionary analysis of differentiated drought-regulatory gene families.

Figure (a) illustrates the phylogenetic tree of the RbcS gene family (Ribulose biphosphate carboxylase/oxygenase activase-like) within the drought-regulatory network, providing detailed information on motifs, domains, cis-acting elements, and gene structures. Figure (b) shows the syntenic positions of the PELP (left panel) and RbcS (right panel) gene families across various species, highlighting their conserved genomic locations. Figure (c) presents the expansion and contraction distributions of the DREB (left panel) and MADS (right panel) gene families based on gene family clustering. Violin plots show distributions, centre dots indicate medians, and vertical lines indicate minima and maxima. Counts were compared among species using two-sided Kruskal-Wallis tests followed by two-sided Mann-Whitney U tests for pairwise comparisons; Benjamini-Hochberg and Bonferroni corrections were applied within each family. Exact raw and adjusted P values are provided in the Source Data. Figure (d) illustrates the Ka/Ks distribution for both orthologs (one-to-one genes between species within the same orthogroup) and paralogs (multiple genes from the same species within the same orthogroup) in the DREB (left panel), SPL (right panel) gene families. Horizontal box plots show medians (centre line), interquartile ranges (box), $1.5 \times \text{IQR}$ whiskers, and individual data points. Groups were compared using Kruskal-Wallis tests followed by two-sided Mann-Whitney U tests with Benjamini-Hochberg and Bonferroni correction. Exact raw and adjusted P values are provided in the Source Data. Figure (e) displays the distribution of synonymous substitution rates (Ks) values for DREB and SPL families within homolog clusters, offering insights into the selective pressures shaping their evolution. Figure (f) examines the relationship between Ka/Ks values and gene expression levels ($\log(\text{FPKM})$) for drought-responsive DREB and SPL family genes across the five species. Lines indicate fitted linear regressions and shaded bands indicate 95% confidence intervals. Associations were tested using two-sided Spearman rank correlation; effect sizes are reported as Spearman's ρ with 95% confidence intervals estimated by Fisher's z transformation. Exact raw and adjusted P values are provided in the Source Data."

Line 456-461:

"Gene expansion and contraction patterns of the DREB and MADS families differed significantly among species (Kruskal-Wallis test, $p < 0.05$; Fig. 8c). Post hoc pairwise comparisons (Mann-Whitney U test with Benjamini-Hochberg correction) revealed heterogeneous interspecies differences, with clearer pairwise divergence in the DREB family and a more variable pattern in the MADS family (Supplementary Table 5). These patterns may reflect lineage-specific evolutionary histories and selective pressures."

Supporting information:

Supplementary Table 5 Results of pairwise interspecies comparisons for the 2OGD, DREB, MADS, PELP, and PP2C gene families shown in Figs. 6c and 8c, including summary statistics, Mann-Whitney U test results, and Benjamini-Hochberg FDR-adjusted p values.

gene family	species1	species2	mean1	mean2	median1	median2	statistic	p value
2OGD	indica	japonica	2.33E-01	-0.14	0	0	1939.5	0.000316948
2OGD	indica	maize	2.33E-01	-0.06	0	0	1227.5	0.017485542
2OGD	indica	sorghum	2.33E-01	0.19	0	0	1018	0.61286582
2OGD	indica	wheat	2.33E-01	1.48	0	1	404	3.77E-06
2OGD	japonica	maize	-1.40E-01	-0.06	0	0	791.5	0.636643671
2OGD	japonica	sorghum	-1.40E-01	0.1875	0	0	637.5	0.047232833
2OGD	japonica	wheat	-1.40E-01	1.483870968	0	1	200.5	3.16E-09
2OGD	maize	sorghum	-0.060606061	0.1875	0	0	444.5	0.207739418
2OGD	maize	wheat	-6.06E-02	1.48	0	1	157	1.81E-06
2OGD	sorghum	wheat	1.88E-01	1.48	0	1	217	0.000146211
DREB	indica	japonica	2.87E-01	-0.38	0	0	7737	0.000515326
DREB	indica	maize	2.87E-01	0.36	0	0	14211	0.274352029
DREB	indica	sorghum	2.87E-01	0.42	0	0	12750	0.453652966
DREB	indica	wheat	0.286956522	-0.556521739	0	-2	9037.5	1.56E-06
DREB	japonica	maize	-0.382608696	0.360869565	0	0	12601.5	0.453652966
DREB	japonica	sorghum	-0.382608696	0.417391304	0	0	10668.5	0.000197237
DREB	japonica	wheat	-0.382608696	-0.556521739	0	-2	8430.5	0.000320344
DREB	maize	sorghum	0.360869565	0.417391304	0	0	24067.5	0.094439684
DREB	maize	wheat	0.360869565	-0.556521739	0	-2	16223	0.000825894
DREB	sorghum	wheat	0.417391304	-0.556521739	0	-2	17871.5	7.84E-08
MADS	indica	japonica	1.027777778	-0.5	0	0	856	0.005260184
MADS	indica	maize	1.027777778	0.333333333	0	0	761	0.092762756
MADS	indica	sorghum	1.027777778	-0.638888889	0	0	864	0.005260184
MADS	indica	wheat	1.027777778	2.166666667	0	2	490	0.092762756
MADS	japonica	maize	-0.5	0.333333333	0	0	543	0.117442713
MADS	japonica	sorghum	-0.5	-0.638888889	0	0	669	0.773571412
MADS	japonica	wheat	-0.5	2.166666667	0	2	347	0.002060799
MADS	maize	sorghum	0.333333333	-0.638888889	0	0	773	0.092762756
MADS	maize	wheat	0.333333333	2.166666667	0	2	422.5	0.013390627
MADS	sorghum	wheat	-0.638888889	2.166666667	0	2	347	0.002060799
PP2C	indica	japonica	-0.017857143	-0.039215686	0	0	1457.5	0.731310366
PP2C	indica	maize	-0.017857143	0.261904762	0	0	889	0.001829754
PP2C	indica	sorghum	-0.017857143	-0.2	0	0	850	0.039028602
PP2C	indica	wheat	-0.017857143	1.25	0	1	152.5	1.27E-10
PP2C	japonica	maize	-0.039215686	0.261904762	0	0	793	0.001829754
PP2C	japonica	sorghum	-0.039215686	-0.2	0	0	761.5	0.073026185

PP2C	japonica	wheat	-0.039215686	1.25	0	1	137.5	4.85E-10
PP2C	maize	sorghum	0.261904762	-0.2	0	0	727.5	0.002317197
PP2C	maize	wheat	0.261904762	1.25	0	1	213	2.75E-05
PP2C	sorghum	wheat	-0.2	1.25	0	1	78	1.04E-05

The discussion section repeats a lot of the details from the results section. This is necessary to some extent, but I think reducing the specificity in the discussion to more clearly communicate a broad, high-level overview of the results would make this section more useful.

Response:

Thank you for this helpful suggestion. We agree that the original Discussion contained excessive result-level detail in several places. In the revised manuscript, we have streamlined this section to provide a broader, higher-level interpretation of the findings while retaining the key biological implications and supporting references. Specifically, we reduced repetitive descriptions of specific modules, gene pairs, and figure-level observations, and instead emphasized the major conclusions regarding network rewiring, modular organization, conserved drought-responsive cores, and divergent adaptive strategies across species.

Line 480-584:

“In our study, an innovative graph network algorithm was developed to construct drought-responsive gene regulatory networks. We delved into the topological structure of these GRNs, which are fundamental to how crops respond to intrinsic and environmental signals, reflecting the strategies species employ to adapt to environmental changes ^{45,46}. We found that transcription factor activity, which corresponds to edge strength in these networks ⁴⁷, is significantly altered under drought stress, shifting the activity of most transcription factor families and reshaping overall network topology (Fig. 3d). This shift leads to changes in network topological properties, with most species showing an increase in average degree under drought stress, a pattern also observed in human cell GRNs under tumor conditions ⁴⁸ and soil bacterial networks under drought ⁴⁹. Overall, changes in co-expression and co-regulation network connectivity suggest that species with weaker drought resistance tend to undergo more extensive network rewiring under drought stress. The exponential relationship between transcription factor activity and average degree indicates that, similar changes in transcription factor activity cause C₃ crops to experience larger average degree shifts than C₄ crops under drought stress in our representative species, pointing to different levels of network sensitivity to drought perturbation. This implies that maize and sorghum, with their evolutionarily stabilized network architectures, maintain greater robustness under drought, reducing sensitivity to environmental fluctuations. In contrast, rice adapts to drought by increasing regulatory complexity,

reflecting its relatively weaker drought resistance and greater need for network reprogramming. This suggests that the network structure could be intrinsically linked to its drought adaptation capacity. Similar patterns are found in microbial networks, where bacterial networks show greater instability and reorganization compared to fungal networks under drought ⁴⁹, suggesting that different biological networks may share common mechanisms to cope with environmental stress.

Understanding gene modularity provides deeper insights into the organization and functionality of drought-responsive networks. The gene modularity distribution curves reveal that japonica and indica rice rely heavily on a few core gene modules for drought response, indicated by steep curves (Fig. 4d). This reliance on highly connected modules may make their networks more vulnerable under drought stress, as damage to key modules could lead to system collapse ⁵⁰. In contrast, wheat, maize, and sorghum display more gradual curves, suggesting that their drought response systems are supported by multiple independent modules, enhancing network robustness. This further underscores the potential link between network structure and drought adaptation capacity. Theoretical studies corroborate these observations, indicating that networks with weakly correlated components are more stable under stress ^{51,52}, while adding key nodes with strong interactions can destabilize systems ⁵³, similar to how bacterial networks with tighter node connections are less stable under stress compared to fungal networks ⁴⁹. Genes within the same module often share related functions, highlighting the importance of co-regulation in drought response. For instance, hormone response pathways such as abscisic acid and ethylene, along with responses to abiotic stresses like water deficit and oxidative stress, are co-regulated within modules (Fig. 4b, Fig. 4c), indicating that drought adaptation depends on coordinated regulation of multiple stress-response processes rather than isolated pathways. Additionally, cross-talk between abiotic stress responses and redox enzyme activity, observed in maize and sorghum modules, further supports the idea that antioxidant defense is an integral part of drought-responsive regulation ⁵⁴. Furthermore, cross-talk between redox enzyme activity and photosynthesis may arise from drought-induced photosynthesis suppression, leading to the accumulation of H₂O₂ and malondialdehyde (MDA), triggering transcriptomic reprogramming that activates antioxidant defense genes to alleviate oxidative damage ³⁷. Moreover, gene expression patterns under drought stress become more similar across species (Fig. 4e), suggesting the activation of conserved hormone-responsive gene modules involved in drought resistance ⁵⁵. The majority of genes within these modules are upregulated during drought, emphasizing their role in adaptation (Supplementary Fig. 4).

Conserved interaction mechanisms across species underpin the core network essential for drought resistance. Drought-resistant gene families such as CYP450, MYB, PP2C, TCP, Cyclin, and ERF form a stable core network across species (Fig. 5b). The conserved combinations within these networks, such as ERF-2OGD, ERF-Cyclin, TCP-PP2C, and MYB-CYP450, provide a stable framework for drought response, suggesting that crops rely on similar core modules for drought adaptation ⁵⁶. Consistent expression patterns across species under both normal and drought conditions (Fig. 5c)

demonstrate the ability of these gene families to maintain functional stability across varying environments^{34,57}. Specifically, gene families like PP2C and 2OGD regulate essential metabolic processes, including hormone signaling and redox balance (Fig. 5e)^{58,59}. Their coordinated expression, exemplified by interactions such as TCP-PP2C and MYB-CYP450, enhances the efficiency of drought resistance mechanisms (Fig. 5d)^{60,61}. Conserved motifs, domains, and cis-regulatory elements across these gene families (Fig. 6a) underline their essential role in drought regulation, with ABRE and DRE elements reinforcing regulatory modules⁶². Additionally, the stable chromosomal positions of 2OGD and PP2C further affirm the robustness of the gene network (Fig. 6b)^{59,63}, while the expansion of these families in wheat, maize, and sorghum enhances their drought adaptability (Fig. 6c)⁶⁴. The PP2C genes in wheat and sorghum, with higher Ka/Ks ratios, suggest they have undergone evolution that supports drought adaptation⁶⁵. Most drought-resistant families undergo strong purifying selection (Fig. 6d), ensuring functional preservation and stability over time⁶⁶. This selective pressure supports the efficiency of the drought response, especially for genes with high expression levels, which tend to have lower Ka/Ks ratios (Fig. 6f, Supplementary Fig. 21g)⁶⁷.

Divergent interaction patterns between C₃ and C₄ crops in our representative species reflect their ecological adaptations to drought. C₄ crops like maize and sorghum enhance photosynthesis and nitrogen use through regulatory modules such as ERF-Psb28 and TCP-ASN, while C₃ crops like rice prioritize drought avoidance via growth regulation through modules like SPL-PELP (Fig. 7b, Fig. 7d, Fig. 7e)^{23,25,68}. Species-specific functional annotations further support interspecies differences in drought-response strategies, suggesting that conserved gene families are recruited into distinct functional contexts in different crops (Fig. 7e). These patterns indicate that species with stronger drought resistance tend to favor drought tolerance mechanisms, whereas those with weaker resistance rely more on drought avoidance or escape strategies^{18,69}. Gene expression patterns further underscore this, with C₄ crops showing enhanced drought-induced expression of genes like ASN and Psb28, while C₃ crops adopt a more muted expression profile in all representative crops (Fig. 7d)⁷⁰. Despite these differences, co-regulation within gene modules facilitates a coordinated drought response (Supplementary Fig. 16). Chromosomal divergence and gene rearrangements in C₄ crops like maize optimize drought regulation⁷¹, while C₃ crops retain conserved gene structures (Fig. 8a, Fig. 8b)⁷². The expansion of drought-resistance gene families in wheat and maize provides redundancy and resilience^{64,73}, whereas rice shows limited expansion, relying on diversified regulation (Fig. 8c)⁷⁴. The negative correlation between evolutionary selection pressure and gene expression (Fig. 8f) indicates that highly expressed genes under stronger purifying selection are conserved, which strengthens drought tolerance⁷⁵. Most family-level drought-resistance genes are subject to purifying selection, ensuring functional conservation⁷⁶. Japonica rice has the fewest positively selected genes in DREB, while wheat, maize, and sorghum have more, reflecting adaptive evolution that boosts their drought resistance⁶⁵. The evolutionary timescale of a species appears to correlate with the complexity of its regulatory network (Fig. 6e, Fig. 8e, and Fig. 3a). Species with later divergence likely lead to the

accumulation of more mutations ⁷⁷, resulting in greater diversification of transcription factors and their target gene connections.”

Line 841-872: These two sections would benefit from being more closely tied together. In particular, the section on baseline methods would be improved by a clearer statement of what training and testing data are used and how they were divided into training and test sets for the supervised methods.

Response :

Thank you for this helpful comment. We agree that the original text did not sufficiently connect the acquisition of regulatory relationship data with the benchmark settings used for model comparison. In the revised manuscript, we added a new paragraph to clarify how the integrated PlantRegMap, STRING, and KEGG relationships were matched with the processed expression matrices to construct species- and layer-specific benchmark datasets. We now explicitly state that, for supervised learning, the datasets were constructed with a positive-to-negative ratio of 1:3 and split into training, validation, and test sets at 60%, 20%, and 20%, respectively.

Line 724-731 :

“To construct benchmark datasets for GRN inference, the integrated regulatory relationships described above were matched with the processed gene expression matrices for each species and regulatory layer (TRN, PPI, and KEGG). For the transcriptional and metabolic layers, labels were treated as directed, whereas the PPI layer was treated as undirected. For supervised learning, the resulting datasets contained positive and negative samples at a ratio of 1:3 and were split into training, validation, and test sets at 60%, 20%, and 20%, respectively. The same data partition was used for all supervised baseline methods to ensure a fair comparison.”

Minor Comments:

Line 119-122: It would strengthen the merit of this novel computational approach to explain the limitations in existing approaches that require the development of a new approach to address this analytical task.

Response :

Thank you for this helpful suggestion. We agree that the novelty and merit of the computational framework are better justified when the limitations of existing approaches are stated more explicitly. In the revised Introduction, we have added text to clarify why a new method was needed for this study. Specifically, we now explain that existing GRN inference methods often rely mainly on gene expression similarity or regression-based associations, which limits their ability to capture higher-order topological dependencies. We also clarify that current deep learning-based approaches

are often insufficient for our task because they may not effectively integrate global graph structure, may not scale well to very large gene networks, and are generally not designed to model condition-specific regulatory changes across multiple omics layers in a unified framework. We then explicitly state that these limitations motivated the development of our new framework for large-scale, cross-species, multi-layer drought-responsive GRN inference.

Line 114-125:

“Our approach to understanding drought resistance shifts from individual gene analysis to a holistic network perspective, revealing insights that are invisible at the gene level alone. However, existing GRN inference methods are limited in their ability to address this task. Many rely primarily on expression similarity or local associations, which may overlook higher-order network topology, while current deep learning-based approaches often struggle to scale to large graphs and to model condition-specific regulation across multiple layers. These limitations hinder cross-species inference of drought-responsive GRNs from large transcriptomic datasets. To overcome these challenges, we constructed a cross-species drought-responsive GRN using innovative deep learning algorithm, integrating large-scale transcriptome data to validate multidimensional regulatory interactions across transcriptional, proteomic, and metabolic layers.”

Line 164: The phrase “is likely attributable” seems too strong for the level of support provided. Additionally, a citation here to support the discussion of the impact of ploidy on regulatory interactions would be beneficial.

Response:

Thank you for this helpful suggestion. We agree that the original wording was too strong for the level of support provided. In the revised manuscript, we replaced “is likely attributable to” with more cautious language and added an appropriate reference to support the possible influence of polyploidy on regulatory complexity and stress-responsive transcription.

Line 153-157:

“Wheat displayed the greatest number of differentially expressed genes (Fig. 1a). This pattern may partly reflect its hexaploid nature, which has been suggested to increase regulatory complexity and transcriptional responsiveness under stress through the presence of multiple homoeologous gene copies⁴⁴.”

References:

“44. Evans, C. E., Arunkumar, R. & Borrill, P. Transcription factor retention through multiple polyploidization steps in wheat. *G3* **12**, jkac147 (2022).”

Line 234-239: This first discussion of metrics such as AUC and AUPR should include

some context as to the ground truth labels used to construct them. This full context is eventually provided in methods, but a brief mention of where the labels for these metrics come from would be appropriate here.

Response:

Thank you for this helpful suggestion. We agree that the first mention of AUC and AUPR should briefly clarify the source of the ground-truth labels used for evaluation. In the revised Results section, we now state that curated regulatory interactions from PlantRegMap, STRING, and KEGG were used as the reference labels for performance evaluation, with full details provided in the Methods. This addition clarifies the basis on which AUC and AUPR were calculated without interrupting the flow of the Results section.

Line 203-212:

“Within a consistent model framework, gene expression data, along with transcriptional, protein, and metabolic regulatory relationships derived from database integration (Methods), were input to construct the regulatory network for each respective layer. Using curated regulatory interactions from PlantRegMap, STRING, and KEGG as ground-truth labels for performance evaluation, we found that our model outperforms existing tools, achieving improvements in AUC (a measure of classification accuracy) ranging from 0.02 to 0.19 and in AUPR (a measure of classification accuracy) ranging from 0.01 to 0.24, establishing a reliable platform for dissecting drought adaptation mechanisms (Supplementary Table 1, Supplementary Table 2, Fig. 2d).”

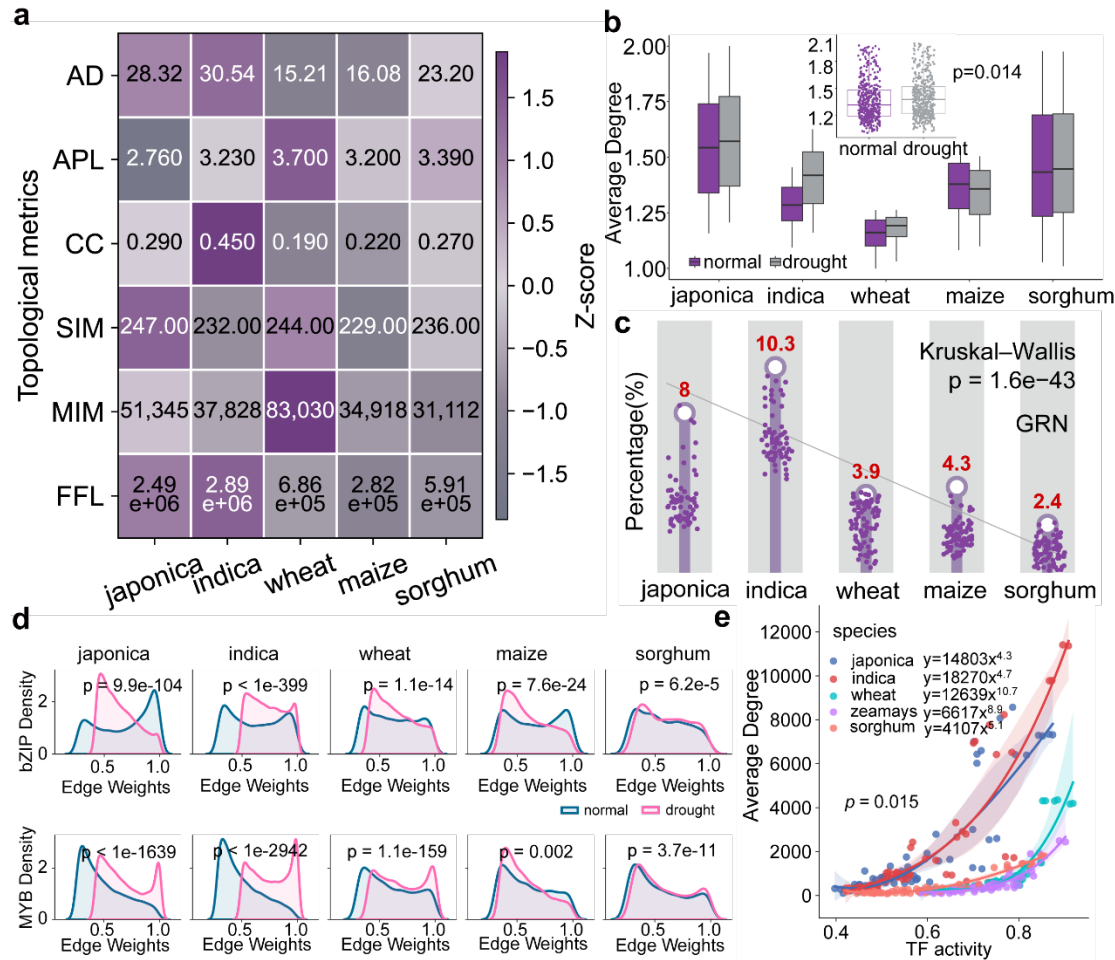
Fig 3a: The Y axis on this plot is somewhat confusing. From the text, it seems that these network metrics for each plant have been scaled so that they sum to 1. Presenting these values as a stacked plot with the Y axis labeled as percentage would seem to imply that these are proportions of an overall whole rather than distinct values for each network. This is a very compact representation, but I think sacrifices some clarity. Particularly as the discussion at 273-286 compares these measures between species, stacking the bars makes the comparison difficult. It may also be worth considering which of these metrics are necessary to support the conclusions that are being drawn later in the text. Plotting fewer metrics would allow for a more interpretable figure.

Response:

Thank you for this helpful comment. We agree that the previous stacked plot could be misleading, because these topological metrics are independent network properties rather than proportions of a common whole. We have therefore revised Fig. 3a by replacing the stacked bar plot with a heatmap, in which rows represent species and columns represent the six topological metrics. To account for the large differences in numerical scale among these metrics, the color of each cell represents the z-score normalized value calculated across species for that metric, whereas the number shown in each cell is the original metric value. This revised presentation improves clarity and

makes cross-species comparisons more direct. We also revised the figure legend accordingly to clarify that the heatmap colors represent relative normalized values rather than percentages. In addition, we retained all metrics but revised the visualization to make cross-species comparisons more interpretable.

Figure 3:



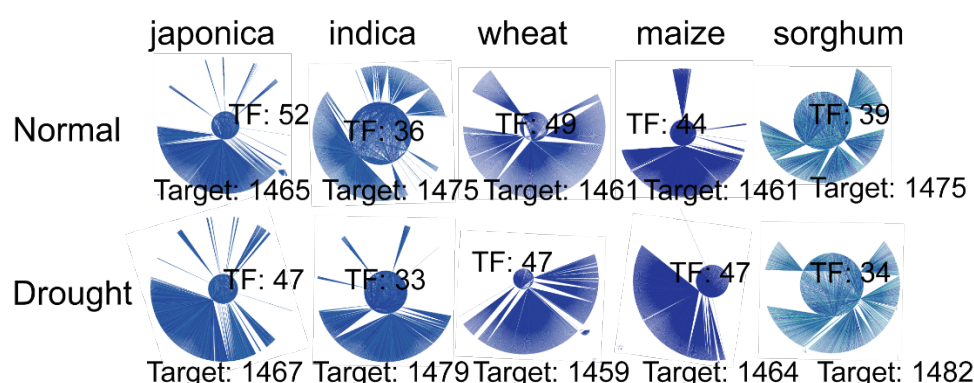
“Fig. 3 Topological analysis of gene regulatory networks under drought across species. Figure (a) shows heatmap of species-specific initial topological properties. Cell colors indicate Z-scores standardized independently across species for each metric; labels show raw values. Figure (b) compares changes in average network degree between normal and drought conditions. Box plots show the median (centre line), the interquartile range (box bounds), and the whiskers extend to $1.5 \times \text{IQR}$; points represent individual observations. Significance was assessed using a two-sided Wilcoxon rank-sum test. Figure (c) shows the percentage distribution of absolute changes in average degree in the co-regulatory network across species between normal and drought conditions. Red numbers indicate the maximum absolute percentage change for each species. Points represent networks with different gene numbers. Statistical significance was assessed using the Kruskal–Wallis test ($H = 206.36$, exact $P = 1.61 \times 10^{-43}$). Figure (d) examines changes in transcription factor family activity between normal and drought conditions, in which blue represents the normal condition and red represents

the drought condition. The top row represents the bZIP transcription factor family, and the bottom row represents the MYB transcription factor family. From left to right, the species are japonica rice, indica rice, wheat, maize, and sorghum. Statistical significance was evaluated using a two-sided Mann-Whitney U test with continuity correction. Source data list U, Z, exact/approximate P values, rank-biserial effect sizes, and BH/Bonferroni-adjusted P values. Figure (e) explores the relationship between the average degree and transcription factor family activity, which corresponds to edge strength in these networks, across species. This comprehensive analysis links topological properties to drought-induced regulatory changes. Associations were analyzed separately for each species using two-sided Spearman rank correlation tests, with effect size reported as Spearman's ρ and 95% confidence intervals estimated by Fisher's z transformation. Differences among species were evaluated using a Fisher z-based heterogeneity Q test, and pairwise comparisons of correlation coefficients used two-sided Fisher z tests with Benjamini-Hochberg correction for multiple comparisons. Exact test statistics, degrees of freedom, confidence intervals, raw P values and adjusted P values are provided in Source Data."

Figure 3c: I think that Fig 3d alone sufficiently supports the claims made using 3c, since 3c is a more qualitative view of the same phenomenon as 3d provides. I would consider moving 3c to supplementary to provide more space for the other elements of figure 3.

Response :

Thank you for this suggestion. We agree that Fig. 3c is largely qualitative and that Fig. 3d provides sufficient quantitative support for the corresponding conclusion. Accordingly, we have moved Fig. 3c to the Supplementary Information, which also allows more space for the other components of Fig. 3.



Supplementary Figure 5 Schematic subnetwork of the top 1,500 genes showing the largest changes in degree centrality under drought stress. Transcription factors are arranged in the inner circle and target genes in the outer circle. Species are shown from left to right as japonica, indica, wheat, maize, and sorghum.

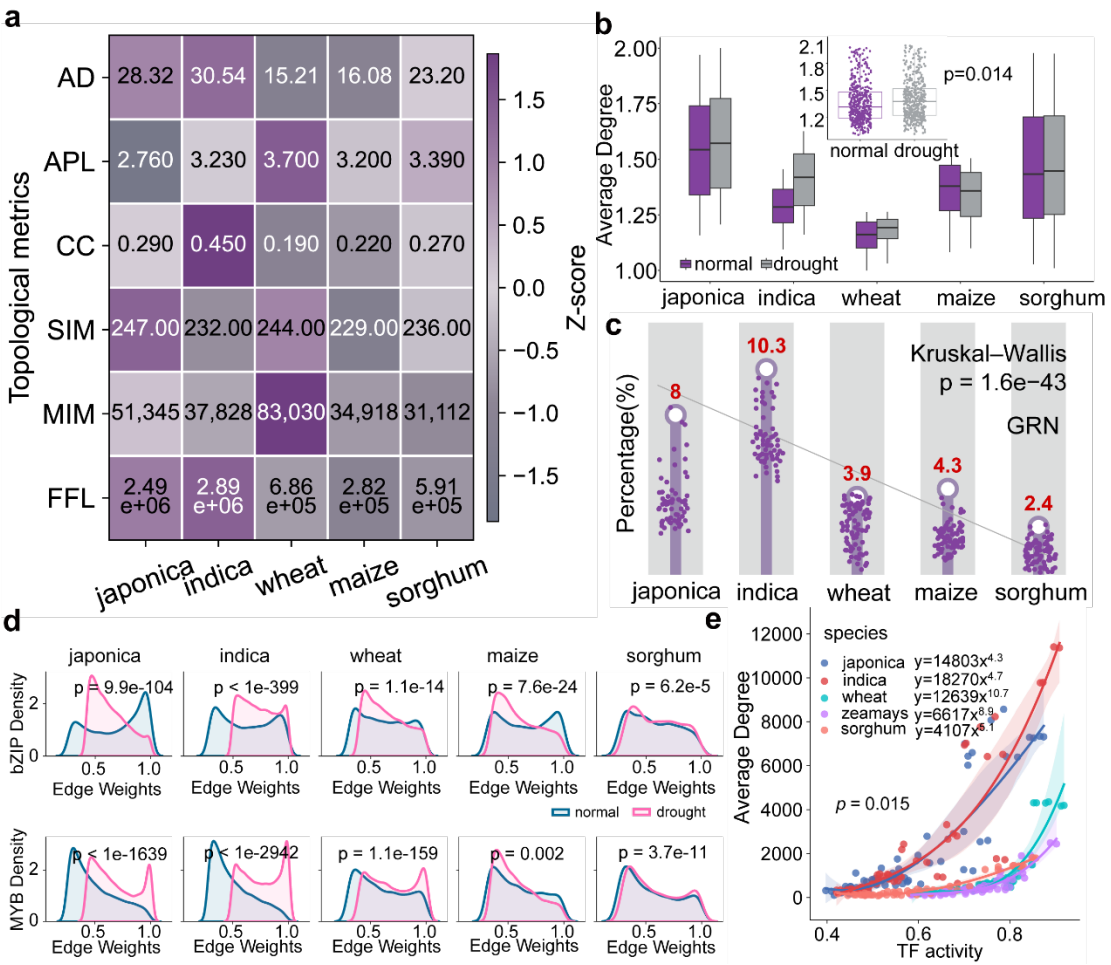
Fig 3c, 3e: these elements are not well labeled. The column labels for species and row

labels for transcription factor family should be in the figure element rather than requiring reference to the figure caption.

Response:

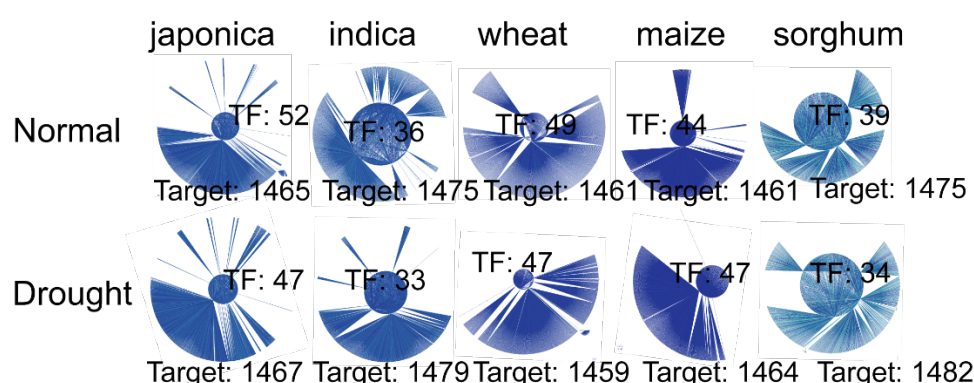
Thank you for this helpful comment. We agree that the original labeling in Fig. 3c and Fig. 3e was not sufficiently clear. In the revised figures, we have added the species column labels and transcription factor family row labels directly within the corresponding panels.

Figure 3:



“Fig. 3 Topological analysis of gene regulatory networks under drought across species. Figure (a) shows heatmap of species-specific initial topological properties. Cell colors indicate Z-scores standardized independently across species for each metric; labels show raw values. Figure (b) compares changes in average network degree between normal and drought conditions. Box plots show the median (centre line), the interquartile range (box bounds), and the whiskers extend to $1.5 \times \text{IQR}$; points represent individual observations. Significance was assessed using a two-sided Wilcoxon rank-sum test. Figure (c) shows the percentage distribution of absolute changes in average degree in the co-regulatory network across species between normal and drought

conditions. Red numbers indicate the maximum absolute percentage change for each species. Points represent networks with different gene numbers. Statistical significance was assessed using the Kruskal–Wallis test ($H = 206.36$, exact $P = 1.61 \times 10^{-43}$). Figure (d) examines changes in transcription factor family activity between normal and drought conditions, in which blue represents the normal condition and red represents the drought condition. The top row represents the bZIP transcription factor family, and the bottom row represents the MYB transcription factor family. From left to right, the species are japonica rice, indica rice, wheat, maize, and sorghum. Statistical significance was evaluated using a two-sided Mann-Whitney U test with continuity correction. Source data list U, Z, exact/approximate P values, rank-biserial effect sizes, and BH/Bonferroni-adjusted P values. Figure (e) explores the relationship between the average degree and transcription factor family activity, which corresponds to edge strength in these networks, across species. This comprehensive analysis links topological properties to drought-induced regulatory changes. Associations were analyzed separately for each species using two-sided Spearman rank correlation tests, with effect size reported as Spearman's ρ and 95% confidence intervals estimated by Fisher's z transformation. Differences among species were evaluated using a Fisher z-based heterogeneity Q test, and pairwise comparisons of correlation coefficients used two-sided Fisher z tests with Benjamini-Hochberg correction for multiple comparisons. Exact test statistics, degrees of freedom, confidence intervals, raw P values and adjusted P values are provided in Source Data."



Supplementary Figure 5 Schematic subnetwork of the top 1,500 genes showing the largest changes in degree centrality under drought stress. Transcription factors are arranged in the inner circle and target genes in the outer circle. Species are shown from left to right as japonica, indica, wheat, maize, and sorghum.

Line 271: what "original database" is being referred to here?

Response:

Thank you for pointing this out. Here, "original database" refers to the PlantRegMap database, from which the transcription factor-target regulatory relationships used to define the initial transcriptional regulatory network were obtained. We have revised the

text to explicitly name PlantRegMap and avoid ambiguity.

Line 222-226:

“We first explored the initial topological features of the transcriptional regulatory network, using regulatory relationships from the PlantRegMap database, and then examined how these features dynamically changed under drought conditions based on networks constructed with the DroGeneNet framework for both normal and drought conditions.”

Line 289-296: The discussion of network properties here would benefit from some interpretation as to what higher centrality under drought conditions might reflect. What does a higher average degree suggest about the system?

Response:

Thank you for this helpful comment. We agree that the biological interpretation of increased network centrality under drought conditions should be stated more clearly. In the revised manuscript, we added text explaining that higher centrality, and particularly higher average degree, suggests denser regulatory connectivity under drought, meaning that genes participate in more regulatory interactions on average and that the network undergoes stronger rewiring in response to stress. This addition clarifies the network-level implication of the observed changes.

Line 240-251:

“Among the compared network centrality measures, the drought group generally exhibited significantly higher values than the normal group based on the mean values, with the difference being statistically significant ($p = 0.014$) (Fig. 3b, Supplementary Fig. 20e). These results suggest that drought stress is accompanied by increased regulatory connectivity at the network level. In particular, a higher average degree indicates that genes are connected by more regulatory interactions on average under drought conditions. Further analysis of the average degree changes under drought conditions across different species’ networks revealed significant differences among species ($p = 1.6e-43$). Notably, rice, potato and soybean exhibited a greater magnitude of change compared to the other species, while maize and sorghum tended to show smaller fluctuations (Supplementary Fig. 5, Fig. 3c, Supplementary Fig. 6, Supplementary Fig. 20f).”

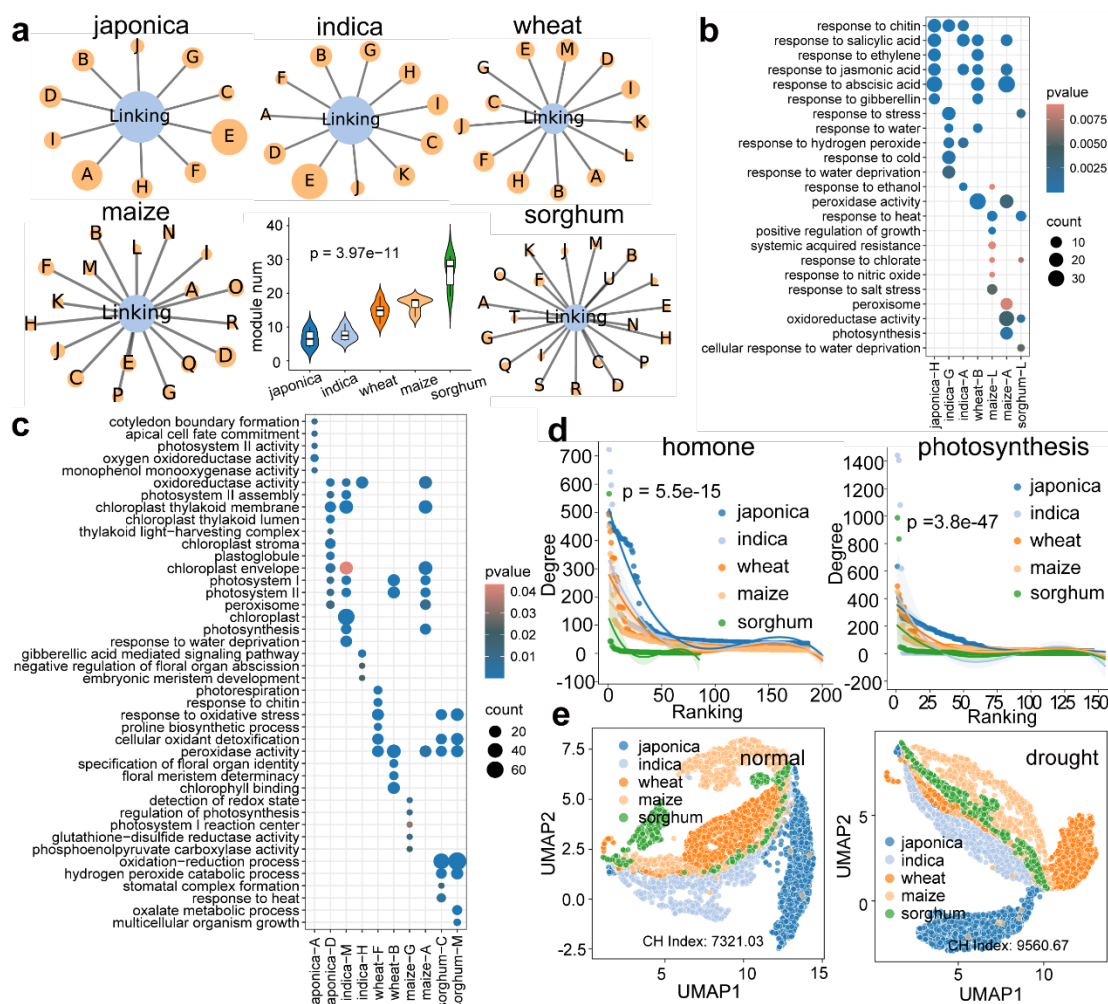
Figure 4a: the individual networks should be labeled by species in the figure, not just in the caption.

Response:

Thank you for this helpful suggestion. We agree that the original labeling was not sufficiently clear. In the revised Fig. 4a, we have added the species labels directly to the

individual network panels to improve readability, and the figure caption has been updated accordingly.

Figure 4:



“Fig. 4 The construction and analysis of a super gene regulatory network, gene modules, and process crosstalk within subnetworks revealed key underlying characteristics. The first row of Figure (a) shows super gene regulatory networks for japonica, indica, and wheat, while the second row depicts those for maize and sorghum, along with the number of gene modules identified under drought conditions. The size of the circles reflects the relative quantity of genes. Violin plots show the data distribution; centre lines indicate the median, boxes indicate the interquartile range (25th to 75th percentiles), and whiskers extend to $1.5 \times \text{IQR}$. Differences in module number among species were assessed using a two-sided Friedman test with network-size thresholds treated as paired blocks. Pairwise comparisons were performed using two-sided exact Wilcoxon signed-rank tests with Benjamini-Hochberg correction for multiple comparisons. Exact P values and effect sizes are provided in Source Data. GO annotations for gene modules participating in similar biological processes are shown in Figure (b), while Figure (c) highlights GO annotations for modules exhibiting biological process crosstalk. Enrichment significance was assessed using Fisher’s exact

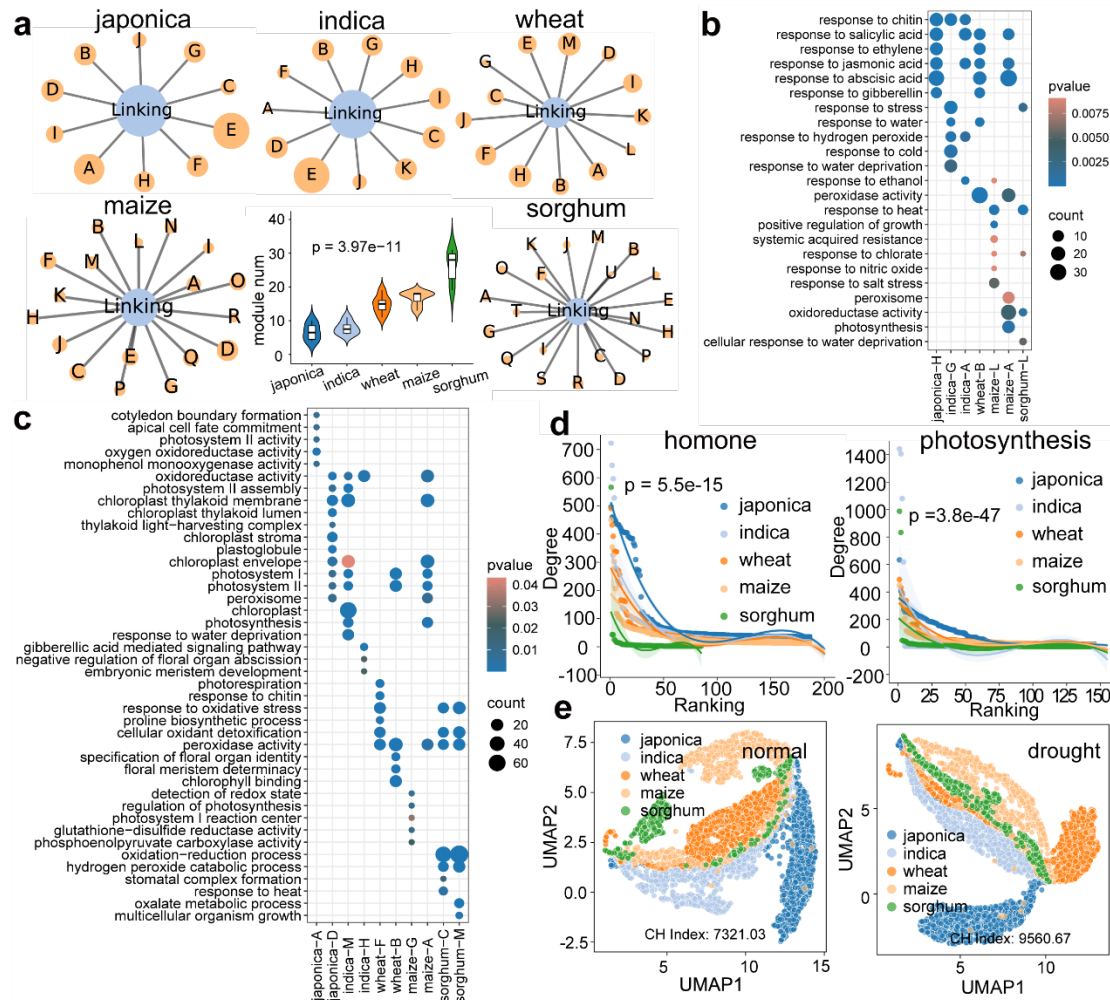
test, and the adjusted P values shown were corrected for multiple comparisons using the Benjamini-Hochberg method. And Figure (d) illustrates degree distributions of modules related to hormone (left panel) and photosynthesis (right panel) response. Species-specific degree distributions were fitted using a power-law model after log-log transformation. For each species, the exponent b was tested against 0 using a two-sided t -test from the log-log linear regression; effect sizes are reported as the fitted power-law exponent b , together with 95% confidence intervals and R^2 values. Differences among species-specific exponents were assessed using a Cochran Q-like heterogeneity test, followed by two-sided z -tests for pairwise comparisons with Benjamini-Hochberg correction for multiple comparisons. Detailed values are provided in Supplementary Table 4 and the Source Data. Figure (e) provides UMAP plots showing gene expression patterns in hormone-responsive modules under normal and drought conditions.”

Figure 4d: the individual panels should be labeled as hormone or photosynthesis in the figure.

Response :

Thank you for this suggestion. We have revised Fig. 4d by adding the corresponding panel labels directly within the figure to improve clarity and make the individual panels easier to interpret.

Figure 4:



“Fig. 4 The construction and analysis of a super gene regulatory network, gene modules, and process crosstalk within subnetworks revealed key underlying characteristics. The first row of Figure (a) shows super gene regulatory networks for japonica, indica, and wheat, while the second row depicts those for maize and sorghum, along with the number of gene modules identified under drought conditions. The size of the circles reflects the relative quantity of genes. Violin plots show the data distribution; centre lines indicate the median, boxes indicate the interquartile range (25th to 75th percentiles), and whiskers extend to $1.5 \times \text{IQR}$. Differences in module number among species were assessed using a two-sided Friedman test with network-size thresholds treated as paired blocks. Pairwise comparisons were performed using two-sided exact Wilcoxon signed-rank tests with Benjamini-Hochberg correction for multiple comparisons. Exact P values and effect sizes are provided in Source Data. GO annotations for gene modules participating in similar biological processes are shown in Figure (b), while Figure (c) highlights GO annotations for modules exhibiting biological process crosstalk. Enrichment significance was assessed using Fisher’s exact test, and the adjusted P values shown were corrected for multiple comparisons using the Benjamini-Hochberg method. And Figure (d) illustrates degree distributions of modules related to hormone (left panel) and photosynthesis (right panel) response. Species-specific degree distributions were fitted using a power-law model after log-log

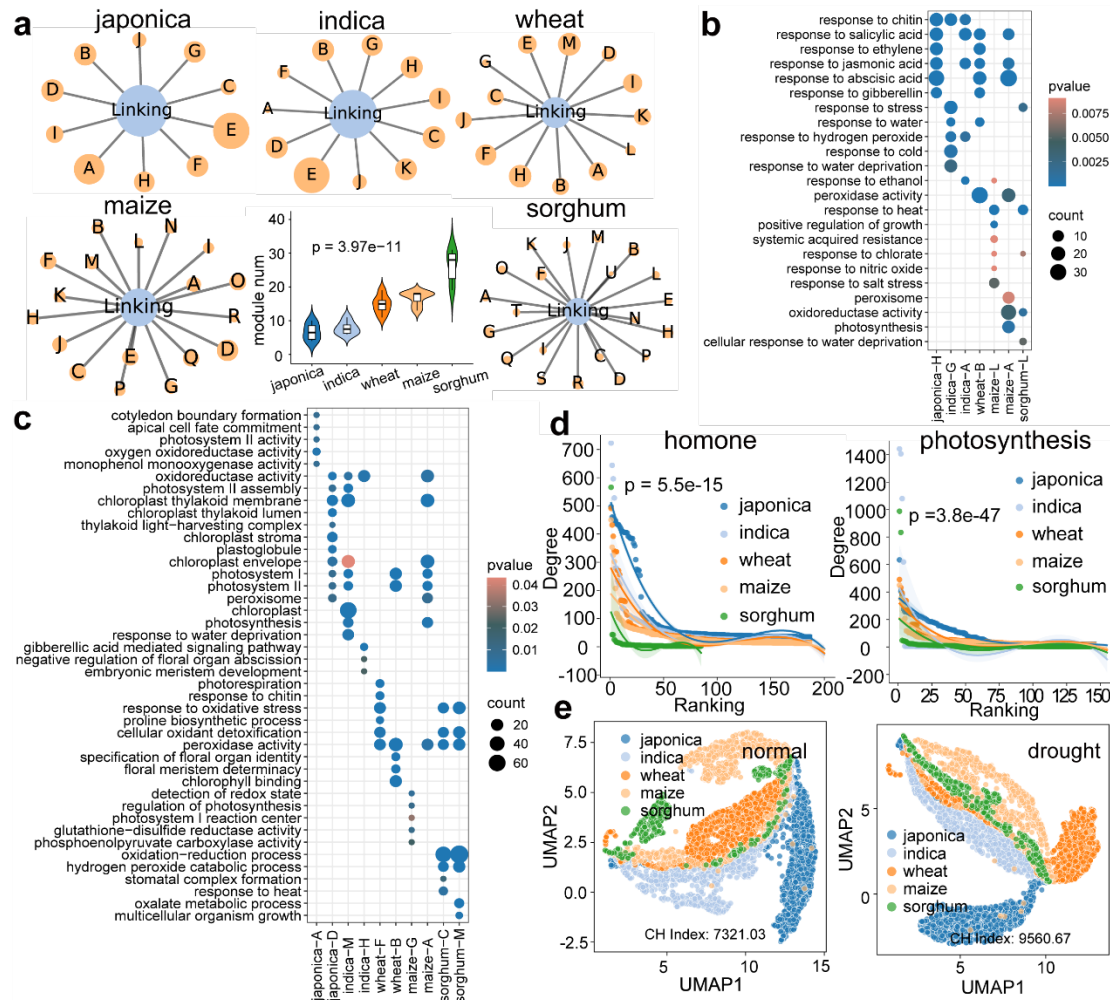
transformation. For each species, the exponent b was tested against 0 using a two-sided t -test from the log-log linear regression; effect sizes are reported as the fitted power-law exponent b , together with 95% confidence intervals and R^2 values. Differences among species-specific exponents were assessed using a Cochran Q-like heterogeneity test, followed by two-sided z -tests for pairwise comparisons with Benjamini-Hochberg correction for multiple comparisons. Detailed values are provided in Supplementary Table 4 and the Source Data. Figure (e) provides UMAP plots showing gene expression patterns in hormone-responsive modules under normal and drought conditions.”

Line 336-338: This claim would be strengthened by a quantitative measure for scale free properties such as power law regression R^2 or that proposed by Li et al: <https://doi.org/10.1080/15427951.2005.10129111>

Response:

Thank you for this helpful suggestion. We agree that the original statement would be strengthened by a quantitative assessment of scale-free properties. In the revised manuscript, we added a power-law regression analysis of the species-specific degree distributions for the hormone-related and photosynthesis-related regulatory networks shown in Fig. 4d, and we now report the corresponding regression statistics in Table S4. Based on these results, we revised the text to use more cautious wording, describing these networks as exhibiting scale-free-like connectivity patterns rather than making a stronger claim of strict scale-free behavior. These analyses support the conclusion that highly connected transcription factors occupy central positions in the modules.

Figure 4:



“Fig. 4 The construction and analysis of a super gene regulatory network, gene modules, and process crosstalk within subnetworks revealed key underlying characteristics. The first row of Figure (a) shows super gene regulatory networks for japonica, indica, and wheat, while the second row depicts those for maize and sorghum, along with the number of gene modules identified under drought conditions. The size of the circles reflects the relative quantity of genes. Violin plots show the data distribution; centre lines indicate the median, boxes indicate the interquartile range (25th to 75th percentiles), and whiskers extend to $1.5 \times \text{IQR}$. Differences in module number among species were assessed using a two-sided Friedman test with network-size thresholds treated as paired blocks. Pairwise comparisons were performed using two-sided exact Wilcoxon signed-rank tests with Benjamini-Hochberg correction for multiple comparisons. Exact P values and effect sizes are provided in Source Data. GO annotations for gene modules participating in similar biological processes are shown in Figure (b), while Figure (c) highlights GO annotations for modules exhibiting biological process crosstalk. Enrichment significance was assessed using Fisher’s exact test, and the adjusted P values shown were corrected for multiple comparisons using the Benjamini-Hochberg method. And Figure (d) illustrates degree distributions of modules related to hormone (left panel) and photosynthesis (right panel) response. Species-specific degree distributions were fitted using a power-law model after log-log

transformation. For each species, the exponent b was tested against 0 using a two-sided t -test from the log-log linear regression; effect sizes are reported as the fitted power-law exponent b , together with 95% confidence intervals and R^2 values. Differences among species-specific exponents were assessed using a Cochran Q-like heterogeneity test, followed by two-sided z -tests for pairwise comparisons with Benjamini-Hochberg correction for multiple comparisons. Detailed values are provided in Supplementary Table 4 and the Source Data. Figure (e) provides UMAP plots showing gene expression patterns in hormone-responsive modules under normal and drought conditions.”

Line 347-350:

“Our results also revealed that gene modules across species display scale-free-like connectivity patterns, as the degree distributions were generally consistent with power-law regression across species (Supplementary Table 4), with highly connected transcription factors dominating the network (Fig. 4d).”

Supporting information:

“Supplementary Table 4 Power-law regression results for species-specific degree distributions of hormone-related and photosynthesis-related regulatory networks shown in Fig. 4d.”

species	slope	intercept	r value	R ²	p value	std_err	class
japonica	-0.79	3.27	-0.94	0.88	1.31E-92	0.02	hormone
indica	-0.82	3.09	-1.00	0.99	1.04E-206	0.01	hormone
wheat	-0.83	3.03	-0.99	0.98	6.29E-182	0.01	hormone
maize	-0.77	2.84	-0.99	0.98	3.68E-166	0.01	hormone
sorghum	-0.98	2.08	-0.97	0.94	5.70E-54	0.03	hormone
japonica	-0.96	3.39	-0.93	0.86	2.41E-68	0.03	photosynthesis
indica	-0.62	1.80	-0.67	0.44	3.00E-21	0.06	photosynthesis
wheat	-0.83	3.03	-0.99	0.98	2.73E-132	0.01	photosynthesis
maize	-0.76	2.82	-0.99	0.97	1.75E-120	0.01	photosynthesis
sorghum	-1.06	2.56	-0.98	0.95	4.61E-104	0.02	photosynthesis

Line 340: What kind of differences might this steeper decline in C_4 vs C_3 crops reflect? Is there a broader connection to the regulatory processes governing drought response in these groups that this observation can connect to? I’m not sure if there is basis to make that kind of statement, but if there is it would be beneficial.

Response:

Thank you for this helpful comment. We agree that this pattern may reflect broader differences in drought-responsive regulatory organization between representative C_3 and C_4 crops. Because this interpretation is more appropriate for Discussion than for the Results section, we have addressed it there rather than adding further speculative

explanation to the Results. In the revised Discussion, we now explicitly connect the steeper decline in representative C₃ crops with a more centralized and strongly rewired network architecture, and the shallower decline in representative C₄ crops with a more distributed and stable regulatory structure.

Line 505-512:

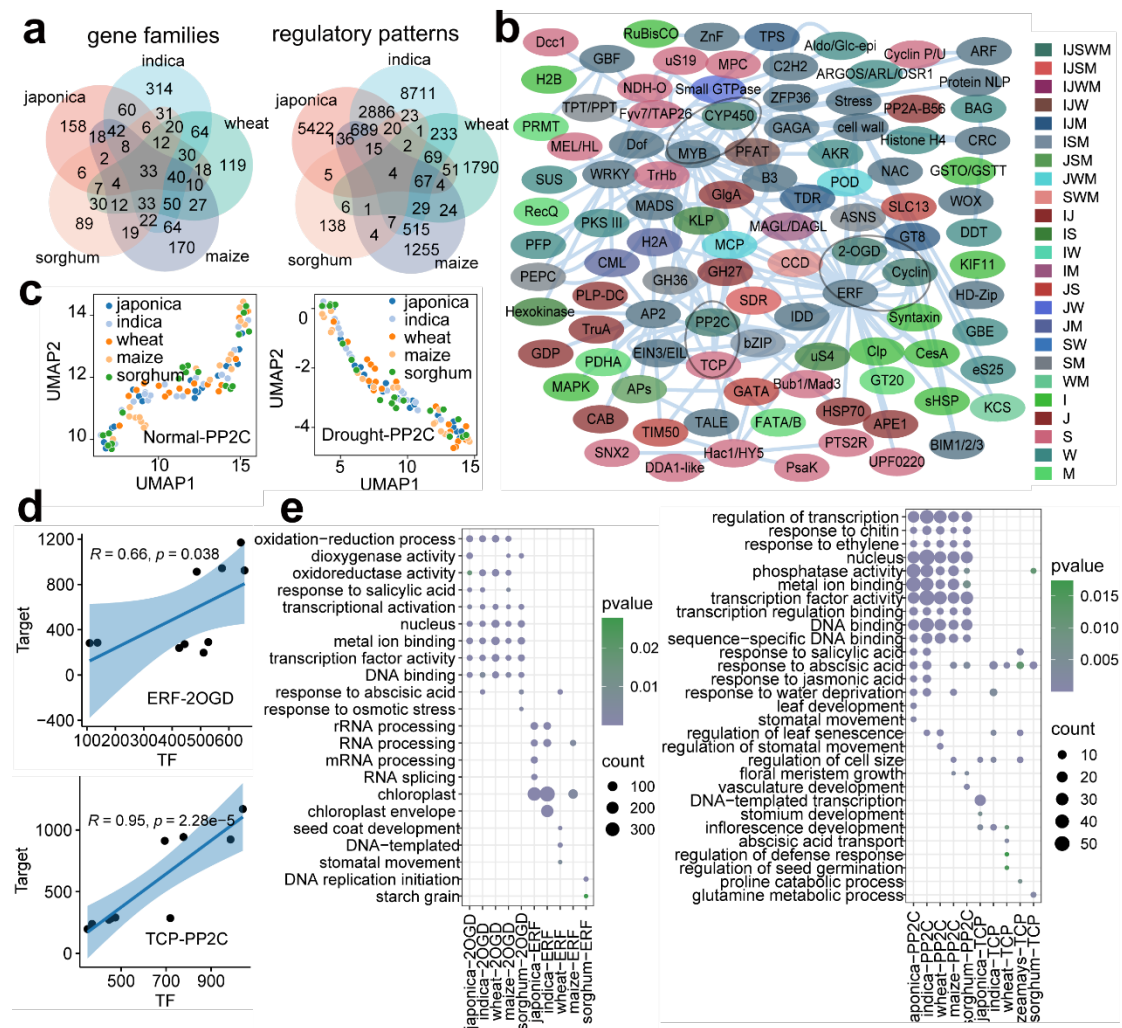
“Understanding gene modularity provides deeper insights into the organization and functionality of drought-responsive networks. The gene modularity distribution curves reveal that japonica and indica rice rely heavily on a few core gene modules for drought response, indicated by steep curves (Fig. 4d). This reliance on highly connected modules may make their networks more vulnerable under drought stress, as damage to key modules could lead to system collapse⁵⁰. In contrast, wheat, maize, and sorghum display more gradual curves, suggesting that their drought response systems are supported by multiple independent modules, enhancing network robustness.”

Fig 5a: I’m not sure this subfigure is legible enough to be useful. For the first 2 elements, UpSet plots might be more readable than the current Venn diagrams, and these elements should be labeled as gene families and regulatory patterns in the figure as well as in the caption. The visualization for chi squared tests seems unnecessary in a main body figure and might be better moved to the supplements.

Response:

Thank you for this helpful suggestion. We agree that the original presentation of Fig. 5a could be improved for clarity. In the revised figure, we have added explicit labels within the figure to indicate the gene families and regulatory patterns in the corresponding panels, in addition to revising the caption accordingly. We have also moved the chi-squared test panel to the Supplementary Information, as suggested, to improve the clarity and space balance of the main figure. We retained the current Venn diagrams, as we believe they remain a concise and intuitive way to visualize the overlap relationships in this dataset.

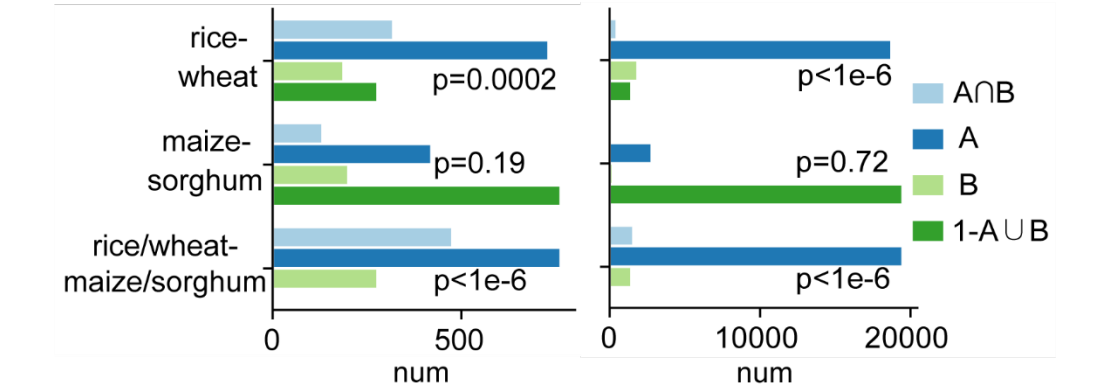
Figure 5:



“Fig. 5 The conserved drought-regulation mechanism in gene regulatory networks across different species. Figure (a) includes two subfigures: the first Venn diagram represents drought-responsive gene families across species, and the second illustrates regulatory patterns among these families. The corresponding chi-squared test results are provided in Supplementary Fig. 11. Figure (b) displays a partial view of the regulatory patterns of drought-responsive gene families, with all gene size families detailed in supplementary data. Different colors indicate drought responses in different species. IJSWM represents indica-japonica-sorghum-wheat-maize, IJWM represents indica-japonica-wheat-maize, IJSM represents indica-japonica-sorghum-maize, IJM represents indica-japonica-maize, ISM represents indica-sorghum-maize, JSM represents japonica-sorghum-maize, JWM represents japonica-wheat-maize, IJW represents indica-japonica-wheat, SWM represents sorghum-wheat-maize, IJ represents indica-japonica, IM represents indica-maize, IW represents indica-wheat, SW represents sorghum-wheat, SM represents sorghum-maize, JM represents japonica-maize, JW represents japonica-wheat, JS represents japonica-sorghum, IS represents indica-sorghum, WM represents wheat-maize, J represents japonica, I represents indica, W represents wheat, M represents maize, and S represents sorghum. Figure (c) shows UMAP plots of gene expression for conserved drought-responsive gene families PP2C (Protein Phosphatase 2C) under normal and drought conditions. Figure (d) presents a

correlation analysis of gene expression in conserved drought-regulatory mechanisms among related gene families. Each datapoint represents the expression of a given gene family in one species under a specific condition (control or drought). For each species, two groups of values are shown, corresponding to expression under normal (well-watered) and drought conditions. Gene expression levels are quantified as FPKM. Lines indicate the fitted linear regression, and shaded bands indicate the 95% confidence interval. Spearman correlation coefficients and two-sided p values are shown. Effect sizes are reported as Spearman’s rho (ρ), with 95% confidence intervals estimated by Fisher’s z transformation. Exact test statistics, confidence intervals, and raw and adjusted P values are provided in the Source Data. Figure (e) provides GO annotations for drought-responsive genes within the ERF (Ethylene-responsive transcription factor)-2OGD (left panel) and TCP (Teosinte branched1/Cycloidea/Proliferating cell factor)-PP2C (right panel) regulatory gene families, respectively. Enrichment significance was assessed using Fisher’s exact test, and the adjusted P values shown were corrected for multiple comparisons using the Benjamini-Hochberg method.”

Supporting information:



“Supplementary Figure 11 Chi-squared test results corresponding to the contingency tables derived from Fig. 5a. There are two rice subspecies, and we merged them into a single set by taking unique gene families. A and B denote gene families that are uniquely drought-responsive in each of the two species being compared, and $A \cap B$ denotes the gene families that are drought-responsive in both species. “1” represents the complete background set of gene families across all five species, and $1 - (A \cup B)$ corresponds to the background gene families that are not drought-responsive in either of the two species in that comparison. The test statistic was χ^2 with 1 degree of freedom; exact P values are provided in the Source Data.”

Fig 5b: The color coding for this subfigure does not have sufficient contrast between categories, and the labeling scheme needs to be revised so it can be more easily interpreted without repeated reference back to the caption. For example, since “4-1” represents indica-japonica-wheat-maize, the grouping might be better labeled as IJWM. This would also allow the figure caption to be substantially simplified. It is also difficult

represents japonica-sorghum-maize, JWM represents japonica-wheat-maize, IJW represents indica-japonica-wheat, SWM represents sorghum-wheat-maize, IJ represents indica-japonica, IM represents indica-maize, IW represents indica-wheat, SW represents sorghum-wheat, SM represents sorghum-maize, JM represents japonica-maize, JW represents japonica-wheat, JS represents japonica-sorghum, IS represents indica-sorghum, WM represents wheat-maize, J represents japonica, I represents indica, W represents wheat, M represents maize, and S represents sorghum. Figure (c) shows UMAP plots of gene expression for conserved drought-responsive gene families PP2C (Protein Phosphatase 2C) under normal and drought conditions. Figure (d) presents a correlation analysis of gene expression in conserved drought-regulatory mechanisms among related gene families. Each datapoint represents the expression of a given gene family in one species under a specific condition (control or drought). For each species, two groups of values are shown, corresponding to expression under normal (well-watered) and drought conditions. Gene expression levels are quantified as FPKM. Lines indicate the fitted linear regression, and shaded bands indicate the 95% confidence interval. Spearman correlation coefficients and two-sided p values are shown. Effect sizes are reported as Spearman's rho (ρ), with 95% confidence intervals estimated by Fisher's z transformation. Exact test statistics, confidence intervals, and raw and adjusted P values are provided in the Source Data. Figure (e) provides GO annotations for drought-responsive genes within the ERF (Ethylene-responsive transcription factor)-2OGD (left panel) and TCP (Teosinte branched1/Cycloidea/Proliferating cell factor)-PP2C (right panel) regulatory gene families, respectively. Enrichment significance was assessed using Fisher's exact test, and the adjusted P values shown were corrected for multiple comparisons using the Benjamini-Hochberg method."

Line 513: If the word "significantly" here refers to a statistical test, it would be appropriate to parenthetically mention the test and threshold used. If it does not, another word should be selected.

Response:

Thank you for this helpful comment. We agree that the term "significantly" may imply a statistical test. As the identification of altered interaction patterns in our analysis was based on ranking by drought response magnitude rather than formal statistical testing, we have revised the text to use "most pronounced" instead, which reflects the most prominent changes without implying statistical significance. This clarification ensures that the language accurately represents the level of support provided by the data.

Line 391-394:

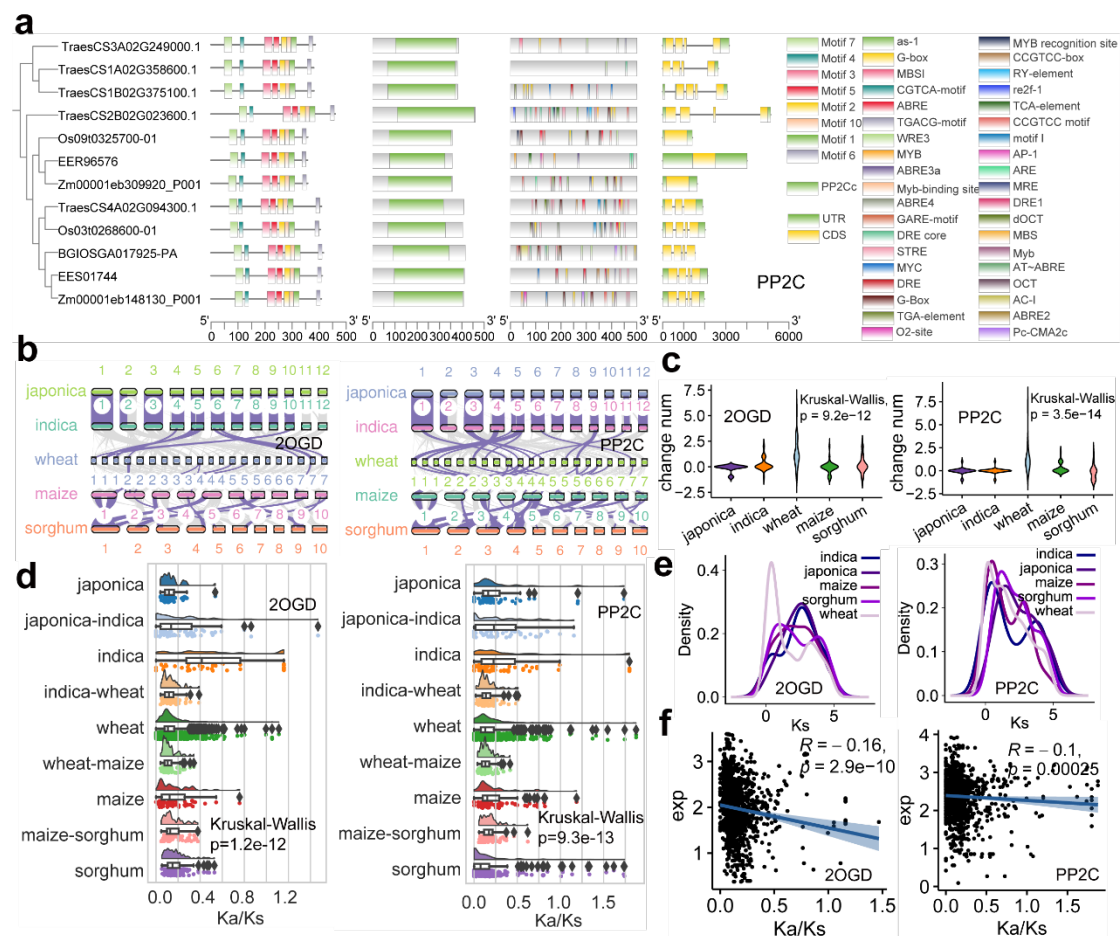
"Further insights into the mechanisms of species divergence were gained through the key interaction relationships of gene families. We analyzed species-specific drought resistance interaction patterns, ranking them by drought response magnitude (Methods), identifying the most pronounced interaction patterns in comparison to other species."

Fig 8b, 8b, 8d, 8e, and 8f: panels should be labeled in the figure, not only in the caption.

Response:

Thank you for this suggestion. We have revised Fig. 8 to add panel labels directly within the figure for 8b, 8c, 8d, 8e, and 8f, so that each panel can be clearly identified without referring back to the caption. We also made the corresponding labeling revisions in Fig. 6 for consistency.

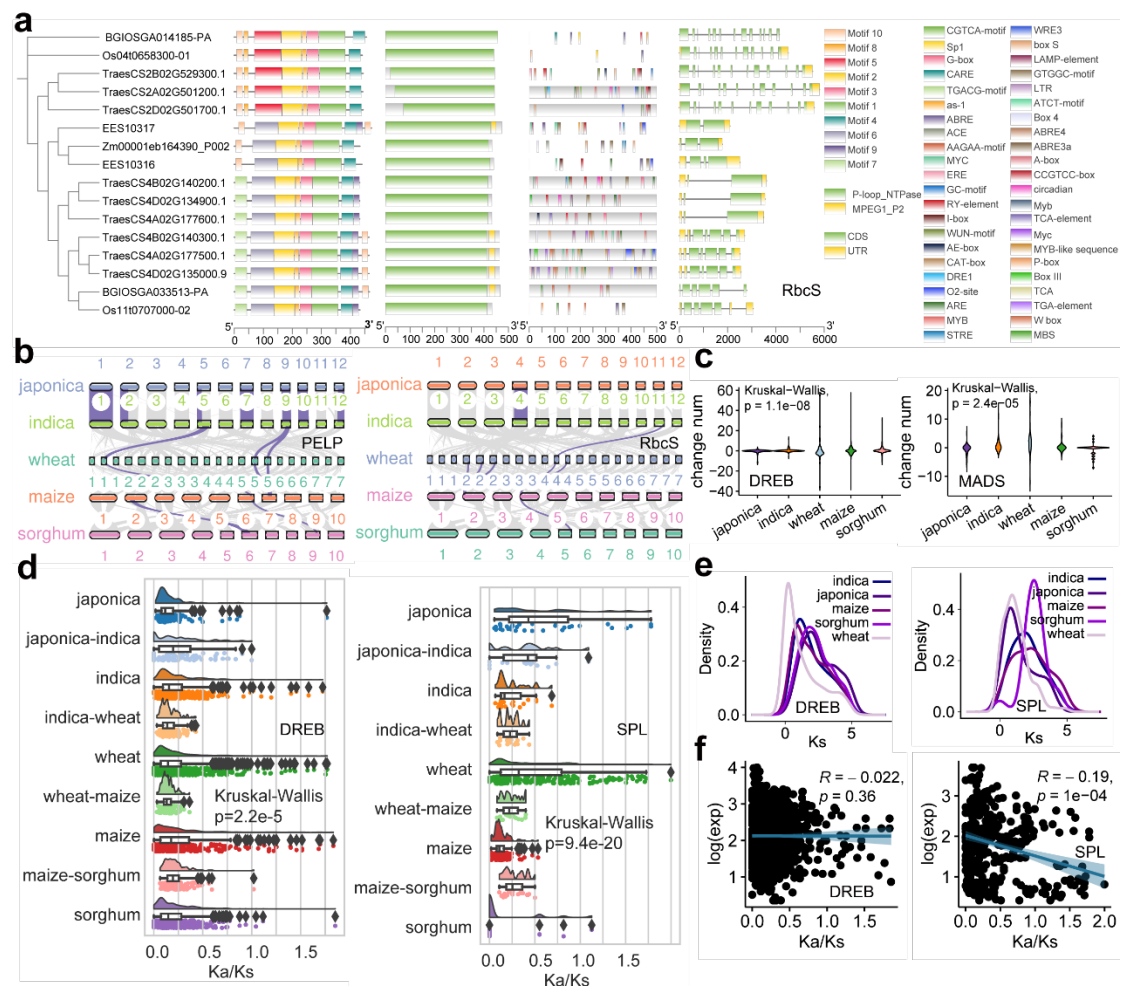
Figure 6:



“Fig. 6 Evolutionary analysis of conserved drought-resistance regulatory gene families. Figure (a) presents a phylogenetic tree for the PP2C gene family within the drought-regulatory network, along with motifs, domains, cis-regulatory elements, and gene structures. Figure (b) illustrates the collinear positions of the 2OGD (left panel) and PP2C (right panel) gene families across species. Figure (c) shows the distribution of gene contraction and expansion counts for the 2OGD (left panel) and PP2C (right panel) families based on clustering results. Violin plots show distributions, centre dots indicate medians, and vertical lines indicate minima and maxima. Counts were compared among species using two-sided Kruskal-Wallis tests followed by two-sided Mann-Whitney U tests for pairwise comparisons; Benjamini-Hochberg and Bonferroni corrections were

applied within each family. Exact raw and adjusted P values are provided in the Source Data. Figure (d) highlights the distribution of Ka/Ks values for both orthologs (one-to-one genes between species within the same orthogroup) and paralogs (multiple genes from the same species within the same orthogroup) in the 2OGD (left panel) and PP2C (right panel) families, providing insights into their divergence. Horizontal box plots show medians (centre line), interquartile ranges (box), $1.5 \times \text{IQR}$ whiskers, and individual data points. Groups were compared using Kruskal-Wallis tests followed by two-sided Mann-Whitney U tests with Benjamini-Hochberg and Bonferroni correction. Exact raw and adjusted P values are provided in the Source Data. Figure (e) displays the distribution of synonymous substitution rates (Ks) values for 2OGD and PP2C families within homolog clusters, providing insights into evolutionary timescales and species-specific differences. Figure (f) explores the correlation between Ka/Ks values and gene expression levels ($\log(\text{FPKM})$) for drought-responsive 2OGD (left panel) and PP2C (right panel) family genes across the five species, indicating evolutionary pressure and functional significance. Lines indicate fitted linear regressions and shaded bands indicate 95% confidence intervals. Associations were tested using two-sided Spearman rank correlation; effect sizes are reported as Spearman's ρ with 95% confidence intervals estimated by Fisher's z transformation. Exact raw and adjusted P values are provided in the Source Data."

Figure 8:



“Fig. 8 The evolutionary analysis of differentiated drought-regulatory gene families. Figure (a) illustrates the phylogenetic tree of the RbcS gene family (Ribulose biphosphate carboxylase/oxygenase activase-like) within the drought-regulatory network, providing detailed information on motifs, domains, cis-acting elements, and gene structures. Figure (b) shows the syntenic positions of the PELP (left panel) and RbcS (right panel) gene families across various species, highlighting their conserved genomic locations. Figure (c) presents the expansion and contraction distributions of the DREB (left panel) and MADS (right panel) gene families based on gene family clustering. Violin plots show distributions, centre dots indicate medians, and vertical lines indicate minima and maxima. Counts were compared among species using two-sided Kruskal-Wallis tests followed by two-sided Mann-Whitney U tests for pairwise comparisons; Benjamini-Hochberg and Bonferroni corrections were applied within each family. Exact raw and adjusted P values are provided in the Source Data. Figure (d) illustrates the Ka/Ks distribution for both orthologs (one-to-one genes between species within the same orthogroup) and paralogs (multiple genes from the same species within the same orthogroup) in the DREB (left panel), SPL (right panel) gene families. Horizontal box plots show medians (centre line), interquartile ranges (box), $1.5 \times$ IQR whiskers, and individual data points. Groups were compared using Kruskal-Wallis tests followed by two-sided Mann-Whitney U tests with Benjamini-Hochberg and Bonferroni correction. Exact raw and adjusted P values are provided in the Source

Data. Figure (e) displays the distribution of synonymous substitution rates (K_s) values for DREB and SPL families within homolog clusters, offering insights into the selective pressures shaping their evolution. Figure (f) examines the relationship between K_a/K_s values and gene expression levels ($\log(\text{FPKM})$) for drought-responsive DREB and SPL family genes across the five species. Lines indicate fitted linear regressions and shaded bands indicate 95% confidence intervals. Associations were tested using two-sided Spearman rank correlation; effect sizes are reported as Spearman's ρ with 95% confidence intervals estimated by Fisher's z transformation. Exact raw and adjusted P values are provided in the Source Data."

625-637: This is a very interesting result, and the kind of discussion I think could be beneficial to introduce in an abbreviated form at lines 289-296 and 340 (see comments there) to contextualize the results and help convey their importance to the overall conclusions of the paper.

Response:

Thank you for this helpful comment. We have incorporated an abbreviated discussion of these results at lines 289–296 and 340 to provide additional context and highlight their relevance to the overall conclusions of the study.

Line 240-251:

"Among the compared network centrality measures, the drought group generally exhibited significantly higher values than the normal group based on the mean values, with the difference being statistically significant ($p = 0.014$) (Fig. 3b, Supplementary Fig. 20e). These results suggest that drought stress is accompanied by increased regulatory connectivity at the network level. In particular, a higher average degree indicates that genes are connected by more regulatory interactions on average under drought conditions. Further analysis of the average degree changes under drought conditions across different species' networks revealed significant differences among species ($p = 1.6e-43$). Notably, rice, potato and soybean exhibited a greater magnitude of change compared to the other species, while maize and sorghum tended to show smaller fluctuations (Supplementary Fig. 5, Fig. 3c, Supplementary Fig. 6, Supplementary Fig. 20f)."

Line 282-286:

"The degree distribution curves show steeper declines in representative C_3 crops compared to representative C_4 crops ($p < 0.05$), reflecting differences in network connectivity and regulation (Fig. 4d), suggesting that representative C_3 networks rely more on central hub genes while representative C_4 networks maintain more distributed connectivity."

The code is well-documented and organized. It is possible to reproduce the results of the paper with the code provided, and the procedure to create a working environment in which to run the code is clear. The authors have provided an exceptionally transparent codebase for this work.

Response:

We sincerely thank the reviewer for the positive feedback regarding our code. We are glad that the documentation, organization, and reproducibility of the code are clear and useful. We appreciate the recognition of our efforts to provide a transparent and accessible codebase for this work.