

## Supplementary Information

The following Supporting Information is available for this article:

**Supplementary Table 1** The performance metrics (AUC and AUPR) for different models across different species, with results presented for transcriptional regulatory networks (TRN)-grey, protein-protein interaction networks (PPI)-blue, and KEGG pathway networks-green.

| species    | Japonica |      | Indica |      | Wheat |      | Maize |      | Sorghum |      |
|------------|----------|------|--------|------|-------|------|-------|------|---------|------|
| model      | AUC      | AUPR | AUC    | AUPR | AUC   | AUPR | AUC   | AUPR | AUC     | AUPR |
| GENIE3     | 0.50     | 0.50 | 0.50   | 0.50 | 0.50  | 0.49 | 0.50  | 0.49 | 0.52    | 0.52 |
| GNE        | 0.92     | 0.92 | 0.91   | 0.93 | 0.85  | 0.82 | 0.88  | 0.80 | 0.83    | 0.82 |
| CNNC       | 0.75     | 0.69 | 0.91   | 0.90 | 0.84  | 0.82 | 0.75  | 0.69 | 0.84    | 0.83 |
| GENELink   |          |      |        |      |       |      |       |      |         |      |
| GNNLink    | 0.75     | 0.69 | 0.75   | 0.69 | 0.82  | 0.86 | 0.74  | 0.68 | 0.76    | 0.66 |
| DroGeneNet | 0.94     | 0.93 | 0.94   | 0.93 | 0.89  | 0.88 | 0.86  | 0.84 | 0.85    | 0.83 |
| GENIE3     | 0.53     | 0.52 | 0.55   | 0.54 | 0.57  | 0.55 | 0.60  | 0.54 | 0.57    | 0.52 |
| GNE        | 0.90     | 0.86 | 0.90   | 0.86 | 0.88  | 0.86 | 0.87  | 0.83 | 0.87    | 0.85 |
| CNNC       | 0.76     | 0.71 | 0.75   | 0.69 | 0.78  | 0.70 | 0.79  | 0.77 | 0.76    | 0.72 |
| GENELink   | 0.91     | 0.87 | 0.92   | 0.87 | 0.90  | 0.85 | 0.90  | 0.79 | 0.89    | 0.81 |
| GNNLink    | 0.75     | 0.69 | 0.75   | 0.69 | 0.76  | 0.68 | 0.75  | 0.69 | 0.75    | 0.69 |
| DroGeneNet | 0.93     | 0.88 | 0.93   | 0.88 | 0.97  | 0.97 | 0.90  | 0.84 | 0.94    | 0.89 |
| GENIE3     | 0.52     | 0.50 | 0.54   | 0.52 | 0.54  | 0.52 | 0.56  | 0.54 | 0.54    | 0.50 |
| GNE        | 0.85     | 0.88 | 0.88   | 0.90 | 0.88  | 0.87 | 0.77  | 0.79 | 0.82    | 0.77 |
| CNNC       | 0.74     | 0.71 | 0.79   | 0.77 | 0.79  | 0.72 | 0.75  | 0.69 | 0.77    | 0.72 |
| GENELink   | 0.94     | 0.94 | 0.95   | 0.95 | 0.94  | 0.91 | 0.92  | 0.92 | 0.85    | 0.84 |
| GNNLink    | 0.75     | 0.69 | 0.75   | 0.69 | 0.75  | 0.68 | 0.76  | 0.70 | 0.74    | 0.69 |
| DroGeneNet | 0.96     | 0.96 | 0.96   | 0.96 | 0.97  | 0.96 | 0.93  | 0.93 | 0.89    | 0.88 |

**Supplementary Table 2** The model is tested in TRN under different conditions.

| treat                                | Japonica |       | Indica |       | Wheat |       | Maize |       | sorghum |       |
|--------------------------------------|----------|-------|--------|-------|-------|-------|-------|-------|---------|-------|
|                                      | AUC      | AUPR  | AUC    | AUPR  | AUC   | AUPR  | AUC   | AUPR  | AUC     | AUPR  |
| DroGeneNet                           | 0.94     | 0.93  | 0.94   | 0.93  | 0.89  | 0.88  | 0.86  | 0.84  | 0.85    | 0.83  |
| DroGeneNet<br>- spatial<br>attention | 0.932    | 0.929 | 0.934  | 0.929 | 0.851 | 0.831 | 0.879 | 0.867 | 0.828   | 0.824 |

|                              |       |       |       |       |       |       |       |       |       |       |
|------------------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| DroGeneNet<br>-gate residual | 0.937 | 0.932 | 0.937 | 0.932 | 0.853 | 0.827 | 0.889 | 0.875 | 0.823 | 0.816 |
| Genenum-<br>2000             | 0.915 | 0.91  | 0.896 | 0.887 | 0.856 | 0.825 | 0.845 | 0.831 | 0.824 | 0.815 |
| Genenum-<br>5000             | 0.916 | 0.91  | 0.921 | 0.915 | 0.861 | 0.834 | 0.861 | 0.843 | 0.836 | 0.832 |
| Genenum-<br>10000            | 0.925 | 0.919 | 0.929 | 0.925 | 0.860 | 0.820 | 0.861 | 0.844 |       |       |
| dot                          | 0.939 | 0.935 | 0.94  | 0.934 | 0.861 | 0.831 | 0.9   | 0.888 | 0.829 | 0.824 |
| cosine                       | 0.926 | 0.917 | 0.927 | 0.91  | 0.849 | 0.834 | 0.898 | 0.889 | 0.824 | 0.828 |
| mlp                          | 0.909 | 0.9   | 0.916 | 0.907 | 0.856 | 0.830 | 0.872 | 0.857 | 0.822 | 0.830 |
| adj                          | 0.939 | 0.935 | 0.94  | 0.934 | 0.856 | 0.832 | 0.9   | 0.888 | 0.834 | 0.826 |
| adj-drop50                   | 0.936 | 0.934 | 0.938 | 0.935 | 0.852 | 0.817 | 0.9   | 0.891 | 0.838 | 0.814 |
| adj-weight60                 | 0.938 | 0.935 | 0.94  | 0.936 | 0.844 | 0.819 | 0.897 | 0.885 | 0.836 | 0.826 |

**Supplementary Table 3** Core genes indicated by attention scores and their linking number, degree distribution in japonica.

| Gene         | Attention score | Linking | Degree |
|--------------|-----------------|---------|--------|
| Os07g0236700 | 690.62          | 7578    | 0.38   |
| Os05g0121400 | 688.70          | 9090    | 0.46   |
| Os03g0821200 | 631.05          | 6286    | 0.32   |

**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 <sup>2</sup> | p value   | std err | class          |
|----------|-------|-----------|---------|----------------|-----------|---------|----------------|
| japonica | -0.79 | 7.52      | -0.94   | 0.88           | 1.31E-92  | 0.02    | hormone        |
| indica   | -0.82 | 7.12      | -1.00   | 0.99           | 1.04E-206 | 0.01    | hormone        |
| wheat    | -0.83 | 6.98      | -0.99   | 0.98           | 6.29E-182 | 0.01    | hormone        |
| maize    | -0.77 | 6.54      | -0.99   | 0.98           | 3.68E-166 | 0.01    | hormone        |
| sorghum  | -0.98 | 4.79      | -0.97   | 0.94           | 5.70E-54  | 0.03    | hormone        |
| japonica | -0.96 | 7.8       | -0.93   | 0.86           | 2.41E-68  | 0.03    | photosynthesis |
| indica   | -0.62 | 4.16      | -0.67   | 0.44           | 3.00E-21  | 0.06    | photosynthesis |
| wheat    | -0.83 | 6.98      | -0.99   | 0.98           | 2.73E-132 | 0.01    | photosynthesis |
| maize    | -0.76 | 6.5       | -0.99   | 0.97           | 1.75E-120 | 0.01    | photosynthesis |
| sorghum  | -1.06 | 5.9       | -0.98   | 0.95           | 4.61E-104 | 0.02    | photosynthesis |

**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    |

**Supplementary Table 6.** The number of transcriptomes, genes, and regulatory relationships at different levels across different species. TRN represents the transcriptional regulatory layer, PPI represents the protein-protein interaction layer, and KEGG represents the metabolic regulatory layer. Gene1 represents the source node, while Gene2 represents the target node.

| type | species  | RNA-Seq | Gene1 | Gene2 | Regulatory |
|------|----------|---------|-------|-------|------------|
| TRN  | japonica | 1547    | 184   | 20303 | 228994     |
|      | indica   | 655     | 182   | 22658 | 230777     |
|      | wheat    | 733     | 102   | 26622 | 717454     |
|      | maize    | 1536    | 112   | 13431 | 51095      |
|      | sorghum  | 580     | 68    | 5730  | 382982     |
| PPI  | japonica | 1547    | 8462  | 8227  | 272604     |
|      | indica   | 655     | 9583  | 9393  | 313097     |
|      | wheat    | 733     | 13838 | 13838 | 824736     |
|      | maize    | 1536    | 2805  | 2768  | 28390      |
|      | sorghum  | 580     | 2988  | 2990  | 50341      |
| KEGG | japonica | 1547    | 1384  | 1404  | 16476      |
|      | indica   | 655     | 1472  | 1503  | 17966      |
|      | wheat    | 733     | 4452  | 4550  | 157573     |
|      | maize    | 1536    | 845   | 870   | 5713       |
|      | sorghum  | 580     | 940   | 975   | 7091       |

**Supplementary Table 7** HMM model numbers corresponding to each gene family.

| Gene family                                         | Gene family abbreviation | PFAM ID |
|-----------------------------------------------------|--------------------------|---------|
| Protein phosphatase 2C                              | PP2C                     | PF00481 |
| Plant 2-oxoglutarate-dependent oxidoreductases      | 2OGD                     | PF03171 |
| Dehydration-responsive element-binding protein 1A-I | DREB(ERF)                | PF00847 |
| Transcription factor RADIALIS-like                  | RAD(MYB)                 | PF00249 |
| Squamosa promoter-binding-like protein              | SPL                      | PF03110 |
| Pistil-specific extensin-like protein               | PELP                     | PF01190 |
| Photosystem II Psb28, class 1                       | Psb28                    | PF03912 |
| Asparagine synthase, glutamine-hydrolyzing          | ASN                      | PF00733 |
| Transcriptional activator Hac1/HY5                  | Hac1                     | PF07716 |

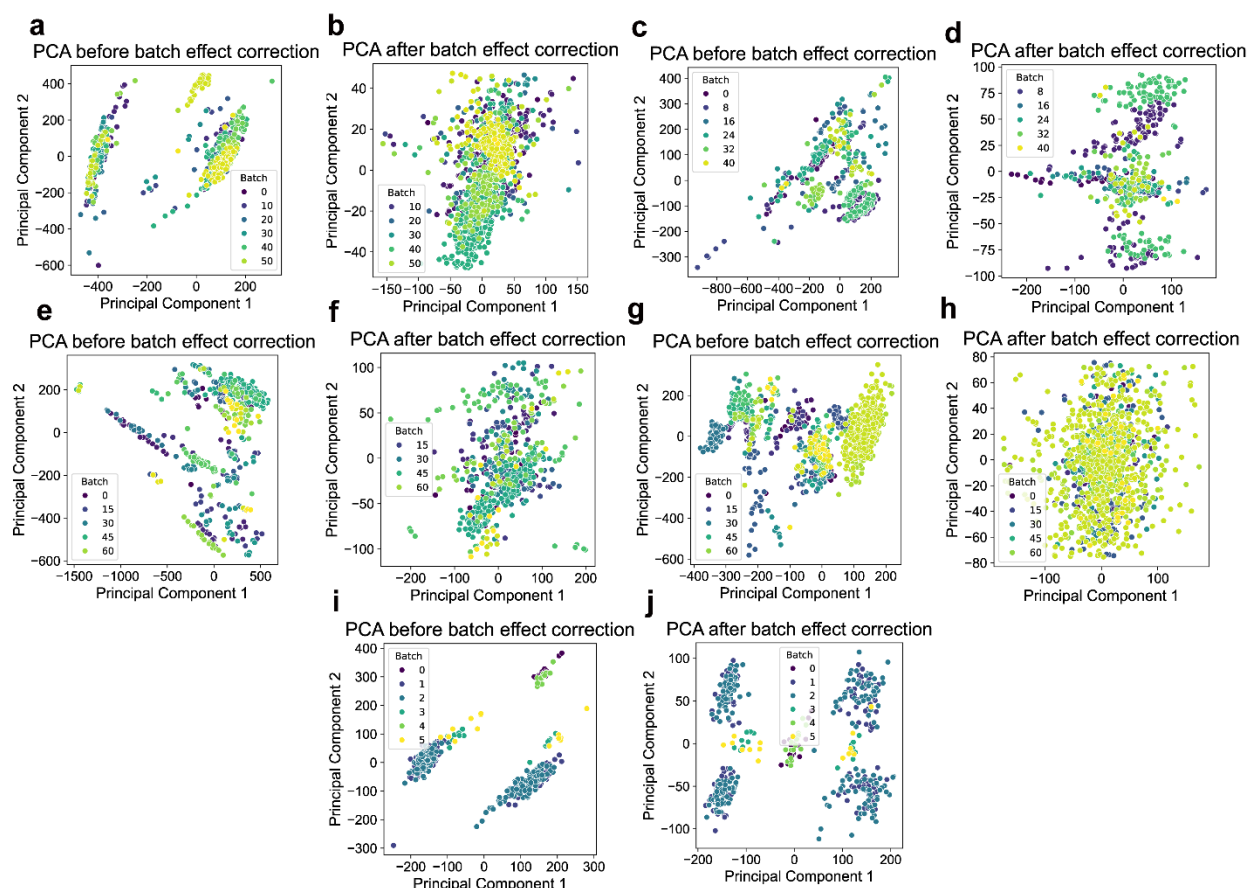
|                                                                |      |         |
|----------------------------------------------------------------|------|---------|
| Ribulose biphosphate<br>carboxylase/oxygenase<br>activase-like | RbcS | PF12338 |
|----------------------------------------------------------------|------|---------|

**Supplementary Table 8** All metadata for the data and all biological information data associated with the images.

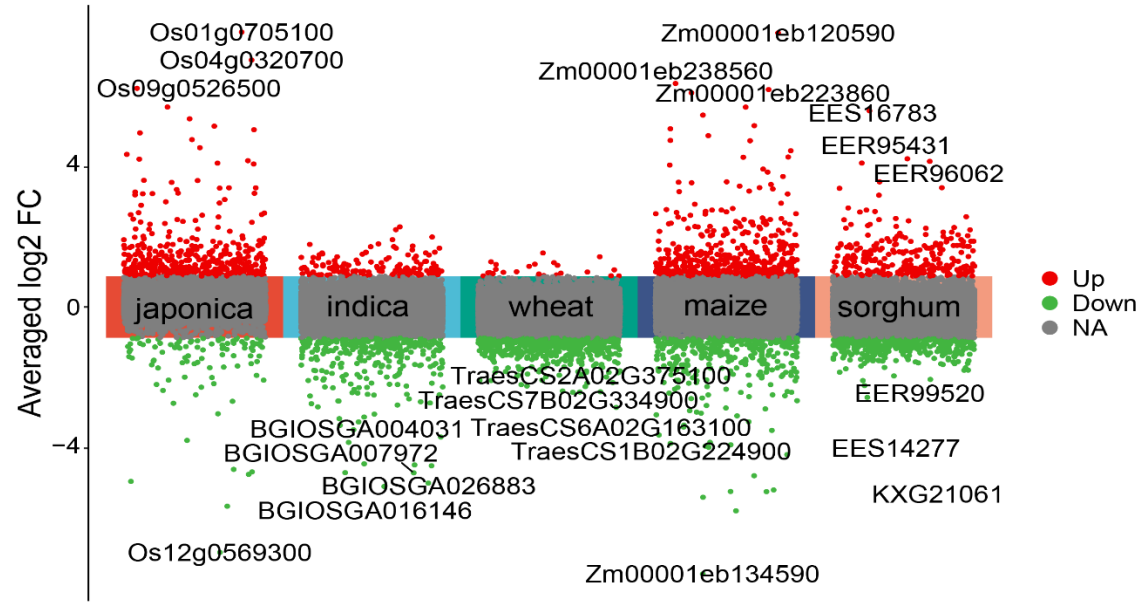
**Supplementary Note 1** As shown in Supplementary Table 1, the DroGeneNet model consistently achieves the highest accuracy across all levels of gene regulatory networks in the three species. For transcriptional regulatory networks, DroGeneNet demonstrated high performance across all species. A key limitation of GENELink is its inability to handle large datasets due to memory constraints, an issue that DroGeneNet effectively resolves. Overall, DroGeneNet demonstrates high accuracy and stability in constructing gene regulatory networks, regardless of the network type or species. Additionally, Supplementary Table 2 presents results from various tests conducted on the DroGeneNet model, including module ablation studies, tests with varying gene quantities, perturbation of the prior network, and different methods for calculating gene pair interaction probabilities. As observed from Supplementary Table 2, the integration of topological attributes into the attention matrix is critical for model performance, as indicated by the significant drop in accuracy when these features were removed. The accuracy of DroGeneNet improves as the number of genes increases, suggesting a scaling effect that enhances the robustness of the model. The dot product method provided the best performance in calculating gene pair interaction probabilities, outperforming cosine similarity and MLP methods. This is likely due to the ability of dot product to directly reflect the shared properties of two vectors in terms of both direction and magnitude. When using a prior network with relatively high noise levels, the average AUC and AUPR scores decreased only slightly, suggesting that the gene expression data contains most of the necessary information for inferring the GRN.

**Supplementary Note 2** DroGeneNet offers several key advantages over traditional and deep learning-based gene regulatory network (GRN) inference models. Compared to the traditional regression tree-based method GENIE3, deep learning models offer superior learning and fitting capabilities. Unlike CNNC, a model based on traditional image processing techniques, DroGeneNet not only considers the similarity information of gene expression but also accounts for the complex topological relationships in gene regulation. In contrast to the GNE model, which converts spatial information directly into feature vectors for concatenation, DroGeneNet can directly incorporate neighboring gene expression data into its model, providing a more integrated approach. When compared to GENELink and GNNLink, both of which rely on traditional graph attention and graph convolution, DroGeneNet goes further by incorporating spatial information at multiple levels. This includes encoding node degree centrality, mapping edge features based on source and target gene expression, and utilizing global topological information through shortest path calculations. This multi-dimensional spatial integration helps DroGeneNet avoid false correlations that can arise when only gene expression similarity is considered, as seen in co-expression networks like WGCNA. Moreover, compared to traditional GNNs, DroGeneNet introduces a novel approach to large-scale graph processing. Instead of loading the entire adjacency matrix, which can be computationally expensive for large gene networks, DroGeneNet converts the matrix into an edge index format. This makes the model

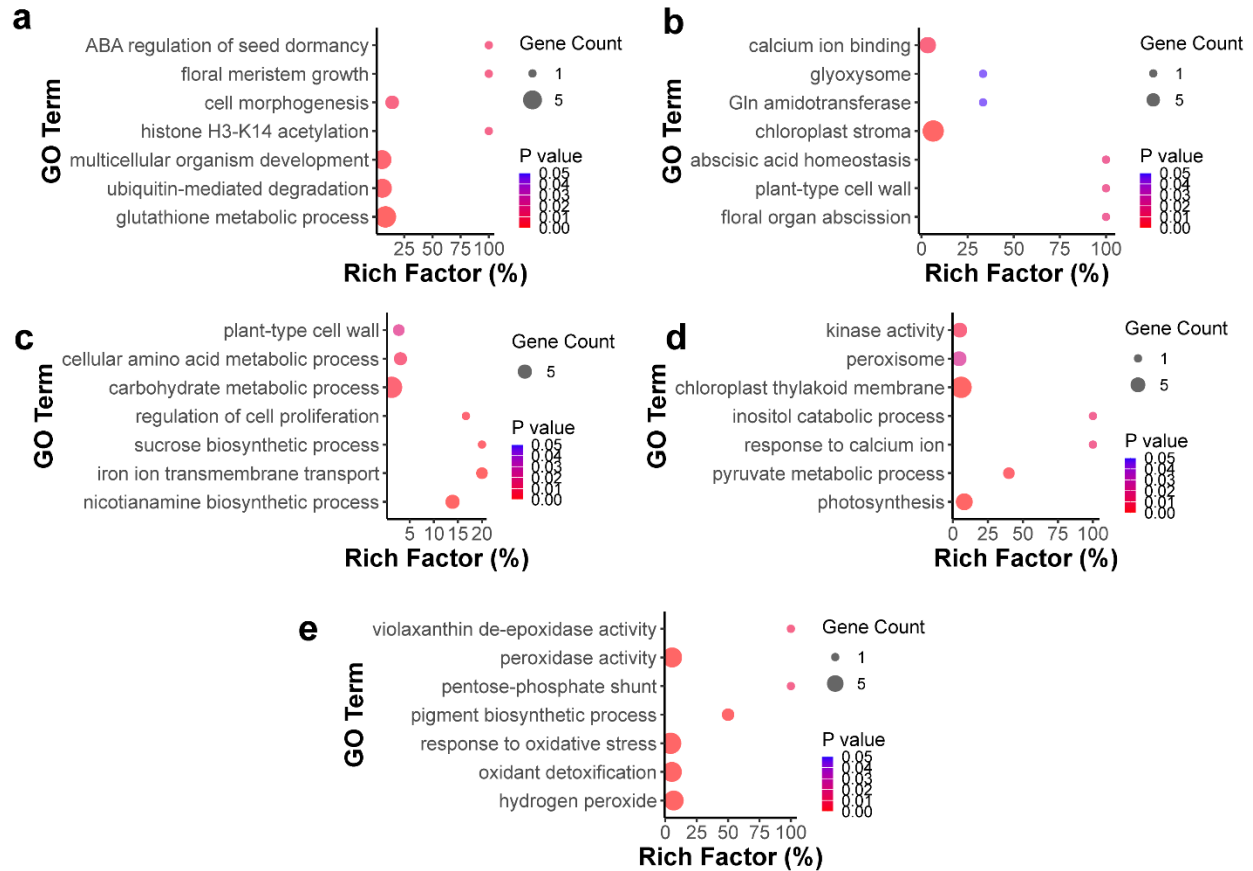
scalable, allowing it to handle large gene regulatory networks efficiently, even as the number of genes increases. One of the standout features of DroGeneNet is its ability to handle gene expression data across different experimental conditions. By setting separate encoding layers for each condition, the model allows joint training while maintaining independent inference capabilities for each network. This enables DroGeneNet to generate gene regulatory networks that specifically respond to environmental changes, such as varying water conditions. In summary, DroGeneNet represents a significant advancement over existing methods by incorporating global network connectivity, efficiently processing large-scale networks, and providing a more accurate and scalable approach to GRN prediction. Its ability to respond to environmental changes further enhances its practical utility in understanding gene regulation under varying conditions.



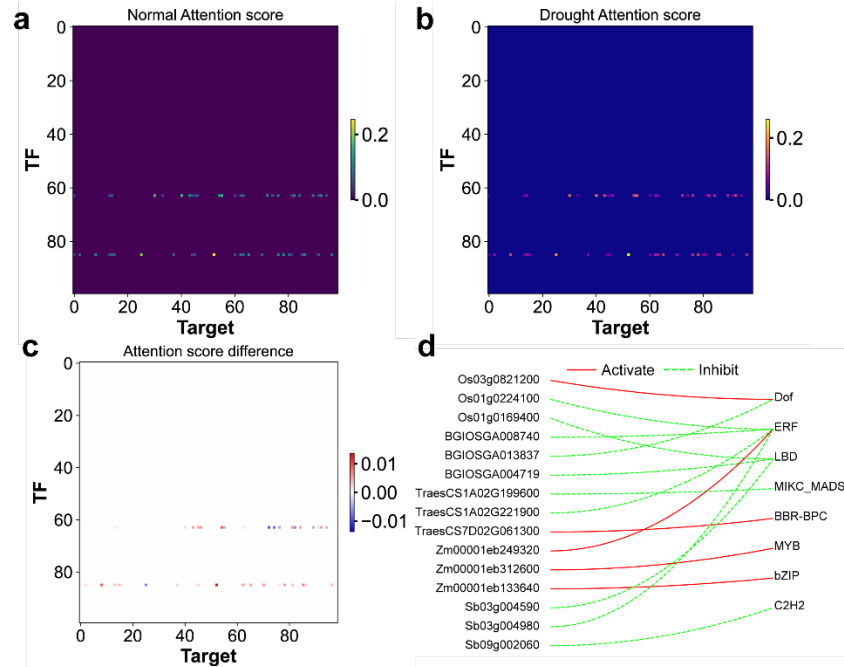
**Supplementary Figure 1** PCA plots of batch-corrected transcriptome data before and after correction. Panels a-b represent japonica, c-d represent indica, e-f represent wheat, g-h represent maize, and i-j represent sorghum. Each point represents one sample, and colours indicate different batches.



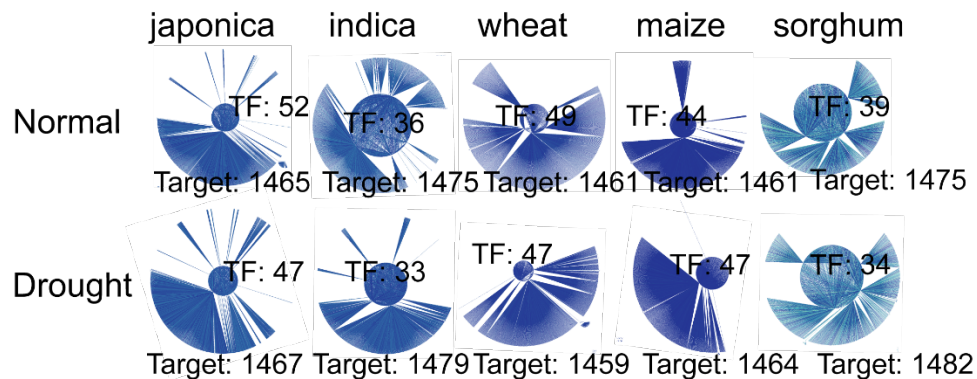
**Supplementary Figure 2** A volcano plot of differentially expressed genes (DEGs) under drought and normal conditions across different species. Red and green dots indicate upregulated and downregulated genes, respectively, and grey dots indicate non-significant genes. Differential expression analysis was performed using DESeq2, with significance defined as adjusted  $q$  value  $< 0.05$  and  $|\log_2FC| > 1.3$ .



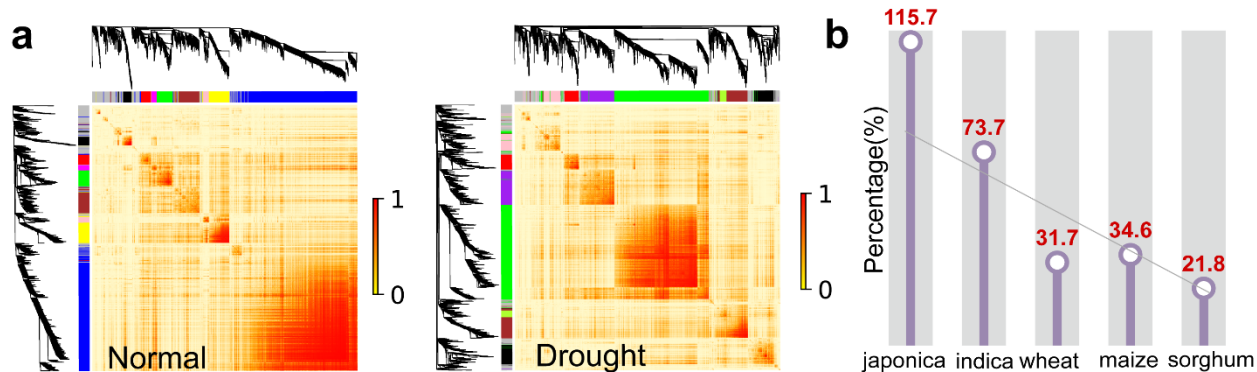
**Supplementary Figure 3** Gene annotation of differentially expressed genes with no orthologs across five species. Panels a-e represent japonica, indica, wheat, maize, and sorghum, respectively. GO enrichment analysis was performed using Enrichr implemented in Python (gseapy), based on Fisher's exact test with Benjamini-Hochberg correction for multiple comparisons. Dot size indicates gene count and colour indicates the adjusted P value.



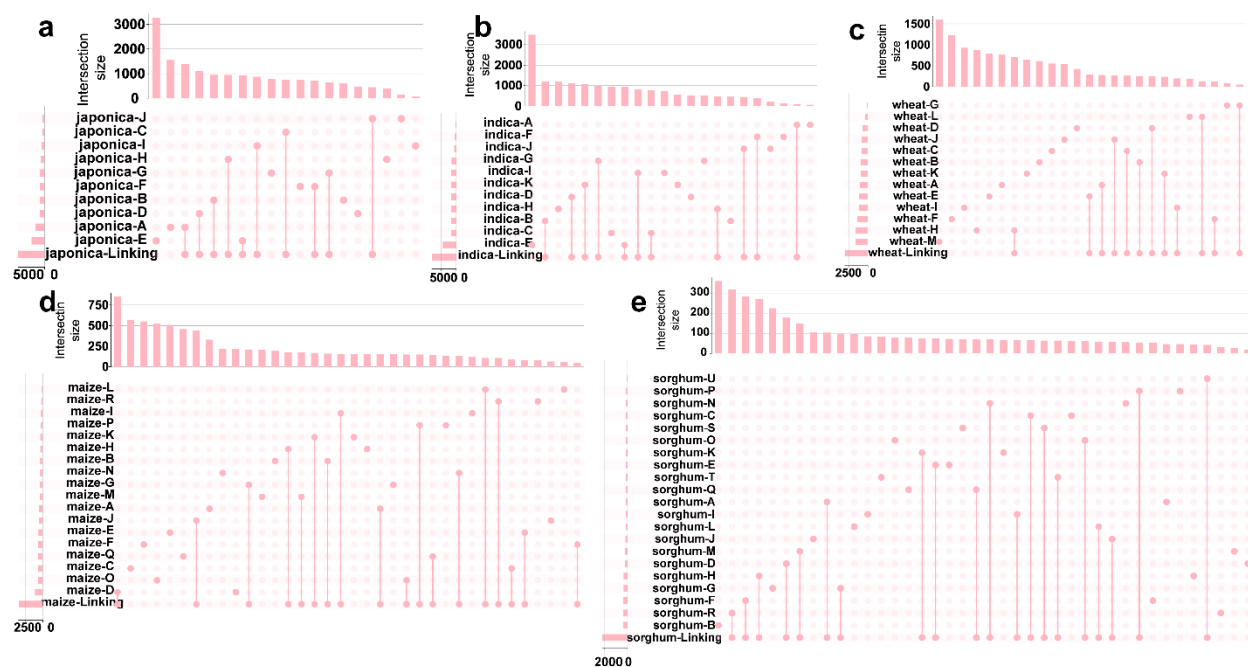
**Supplementary Figure 4** The core regulatory genes explained by the attention scores of the model. Figure (a) shows the network attention scores under normal conditions. Figure (b) shows the network attention scores under drought conditions. Higher x-axis and y-axis values indicate greater confidence 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.



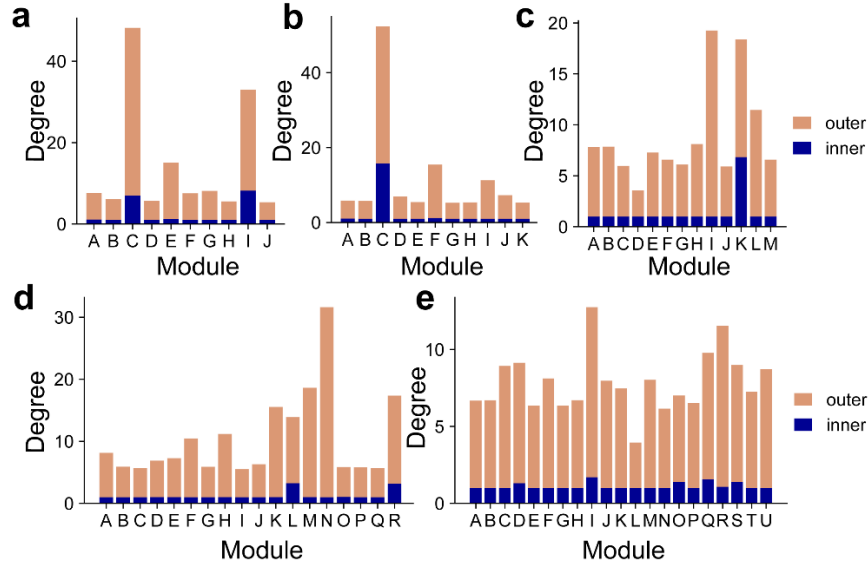
**Supplementary Figure 5** Schematic subnetwork of the top 1,500 genes showing the largest changes in degree centrality under drought stress. Transcription factors are arranged in the inner circle and target genes in the outer circle. Species are shown from left to right as japonica, indica, wheat, maize,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



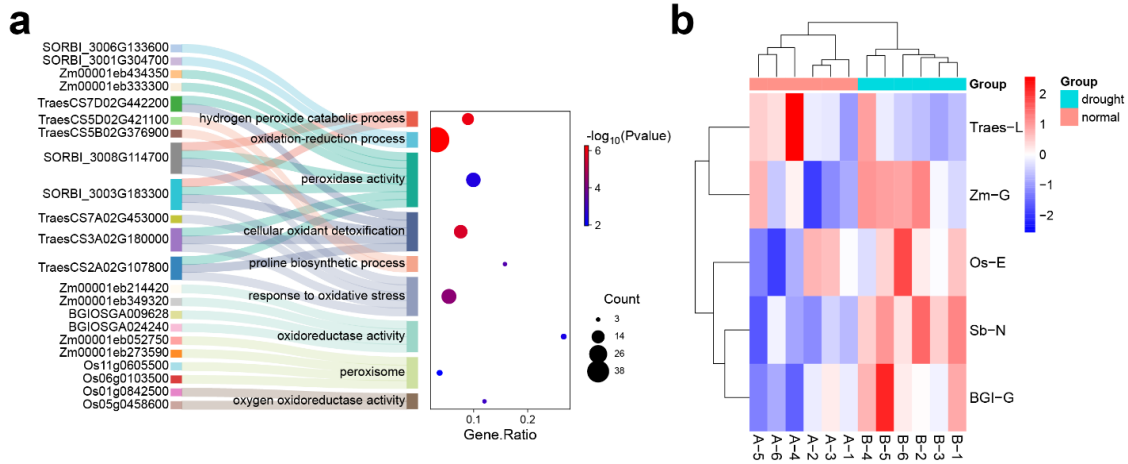
**Supplementary Figure 6** Clustering of sorghum co-expression networks and cross-species changes in average degree under drought stress. Figure (a) provides a schematic clustering of the co-expression network (GCN) in sorghum under normal and drought conditions. Colour scales indicate pairwise co-expression similarity. Figure (b) illustrates the percentage distribution of changes in average degree within co-expression networks across species under drought stress. Red numbers indicate the maximum absolute percentage change for each species.



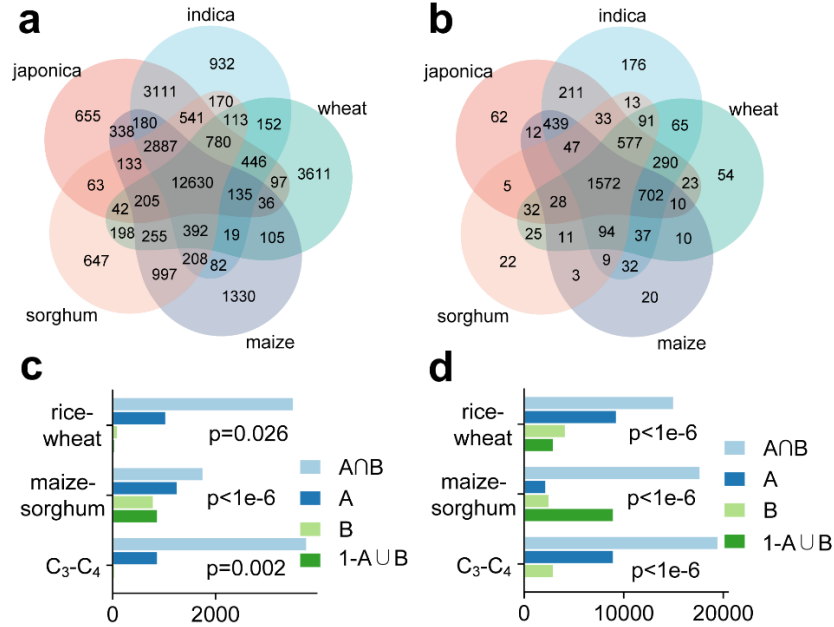
**Supplementary Figure 7** UpSet plots of intersecting gene module counts across species. a–e, UpSet plots for japonica (a), indica (b), wheat (c), maize (d), and sorghum (e). Vertical bars indicate the size of each module intersection, connected dots indicate the module combinations included in each intersection, and horizontal bars indicate the total size of each module.



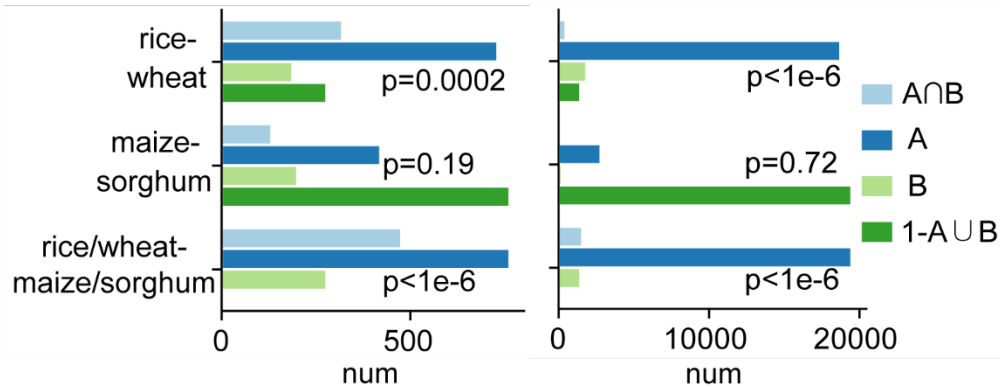
**Supplementary Figure 8** Connectivity degrees across modules in different species. a–e, Connectivity degrees across modules in japonica (a), indica (b), wheat (c), maize (d), and sorghum (e). Blue and orange bars indicate inner modules (without external connections) and outer modules (connected to other modules), respectively. Bar heights represent module connectivity degree.



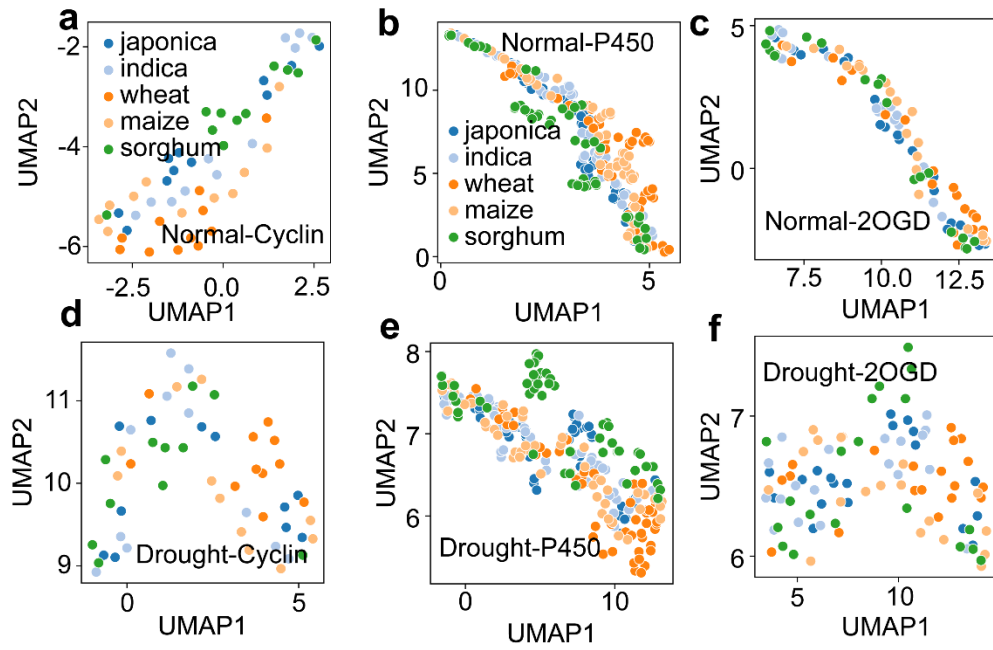
**Supplementary Figure 9** Gene annotation and expression patterns of the antioxidant and hormone modules. Figure (a) highlights GO annotations and core responsive genes for modules associated with hydrogen peroxide response and redox processes. Enrichment significance was assessed using Fisher's exact test with Benjamini–Hochberg correction for multiple comparisons. Dot size indicates gene count and colour indicates the adjusted P value. Figure (b) presents the expression of core genes within hormone-responsive modules, including Os01g0224100 (Os-E), BGIOSGA011298 (BGI-G), TraesCS4D02G026100 (Traes-L), Zm00001eb249320 (Zm-G), and EER97011 (Sb-N). Colour scale indicates scaled expression levels, and the top annotation bar indicates sample group.



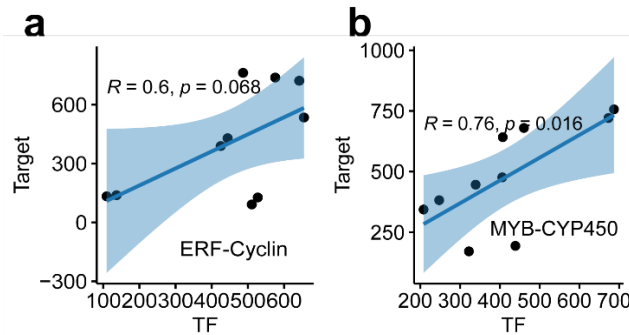
**Supplementary Figure 10** Venn analysis of gene families across different species. Figure a shows the number of gene families across species based on OrthoFinder, while Figure c presents the corresponding Pearson's chi-squared tests on  $2 \times 2$  contingency tables for the indicated comparisons. Figure b shows biologically meaningful gene families annotated by InterPro, while Figure d presents the corresponding Pearson's chi-squared tests on  $2 \times 2$  contingency tables for the indicated comparisons. C<sub>3</sub> denotes rice/wheat lineage, while C<sub>4</sub> denotes maize/sorghum lineage. The test statistic was  $\chi^2$  with 1 degree of freedom; exact P values are provided in the Source Data.



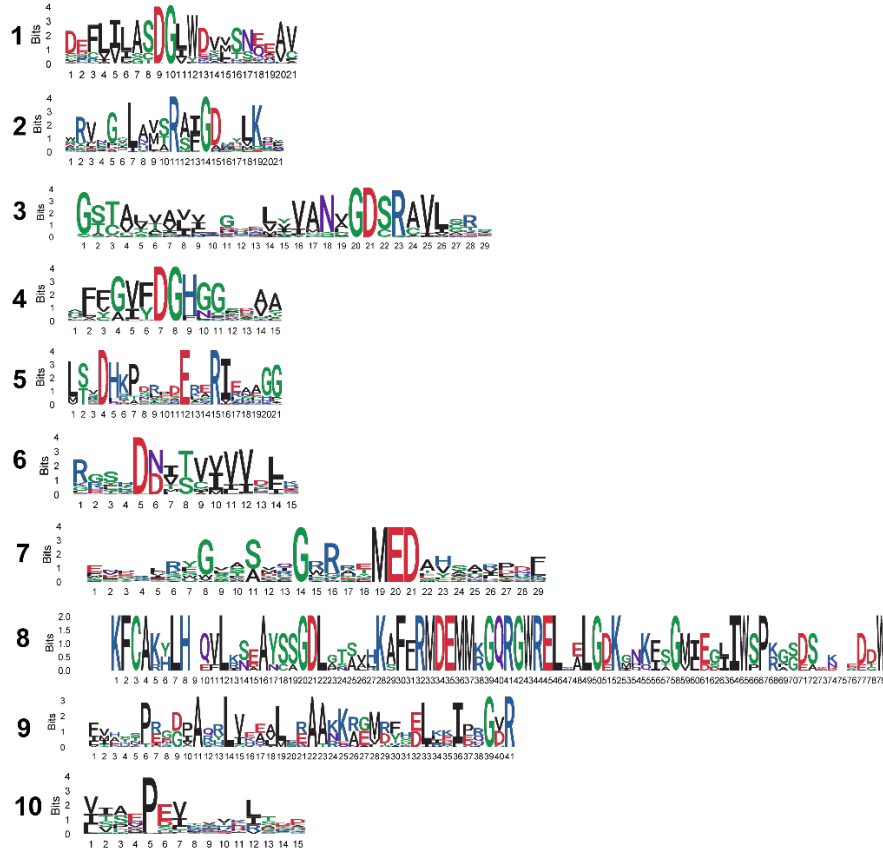
**Supplementary Figure 11** Chi-squared test results corresponding to the contingency tables derived from Fig. 5a. There are two rice subspecies, and we merged them into a single set by taking unique gene families. A and B denote gene families that are uniquely drought-responsive in each of the two species 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  $\chi^2$  with 1 degree of freedom; exact P values are provided in the Source Data.



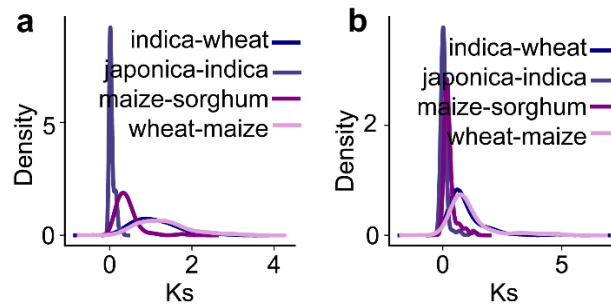
**Supplementary Figure 12** UMAP plots of gene expression for conserved drought-responsive gene families under normal and drought conditions. a–c, UMAP plots of Cyclin, P450 and 2OGD gene families under normal conditions. d–f, UMAP plots of Cyclin, P450 and 2OGD gene families under drought conditions. Each point represents one gene, and colours indicate species identity (japonica, indica, wheat, maize and sorghum).



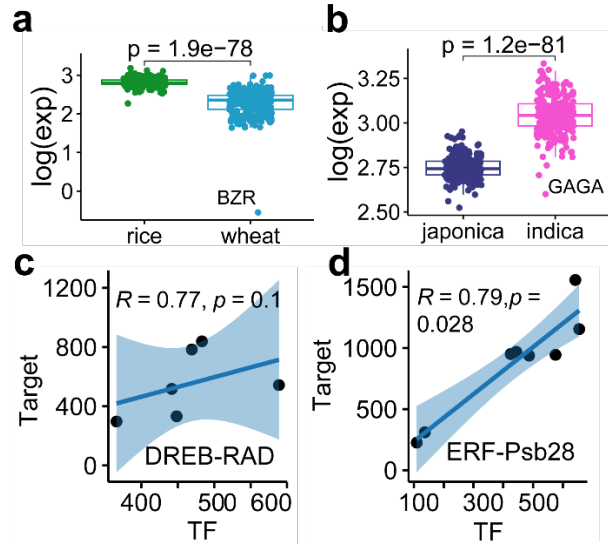
**Supplementary Figure 13** A correlation analysis of gene expression in conserved drought-regulatory mechanisms among related gene families. a, Correlation between ERF and Cyclin family genes. b, Correlation between MYB and CYP450 family genes. Points represent individual observations. Lines indicate the fitted linear regression, and shaded bands indicate the 95% confidence interval. Statistical significance was assessed using two-sided Spearman correlation analysis, and the correlation coefficients and exact P values are shown in the panels.



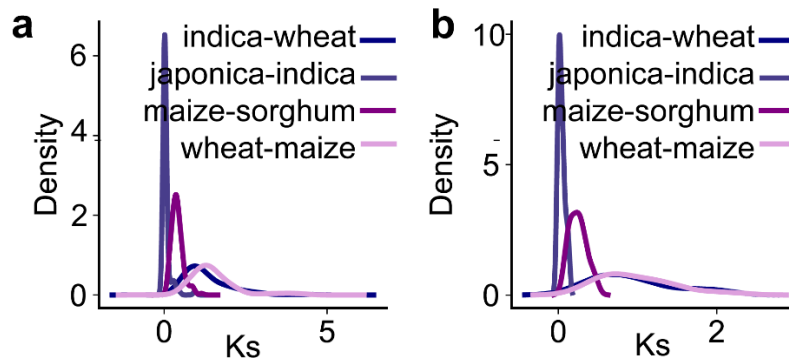
**Supplementary Figure 14** Conserved motif diagram of PP2C gene family. Motifs 1–10 identified in PP2C proteins are shown as sequence logos. The height of each letter indicates its relative frequency at that position, and the total height at each position is measured in bits.



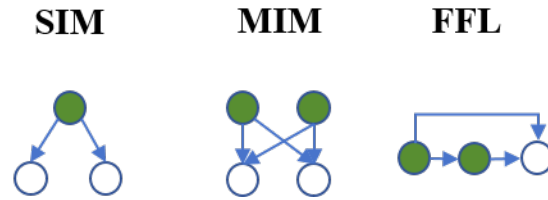
**Supplementary Figure 15** Distribution of synonymous substitution rate ( $K_s$ ) values within orthologous gene pairs. a,  $K_s$  distributions for the 2OGD family. b,  $K_s$  distributions for the PP2C family. Coloured curves indicate pairwise species comparisons.



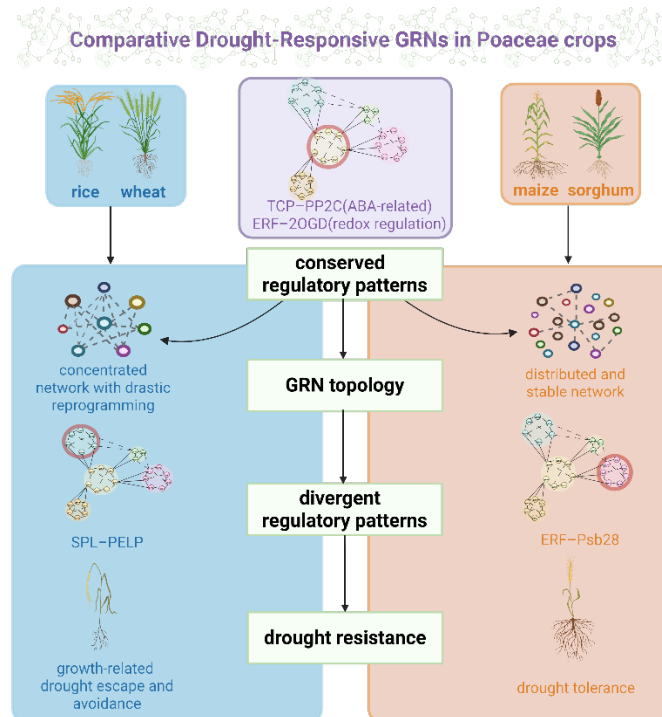
**Supplementary Figure 16** Expression differentiation and regulatory correlations of selected drought-responsive gene families. a, Expression distribution of the BZR family for wheat–rice differentiation. b, Expression distribution of the GAGA (GAGA-binding transcriptional activator) family for japonica–indica differentiation. Box plots show the median (centre line), the interquartile range (box bounds; 25th to 75th percentiles), and whiskers extend to  $1.5 \times \text{IQR}$ ; points represent individual data points. Statistical significance was assessed using a two-sided Wilcoxon rank-sum test, and exact P values are shown in the panels. c, Correlation analysis of DREB and RAD (Transcription factor RADIALIS-like) family genes. d, Correlation analysis of ERF and Psb28 (Photosystem II Psb28) family genes. Points represent individual observations. Lines indicate the fitted linear regression, and shaded bands indicate the 95% confidence interval. Statistical significance was assessed using two-sided Spearman correlation analysis, and the correlation coefficients and exact P values are shown in the panels.



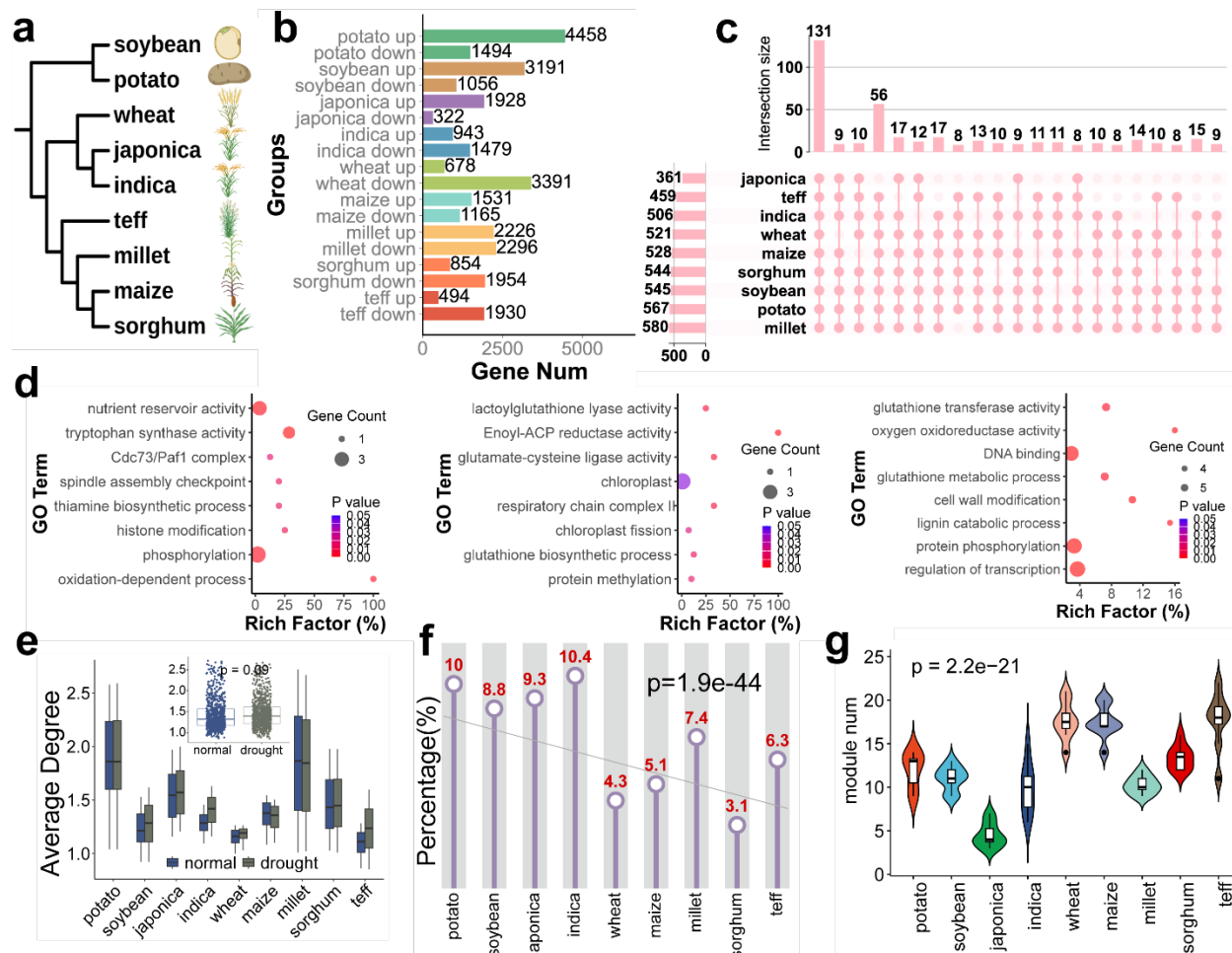
**Supplementary Figure 17** Distribution of synonymous substitution rate ( $K_s$ ) values within orthologous gene pairs. a,  $K_s$  distributions for the DREB family. b,  $K_s$  distributions for the SPL family. Coloured curves indicate pairwise species comparisons.



**Supplementary Figure 18** Schematic representation of regulatory network motifs. SIM denotes the single-input motif, MIM denotes the multiple-input motif, and FFL denotes the feed-forward loop. Filled and open circles represent regulator and target nodes, respectively, and arrows indicate regulatory direction.

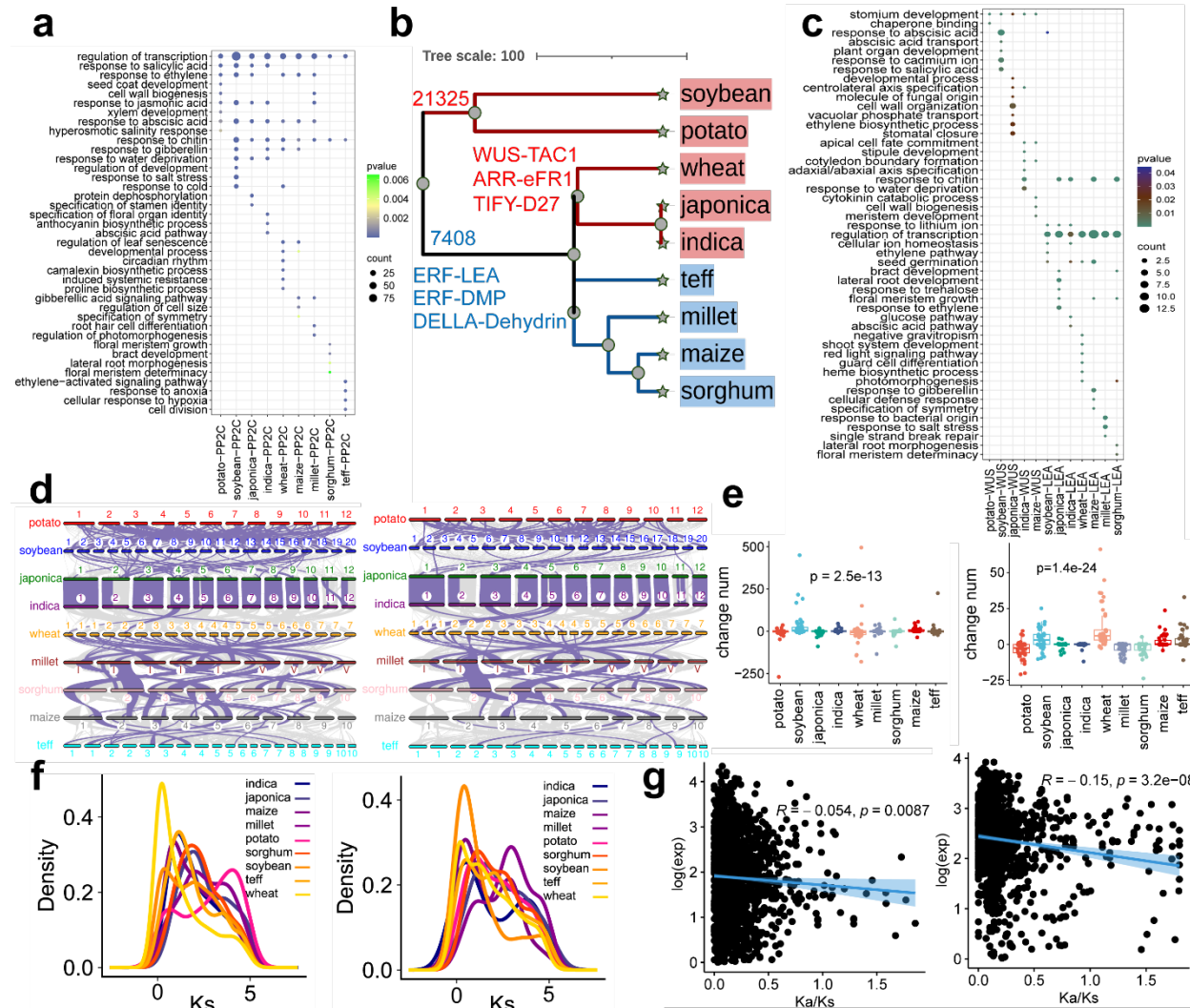


**Supplementary Figure 19** Schematic model of conserved and divergent gene regulatory networks in Poaceae crops under drought stress. The left panel summarizes concentrated network reprogramming in representative C<sub>3</sub> crops (rice and wheat), and the right panel summarizes distributed and stable network organization in representative C<sub>4</sub> crops (maize and sorghum). The central panels highlight conserved regulatory patterns, GRN topology, divergent regulatory patterns, and their associations with drought resistance. Created in BioRender. deng, X. (2026) <https://BioRender.com/fup3ym2>.



**Supplementary Figure 20** Comparative analysis of the C<sub>3</sub> crops (potato, soybean, japonica, indica, wheat) and C<sub>4</sub> 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 C<sub>3</sub> crops (soybean, potato, japonica rice, indica rice, and wheat), the center panel on C<sub>4</sub> crops (millet, teff, maize, and sorghum), and the right panel shows GO terms of common orthologous DEGs shared by all species under drought stress. GO enrichment significance was assessed using Fisher's exact test with Benjamini–Hochberg correction for multiple comparisons; dot size indicates gene count and colour indicates the adjusted P value. Figure (e) 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 data points. Statistical significance between normal and drought conditions was assessed using a two-sided Wilcoxon rank-sum test. Exact P values are shown in the panel. 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. Statistical significance was assessed using the Kruskal–Wallis test.

Figure (g) shows the number of gene modules identified under drought conditions. Violin plots show the data distribution; centre lines indicate the median, boxes indicate the interquartile range (25th to 75th percentiles), and whiskers extend to  $1.5 \times \text{IQR}$ . Statistical significance among groups was assessed using the Kruskal–Wallis test.



**Supplementary Figure 21** Functional annotation and phylogenetic analysis of conserved and divergent gene regulatory networks under drought stress, with the inclusion of newly added species. 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 C<sub>3</sub> crops, and blue branches represent C<sub>4</sub> 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. Enrichment significance was assessed using Fisher's exact test with Benjamini–Hochberg correction for multiple comparisons; dot size indicates gene count and colour indicates the adjusted P value. 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. Statistical significance among groups was assessed using the Kruskal–Wallis test. 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. Points represent individual observations. Lines indicate the fitted linear regression, and shaded bands indicate the 95% confidence interval. Statistical significance was assessed using two-sided Spearman correlation analysis, and the correlation coefficients and exact P values are shown in the panels.

### Supplementary Method 1

In general, standard graph attention methods calculate self-attention by computing the semantic similarity between node  $i$  and other nodes, without considering the structural information of the graph reflected in the nodes or the relationships between node pairs. Here, we made further adjustments based on the spatial information encoding method used in Graphormer for node classification <sup>1</sup>. Specifically, we encoded the degree centrality features of expression vector of each gene, enhancing the ability of the model to capture key structural information in the gene regulatory network.

$$h_i = x_i + z_{\text{deg}^-}(v_i) + z_{\text{deg}^+}(v_i) \quad (1)$$

Here,  $x_i$  represents the gene expression vector of the source node,  $z_{\text{deg}^-}(v_i)$  denotes the out-degree feature, and  $z_{\text{deg}^+}(v_i)$  represents the in-degree feature.

Additionally, we incorporated spatial encoding of the topological information reflecting the overall network connectivity. Existing neural networks primarily focus on the local connections of nodes and often overlook the global position of nodes as well as their relationships with non-neighboring nodes within the graph. To address this issue, we calculated the shortest connection paths between nodes. If two nodes are not connected, the path is set to -1. This approach embeds the entire graph structure into a matrix that captures the spatial relationships between nodes. Unlike the adjacency matrix, the resulting matrix reflects the global connectivity of the graph.

The shortest path distance (SPD) between two nodes was computed using the Floyd–Warshall algorithm, encoding the spatial properties of the network. To avoid memory limitations due to the large number of genes, this spatial encoding was also converted into an edge index format. This approach ensures that the spatial relationships in the graph are effectively captured without exceeding memory constraints.

$$Q = h_i W_Q \quad (2)$$

$$K = h_j W_K \quad (3)$$

$$V = h_j W_V \quad (4)$$

$$A_{ij} = \frac{(h_i W_Q)(h_j W_K + e_{ij})^T}{\sqrt{d}} + b_{ij} \quad (5)$$

$$h_{i+1} = h_i + \text{softmax}(A_{ij})(V + e_{ij}) \quad (6)$$

Here,  $b_{ij}$  represents the shortest path distance between nodes, based on the spatial properties encoded from the adjacency matrix. The edge feature  $e_{ij}$  is obtained by mapping the sum of the gene expression values of the source gene and the target gene. The calculation formula is as follows:

$$e_{ij} = W(x_i + x_j) \quad (7)$$

Here,  $x_i$  represents the gene expression vector of the source gene,  $x_j$  is the gene expression vector of the target gene, and  $W$  is a learnable parameter.

Following the TopoGAT layer, we apply layernorm and the relu activation function, similar to the transformer architecture. Additionally, we incorporate gated residual connections to prevent the model from becoming overly smooth, ensuring the preservation of important features and gradients during training.

## Supplementary Method 2

For the transcriptional and metabolic regulation layers, both the predicted labels and the adjacency matrix of the model are directed, meaning that the adjacency matrix is asymmetric and the labels are represented by two-dimensional vectors to indicate directionality. For the protein interaction layer, both the predicted labels and the adjacency matrix are undirected, with the adjacency matrix being symmetric and the labels represented by one-dimensional vectors.

The model utilizes binary cross-entropy loss <sup>2</sup>, calculated as follows:

$$L = -\frac{1}{N} \sum_{i=1}^N [y_i \log(p_i) + (1 - y_i) \log(1 - p_i)] \quad (8)$$

where  $N$  is the total number of samples,  $y_i$  represents the true label for the  $i$ -th sample, with values of (0,1) or (1,0). (0,1) indicates the presence of an edge from the source node to the target node, while (1,0) indicates the absence of an edge from the source node to the target node.  $p_i$  represents the predicted probability that the  $i$ -th sample belongs to the positive class, and  $\log$  represents the natural logarithm.

For model training, gene expression data was not simply concatenated across normal, drought, or recovery conditions. Instead, gene expression data for each batch was randomly selected from normal, drought, or recovery conditions, ensuring the model effectively learns and generalizes across different environmental contexts, thereby enhancing its ability to predict gene regulatory interactions accurately under varying conditions.

The dataset had a positive-to-negative sample ratio of 1:3, and was divided into training, validation, and test sets in proportions of 60%, 20%, and 20%, respectively. Validation set results were used to evaluate model performance. The model was optimized using the NAdam optimizer, with a learning rate decay strategy implemented via MultiStepLR. The initial learning rate ( $lr$ ) was set to 0.001, with a weight decay parameter of 0.001. During training, mini-batches included edge data and gene expression data under normal and drought conditions, as well as the adjacency matrix of the training set. The construction of gene regulatory networks (GRNs) under normal and drought conditions is inferred using gene expression matrices corresponding to conditions, respectively.

Model training and testing were performed using the PyTorch framework on an NVIDIA GeForce RTX 4080 SUPER with 16GB of VRAM.

The evaluation metrics used were AUC (area under the receiver operating characteristics curve) <sup>3</sup> and AUPR (area under the precision-recall curve) <sup>4</sup>. These metrics are widely used to

assess classification models, particularly for imbalanced datasets. The formulas for true positive rate (TPR) and false positive rate (FPR) are given by:

$$\text{TPR} = \frac{TP}{TP + FN} \quad (9)$$

$$\text{FPR} = \frac{FP}{FP + TN} \quad (10)$$

where  $TP$  (True Positives) represents correctly predicted positive samples,  $FN$  (False Negatives) represents positive samples incorrectly classified as negative,  $FP$  (False Positives) represents negative samples incorrectly classified as positive, and  $TN$  (True Negatives) represents correctly predicted negative samples. The ROC (Receiver Operating Characteristic) curve is plotted by graphing TPR against FPR at different thresholds. The AUC measures the area under the ROC curve, with values ranging from 0 to 1, where a higher value indicates better classification performance.

Precision and Recall are defined as:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (11)$$

$$\text{Recall} = \text{TPR} = \frac{TP}{TP + FN} \quad (12)$$

Precision measures the proportion of correctly predicted positive samples among all samples predicted as positive, while Recall (equivalent to TPR) measures the proportion of actual positive samples correctly identified by the model.

**Supplementary Method 3** The model parameters are set as follows: for the linear mapping layer, the input dimensions under normal, drought, and rehydration conditions are [702, 760, 85] for japonica, [333, 263, 59] for indica, [256, 452, 25] for wheat, [743, 777, 16] for maize, and [255, 325, 0] for sorghum, with the output size dimension uniformly set to 1000. The attention heads of the three-layer TopoGAT layers are [3, 3, 1], and the input and output dimensions are [1000, 128], [128, 64], and [64, 32], respectively. The final fully connected mapping layer has input and output dimensions of [32, 2] (for the directed gene regulatory network) or [32, 1] (for the undirected gene regulatory network).

#### Supplementary Method 4

The topological characteristics of a network include several descriptive metrics, such as average degree, average path length, and clustering coefficient, which represent its fundamental properties. The average degree (AD,  $\langle k \rangle$ ) is the mean number of connections per node <sup>5</sup>. The average path length (APL,  $L$ ) <sup>6</sup> is the mean shortest distance between all pairs of nodes, reflecting overall connectivity. The average clustering coefficient (CC,  $C$ ) measures the tendency of nodes to form clusters <sup>7</sup>. The respective formulas are as follows:

$$\langle k \rangle = \frac{1}{n} \sum_{i=1}^n k_i \quad (13)$$

$$L = \frac{1}{\frac{n(n-1)}{2}} \sum_{i \neq j} d_{ij} \quad (14)$$

$$C = \frac{1}{n} \sum_{i=1}^n C_i \quad (15)$$

$$C_i = \frac{2T_i}{k_i(k_i - 1)} \quad (16)$$

where  $n$  represents the number of nodes in the network, and  $k_i$  is the degree of node  $i$ .  $d_{ij}$  represents the shortest path length between nodes  $i$  and  $j$ .  $T_i$  represents the number of actual edges between the neighbors of node  $i$ , and  $k_i(k_i - 1)$  is the maximum possible number of edges among these neighbors, assuming full connectivity. The clustering coefficient for each node ( $C_i$ ) is calculated as the ratio of existing edges between neighbors to the total possible edges, and the average clustering coefficient ( $C$ ) is the mean across all nodes.

Network motifs are fundamental units with simple structures within the network<sup>8</sup>. They represent small-scale patterns of regulatory relationships between transcription factors and target genes, typically involving a few genes arranged in specific topological structures. Here, we analyzed single-input motifs (SIM), multi-input motifs (MIM), and feed-forward loops (FFL), with specific motif structures detailed in the Fig. S11. The dynamic characteristics of these motifs help explain the adaptive responses of crop drought-resistance gene networks under drought conditions.

### Supplementary Method 5

The algorithm initializes by assigning each node to its own community. For each node, the algorithm attempts to move it to the community of its neighbors, calculating the resulting change in modularity. If modularity increases, the move is accepted. This process continues until no further moves can improve modularity.

Once communities are formed, they are aggregated into new nodes (supernodes), where the edge weight between any two supernodes equals the sum of the weights of all edges connecting the original communities. This produces a reduced network, and the steps of local optimization and network aggregation are repeated until modularity is maximized.

Modularity ( $Q$ ) measures how well edges are clustered into modules compared to a random edge distribution, ignoring community structure. A high modularity value indicates that nodes within the same community are more interconnected than would be expected by chance. Mathematically, modularity  $Q$  can be expressed as:

$$Q = \frac{1}{2m} \sum_{ij} \left[ A_{ij} - \frac{k_i k_j}{2m} \right] \delta(c_i, c_j) \quad (17)$$

where  $A_{ij}$  is the adjacency matrix of the network.  $k_i$  and  $k_j$  represent the degrees of nodes  $i$  and nodes  $j$ , and  $m$  is the total number of edges in the network.  $\delta(c_i, c_j)$  is an indicator function that equals 1 when nodes  $i$  and  $j$  are in the same community, and 0 otherwise. For community detection in the GRNs of rice, wheat, maize, and sorghum, a resolution parameter of 1.2 was used.

We visualized the expression of gene modules and applied UMAP<sup>9</sup> for dimensionality reduction to reveal changes in crop gene expression patterns under different water conditions. The Calinski-Harabasz (CH) index<sup>10</sup> was used to evaluate changes in co-expression patterns within UMAP clustering. The CH index, also known as the variance ratio criterion, is a metric for assessing the quality of clustering results. It measures the compactness within clusters and the separation between clusters. The formula for calculating the CH index is as follows:

$$CH = \frac{SST - SSW}{SSW} \times \frac{N - k}{k - 1} \quad (18)$$

where  $SST$  (Total Sum of Squares) is the sum of squared distances between all data points and the centroid of the entire dataset, representing the overall dispersion of the data.  $SSW$  (Within-cluster Sum of Squares) is the sum of squared distances between data points and the centroid of their respective clusters, representing the compactness within clusters.  $N$  is the total number of samples in the dataset.  $k$  is the number of clusters.

### Supplementary References

1. Ying, C. *et al.* Do transformers really perform badly for graph representation? *Advances in neural information processing systems* **34**, 28877–28888 (2021).
2. Ho, Y. & Wookey, S. The real-world-weight cross-entropy loss function: Modeling the costs of mislabeling. *IEEE access* **8**, 4806–4813 (2019).
3. Huang, J. & Ling, C. X. Using AUC and accuracy in evaluating learning algorithms. *IEEE Transactions on knowledge and Data Engineering* **17**, 299–310 (2005).
4. Boyd, K., Eng, K. H. & Page, C. D. Area under the precision-recall curve: point estimates and confidence intervals. in *Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2013, Prague, Czech Republic, September 23-27, 2013, Proceedings, Part III* 13 451–466 (Springer, 2013).
5. Damas, M. P., Markenzon, L. & de Abreu, N. M. M. New concepts and results on the average degree of a graph. *Applicable Analysis and Discrete Mathematics* 284–292 (2007).
6. Mao, G. & Zhang, N. Analysis of average shortest-path length of scale-free network. *Journal of Applied Mathematics* **2013**, 865643 (2013).
7. Li, Y., Shang, Y. & Yang, Y. Clustering coefficients of large networks. *Information Sciences* **382**, 350–358 (2017).
8. Babu, M. M., Luscombe, N. M., Aravind, L., Gerstein, M. & Teichmann, S. A. Structure and evolution of transcriptional regulatory networks. *Current opinion in structural biology* **14**, 283–291 (2004).
9. McInnes, L., Healy, J. & Melville, J. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426* (2018).
10. Wang, X. & Xu, Y. An improved index for clustering validation based on Silhouette index and Calinski-Harabasz index. in *IOP Conference Series: Materials Science and Engineering* vol. 569 052024 (IOP Publishing, 2019).