



## OPEN ACCESS

## EDITED BY

Haifei Hu,  
Guangdong Academy of Agricultural Sciences  
(GDAAS), China

## REVIEWED BY

Shoukun Chen,  
Chinese Academy of Agricultural Sciences,  
China  
Huadong Zhan,  
Nanjing Agricultural University, China  
Baopeng Ding,  
Shanxi Datong University, China

## \*CORRESPONDENCE

Jiang-hui Cui  
✉ 13663123545@126.com

RECEIVED 06 October 2024

ACCEPTED 23 December 2024

PUBLISHED 11 February 2025

## CITATION

Man Q-c, Wang Y-q, Gao S-j, Gao Z-c,  
Peng Z-p and Cui J-h (2025) Pan-genome  
analysis and expression verification  
of the maize ARF gene family.  
*Front. Plant Sci.* 15:1506853.  
doi: 10.3389/fpls.2024.1506853

## COPYRIGHT

© 2025 Man, Wang, Gao, Gao, Peng and Cui.  
This is an open-access article distributed under  
the terms of the [Creative Commons Attribution  
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or  
reproduction in other forums is permitted,  
provided the original author(s) and the  
copyright owner(s) are credited and that the  
original publication in this journal is cited, in  
accordance with accepted academic  
practice. No use, distribution or reproduction  
is permitted which does not comply with  
these terms.

# Pan-genome analysis and expression verification of the maize ARF gene family

Quan-cai Man<sup>1</sup>, Yan-qun Wang<sup>2</sup>, Shun-juan Gao<sup>1</sup>,  
Zhi-chang Gao<sup>1</sup>, Zheng-ping Peng<sup>2</sup> and Jiang-hui Cui<sup>1\*</sup>

<sup>1</sup>College of Agriculture, Hebei Agricultural University, Baoding, China, <sup>2</sup>College of Resource and Environmental Sciences, Hebei Agricultural University, Baoding, China

Auxin transcription factors regulate auxin responses and play crucial roles in plant growth, development, and responses to abiotic stress. Utilizing the maize pan-genome data, this study identified 35 ARF family members in maize, comprising 21 core genes, 10 near-core genes, and 4 non-essential genes; no private genes were detected. The construction of a phylogenetic tree using *Arabidopsis thaliana* revealed that the G3 subfamily comprises the highest number of core genes, with a total of 10, and exhibits relative stability throughout the evolution of maize. The calculation of the Ka/Ks ratios for ARF family members across 26 genomes indicated that, aside from *ARF8* and *ARF11*, which were subjected to positive selection, the remaining genes underwent purifying selection. Analysis of structural variation revealed that the expression level of the *ARF4* gene significantly differed as a result of this variation. Simultaneously, the structural variation also influenced the conserved domain and cis-acting elements of the gene. Further combining the transcriptome data and RT-qPCR found that, The expression levels of ARF family members in maize were higher at the early stage of embryo and grain development, and the expression levels of each member in embryo and grain were complementary, and the *ARF4* plays an important role in abiotic stress. In summary, this study utilizes the maize pan-genome and bioinformatics methods to investigate the evolutionary relationships and functional roles of ARF family members in maize, thereby providing a novel theoretical framework for further research on the maize ARF family.

## KEYWORDS

maize pan-genome, auxin transcription factors, structure variation, selection pressure, transcriptome data analysis

## Introduction

Auxin is a crucial phytohormone that plays a significant role in plant growth and development, regulation of plant metabolism, and response to abiotic stress (Casanova-Sáez et al., 2021; Gomes and Scortecci, 2021; Verma et al., 2022). Auxin is distributed through tissue transport mechanisms to ensure the proper functioning of signal

transduction pathways, thereby regulating the development of the plant root system (Roychoudhry and Kepinski, 2022). Auxin controls cell growth through the regulation of cell wall remodeling enzyme expression in plants, thereby influencing pectin biosynthesis (Jobert et al., 2023). Changes in auxin content and transport in plants can mitigate plant damage caused by salt stress (Ribba et al., 2020).

The TAA/YUC pathway is the most well-defined auxin synthesis pathway and includes two key components, tryptophan aminotransferase (TAA) and Yucca sativa (YUC) flavin-dependent monooxygenase, these two enzymes catalyze the conversion of tryptophan (Trp) to auxin (IAA), and the pathway plays an important role in regulating plant development (Blakeslee et al., 2019). In addition to this, different auxin concentrations lead to different auxin responses (Tan et al., 2021). In plants, auxin signal transduction is divided into classical signaling pathways and non-classical signaling pathways. The classical pathway primarily refers to the widely recognized TIR1/AFB-Aux/IAA-ARF signal transduction system. The ETT-mediated auxin conduction system, discovered in recent years, is referred to as a non-classical signaling pathway (Leyser, 2018; McLaughlin et al., 2021). Nuclear auxin signal transduction primarily occurs via the classical pathway, namely the TIR1/AFB-Aux/IAA-ARF signal transduction system, which relies on three key components: TIR1/AFB, Aux/IAA, and ARF (Morffy and Strader, 2022). TIR1/AFB primarily functions as an auxin receptor, mediating its function by binding to Aux/IAA to form distinct complexes (Powers and Strader, 2020). After Aux/IAA is polyubiquitinated and degraded by the proteasome, ARF releases its inhibition and activates the expression of auxin response genes. ARF is the key regulator of auxin-responsive gene expression (Rienstra et al., 2023; Liu et al., 2024b). ARF mainly contains three domains with different functions: DNA binding domain (DBD), middle region (MR) and C-terminal domain (PB1) (Cancé et al., 2022). ARF can be categorized into three major groups based on function: transcription activator A, and transcription repressors B and C. Various ARF transcription factor types regulate distinct auxin responses (Freire-Rios et al., 2020). Currently, ARF has been identified in many plants, and their functions have been elucidated. 25 ARF family members were identified in rice. RT-pcr results showed that the expression of rice ARF genes was induced by auxin, and these genes play important roles in metabolic pathways and cellular processes in rice (Wang et al., 2007). 67 ARF family members were identified in wheat, which were unevenly distributed in six subfamilies. GO enrichment analysis showed that ARF family members played important roles in the growth and development, metabolic process, and response to auxin stimulation in wheat (Chaudhary et al., 2023).

The previous study have identified 31 ARF family members in maize genome and divided them into six subfamilies. ZmARFs proteins range in length from 462 to 1192 amino acids, the expression of ZmARF gene is regulated by auxin and small RNA. It may play a key role in seed development and germination (Xing et al., 2011). However, traditional gene family identification methods typically rely on a single reference genome, which exhibit notable limitations and are unable to discern variations among distinct reference genomes. With the publication of the

maize pan-genome, the presence-absence and structural variation (SV) information between the genomes of different maize strains have been made available, offering enhanced support for gene family analysis and research (Hufford et al., 2021).

Based on 26 maize pan-genomes, this study screened and identified 35 ARF gene family members, including 21 core genes, 10 near-core genes, 4 non-essential genes, and found no private genes. By analyzing the nonsynonymous to synonymous substitution ratios (Ka/Ks) of ARF members in 26 genomes, the effects of structural variation on gene expression, gene structure, and conserved domains were revealed. Further analysis of the transcriptome expression data of ARF family members revealed that ARF is crucial for the development of maize embryos and seeds.

## Materials and methods

### Plant material and stress treatment

Maize inbred line B73 was selected as the experimental material. Maize seeds were sown in black pot (7cm×8cm×5cm) soil. Only one corn was sotted in each pot and then placed in a 25 °C thermostatic chamber for cultivation. When maize grew to the trifoliate stage, maize plants were subjected to simulated drought stress and salt stress (Drought stress:20% PEG6000; Salt stress: 200mM NaCl). Samples were then taken at 12 h and 24 h of drought and salt stress treatments, respectively. The samples were flash frozen in liquid nitrogen and then stored in an ultra-low temperature refrigerator at -80 °C.

### Identification of maize ARF gene family

26 maize pan-genome data sets were derived from the research conducted by Hufford et al (Hufford et al., 2021). Subsequently, the hidden Markov model of ARF (PF06507) was retrieved from the Pfam database (<https://www.ebi.ac.uk/interpro/entry/pfam/>), and then the ARF domain was searched using HMMER 3.1 software, with the threshold set to  $1e^{-5}$ . The preliminary ARF members were submitted to the SMART (<http://smart.embl.de/>) and Pfam databases (<https://www.ebi.ac.uk/interpro/>) for identification, and genes lacking the ARF domain were removed. Ultimately, the maize ARF members were characterized, and the protein lengths of the identified ARF members were evaluated using BioPerl (Stajich et al., 2002).

### ARF presence-absence variation analysis

The presence-absence information of ARF members was derived from the research conducted by Hufford et al (Hufford et al., 2021). Subsequently, a built-in script was utilized to generate the list of presence-absence genes, and the ggplot2 package in R scripts was employed to create the presence-absence heat map of ARF members in 26 genomes (Ito and Murphy, 2013).

## Phylogenetic analysis

Arabidopsis protein sequences were retrieved from the Arabidopsis database, and the conserved domain information for Arabidopsis ARF family members was extracted. Then, muscle was used to align the conserved domains of maize and Arabidopsis ARFs, and then the maximum likelihood method (Maximum Likelihood, Bootstrap repeated 1000 times) was used to construct the phylogenetic tree of Arabidopsis and maize ARF family members through IQTREE (Nguyen et al., 2015). The resultant evolutionary tree was submitted to ITOL v6 (<https://itol.embl.de/>) for enhancement and visualization.

## Ka/Ks calculation

The coding sequences (CDS) and protein sequences of ARF family members were retrieved from 26 maize genomes, and the Ka/Ks values for each ARF family member were computed using KaKs Calculator 2.0 (Wang et al., 2010a). Subsequently, the R software packages ggridges and ggplot2 were employed to create the Ridgeline plot of Ka/Ks, and the R software package pheatmap was used to create the heat map for ARF members with Ka/Ks values exceeding 1.

## Analysis of the expression of ARFs overlapped with SVs

The structural variation analysis results of 26 maize pan-genomes were derived from the research conducted by Hufford et al (Hufford et al., 2021). The B73 genome was chosen as the reference genome for library construction; subsequently, ANNOVAR was employed to annotate the structural variation (Wang et al., 2010b). Then the information that needs to be retained and structural variation is extracted, and then the structural variation information of ARF family members is extracted. Subsequently, a correlation analysis was performed on the structural variation of ARF family members and gene expression levels, and a histogram was generated for genes exhibiting significant differences.

## Analysis of the SV and gene structure

Download the genome annotation file from the maize database website (<http://maize-pange.nome.gramene.org/>), then extract the significantly different gene structure information identified in the previous step, and submit the protein sequence to the MEME program (<http://meme-suite.org/>) for protein conserved motif analysis (parameter settings: the maximum number of motifs is 10, the width of the motif is set to 6-50, and the rest are default parameter values). Use TBtools II to visualize gene structure and conserved motifs (Chen et al., 2023).

Subsequently, the protein sequence of the reference genome B73, along with the protein sequence of the strain exhibiting the greatest overlap with its structural variation (SV), was submitted to the MEME program for conserved motif analysis (parameter settings are the same as above), and webLogos representations of the ARF member protein sequences for both strains were generated.

## Analysis of cis-acting elements of the promoter

The reference genome B73 and the 2000bp promoter sequence of the strain with the most overlap with SV were extracted. The promoter sequences were submitted to the PlantCARE database (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>) for cis-acting element analysis (Lescot et al., 2002). Then, TBtools II software was used to visualize the cis-acting elements (Chen et al., 2023).

## Protein-protein interaction network analysis of auxin signaling pathway

From phytozome (JGI) database (<https://phytozome.jgi.doe.gov>) to download the members of the family of IAA corn protein sequences, it was uploaded to the STRING database together with the ARF family member sequences of B73 (Szklarczyk et al., 2023), the interaction between ARF family and IAA family was analyzed, and the results were visualized by cytoscape software (Shannon et al., 2003).

## RNA-seq data analysis

Data were downloaded from the NCBI database based on the RNA-seq data (PRJNA237837), with B73 selected as the reference genome, and Transcripts Per Million (TPM) were utilized as the metric to quantify the expression level. The expression profiles of ZmARF family members in embryos and seeds across various developmental stages were analyzed, and TBtools II was utilized to generate the expression heatmap for ZmARFs (Chen et al., 2023).

## RT-qPCR analysis

Total RNA was extracted from the samples using Promega (LS1040), cDNA was synthesized by reverse transcription using Kangwei Century RT gDNA (CW2020M), RT-qPCR was performed on CFX96 (Bio-Rad, USA) using the US EVERBRIGHT (AugeGreen qPCR Master Mix S2008L) (Three biological replicates and three technical replicates were used). Relative expression was calculated using the  $2^{-\Delta\Delta ct}$  method (Maren et al., 2023). The primers used were synthesized by Sangon Bioengineering (Shanghai) Co., LTD.

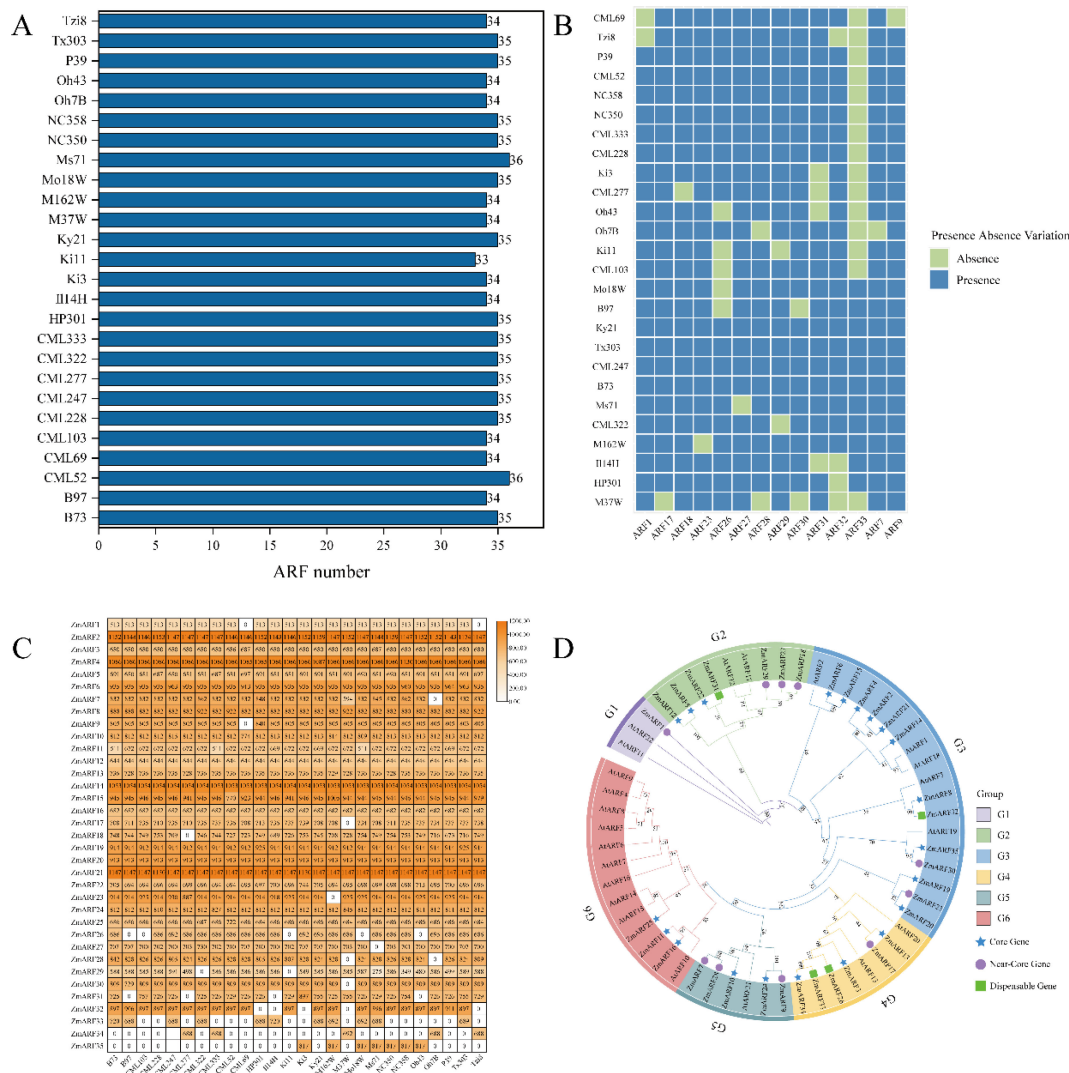
## Result

### Identification and phylogenetic analysis of ZmARFs gene-based pan-genome

In total, 35 ZmARF family members were identified from the maize pan-genome. B73 was utilized as the reference genome to compare the quantity of ARF family members and the lengths of protein sequences across various maize lines, revealing significant differences among the lines. Regarding the number of ARF family members, CML52 and MS71 both identified 36 ARF members, the highest number among the lines, whereas Ki1 identified 33 ARF members, the lowest count (Figure 1A). The presence-absence analysis of the identified ZmARF family members revealed that the ZmARF family comprises 21 core genes, 10 near-core genes, and 4 non-essential genes. No private genes were identified within the ZmARF family (Figure 1B). A further analysis of the protein lengths

of ARF family members across different strains indicated that only *ARF12*, *ARF16*, and *ARF20* were consistently present in all strains with identical protein lengths, whereas *ARF1* was absent in CML69 and Tzi8, although the protein lengths of *ARF1* in the other strains remained consistent. With the exception of these four genes, the protein lengths of the other family members exhibited variability (Figure 1C). Complete information of ARF family members is provided in Supplementary Table S1.

A phylogenetic tree was constructed based on the domain sequences of maize and Arabidopsis and classified into six subtribes according to the study by Xing et al (Xing et al., 2011). The G1 subtribe includes only one near-core gene, the lowest count, while the G3 subtribe comprises 13 ARF family members, the highest count, comprising 10 core genes, 2 near-core genes, and 1 non-essential gene. The G2, G4, G5, and G6 subtribes comprise 7, 6, 5, and 3 ARF family members, respectively, and all ARF family members in the G6 subtribe are core genes (Figure 1D).



**FIGURE 1** Identification and phylogenetic analysis of ZmARFs in pan-genome. (A) Number of ZmARFs. (B) Heatmap of the presence and absence of 14 ZmARFs in 26 maize varieties except for the core genes. (C) Heatmap of ZmARFs protein length. (D) Phylogenetic tree of ARFs from Arabidopsis and Maize.

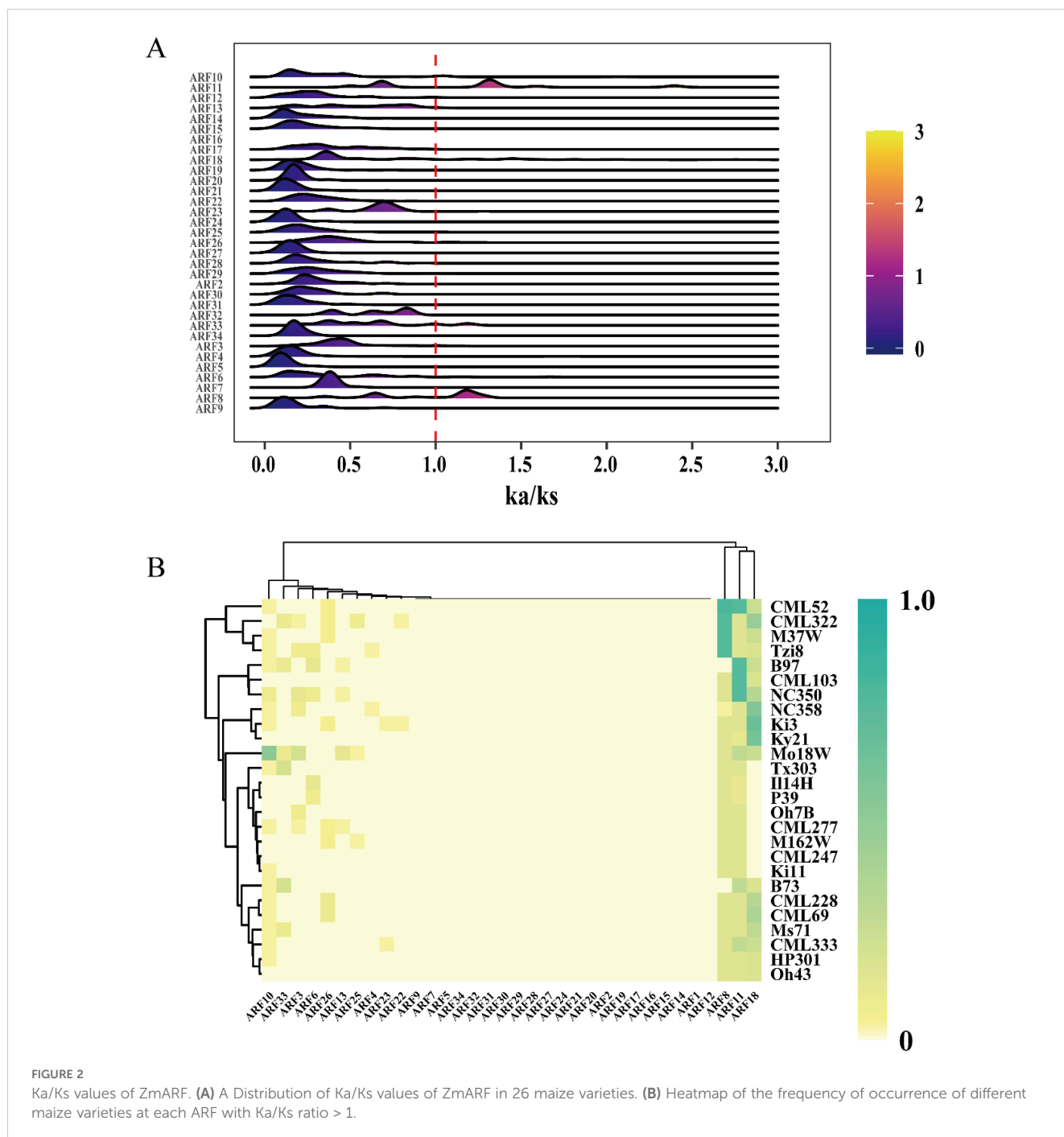
## ZmARF is subjected to different selection pressures among maize varieties

The Ka/Ks ratio serves as a significant metric for investigating the mechanisms of genetic evolution (Hurst, 2002). To investigate the selection pressure on ARF family members across 26 maize genomes, the Ka/Ks values for each ARF family member were computed. The results indicated that the Ka/Ks values for *ARF1* and *ARF35* could not be determined. With the exception of the peak values for *ARF8* and *ARF11*, which exceeded 1, the Ka/Ks peak values of the majority of other ARF family members were below 1 (Figure 2A). Some maize strains exhibit positive selection effects on

*ARF8* and *ARF11*, whereas the other genes are subject to purifying selection. Further analysis of the heatmap for Ka/Ks ratios greater than 1 revealed that only *ARF8*, *ARF11*, and *ARF18* exhibited high Ka/Ks ratios, indicating that these genes were subjected to selection pressure during maize development (Figure 2B).

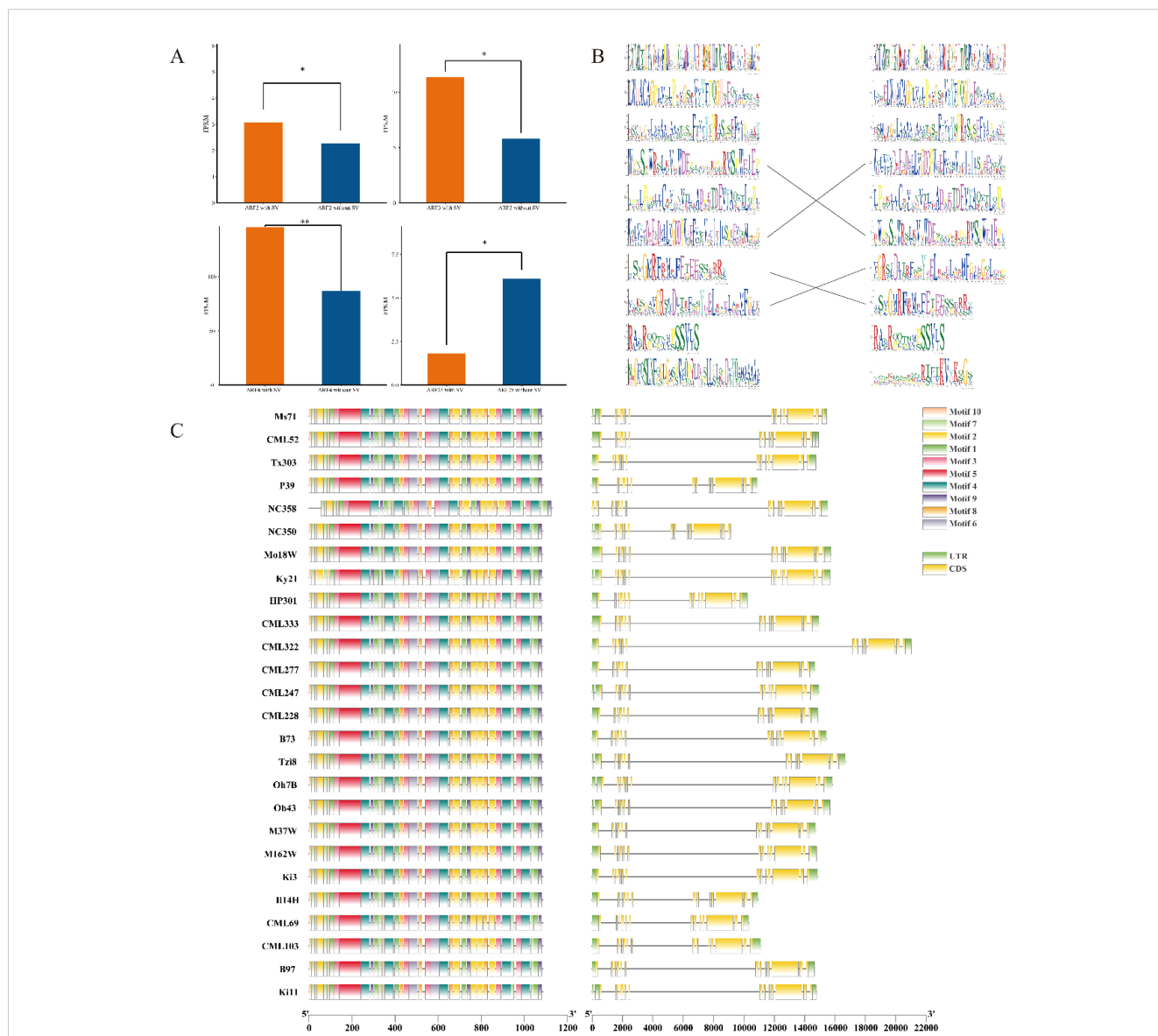
## Expression, structure and motif of ZmARF genes are affected by SV

The variation data of ARF family members were extracted from the maize structural variation file; subsequently, the Pearson



correlation coefficients of the expression levels of genes both overlapping and not overlapping with SV were calculated. The results indicated that *ARF2*, *ARF3*, *ARF4*, and *ARF25* exhibited significant differences in expression levels between the presence of structural variation (SV) and its absence. These significant differences suggest that SV markedly influences the expression of these four genes (Figure 3A). Subsequently, we calculated the number of genes exhibiting structural variation (SV) overlap between the reference genome B73 and other genomes, and selected the strain Ki3, which exhibited the highest degree of overlap, to generate a conserved domain alignment diagram (Figure 3B). The results indicated that in the Ki3 strain genome,

three conserved domains were identified as corresponding to the reference genome B73, whereas the remaining seven did not align. Furthermore, the amino acids within the corresponding conserved domains of each group do not exhibit complete alignment, suggesting that structural variation (SV) substantially influences the conserved domains of ZmARF. In the gene expression analysis, *ARF4* exhibited the most significant difference; therefore, we delineated its gene structure and conserved domains across 26 maize genomes. Our analysis revealed that the gene structure and conserved domains in the remaining genomes were in alignment with those of the reference genome B73, showing no significant discrepancies (Figure 3C).



**FIGURE 3** SVs affecting the expression, structure and motif in 26 maize genomes. **(A)** The expression of *ARF2*, *ARF3*, *ARF4*, *ARF25* was significantly affected by SVs (\*:  $P < 0.05$ , \*\*:  $P < 0.01$ ). **(B)** A The weblogs of the Ki3 ZmARF and the reference genome are shown on the left and right, respectively. The weblogs connected by the lines indicate that they are corresponding. Weblogs are arranged in the order of E-value. **(C)** The structure and motifs of ZmARF4 in maize pan-genome.

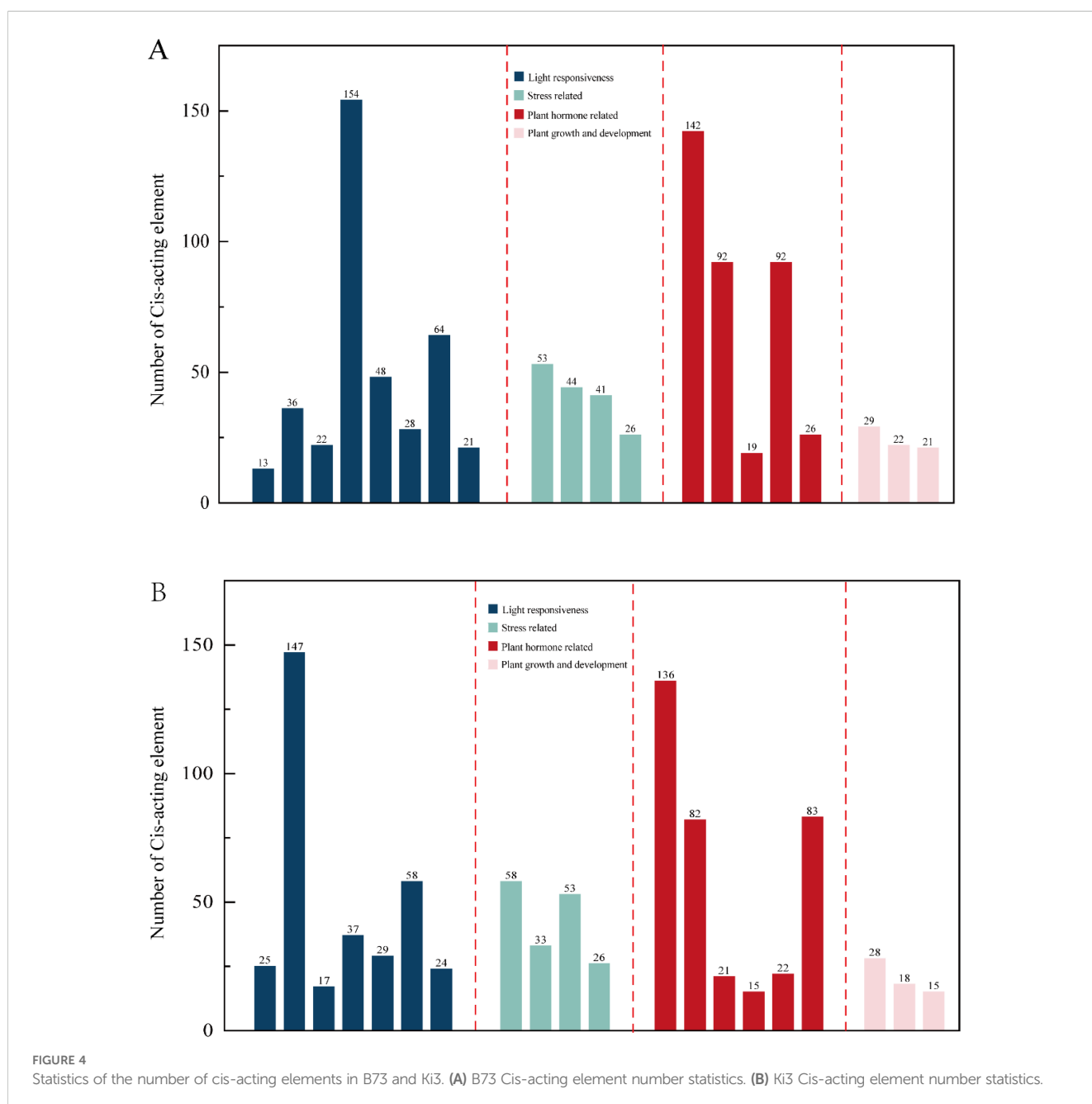
## Analysis of cis-acting elements of the promoter

The Ki3 strain with the most overlap with the reference genome SV was selected, analysis of cis-acting elements in the promoter, the top 20 elements with the largest number of cis-acting elements were selected for statistics and drawing. It was found that in the reference genome B73 and Ki3, there are differences between cis-acting elements related to light response and cis-acting elements related to plant hormones, B73 contains 8 cis-acting elements related to light response and 5 cis-acting elements related to plant hormones (Figure 4A); Ki3 contains 7 cis-acting elements related to light response and 6 cis-acting elements related to plant hormones (Figure 4B). Both B73 and Ki3 contained four stress-related cis-

acting elements and three growth-related cis-acting elements. It is concluded that structural variation alters the cis-acting element composition of ARF members in different strains, which may further affect physiological processes such as photosynthesis and hormone response in plants. The composition of ARF cis-acting elements in B73 and Ki3 is detailed in Supplementary Figure S1.

## IAA-ARF protein-protein interaction network analysis

By analyzing the interaction between IAA family and ARF family in maize, it was found that they were closely related (Figure 5). There are 9 members of the ARF family, *ARF3*, *ARF6*,



*ARF11*, *ARF13*, *ARF15*, *ARF16*, *ARF17*, *ARF26* and *ARF33*, play key roles in protein-protein interactions with 28 members of the IAA family. It is concluded that these nine members of the ARF family are more closely related to auxin during auxin signaling and thus play an important role in auxin signaling and regulating plant growth and development.

## Effects of ARF on maize seed development

Transcriptome data pertaining to maize embryos and seeds across various developmental stages were retrieved from the SRA database, and the expression profiles of ARF family members at these stages were subsequently analyzed. The results showed that 11 ARF family members were highly expressed in E1 (embryo-10day), 12 ARF family members were highly expressed in E2 (embryo-16day), 16 ARF family members were highly expressed in S1 (seed-0day), 7 ARF family members were highly expressed in S2 (seed-6day), and only 2 ARF family members were highly expressed in S5 (seed-32day), while the expression of ARF family members in the rest of the developmental stages was at a low level (Figure 6). These findings suggest that ARF family members play a crucial role in the early development of maize embryos and seeds.

## ARF in response to abiotic stress

The response of the ZmARF gene to drought and salt stress was further analyzed. The results showed that under 20% PEG6000 drought stress treatment, ZmARF gene was significantly differentially expressed at different treatment times (Figure 7A). The expression levels of *ARF2*, *ARF6*, *ARF8*, *ARF9*, *ARF15*, *ARF19*, *ARF23*, *ARF29*, *ARF30* and *ARF32* reached the peak at 12 h of stress, and then decreased to different degrees at 24 h of stress. *ARF4* and *ARF25* peaked at 24 h after stress treatment. Compared with CK, *ARF4* was up-regulated at 12 h of stress, while *ARF25* was down-regulated at 12 h of stress and then up-regulated at 24 h of stress. Under the treatment of 200 mM NaCl, the expression of ZmARF gene was also significantly different under different treatment time (Figure 7B). *ARF2* and *ARF6* reached the peak at 12 h, and then were down-regulated at 24 h compared with 12 h. *ARF4*, *ARF15*, *ARF19*, *ARF23*, *ARF30* and *ARF32* peaked at 24 h of stress treatment. Among them, *ARF4* and *ARF15* were up-regulated, while *ARF19*, *ARF23*, *ARF30* and *ARF32* were down-regulated after 12 h of stress compared with CK; The expression levels of *ARF8*, *ARF9*, *ARF25* and *ARF29* were down-regulated under salt stress. Among them, *ARF8* and *ARF29* were up-regulated at 24 h compared with 12 h, while *ARF9* and *ARF25* were gradually

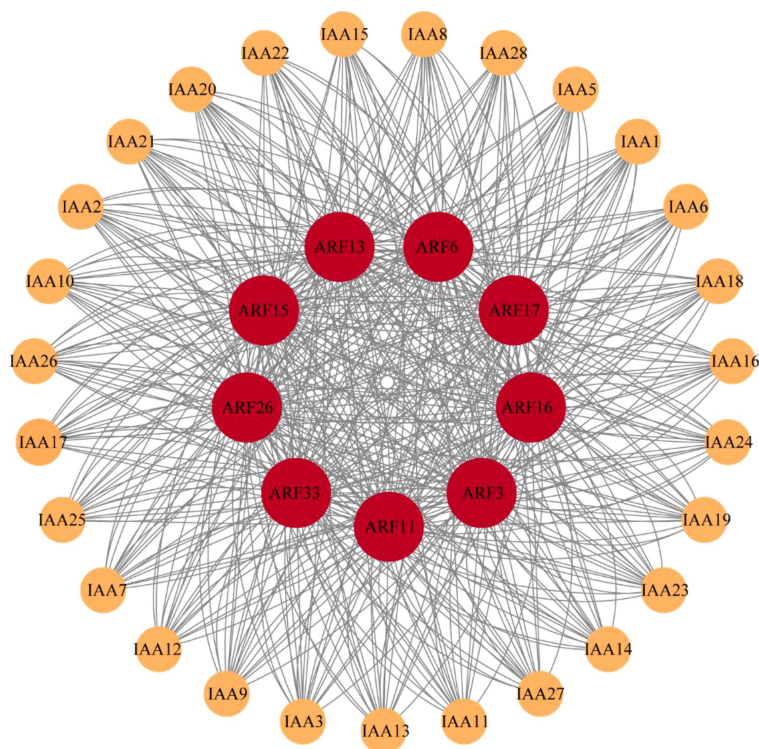


FIGURE 5  
ARF- IAA family protein-protein interaction network.



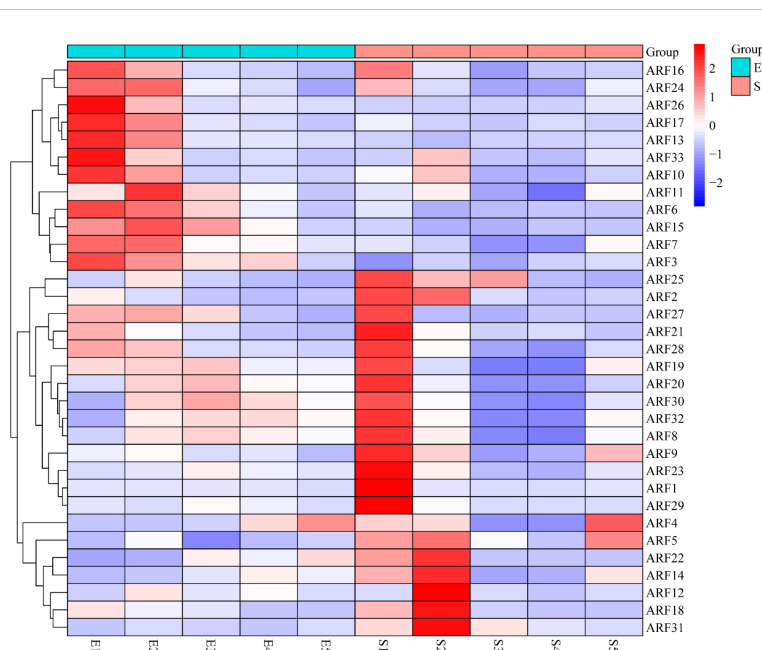


FIGURE 6

Heatmap of maize embryo and seed development with ARF family members (E1: embryo-10day; E2: embryo-16day; E3: embryo-24day; E4: embryo-30day; E5: embryo-36day; S1: seed-0day; S2: seed-6day; S3: seed-12day; S4: seed-22day; S5: seed-32day).

down-regulated over time under salt stress. In conclusion, ZmARF members responded to drought stress and salt stress to different degrees. Among them, only *ARF4* expression increased gradually with the increase of stress time under drought stress and salt stress, suggesting that it plays an important role in the response to abiotic stress.

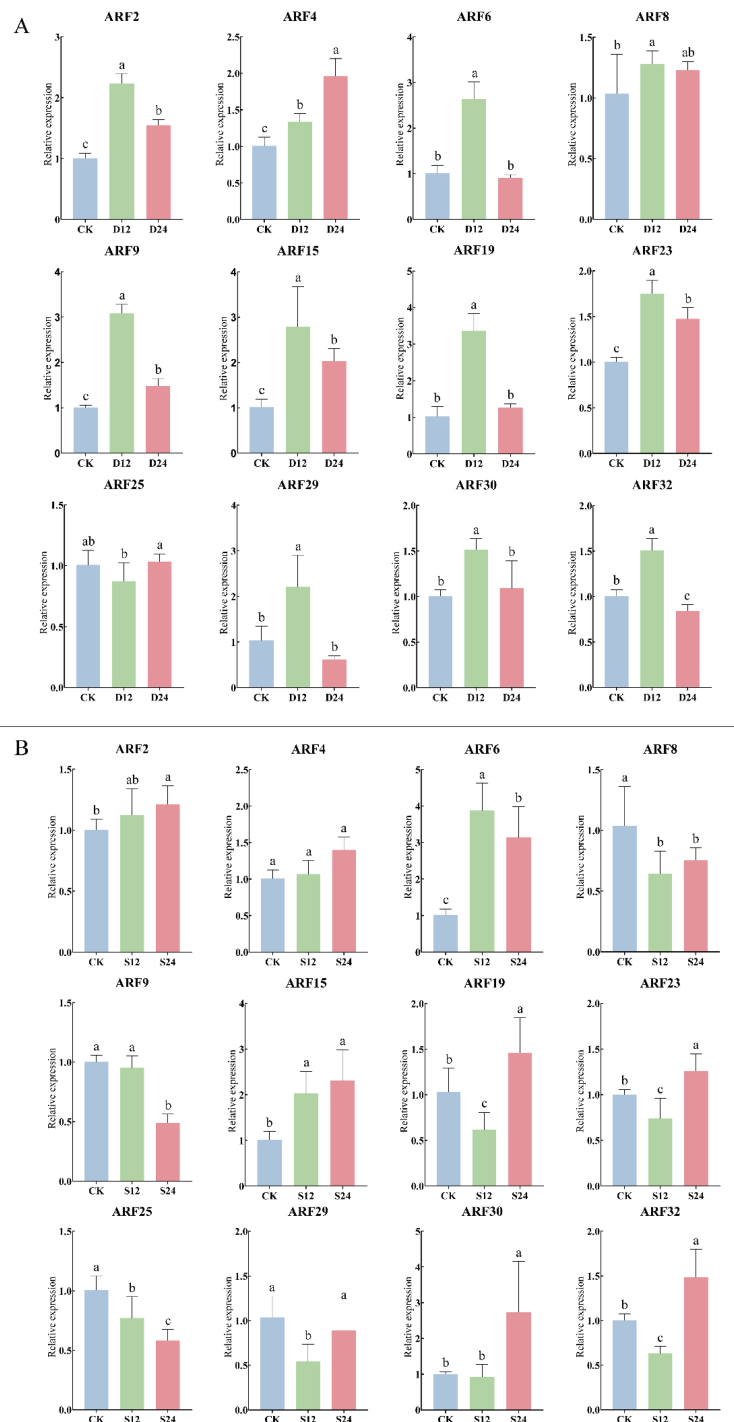
## Discussion

Traditional gene family analysis is usually based on a single reference genome. Given the limitations inherent to a single reference genome, it becomes challenging to ascertain the presence or absence of genes across multiple reference genomes. Following the publication of the maize pan-genome, a more thorough and extensive analysis of the maize gene family can be conducted. The maize pan-genome encompasses the genomes of 26 distinct maize lines constructed using B73 as the reference genome, thereby incorporating genes that are absent from the reference genome. Utilizing the maize pan-genome, this investigation identified 35 members of the maize ARF family. In comparison to the 31 members identified by Xing et al. (Xing et al., 2011) the information regarding ARF family members derived from the pan-genome is more comprehensive. Of the 35 ZmARF family members, only 33 genes were discerned within the reference genome, a phenomenon commonly observed across various species. For instance, the maize pan-genome identified 32 ZmTPS genes, representing an increase of three compared to the 29 genes found in the reference genome (Sun et al., 2023). The number of TPS members present in the rice reference genome is fewer in comparison to the pan-genome (Sun et al., 2022). The PAV analysis

revealed that only 21 out of the 35 gene family members were consistently present across all maize varieties. These core genes are likely to play crucial roles in the growth and development of maize. Furthermore, the remaining genes were not universally absent across all varieties, thereby ensuring the complementarity of the genomes among the different varieties.

Gene structural variations typically encompass deletions, insertions, copy number variations (CNVs), inversions, and translocations, which significantly influence plant phenotypic traits, gene expression, and protein functionality (Gabur et al., 2019; Yuan et al., 2021). Furthermore, structural variation significantly influences crop resistance to both biotic and abiotic stresses (Tao et al., 2019). For instance, the structural variation in the rice restorer gene *Rf20* enables the restoration of fertility in wild abortive cytoplasmic male sterility (CMS-WA) rice lines (Song et al., 2024); The maize *ZmWUS1* gene promotes the occurrence of duplication events in the *Bif3* mutant. The expression of the *ZmWUS1* gene is inhibited, resulting in altered maize meristem size (Chen et al., 2021). Based on the maize pan-genome structural variation file, this study identified the structural variation information of ARF family members in different genomes. The results indicated significant differences in the expression levels of four genes, namely *ARF2*, *ARF3*, *ARF4*, and *ARF25*. Further verification revealed that structural variation (SV) significantly affects the conserved domain of the gene. Further analysis of structural variants on cis-acting elements revealed that, Structural variants lead to differences in the composition of cis-acting elements between the Ki3 strain and the reference genome B73, may further lead to changes in some physiological activities of ARF family members in these two strains.

Maize kernels are composed of two main parts: the embryo and the endosperm. These kernels typically undergo three maturation



**FIGURE 7** Expression of 12 ARF genes in response to abiotic stress (The horizontal axis represents the different stress treatment times; The vertical axis represents the relative expression of gene). **(A)** Expression of ARF genes under drought stress ( $P < 0.05$ , The error bars indicate standard deviation. D12, Drought stress for 12 h; D24, Drought stress for 24 h). **(B)** Expression of ARF genes under salt stress ( $P < 0.05$ , The error bars indicate standard deviation. S12, Salt stress for 12 h; S24, Salt stress for 24 h).

processes: early development, filling, and dehydration (Doll et al., 2017). Research has indicated that the development of maize kernels is typically influenced by plant hormones, sugars, receptors, and transcription factors (Li et al., 2021). Auxin, as a vital plant hormone, plays a crucial role in the development of maize kernels. Auxin regulates physiological processes, including

plant growth and development, by interacting with auxin response factors (Chen et al., 2020; Cancé et al., 2022). Analysis of transcriptome expression in seeds and embryos at various developmental stages revealed that ZmARF members were highly expressed during the early developmental phase of maize seeds and embryos. ARF family members may function as the initial signal for

maize kernel maturation and facilitate the signaling process involved in this maturation. Furthermore, our analysis revealed that during the development of maize embryos and kernels, the highly expressed ARF family members exhibited a complementary relationship; specifically, members that were highly expressed in embryos demonstrated lower expression levels in seeds, while those highly expressed in seeds exhibited lower levels in embryos. Moreover, among the 12 ARF members that were highly expressed in embryos, 8 were identified as core genes; similarly, among the 16 members highly expressed in kernels, 13 were classified as core genes. These findings suggest that ZmARF plays a crucial role in the development of both maize embryos and kernels.

Previous studies have found that ARF family members are involved in plant response to abiotic stress. For example, overexpression of *ZmARF1* can significantly enhance the tolerance of transgenic Arabidopsis plants to abiotic stresses such as low phosphorus stress, drought stress and salt stress (Liu et al., 2024a); Transcriptome studies showed that *ARF10* and *ARF14* were involved in the response of maize to drought stress and were significantly up-regulated under drought stress (Zou et al., 2024). Based on this, further study on the responses of ARF family members to abiotic stress revealed that ARF family members had different responses to drought stress and salt stress. 12 ARF family members were up-regulated under drought stress, 8 members were up-regulated under salt stress, and 4 members were down-regulated under salt stress. In conclusion, ARF members are involved in maize responses to drought stress and salt stress, and ARF members may be more sensitive to drought stress.

In summary, this study provides an in-depth exploration of maize ARF family based on high-quality pan-genome data. The impact of structural variation on the number of ARF family members, gene structure, conserved domains and cis-acting elements in maize was revealed. In addition, the important roles of ARF family members in maize physiological processes such as grain maturation and response to abiotic stress were further elucidated. These results provide effective theoretical support for the growth and development mechanism of maize.

## Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: PRJNA237837.

## References

- Blakeslee, J. J., Spatola Rossi, T., and Kriebchaumer, V. (2019). Auxin biosynthesis: spatial regulation and adaptation to stress. *J. Exp. Bot.* 70, 5041–5049. doi: 10.1093/jxb/erz283
- Cancé, C., Martin-Arevalillo, R., Boubekeur, K., and Dumas, R. (2022). Auxin response factors are keys to the many auxin doors. *New Phytol.* 235, 402–419. doi: 10.1111/nph.18159
- Casanova-Sáez, R., Mateo-Bonmati, E., and Ljung, K. (2021). Auxin metabolism in plants. *Cold Spring Harb. Perspect. Biol.* 13, a039867. doi: 10.1101/cshperspect.a039867
- Chaudhary, C., Sharma, N., and Khurana, P. (2023). Genome-wide identification of Aux/IAA and ARF gene families in bread wheat (*Triticum aestivum* L.). *Protoplasma* 260, 257–270. doi: 10.1007/s00709-022-01773-y
- Chen, C., Wu, Y., Li, J., Wang, X., Zeng, Z., Xu, J., et al. (2023). TBtools-II: A “one for all, all for one” bioinformatics platform for biological big-data mining. *Mol. Plant* 16, 1733–1742. doi: 10.1016/j.molp.2023.09.010
- Chen, L., Li, Y.-X., Li, C., Shi, Y., Song, Y., Zhang, D., et al. (2020). The retromer protein ZmVPS29 regulates maize kernel morphology likely through an auxin-dependent process(es). *Plant Biotechnol. J.* 18, 1004–1014. doi: 10.1111/pbi.13267

## Author contributions

Q-CM: Data curation, Software, Writing – original draft. Y-QW: Formal analysis, Supervision, Writing – review & editing. S-JG: Validation, Writing – review & editing. Z-CG: Validation, Writing – review & editing. Z-PP: Funding acquisition, Resources, Writing – review & editing. J-HC: Conceptualization, Project administration, Resources, Writing – review & editing.

## Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the National Key Research and Development Program (2023YFD2301500).

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2024.1506853/full#supplementary-material>

### SUPPLEMENTARY FIGURE 1

Analysis of cis-acting elements of the promoter. (A) Cis-acting elements of the B73 promoter. (B) Cis-acting elements of the Ki3 promoter.

- Chen, Z., Li, W., Gaines, C., Buck, A., Galli, M., and Gallavotti, A. (2021). Structural variation at the maize WUSCHEL1 locus alters stem cell organization in inflorescences. *Nat. Commun.* 12, 2378. doi: 10.1038/s41467-021-22699-8
- Doll, N. M., Depège-Fargeix, N., Rogowsky, P. M., and Widiez, T. (2017). Signaling in early maize kernel development. *Mol. Plant* 10, 375–388. doi: 10.1016/j.molp.2017.01.008
- Freire-Rios, A., Tanaka, K., Crespo, I., van der Wijk, E., Sizentsova, Y., Levitsky, V., et al. (2020). Architecture of DNA elements mediating ARF transcription factor binding and auxin-responsive gene expression in Arabidopsis. *Proc. Natl. Acad. Sci. U.S.A.* 117, 24557–24566. doi: 10.1073/pnas.2009554117
- Gabur, I., Chawla, H. S., Snowdon, R. J., and Parkin, I. A. P. (2019). Connecting genome structural variation with complex traits in crop plants. *Theor. Appl. Genet.* 132, 733–750. doi: 10.1007/s00122-018-3233-0
- Gomes, G. L. B., and Scortecci, K. C. (2021). Auxin and its role in plant development: structure, signalling, regulation and response mechanisms. *Plant Biol. (Stuttg)* 23, 894–904. doi: 10.1111/plb.13303
- Hufford, M. B., Seetharam, A. S., Woodhouse, M. R., Chougule, K. M., Ou, S., Liu, J., et al. (2021). *De novo* assembly, annotation, and comparative analysis of 26 diverse maize genomes. *Science* 373, 655–662. doi: 10.1126/science.abg5289
- Hurst, L. D. (2002). The Ka/Ks ratio: diagnosing the form of sequence evolution. *Trends Genet.* 18, 486. doi: 10.1016/s0168-9525(02)02722-1
- Ito, K., and Murphy, D. (2013). Application of ggplot2 to pharmacometric graphics. *CPT Pharmacom. Syst. Pharmacol.* 2, e79. doi: 10.1038/psp.2013.56
- Jobert, F., Yadav, S., and Robert, S. (2023). Auxin as an architect of the pectin matrix. *J. Exp. Bot.* 74, 6933–6949. doi: 10.1093/jxb/erad174
- Lescot, M., Déhais, P., Thijs, G., Marchal, K., Moreau, Y., Van de Peer, Y., et al. (2002). PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res.* 30, 325–327. doi: 10.1093/nar/30.1.325
- Leyser, O. (2018). Auxin signaling. *Plant Physiol.* 176, 465–479. doi: 10.1104/pp.17.00765
- Li, P., Li, Z., Xie, G., and Zhang, J. (2021). Trihelix transcription factor zmThx20 is required for kernel development in maize. *Int. J. Mol. Sci.* 22, 12137. doi: 10.3390/ijms222212137
- Liu, L., Gong, Y., Yahaya, B. S., Chen, Y., Shi, D., Liu, F., et al. (2024a). Maize auxin response factor ZmARF1 confers multiple abiotic stresses resistances in transgenic Arabidopsis. *Plant Mol. Biol.* 114, 75. doi: 10.1007/s11103-024-01470-9
- Liu, L., Yahaya, B. S., Li, J., and Wu, F. (2024b). Enigmatic role of auxin response factors in plant growth and stress tolerance. *Front. Plant Sci.* 15. doi: 10.3389/fpls.2024.1398818
- Maren, N. A., Duduit, J. R., Huang, D., Zhao, F., Ranney, T. G., and Liu, W. (2023). Stepwise optimization of real-time RT-PCR analysis. *Methods Mol. Biol.* 2653, 317–332. doi: 10.1007/978-1-0716-3131-7\_20
- McLaughlin, H. M., Ang, A. C. H., and Østergaard, L. (2021). Noncanonical auxin signaling. *Cold Spring Harb. Perspect. Biol.* 13, a039917. doi: 10.1101/cshperspect.a039917
- Morffy, N., and Strader, L. C. (2022). Structural aspects of auxin signaling. *Cold Spring Harb. Perspect. Biol.* 14, a039883. doi: 10.1101/cshperspect.a039883
- Nguyen, L.-T., Schmidt, H. A., von Haeseler, A., and Minh, B. Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* 32, 268–274. doi: 10.1093/molbev/msu300
- Powers, S. K., and Strader, L. C. (2020). Regulation of auxin transcriptional responses. *Dev. Dynam.* 249, 483. doi: 10.1002/dvdy.139
- Ribba, T., Garrido-Vargas, F., and O'Brien, J. A. (2020). Auxin-mediated responses under salt stress: from developmental regulation to biotechnological applications. *J. Exp. Bot.* 71, 3843–3853. doi: 10.1093/jxb/eraa241
- Rienstra, J., Hernández-García, J., and Weijers, D. (2023). To bind or not to bind: how AUXIN RESPONSE FACTORS select their target genes. *J. Exp. Bot.* 74, 6922–6932. doi: 10.1093/jxb/erad259
- Roychoudhry, S., and Kepinski, S. (2022). Auxin in root development. *Cold Spring Harb. Perspect. Biol.* 14, a039933. doi: 10.1101/cshperspect.a039933
- Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504. doi: 10.1101/gr.1239303
- Song, S., Li, Y., Qiu, M., Xu, N., Li, B., Zhang, L., et al. (2024). Structural variations of a new fertility restorer gene, RF20, underlie the restoration of wild abortive-type cytoplasmic male sterility in rice. *Mol. Plant* 17, 1272–1288. doi: 10.1016/j.molp.2024.07.001
- Stajich, J. E., Block, D., Boulez, K., Brenner, S. E., Chervitz, S. A., Dagdigian, C., et al. (2002). The Bioperl toolkit: Perl modules for the life sciences. *Genome Res.* 12, 1611–1618. doi: 10.1101/gr.361602
- Sun, Y., Xiao, W., Wang, Q.-N., Wang, J., Kong, X.-D., Ma, W.-H., et al. (2023). Multiple variation patterns of terpene synthases in 26 maize genomes. *BMC Genomics* 24, 46. doi: 10.1186/s12864-023-09137-3
- Sun, Y., Zhang, P.-T., Kou, D.-R., Han, Y.-C., Fang, J.-C., Ni, J.-P., et al. (2022). Terpene synthases in rice pan-genome and their responses to chilo suppressalis larvae infesting. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.905982
- Szklarczyk, D., Kirsch, R., Koutrouli, M., Nastou, K., Mehryary, F., Hachilif, R., et al. (2023). The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res.* 51, D638–D646. doi: 10.1093/nar/gkac1000
- Tan, S., Luschnig, C., and Friml, J. (2021). Pho-view of auxin: reversible protein phosphorylation in auxin biosynthesis, transport and signaling. *Mol. Plant* 14, 151–165. doi: 10.1016/j.molp.2020.11.004
- Tao, Y., Zhao, X., Mace, E., Henry, R., and Jordan, D. (2019). Exploring and exploiting pan-genomics for crop improvement. *Mol. Plant* 12, 156–169. doi: 10.1016/j.molp.2018.12.016
- Verma, S., Negi, N. P., Pareek, S., Mudgal, G., and Kumar, D. (2022). Auxin response factors in plant adaptation to drought and salinity stress. *Physiol. Plant* 174, e13714. doi: 10.1111/ppl.13714
- Wang, D., Pei, K., Fu, Y., Sun, Z., Li, S., Liu, H., et al. (2007). Genome-wide analysis of the auxin response factors (ARF) gene family in rice (*Oryza sativa*). *Gene* 394, 13–24. doi: 10.1016/j.gene.2007.01.006
- Wang, D., Zhang, Y., Zhang, Z., Zhu, J., and Yu, J. (2010a). KaKs\_Calculator 2.0: a toolkit incorporating gamma-series methods and sliding window strategies. *Genomics Proteomics Bioinf.* 8, 77–80. doi: 10.1016/S1672-0229(10)60008-3
- Wang, K., Li, M., and Hakonarson, H. (2010b). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* 38, e164. doi: 10.1093/nar/gkq603
- Xing, H., Pudake, R. N., Guo, G., Xing, G., Hu, Z., Zhang, Y., et al. (2011). Genome-wide identification and expression profiling of auxin response factor (ARF) gene family in maize. *BMC Genomics* 12, 178. doi: 10.1186/1471-2164-12-178
- Yuan, Y., Bayer, P. E., Batley, J., and Edwards, D. (2021). Current status of structural variation studies in plants. *Plant Biotechnol. J.* 19, 2153–2163. doi: 10.1111/pbi.13646
- Zou, C., Tan, H., Huang, K., Zhai, R., Yang, M., Huang, A., et al. (2024). Physiological Characteristic Changes and Transcriptome Analysis of Maize (*Zea mays* L.) Roots under Drought Stress. *Int. J. Genomics* 2024, 5681174. doi: 10.1155/2024/5681174