

RESEARCH

Open Access



Pan-genome identification and expression analysis of the ARF gene family in potato

Quancai Man¹, Wei Li¹, Shunjuan Gao¹, Fangming Liu², Shuo Sun², Xiaojing Liu², Baisong Lin², Lei Wang², Mingya Ding², Xiaotian Chen^{2*} and Jianghui Cui^{1*}

Abstract

Background The auxin response factor (ARF) is a key regulator involved in plant growth, development, and stress response. In this study, we systematically identified the ARF gene family using a pan-genomic approach based on 47 high-quality potato genomes.

Results Among the 28 members of the potato ARF pan-gene family, no core or private genes were identified. Only one gene was classified as near-core, while the remaining genes were considered non-essential. Ka/Ks analysis indicated that a single gene was under positive selection, whereas the rest of the family members were subject to purifying selection. Structural variations across the 21 potato genomes did not significantly alter the expression of ARF pan-gene family members, and no notable expression differences were detected. However, these structural variations did lead to changes in conserved domains in certain strains. RNA-seq analysis revealed differential expression of ARF family members under drought stress compared to control conditions, with variation in expression levels among different genes. Further investigation suggested that transcription factors are involved in regulating the expression of ARF gene family members.

Conclusions Our findings revealed the function of the auxin response-factor family members in biological processes in potato, and provide a new theoretical reference for drought resistance-breeding in this important crop species.

Keywords Auxin response factor, Differentially expressed genes, Drought stress, Pan-gene family, Potato

Background

Auxin is an indispensable plant hormone, and changes in its concentration regulate plant growth and development, and various life activities [1]. For example, changes in auxin concentration can control cell division and differentiation, thereby regulating the development of seed embryos [2]. Indeed, auxin polar transport and local biosynthesis are essential for plant meristem formation and lateral organ development [3]. Further, the interaction between auxin and ethylene plays an important role in the developmental transition from flowering to fruit ripening [4]. Additionally, auxin plays a significant role in plant stress responses. Thus, for example, auxins can reduce plant damage caused by heavy metals by reducing

*Correspondence:

Xiaotian Chen
xiaotian0301@163.com

Jianghui Cui
13663123545@126.com

¹College of Agriculture, Hebei Agricultural University, Baoding, Hebei 071000, China

²Weichang county manchu and Mongolian potato research institute, Chengde, Hebei 067000, China



their uptake and promoting their chelation and sequestration by plant tissues [5]. Furthermore, auxin biosynthesis plays an important role in plant resistance to drought and salinity [6].

Auxin signal transduction comprises two components: the classic, nuclear auxin signal-transduction system and the recently discovered extranuclear auxin signal-transduction system [7]. In turn, the core auxin signal-transduction system consists of three parts: receptors, inhibitors, and transcriptional export regulators [8]. In this context, the auxin response factor (ARF) is an important component of auxin signal transduction that plays an important role in determining the specificity of the auxin response [9]. Particularly, ARF transcription factors (TFs) include three categories: transcription activators A, and transcription inhibitors B and C. Different types of ARF TFs regulate distinct auxin responses [10]. Thus, class A ARFs are the core components of nuclear auxin signaling, and their functions are closely related to auxin concentration. In this case, ARF activates gene expression under the presence of auxin and suppresses gene expression in its absence [11]. Studies have shown that the special relationship between class A ARFs and auxin plays an important role in specific developmental processes, such as embryogenesis and lateral root formation [12]. Furthermore, ARF expression regulates plant responses to abiotic stress factors, such as drought and salinity, and enhances plant resistance to stress [13].

Auxin response factors have been identified as key TFs that regulate the auxin response, and their functions in many plants have been elucidated. Further, the expression patterns of 25 ARF family members identified in sweet potatoes indicate that they play an important role in regulating root development in this species [14]. Similarly, 81 ARF family members identified in alfalfa (*Medicago sativa*) were specifically expressed under abiotic stress, including drought, salt, and high and low temperatures, and their expression was closely related to stress duration [15]. In all, 660 ARF family members were identified in the blueberry genome as playing important roles in fruit development and the pH stress response [16]. These reports highlight the fact that ARFs are widely involved in plant physiological processes such as organ development and responses to various abiotic stress factors.

Previously, 20 ARF members were identified in a single reference potato genome as playing important roles in growth and development, and plant stress responses [17]. These findings revealed the evolutionary relationship between ARF family members in a single potato reference genome; however, the presence or absence of ARF family members in different potato genomes were not verified. The potato pan-genome contains the presence-deletion and structural variation (SV) information of 46

potato reference genomes. However, it has yet to become a valuable resource for potato gene family research [18].

A comprehensive analysis of the ARF family in potatoes is necessary to further verify the functions of ARF family members in this crop species. Therefore, this study aimed to identify ARF family members based on the potato pan-genome and compare the results with those of previous studies. In particular, we determined the presence or absence of ARF family members in each genome and analyzed the effects of gene SV on gene structure, gene expression, conserved domains, and the expression of ARF family members under drought stress.

Results

Pan-genome identification and presence/absence variation analysis

In this study, 28 ARF family members were identified in the potato pan-genome, with significant differences detected in the number of ARF family members among the genomes. Among them, PG6359 contained the largest number of members (14), whereas PG1013 contained the smallest (6) number (Fig. 1A). Further analysis of the presence/absence of ARF family members revealed that ARF5 and ARF6 are soft-core genes(present in 42–46 samples), while all other members were shell genes(present in 2–41 samples). No core(present in all samples) or accession-specific genes(present in 1 sample) were found. ARF21 was present only in PG6002 and PG6241, which may be related to its specific function (Fig. 1B). The potato ARF family member ID and gene names correspond to Table S1.

Phylogenetic analysis of ARF in potato

To further understand the evolutionary relationships among ARF family members, phylogenetic trees of *Arabidopsis* and potato were constructed and divided into six components. In addition to G1, potato ARF family members were unevenly distributed across all components, among which G2 contained 13 potato ARF family members and was the only soft-core gene. *Arabidopsis* and potato ARF family members were most abundant in G6, suggesting that these genes may be more closely related (Fig. 2).

Auxin response-factor pressure analysis

To explore the selective pressure on ARF family members in potatoes during evolution, we calculated the Ka/Ks ratio values for ARF members in the potato pan-genome. The peak of ARF24 in some varieties was between 1 and 2, which had a positive selection effect. The remaining ARF family members had a peak Ka/Ks value between 0 ~ 1, but there were significant differences in the location of the peak, among which ARF22 had the highest peak. (Fig. 3A). Among the 28 ARF family members, *ARF2*,

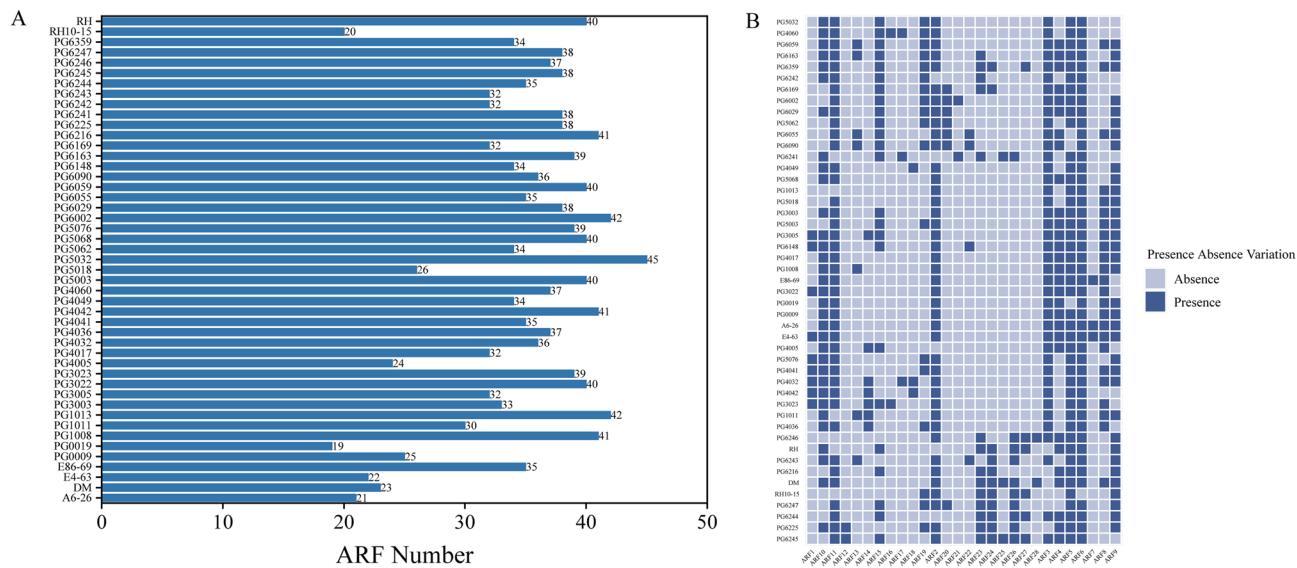


Fig. 1 Number of ARF family members according to presence/absence analysis. (A) ARF family members in different strains; (B) Presence/absence analysis of ARF family members(Dark blue represents presence, Light blue represents absence)

ARF4, *ARF5*, *ARF6*, *ARF8*, *ARF9*, *ARF11*, *ARF15*, *ARF19* and *ARF24* had Ka/Ks values greater than 1 in some lines, indicates that these genes were under the influence of purifying selection, and the Ka/Ks values of the remaining genes were less than 1 in all lines(Fig. 3B).

Influence of structural variants on ARF genes

Huang et al. used read mapping and assembly-based methods to identify structural variants ≥ 50 bp in length, including insertions, deletions, inversions, duplications, and translocations. The analysis of the effect of structural variation on gene expression using the structural variation files provided in the study of Huang et al., showed that no significant difference was found between genes with and without SV, indicates that SV did not affect the expression of ARF family members in the reference genome (DM) in potato stolons. (Table S2). For conserved domain analysis, we selected the lines with the highest SV overlap, PG1011 and PG6241 (Fig. 4). The results showed that there were seven conserved domains corresponding to DM(reference genome) and three non-corresponding domains between DM and PG1011 (Fig. 4A); additionally, eight conserved domains corresponded to DM, and two did not correspond to DM or PG6241 (Fig. 4B). Lastly, the amino acid sequences of the corresponding conserved domains changed, indicating that the conserved domain was significantly affected by SV.

The response of ARFs to drought stress

Owing to their shallow root system, potatoes are extremely sensitive to drought stress [19]. Analysis of the transcriptome expression data showed that ARF family

members were differentially expressed under drought stress. Thus, at 3 h after drought stress, *ARF6*, *ARF8*, *ARF10*, *ARF24* and *ARF26* were significantly downregulated, whereas the expression levels of *ARF5* and *ARF25* were significantly upregulated after 6 h of drought stress, and those of *ARF2*, *ARF4*, *ARF9*, *ARF1*, *ARF23* and *ARF28* were significantly upregulated after 12 h of drought stress (Fig. 5A). Further analysis of the correlation between ARF family members and TFs showed that bZIP, WRKY, bHLH, GATA, and ERF were involved in the regulation of *ARF5* after 3 h of drought stress (Fig. 5B). Meanwhile, at 6 h after drought stress, bZIP, WRKY, bHLH, GATA, and ERF were found to be involved in regulating *ARF5* and *ARF25* (Fig. 5C). However, no TFs involved in the regulation of ARF family members were observed after 12 h of drought treatment.

Discussion

Plant genome sequences provide valuable information for studying the significance of genes or gene families for evolution. Gene family analysis based on a single reference genome includes the identification of family members, gene structure, evolutionary relationships, and expression patterns [20, 21]. However, this analysis has some limitations and cannot fully explain gene differences among different genomes. Thus, the construction of pan-genomes in many species can supplement the lack of information on the presence/absence of genes [22]. In this study, we systematically analyzed the ARF family members based on the potato pan-genome constructed by Huang et al. [18].

Gene presence/absence variation (PAV) is common in plants and plays an important role in the study of

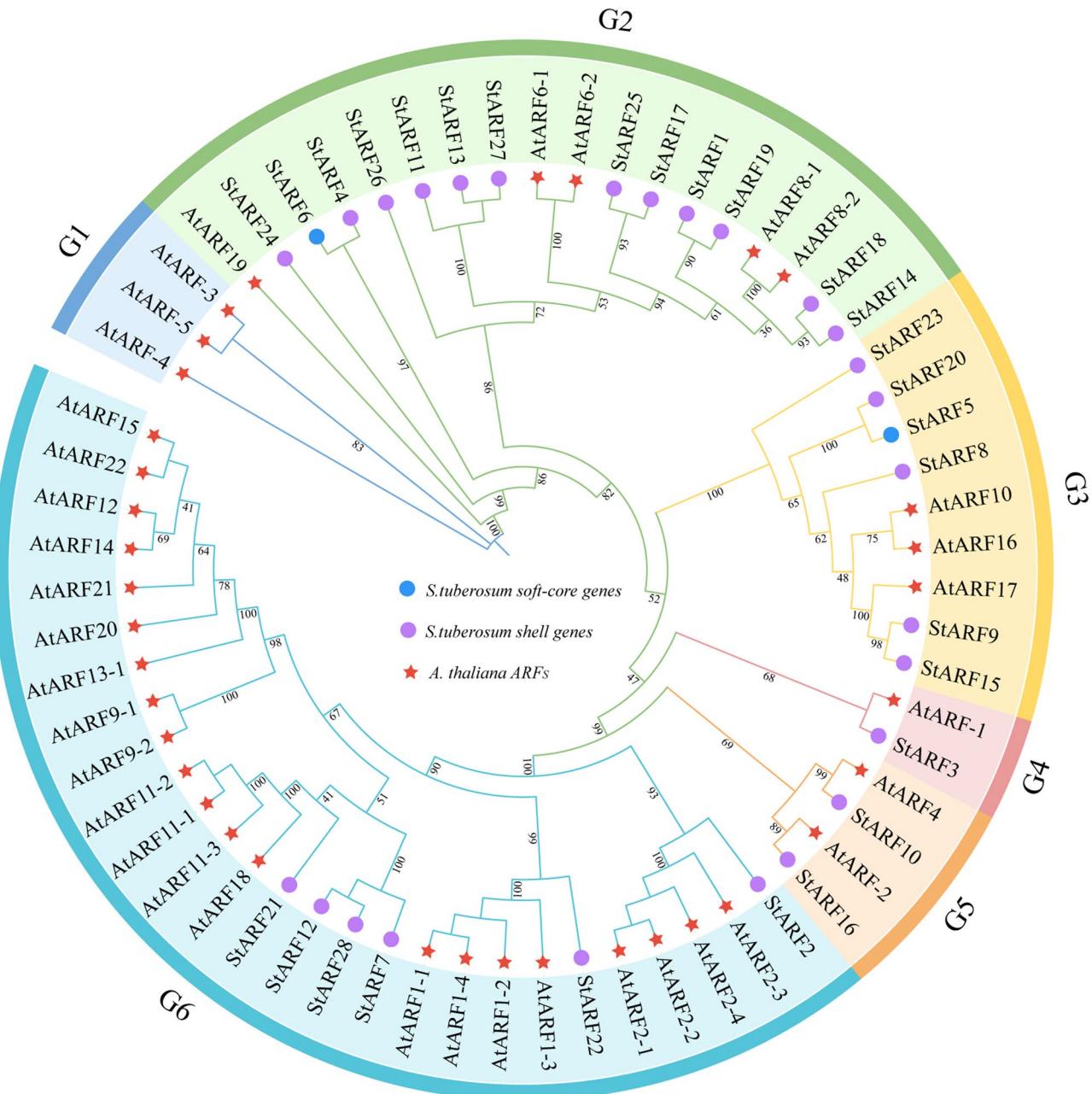


Fig. 2 Phylogenetic tree of ARFs from *Arabidopsis* and potato (Colored circles represent subgroups, bootstrap values are only shown to be greater than 40)

plant genetic diversity, gene identification, and molecular marker development [23]. Furthermore, dispensable genes resulting from PAV variation in Amborellaceae can enhance resistance to abiotic stress and improve environmental adaptability [24]. Further, PAV variants have been identified in other species; thus, for example, in a pan-gene family analysis of maize TPS, only 20 of 32 family members were identified as core genes; that is, they were present in all genomes [25]. Similarly, in a rice TPS pan-gene family analysis, 32 TPS family members were found in all genomes [26]. Meanwhile, in the

analysis of the cucumber WOX pan-gene family, the 9930 genomes reportedly contained 11 family members, which was the most common, and XTMC contained nine family members, which was the least common [27]. As for this study, 28 ARF family members were identified in 47 potato genomes. The reference genome DM contained 13 ARF family members, with PG6359 containing the largest number (14) of ARF family members, whereas PG1013 contained the smallest number (6). At the same time, this study identified a larger number of ARF family members than the single reference genome (8 more members than

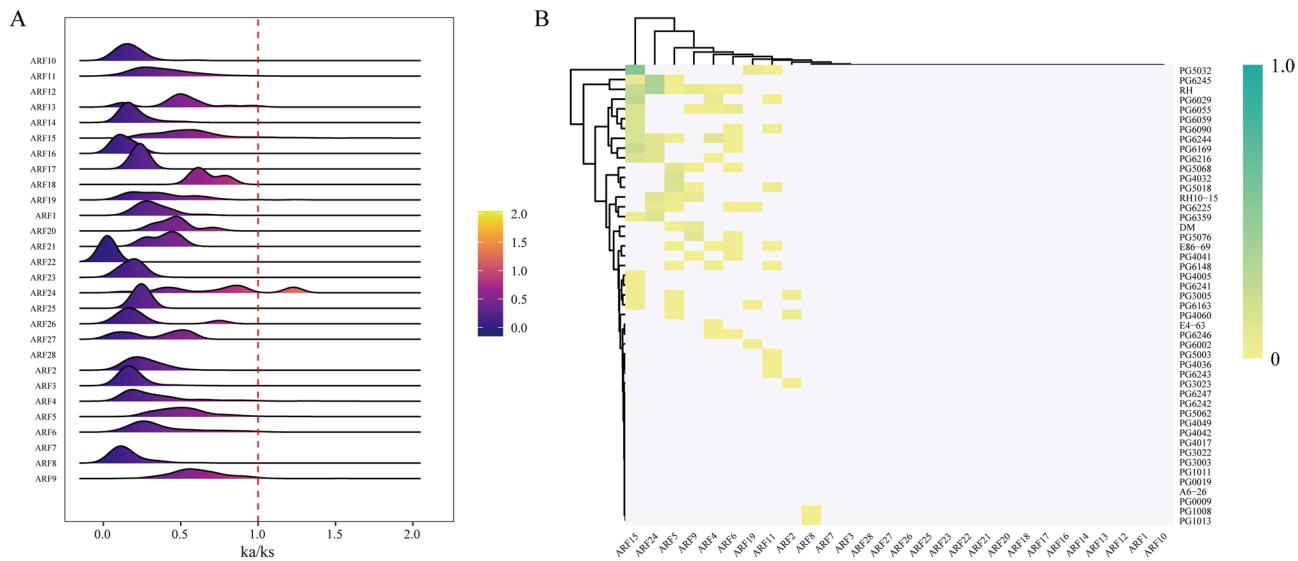


Fig. 3 Ka/Ks ratio values for ARFs. **(A)** Ka/Ks values of ARFs in the potato pan-genome; **(B)** Heatmap of the frequency of ARF family members with Ka/Ks > 1 in the genome(Gray: The Ka/Ks ratio of this gene in this line was less than 1)

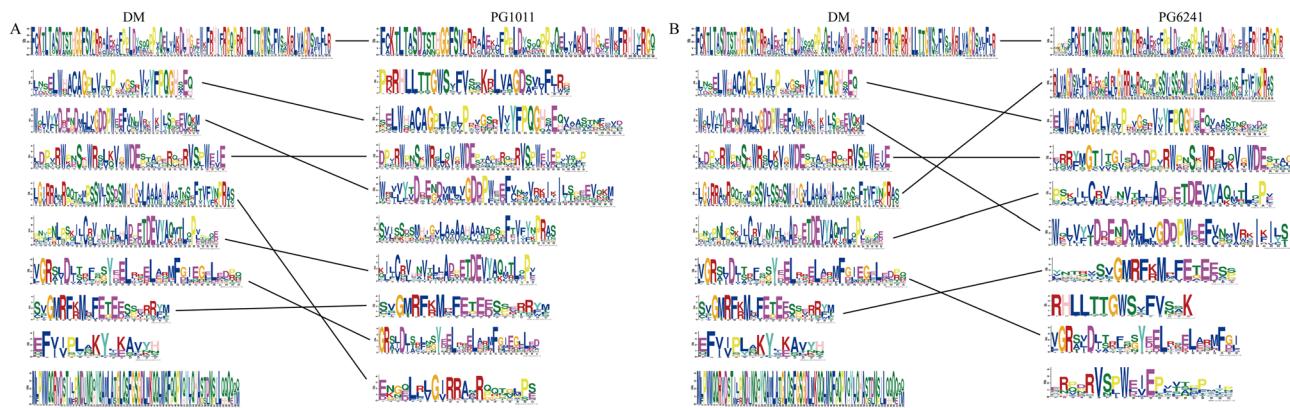


Fig. 4 Effects of structural variation on conserved domains(The same Motif is connected by the black lines). **(A)** Conserved domains of DM and PG1011; **(B)** Conserved domains of DM and PG6042

the single reference genome), and Presence/Absence variants are prominent in the potato pan-genome, these results indicated that ARF family members were lost during evolution, which led to the differences in characteristics among different strains.

As the main source of genetic diversity, including inversions, duplications, deletions, and translocations, SVs play an important role in improving plant agricultural traits and analyzing evolutionary events [28, 29]. Thus, for example, Xie et al. found that structural variation had a significant impact on peach fruit traits, with a 1.67 Mb heterozygous inversion leading to the separation of flat fruit shapes [30], and pan-genome studies of tomato have found that structural variation plays an important role in gene expression, phenotypic differences between different lines, and in improving plant resistance to stress [31]. Comparing the genomes of cultivated and wild potato species, structural variation has been identified

as the key to potato genome diversity. Moreover, structural variation has been determined to play an important role in improving potato resistance to biotic and abiotic stress factors [32]. No significantly differentially expressed genes were found in the stolon expression profile of the 21 potato lines, indicating that structural variation did not affect the expression of ARF family members in stolon of these lines. Therefore, we speculated that the function of ARF family members in regulating stolon development was not affected by SV during the evolution of these potato lines, and ARF was involved in regulating stolon development in these lines to ensure the normal operation of potato growth and development, the effect of structural variation on the expression of other tissues needs to be further verified. However, other studies have reported that SV does affect the conserved domains of ARF. Compared to the reference genome DM, PG1011, and PG6241 had seven and eight conserved domains,

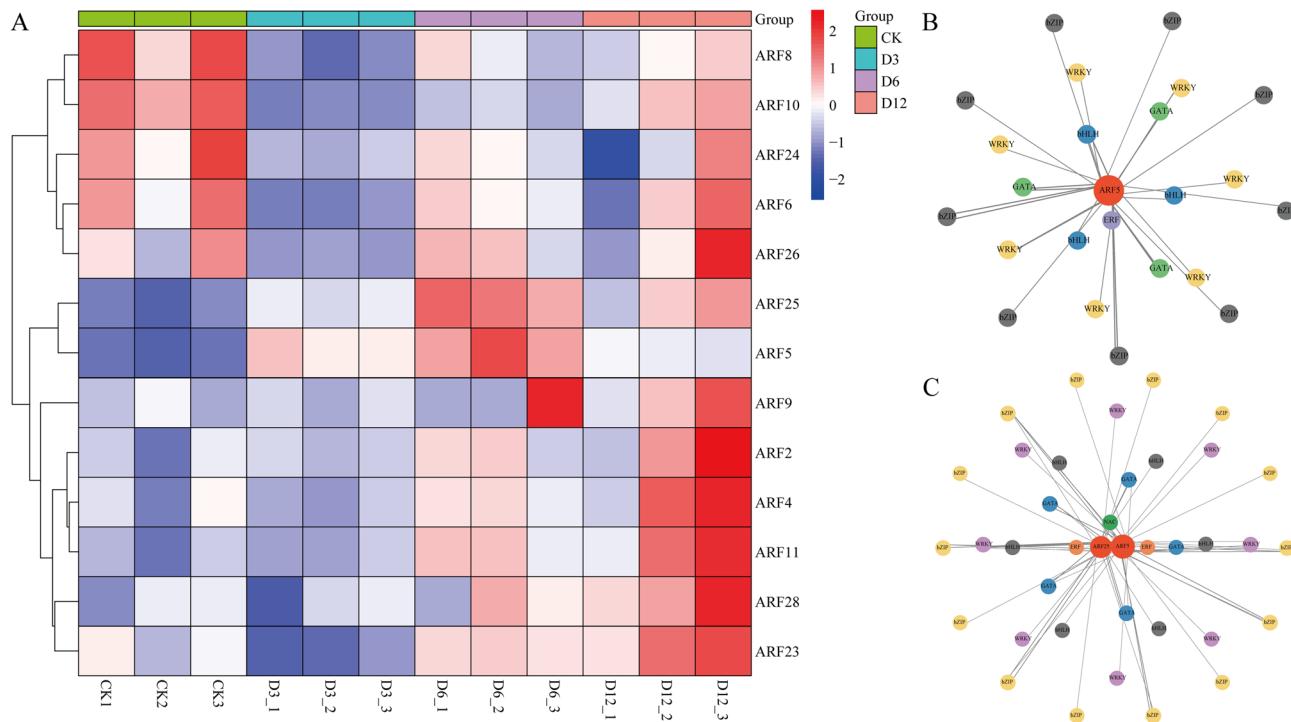


Fig. 5 Response of potato ARF to drought stress. **(A)** Heatmap of differential gene expression in ARF under drought treatment(CK: Drought stress 0 h, D3: Drought stress 3 h, D6: Drought stress 6 h, D12: Drought stress 12 h); **(B)** Interaction network between ARF and transcription factors after a 3 h drought treatment; **(C)** Interaction network between ARF and transcription factors after a 6 h drought treatment

respectively. However, the amino acids of the corresponding conserved domains did not fully match, indicating that the latter were affected by SV. As described in a maize TPS pan-gene family study, the conserved domains of TPS family genes were significantly affected by SV, resulting in a large number of atypical genes [25]. Consistently, a rice TPS pan-gene family study found that the effect of SV on conserved domains led to the generation of many atypical genes in the rice TPS family [26].

As key components of the auxin signaling pathway, ARFs are involved in the regulation of plant growth and development, and stress responses. In particular, when the inhibitory factor ARF2 was knocked out, tomato stress-related genes such as reactive oxygen species (ROS)-scavenging and defense-related genes were significantly upregulated, thereby improving stress resistance [33]. Concomitantly, Hao et al. found that when the ARF4 gene was knocked out, tomato stomatal morphology and vascular bundle development were affected, thereby enhancing drought resistance [34]. In addition, ARF1 acts as a stress response factor, with its overexpression significantly improving crop tolerance to abiotic stress [35]. In short, different types of ARFs play different roles in plant growth and development, and stress resistance. Meanwhile, Based on the transcriptome results, we found that ARF6, ARF8, ARF10 and ARF24 may be

transcriptional represses. ARF5 and ARF25 may be the response factors of ARF family to drought stress.

The function of TFs is to control gene expression in the appropriate cells to ensure normal plant biological processes [36]. TFs such as NAC, MYB, WRKY, and ERF improve soybean resistance to biotic and abiotic stresses by regulating N-acetylserotonin O-methyltransferase expression and soybean auxin activity [37]. The analysis of the interaction network between transcription factors and ARF family(Spearman's correlation coefficients: $p < 0.05, |r| > 0.5$) showed that transcription factors such as bZIP, WRKY, bHLH, GATA and ERF were involved in regulating the expression of ARF family members, which ensured the normal operation of life activities under drought stress and resisted the damage of potato under drought stress.

Conclusions

In this study, ARF family members were identified for the first time within the potato pan-genome, yielding a total of 28 ARF genes. Phylogenetic analysis classified these 28 members into six distinct subfamilies. Further analyses were conducted to assess selection pressure acting on the ARF family and to evaluate the impact of structural variations on these genes. Transcriptome data revealed a potential mechanism through which ARF members respond to drought stress, and indicated the involvement

of transcription factors—including bZIP, WRKY, bHLH, GATA, and ERF—in regulating ARF expression. In summary, this work expands the known repertoire and functional understanding of the ARF gene family in potato, and provides a more comprehensive theoretical foundation for investigating mechanisms of potato development.

Materials and methods

Identification and presence/absence variation in the ARF gene family

From the Pan-Potato Database (<http://solomics.agis.org.cn/potato/>, <http://218.17.88.60/potato/>) to download the Potato 46 pan-genome and gene annotation files, Potato reference genome (DM v6.1) and gene annotation files were obtained from Spud DB (<http://spuddb.uga.edu/>). The ARF domain (PF06507) hidden Markov model file was downloaded from the Pfam database (<http://pfam.xfam.org>), Pan-genome ARF family members were confirmed by searching ARF domains (parameters: e < 1E-5) with the HMMER program and by homology alignment [38]. The preliminary genome-wide ARF members of the family were submitted to the SMART CDD (<http://smart.embl.de/>) database (<https://www.ncbi.nlm.nih.gov/cdd/>) [39, 40]. The presence of ARF domains was verified to confirm the final ARF family members in the pan-genome.

Huang et al. used a reannotation approach to reidentify gene families in the potato pangenome. The ARF family pan-gene list was extracted from the potato pan-gene list provided in the study by Huang et al., to obtain the presence/absence information of ARF family members in different genomes. The ggplot2 package was then used to map the presence and deletion of ARF genes in the potato pan-genome [41].

Phylogenetic tree analysis of ARF gene family

The protein sequences of Arabidopsis and potato ARF family members were extracted to construct phylogenetic trees. Multiple sequence alignments of Arabidopsis and potato protein sequences were performed using Muscle comparison method, and phylogenetic trees were constructed using IQ-TREE (Maximum Likelihood, Bootstrap: 1000) [42]. The obtained evolutionary tree was submitted to iTOL V6 (<https://itol.embl.de/>) for graphical enhancement [43].

Ka/Ks ratio calculation for ARF genes

Auxin response-factor member protein sequences and CDS sequences of each genome were extracted from the potato pan-genome, and Ka/Ks values for the ARF members were calculated using KaKs_Calculator 2.0 [44]. Then, the R packages ggridge and ggplot2 were used to draw Ka/Ks ridge maps and heat maps of the ARF family members with Ka/Ks > 1.

Influence of structural variation among ARF family members on ARF expression

From the Pan - Potato Database (<http://solomics.agis.org.cn/potato/>, <http://218.17.88.60/potato/>) to download the Potato structure variation information file (http://solomics.agis.org.cn/potato/ftp/variation/potato_VCF_44sp.vcf.gz), and the ARF members in the family structure information were extracted. Huang et al. detected the expression levels in different tissues of potato, but did not cover all strains. We used stolon expression data from different potato lines reported by Huang et al. (PRJNA754534) to calculate the effect of structural variation on the expression of ARF family members. Wilcoxon and t-tests were used to determine whether there was a significant difference in ARF expression between structural and non-structural variants. The structure showed that there was no significant difference among ARF family members.

Effect of structural variation on conserved domains

We reasoned that SV affected the conserved domain of the ARF gene. Therefore, we selected the genome lines PG6011 and PG6241 with the most structural variant gene family members and submitted the protein sequences of the reference genome to meme suite//meme-suite.org/meme/tools/meme). weblogos mapping the protein sequences of the three genomic ARF family members (conserved base number: 10) to test our hypothesis [45].

Analysis of the expression of ARF family members under drought stress

We downloaded potato drought-stress treatment RNA-seq data (PRJNA728834) from the NCBI database and calculated ARFs expression using the DM strain as the reference genome [46]. The expression of potato ARFs under drought stress at 0 h(CK), 3 h(D3), 6 h(D6), and 12 h (D12) was analyzed and an expression heatmap was drawn using the pheatmap package. Spearman's correlation coefficients of TFs and their family members were calculated using the Psych software package ($p < 0.05, |r| > 0.5$). Finally, cytoscape v3.10.1 was used to plot the co-expression network of TFs and ARF family members [47].

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12870-025-07447-0>.

Supplementary Material 1.

Supplementary Material 2.

Authors' contributions

Q.M.、J.C. and X.C. conceived and designed the study, Q.M.、W.L. and S.G. analyzed the data and plotted, Q.M.、F.L.、S.S. and X.L. wrote the manuscript, B.L.、L.W. and M.D. reviewed and revised the manuscript.

Funding

Hebei province detoxigenic potato breeding technology innovation center(SG2012016).
 Hebei Potato Seed Potato Industry Technology Research Institute(SC2021028).
 Hebei (Chengde) Potato Seed Potato Industry Technology Research Institute improved innovation ability(202206F013).

Data availability

RNA-seq data were obtained from the NCBI SRA database(PRJNA754534, PRJNA728834).

Declarations**Ethics approval and consent to participate**

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 2 December 2024 / Accepted: 18 September 2025

Published online: 22 October 2025

References

1. Casanova-Sáez R, Voß U. Auxin metabolism controls developmental decisions in land plants. *Trends Plant Sci*. 2019;24:741–54.
2. Verma S, Attuluri VPS, Robert HS. An essential function for auxin in embryo development. *Cold Spring Harb Perspect Biol*. 2021;13:a039966.
3. Wang Y, Jiao Y. Auxin and above-ground meristems. *J Exp Bot*. 2018. <https://doi.org/10.1093/jxb/erx299>.
4. An J, Althiab Almasaud R, Bouzayen M, Zouine M, Chervin C. Auxin and ethylene regulation of fruit set. *Plant Sci*. 2020;292:110381.
5. Moeen-Ud-Din M, Yang S, Wang J. Auxin homeostasis in plant responses to heavy metal stress. *Plant Physiol Biochem*. 2023. <https://doi.org/10.1016/j.plaphy.2023.108210>.
6. Gamalero E, Glick BR. Recent advances in bacterial amelioration of plant drought and salt stress. *Biology*. 2022;11:437.
7. Pérez-Henríquez P, Yang Z. Extranuclear auxin signaling: a new insight into auxin's versatility. *New Phytol*. 2023;237:1115–21.
8. Ma Y, Wolf S, Lohmann JU. Casting the net-connecting auxin signaling to the plant genome. *Cold Spring Harb Perspect Biol*. 2021;13:a040006.
9. Matthes MS, Best NB, Robil JM, Malcomber S, Gallavotti A, McSteen P. Auxin evoevo: conservation and diversification of genes regulating auxin biosynthesis, transport, and signaling. *Mol Plant*. 2019;12:298–320.
10. Jing H, Korasick DA, Emenecker RJ, Morffy N, Wilkinson EG, Powers SK, et al. Regulation of AUXIN RESPONSE FACTOR condensation and nucleo-cytoplasmic partitioning. *Nat Commun*. 2022. <https://doi.org/10.1038/s41467-022-31628-2>.
11. Israeli A, Schubert R, Man N, Teboul N, Serrani Yarce JC, Rosowski EE, et al. Modulating auxin response stabilizes tomato fruit set. *Plant Physiol*. 2023;192(3):2336–55.
12. Israeli A, Capua Y, Shwartz I, Tal L, Meir Z, Levy M, et al. Multiple auxin-response regulators enable stability and variability in leaf development. *Curr Biol*. 2019. <https://doi.org/10.1016/j.cub.2019.04.047>.
13. Marzi D, Brunetti P, Saini SS, Yadav G, Puglia GD, Dello Ilio R. Role of transcriptional regulation in auxin-mediated response to abiotic stresses. *Front Genet*. 2024;15:1394091.
14. Pratt IS, Zhang B. Genome-wide identification of ARF transcription factor gene family and their expression analysis in sweet potato. *Int J Mol Sci*. 2021;22:9391.
15. Chen F, Zhang J, Ha X, Ma H. Genome-wide identification and expression analysis of the Auxin-Response factor (ARF) gene family in medicago sativa under abiotic stress. *BMC Genomics*. 2023;24:498.
16. Li X, Zhang X, Shi T, Chen M, Jia C, Wang J, et al. Identification of ARF family in blueberry and its potential involvement of fruit development and pH stress response. *BMC Genomics*. 2022;23:329.
17. Song S, Hao L, Zhao P, Xu Y, Zhong N, Zhang H, et al. Genome-wide identification, expression profiling and evolutionary analysis of auxin response factor gene family in potato (*Solanum tuberosum* group Phureja). *Sci Rep*. 2019. <http://doi.org/10.1038/s41598-018-37923-7>.
18. Tang D, Jia Y, Zhang J, Li H, Cheng L, Wang P, et al. Genome evolution and diversity of wild and cultivated potatoes. *Nature*. 2022. <https://doi.org/10.1038/s41586-022-04822-x>.
19. Gervais T, Creelman A, Li X-Q, Bizimungu B, De Koeyer D, Dahal K. Potato response to drought stress: physiological and growth basis. *Front Plant Sci*. 2021;12:698060.
20. Safder I, Shao G, Sheng Z, Hu P, Tang S. Genome-wide identification studies - a primer to explore new genes in plant species. *Plant Biol (Stuttg)*. 2022;24:9–22.
21. Bernal-Gallardo JJ, de Folter S. Plant genome information facilitates plant functional genomics. *Planta*. 2024;259:117.
22. Shi J, Tian Z, Lai J, Huang X. Plant pan-genomics and its applications. *Mol Plant*. 2023;16:168–86.
23. Bie H, Li Y, Zhao Y, Fang W, Chen C, Wang X, et al. Genome-wide presence/absence variation discovery and its application in peach (*Prunus persica*). *Plant Sci*. 2023;335:111778.
24. Hu H, Scheben A, Verpaalen B, Tirnaz S, Bayer PE, Hodel RGJ, et al. Amborella gene presence/absence variation is associated with abiotic stress responses that may contribute to environmental adaptation. *New Phytol*. 2022;233:1548–55.
25. Sun Y, Xiao W, Wang Q-N, Wang J, Kong X-D, Ma W-H, et al. Multiple variation patterns of terpene synthases in 26 maize genomes. *BMC Genomics*. 2023;24:46.
26. Sun Y, Zhang P-T, Kou D-R, Han Y-C, Fang J-C, Ni J-P, et al. Terpene synthases in rice Pan-genome and their responses to *Chilo suppressalis* larvae infesting. *Front Plant Sci*. 2022;13:905982.
27. Yin S, Zhao L, Liu J, Sun Y, Li B, Wang L, et al. Pan-genome analysis of WOX gene family and function exploration of CsWOX9 in cucumber. *Int J Mol Sci*. 2023;24:17568.
28. Kou Y, Liao Y, Toivainen T, Lv Y, Tian X, Emerson JJ, et al. Evolutionary genomics of structural variation in Asian rice (*Oryza sativa*) domestication. *Mol Biol Evol*. 2020;37:3507–24.
29. Li H, Wang S, Chai S, Yang Z, Zhang Q, Xin H, et al. Graph-based pan-genome reveals structural and sequence variations related to agronomic traits and domestication in cucumber. *Nat Commun*. 2022;13:682.
30. Guan J, Xu Y, Yu Y, Fu J, Ren F, Guo J, et al. Genome structure variation analyses of Peach reveal population dynamics and a 1.67 mb causal inversion for fruit shape. *Genome Biol*. 2021;22:13.
31. Alonge M, Wang X, Benoit M, Soyk S, Pereira L, Zhang L, et al. Major impacts of widespread structural variation on gene expression and crop improvement in tomato. *Cell*. 2020;182:145–e16123.
32. Kyriakidou M, Achakkagari SR, Gálvez López JH, Zhu X, Tang CY, Tai HH, et al. Structural genome analysis in cultivated potato taxa. *Theor Appl Genet*. 2020;133:951–66.
33. El Mamoun I, Bouzroud S, Zouine M, Smouni A. The knockdown of AUXIN RESPONSE FACTOR 2 confers enhanced tolerance to salt and drought stresses in tomato (*Solanum lycopersicum* L). *Plants*. 2023;12:2804.
34. Chen M, Zhu X, Liu X, Wu C, Yu C, Hu G, et al. Knockout of auxin response factor SIARF4 improves tomato resistance to water deficit. *Int J Mol Sci*. 2021;22:3347.
35. Liu L, Gong Y, Yahaya BS, Chen Y, Shi D, Liu F, et al. Maize auxin response factor ZmARF1 confers multiple abiotic stresses resistances in transgenic Arabidopsis. *Plant Mol Biol*. 2024;114:75.
36. Hussain Q, Asim M, Zhang R, Khan R, Farooq S, Wu J. Transcription factors interact with ABA through gene expression and signaling pathways to mitigate drought and salinity stress. *Biomolecules*. 2021;11:1159.
37. Ma Q, Xia Z, Cai Z, Li L, Cheng Y, Liu J, et al. GmWRKY16 enhances drought and salt tolerance through an ABA-mediated pathway in *Arabidopsis thaliana*. *Front Plant Sci*. 2019;9:1979.
38. Potter SC, Luciani A, Eddy SR, Park Y, Lopez R, Finn RD. HMMER web server: 2018 update. *Nucleic Acids Res*. 2018;46:W200–4.
39. Letunic I, Khedkar S, Bork P. SMART: recent updates, new developments and status in 2020. *Nucleic Acids Res*. 2021;49:D458–60.
40. Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C et al. CDD: a conserved domain database for the functional annotation of proteins. *Nucleic Acids Res*. 2011;39 Database issue:D225–9.
41. Ito K, Murphy D. Application of ggplot2 to pharmacometric graphics. *CPT: Pharmacometrics & Systems Pharmacology*. 2013;2:e79.

42. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. IQ-tree: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol*. 2015;32:268–74.
43. Letunic I, Bork P. Interactive tree of life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Res*. 2024;52:W78–82.
44. Wang D, Zhang Y, Zhang Z, Zhu J, Yu J. Kaks_calculator 2.0: a toolkit incorporating gamma-series methods and sliding window strategies. *Genomics Proteomics Bioinformatics*. 2010;8:77–80.
45. Bailey TL, Johnson J, Grant CE, Noble WS. The MEME suite. *Nucleic Acids Res*. 2015;43:W39–49.
46. Jian H, Sun H, Liu R, Zhang W, Shang L, Wang J, et al. Construction of drought stress regulation networks in potato based on SMRT and RNA sequencing data. *BMC Plant Biol*. 2022;22:381.
47. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003;13:2498–504.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.