

Genome-wide analysis of DnaJA proteins in *Nicotiana tabacum* reveals that NtDnaJA3 responses to drought stress

Qiao Wang

Beijing Life Science Academy

Lijun Meng

China Tobacco Gene Research Center, Zhengzhou Tobacco Research Institute of CNTC

Zechao Qu

China Tobacco Gene Research Center, Zhengzhou Tobacco Research Institute of CNTC

Huan Su

Beijing Life Science Academy

Jiemeng Tao

Beijing Life Science Academy

Peng Lu

Beijing Life Science Academy

Jianfeng Zhang

Beijing Life Science Academy

Peijian Cao

Beijing Life Science Academy

Jingjing Jin

jinjingjing1218@126.com

Beijing Life Science Academy

Research Article

Keywords: Tobacco, DnaJA, Phylogenetic analysis, Regulatory network, Expression pattern, Drought stress

Posted Date: December 6th, 2024

DOI: <https://doi.org/10.21203/rs.3.rs-5051527/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Background

DnaJA proteins, a prominent subfamily of the DnaJ family, function as molecular chaperones that respond to various external stresses. Extensive studies on the DnaJ family have been conducted in plants. However, research on this subfamily in tobacco remains relatively scarce.

Results

In this study, we identified 24 DnaJA genes in tobacco, and classified them into three individual groups. A comprehensive analysis based on gene structure, motif composition, and evolutionary pattern revealed the divergence of tobacco DnaJA genes. For the evolution of the *NtDnaJA* genes, purification selection was the major factor. In addition, the potential regulatory network unveiled that *NtDnaJAs* could be regulated by miRNAs and various transcription factors associated with diverse stress responses. Through expression pattern analysis and qRT-PCR experiments, it was observed that many *NtDnaJAs* displayed tissue-specific expression and might play significant roles in different biotic and abiotic stresses. Additionally, the pivotal role of *NtDnaJA3* in boosting plant drought resistance was confirmed.

Conclusions

This study provides important perspectives on the evolution of *NtDnaJA* genes and their involvement in stress responses, laying the groundwork for future research into the roles of DnaJA regulatory genes in tobacco.

Graphical Abstract

Background

As the largest category of molecular chaperones, heat shock proteins (HSPs) were discovered in the 1960s in *Drosophila melanogaster* [1], participating in regulating processes such as protein folding, degradation, localization, accumulation, and plant signal transduction [2]. In humans, HSPs have been studied for their important roles in cancer, immunity, and ferroptosis [3–5]. HSPs have also been extensively studied in various species including aquatic organisms, viruses, *Arabidopsis* [6–8]. In plants, upon perceiving heat stress (HS) signals, HSPs initiate a series of cellular signaling cascades. These cascades regulate the activation of HSPs and other stress-induced genes, which help mitigate the damages caused by HS and protect other proteins from the detrimental effects of these stressors [9]. Based on their relative molecular weights, HSPs families in plants can be further classified into several sub-families: HSP110, HSP90, HSP70, HSP60, HSP40, and small HSPs [10, 11]. HSP40, also known as

DnaJ, is the most prominent family of HSPs. They strongly stimulate the ATPase activity of HSP70, which is crucial for inducing conformational changes in stable chaperone protein [12].

Commonly, DnaJ proteins typically consist of three domains: the highly conserved N-terminal J-domain, a zinc finger domain, and a C-terminal domain. These conserved domains allow DnaJ proteins to be categorized into four distinct types. Type I DnaJ proteins, also known as DnaJ homolog subfamily A members (DnaJA), contain all three domains, whereas Type II DnaJ (DnaJB) lack the CxxCxGxG zinc-finger domain. Type III DnaJ (DnaJC) only have the J-domain. Proteins in the Type IV DnaJ (DnaJD) are considered as DnaJ-like proteins because they contain an incomplete HPD (histidine, proline, and aspartic acid residues) tripeptide of the J-domain. Despite lacking many common residues, their overall structure is similar to the J-domain, and also refer to as J-like proteins. Interestingly, DnaJD proteins function in an HSP70-independent manner [10]. Moreover, given the evolutionary conservation of DnaJ proteins, they have been identified across different species. These proteins typically play crucial roles in plant growth, development, and stress resistance, highlighting their functional significance in plants [13–15]. For instance, in *Catalpa bungei*, overexpression of *CbuDnaJ49* resulted in transgenic plants exhibiting albino leaves, and significantly reduced chlorophyll and carotenoid contents compared to wild type [16]. In wheat, 16 *TaDnaJ* genes were up-regulated upon wheat yellow mosaic virus (WYMV) infection, indicating that the DnaJ family was involved in plant defense responses [17]. In soybean, DnaJ family could interact with HSP70 to enhance stress tolerance in chimeric soybean hairy roots [18]. These findings have underscored the importance of the DnaJ family.

As the evolutionarily conserved HSP40 type I (DnaJA), there are ample evidences supporting its critical functions. Recent studies [19] have highlighted the importance of other domains beyond the J-domain. Except the core J-domain of DnaJ proteins, the zinc-finger motif facilitates the interaction between DnaJA proteins and their substrate proteins [20]. This interaction is critical for the chaperone function of DnaJA, which helps recognize and bind unfolded or partially folded proteins, guiding them towards proper folding. The zinc-finger motif binds zinc ions, which stabilizes the three-dimensional structure of the DnaJA protein [12]. This stability is essential for the protein to perform its molecular chaperone functions effectively within the cellular environment. Besides, the C-terminal domain is crucial for binding to substrate proteins. This domain recognizes and interacts specifically with unfolded or partially folded proteins, ensuring that DnaJA can capture and stabilize these substrates [21]. These canonical DnaJA proteins are widely present in plants. For instance, eight DnaJA proteins have been described in *Arabidopsis* with exemplary functions including plasma membrane H⁺-ATPase regulation and the integration of flowering signals [22–24]. Overexpression lines of *GmDnaJA6* in soybean showed increased saline and alkaline tolerance [18]. In tobacco, the ectopic expression of *VaDJI* reduced sensitivity to ABA during seed germination and early seedling growth, by improving drought stress tolerance, reducing oxidative damage, and increasing chlorophyll content [25].

Although DnaJA research has been extensively conducted in several plants, studies in tobacco remain relatively limited. Tobacco (*Nicotiana tabacum* L.) is a widely cultivated economic crop across the globe and serves as a model plant for genetic and synthetic biology studies. Additionally, as a classic

allopolyploid, tobacco holds significant importance in evolutionary research. Recent studies have revealed that overexpression of *NtDnaJ1* enhanced drought tolerance in transgenic *Arabidopsis* [26]. However, the specific functions of DnaJ genes, particularly those of the DnaJA subfamily in tobacco, remained elusive. In this study, we aimed to conduct a genome-wide identification and characterization of the DnaJ gene family in tobacco, and specifically focused on DnaJA genes. A comprehensive analysis was performed on the classification, evolution, gene structure, three-dimensional structure, and potential regulatory networks of DnaJA genes in tobacco. Meanwhile, we explored gene expression patterns at different stages of tissue development and in response to various stresses. Additionally, the critical role of *NtDnaJA3* during drought response was confirmed. Our research provided a detailed exploration of the biological functions of DnaJA genes in tobacco, emphasizing their significant roles under diverse stress treatment.

Results

Genome-wide identification and characterization of tobacco DnaJA proteins

Genome-wide analysis revealed 24 DnaJA genes in tobacco, designated as *NtDnaJA1* -*NtDnaJA24* based on their homological relationships (Fig. S1; Table S1). Amino acid sequences of NtDnaJAs ranged from 418 to 535. The predicted isoelectric point and molecular weight ranged from 5.86 to 9.45, and 46,466 to 10,6703, respectively (Table S1). The theoretical isoelectric point was highest for *NtDnaJA22* and *NtDnaJA23* (9.45), whereas *NtDnaJA3* had the lowest value (5.86) (Table S1). Besides, subcellular localization predictions revealed that the majority of NtDnaJA proteins (22 out of 24) were localized in the nucleus, while only small number (2 out of 24) were found in the chloroplast (Table S1). This distribution suggested the primary functional role of the NtDnaJA proteins within the nuclear compartment.

The phylogenetic and structural features of DnaJA family genes in tobacco

The protein sequences of the DnaJA family from *Glycine max*, *Zea mays*, *Oryza sativa*, *Arabidopsis thaliana*, and *Sorghum bicolor* were selected for phylogenetic analysis in order to investigate the evolutionary history of DnaJA genes in tobacco (Fig. 1A; Table S2). As shown in Fig. 1A, DnaJA family could be categorized into three distinct clusters: I, II, and III, whereas tobacco DnaJA genes were unevenly distributed among three clusters, with half of them in cluster I. *NtDnaJAs* in the same cluster harbored similar exon–intron structures, except *NtDnaJA11*, *NtDnaJA14*, and *NtDnaJA18* which lacked distinct UTRs (Fig. 1B). *NtDnaJAs* in cluster I contained 6 or 7 exons, cluster II ranged from 7 to 11 exons, and cluster III comprised 17 to 19 exons. These findings indicated significant sequence conservation and consistent exon–intron structures within each cluster. The motifs analysis revealed 15 conserved structural domains within NtDnaJA family proteins (Fig. 1B). Seven of them (motif 1, 2, 3, 4, 6, 10, and 12) were present among all NtDnaJA proteins. Cluster I showed exactly the same motif pattern,

whereas proteins in cluster II and III had a slightly different motif composition, especially for motif 15, 13, and 9. Interestingly, two unknown motifs (10 and 11) could be found in most *NtDnaJA* members, and motif 12 (HSP40/DnaJ peptide-binding) exhibited tandem replication in most of members among cluster II and III (Fig. 1B; Table S3). All 24 *NtDnaJA* proteins contained α -helixes, β -sheets, and loops (Fig.S2).

Analysis of gene duplication events and collinearity among tobacco DnaJA genes

Synteny analysis was conducted among *N. tabacum*, *S. bicolor*, *Z. mays*, *O. sativa*, *A. thaliana*, and *G. max* (Fig. 2A). The 24 identified tobacco DnaJA genes were distributed across 11 of the 24 chromosomes (Fig. S3), with 1 to 3 genes per chromosome. A comparison of the *NtDnaJA* family with *S. bicolor*, *Z. mays*, *O. sativa*, *A. thaliana*, and *G. max* revealed 20, 3, 1, 10, and 2 pairs of homologous gene pairs, respectively. Numerous highly conserved orthologous gene were revealed between *NtDnaJAs* and *GmDnaJAs*, such as *NtDnaJA4* and *GmDnaJA2/GmDnaJA5*, *NtDnaJA10* and *GmDnaJA8/GmDnaJA16/GmDnaJA14*. In addition, a total of 22 pairs of colinear tobacco DnaJA gene pairs were identified, with the chromosomes 2, 5, 6, 9, and 10 holding the most colinear gene pairs (Fig. 2A). Given the importance of gene duplication in the expansion of gene families, various duplication events were investigated in this study. A total of 21 duplication events were identified by DupGen_finder, and all of them were dispersed duplication (DSD) among the five duplication types (Table S4). Additionally, a large number of genes (18/24) underwent 1–3 rounds of DSD duplication events (Table S4). This suggested that *NtDnaJA* genes had a relatively uniform evolutionary pattern, and their functional differentiation was mainly through multiple rounds of dispersed duplication. The evolutionary trends of duplicated *NtDnaJA* gene pairs were studied by calculating the non-synonymous (K_a) and synonymous (K_s) substitution rates. Commonly, a K_a/K_s ratio greater than 1 generally denoted positive selection, while a ratio of 1 implied neutral evolution, and a value below 1 meant purifying or negative selection. The results indicated that the K_a/K_s ratios for all examined duplication events were under 1, with values spanning from 0.02 to 0.35 (Table S5). This demonstrated that all the examined duplicate gene pairs experienced significant purifying selective pressures throughout their evolutionary history.

Cis-acting elements and regulatory networks for NtDnaJAs

Cis-regulatory elements (CREs) in gene promoter regions were regarded as the important regulators of gene expression, playing crucial roles in shaping the spatiotemporal patterns of protein-coding gene expression. Potential CREs of *NtDnaJA* were identified within the 2,000 bp upstream region of the translation start codon, aiming to gain a more detailed insight into the transcriptional regulation and potential biological functions of the *NtDnaJA* family. These identified CREs could be categorized into five groups based on their types (Fig. 3; Table S6), including light responsive (40%), phytohormone responsive (32%), stress responsive (16%), transcription factor (8%) and plant development (4%). A majority of *NtDnaJA* genes harbored numerous light-responsive elements, suggesting their potential association with light response. Methyl jasmonate (CGTCA-motif and TGACG-motif), gibberellin (GARE-motif and P-box), salicylic acid (TCA-element), and abscisic acid (ABRE) responsive elements

constituted the largest portion of plant hormone responsive elements. The abundance of phytohormone responsive elements indicated that hormone signals might modulate the expression levels of *NtDnaJAs* by either activating or inhibiting them. In addition, the promoter regions also showed an enrichment of CAT-box, MSA-like, and GCN4_motifs CREs linked to plant development, which were related to meristem expression, cell cycle regulation, and endosperm expression. Multiple stress-responsive elements were also observed for *NtDnaJAs*, such as pressure-responsive (TC-rich) elements, CGTCA-motif and TGACG-motif, and ARE (for anaerobic induction). These findings suggested that *NtDnaJA* genes might play significant roles in various external stresses and treatments.

Considering that numerous CREs were enriched in the promoter regions of the *NtDnaJAs*, we speculated that certain TFs might have directly regulated *NtDnaJA* genes. Therefore, we sought to explore the relationship between *NtDnaJAs* and TFs using TFDB and STRING. A total of 97 transcription factors from 10 families (Table S7) were identified as potentially significant regulators of *NtDnaJAs*. Among these, we found that HSF, WRKY, bZIP, and bHLH transcription factors were the most abundant (Fig. 4 and Fig.S4), with *NtDnaJA23* and *NtDnaJA5* having the highest number of interactions with transcription factors. WRKYs were known for their involvement in biotic and abiotic stress signaling pathways, while HSF proteins were crucial for heat stress tolerance. This network highlighted the complex regulation and interactions involved in *NtDnaJA* gene expression. Additionally, we employed the PsRNATarget to explore potential miRNA binding sites for *NtDnaJAs*. Finally, 17 miRNA families, comprising 36 miRNAs, were identified to play potential roles in the regulation of *NtDnaJAs* (Fig. 4 and Fig.S4). Most miRNAs were found to have multiple *NtDnaJA* targets. For instance, nta-miR167b could target three *NtDnaJAs* (Fig. 4 and Fig.S4). Conversely, multiple miRNAs could target certain *NtDnaJAs*; for example, *NtDnaJA19* could be targeted by several miRNAs, including miR167, miR395, and miR6025d. Hence, further investigations were needed to elucidate the complex interactions between *NtDnaJAs* and TFs/miRNAs.

Profiling expression patterns of *NtDnaJA* genes in various tissues and development stages

Eight representative tobacco tissues from different developmental stages were selected to study the expression patterns of *NtDnaJAs* across various tissues (Table S8). As shown in the Fig. 5A, tissue-specific expression patterns for *NtDnaJA* gene family were clearly observed, which could be divided into three distinct clusters. Genes in cluster I predominantly showed high expression levels in stems and roots, while their expression was relatively low in leaves. Interestingly, during the development of leaves, we observed low expression levels in young and mature leaves, while expression increased in senescent leaves. On the contrary, genes in cluster II exhibited specific expression throughout the growth and development of leaves, and their expression was relatively low in other tissues. Different from these two clusters, genes in cluster III were almost expressed only in roots, with relatively low expression in senescent flower, indicating that genes in cluster III might play important function in root development. To further validate the reliability of RNA-seq results, six *NtDnaJA* genes with tissue specific expression patterns were randomly selected for further qRT-PCR analysis, including *NtDnaJA1*, *NtDnaJA3*, *NtDnaJA5*, *NtDnaJA14*, *NtDnaJA16*, and *NtDnaJA23*. The expression profile data from RNA-seq largely aligned with the qRT-PCR results (Fig. 5B). Specifically, *NtDnaJA1*, *NtDnaJA3*, and *NtDnaJA5*, which were

grouped in cluster I, exhibited markedly high expression levels in both the roots and stems, implied they might play critical and potentially significant roles in root and stem development. Comparing with flower tissues, *NtDnaJA14* exhibited a little higher in roots and senescent leaves. Notably, *NtDnaJA14*, from cluster II, showed a 4-fold higher expression level in the young leaves compared to other tissues. These observations indicated that *NtDnaJA14* might be crucial for the growth and maintenance of leaves. Similar with RNA-seq results, *NtDnaJA23*, from cluster III, showed high expression pattern in roots. However, it also exhibited relative high expression in stems, which might be due to different sampling stages between RNA-seq and qRT-PCR samples. In addition, both RNA-seq and qRT-PCR analyses showed that most *NtDnaJAs* had relatively low expression levels in flowers. This finding aligned with previous studies on the DnaJA gene family in other plant species [18], whereas most of them were not highly expressed in reproductive organs during developmental stages.

Profiling expression patterns of *NtDnaJA* genes under various abiotic and biotic stresses

Our previous results have demonstrated that the promoters of tobacco DnaJA genes were enriched with CREs (LTR, ARE, CCAAT-box, MBS, ABRE) and transcription factor binding motifs (ERF, MYB, bZIP, bHLH, Dof), which were associated with various abiotic/biotic stresses. The expression patterns of *NtDnaJAs* under various abiotic and biotic stresses were investigated to reveal their potential roles in responding to abiotic/biotic stresses (Fig. 6A and Table S9). Our results revealed that *NtDnaJA* genes of different clusters demonstrated unique response patterns to various stresses. Most genes in cluster I broadly responded to various stresses, especially cadmium treatment, drought and salt stress. The expression of almost all the *NtDnaJA* genes were significantly reduced after topping treatment, whereas most of them were up-regulated in both leaf and root tissues under cadmium treatment. Notably, the response pattern of *NtDnaJAs* in leaf and root tissues under high and low-temperature stresses showed distinct trend. For instance, *NtDnaJA6*, *NtDnaJA8*, *NtDnaJA16* and *NtDnaJA17* exhibited upregulation in leaves under high temperature, whereas showed downregulation in roots. Interestingly, the expression of most of genes in cluster II were specially increased under high temperature both in leaf and root tissues. This observation suggested that these candidates, functioning as molecular chaperones, might play critical roles in helping the plant cope with heat stress. Different from the genes in cluster I and II, only several members in cluster III showed a certain level of response to abiotic/biotic stresses, with *NtDnaJA23* up-regulated under *Ralstonia solanacearum* (RS) and Potato Virus Y (PVY) treatment, and *NtDnaJA24* up-regulated under high temperature treatment.

Next, we focused on six genes, which showed significant responses under cold, cadmium, RS, salt, and drought stresses. qRT-PCR experiments were conducted to further examine their expression patterns (Fig. 6B). Under drought treatment, the expression levels of *NtDnaJA3* and *NtDnaJA5* were significantly up-regulated, with a more than twenty-fold increase comparing with control. *NtDnaJA14* was just slightly up-regulated under drought treatment. Similar with RNA-seq results, *NtDnaJA14* expression was markedly increased under low-temperature. Both *NtDnaJA1*, *NtDnaJA3*, and *NtDnaJA5* were significantly up-regulated under cadmium treatment in root, whereas similar result was only observed for *NtDnaJA1* in leaf, suggesting different response patterns of *NtDnaJA* in different tissues under cadmium treatment.

However, different from RNA-seq analysis, no significant expression change was observed for *NtDnaJA3* and *NtDnaJA5* under salt treatment, which might be due to differences in concentration or treatment time between different studies. As for biotic stress, the expression of *NtDnaJA2* and *NtDnaJA7* was significantly decreased under *R. solanacearum* infection. Overall, the expression patterns of *NtDnaJAs* under various abiotic/biotic treatment highlighted that they could play significant roles in various stress responses.

Silencing of *NtDnaJA3* reduced tobacco's tolerance to drought

We observed that *NtDnaJA3* was significantly upregulated under drought stress and highly expressed in roots. Similar with the prediction, the subcellular localization analysis revealed that *NtDnaJA3* localized to the nucleus (Fig. 7A). In addition, we also observed cell membrane localization for it. This difference might be due to variations in amino acid sequences among DnaJA family members across species, which could affect their signal peptides or localization signals, leading to distinct localization patterns. Furthermore, VIGS technology was applied to construct *DnaJA3*-silenced plants. In successfully silenced plants, the transcript levels of *DnaJA3* were markedly lower compared to wild-type plants (Fig. 7B). Moreover, after five days of drought treatment, the upper leaves of wild-type plants maintained an intact morphology without wilting, whereas the *DnaJA3*-silenced plants exhibited more severe wilting (Fig. 7C). These results suggested that silencing *NtDnaJA3* could significantly reduce drought tolerance.

Discussion

The DnaJA protein family is essential for multiple aspects of plant growth and development. Recently, due to the quality improvement of tobacco genome [27], it has become possible to comprehensively examine tobacco DnaJA family at the genome-wide scale. In this study, 216 members of the DnaJ family were identified, among which 24 belonged to the DnaJA subfamily (Fig. S1). According to recent studies [18, 28–31], the number of identified DnaJA family members in tobacco was greater than that in soybean (19), *Arabidopsis* (7), wheat (20), sorghum (10), grape (8), and pepper (9). Typically, species with a greater number of DnaJA genes have undergone polyploidization events. Common tobacco (as an allotetraploid: $2n = 4x = 48$) and wheat (as a hexaploid: $2n = 6x = 42$), harbored 2 or 3 times DnaJA genes compared to diploid species like *Arabidopsis*. These results indicated that genome size might be correlated with the number of DnaJA gene family members (Table S10). However, common wheat, with a genome size of 16 GB, had a similar number of DnaJA genes comparing with the tetraploid *N. tabacum*, suggesting that polyploidy alone might not be the sole determinant of DnaJA gene family expansion. Additionally, *C. annuum*, with a genome size of 3.5 GB, belonging to the Solanaceae family, had relative fewer DnaJ and DnaJA genes, indicating that other factors might also influence the number of DnaJA genes. Usually, polyploidization, segmental duplication, tandem duplication and non-allelic homologous recombination were the primary forces behind the expansion of gene families. A total of 21 duplication events were identified for tobacco DnaJA genes, and all of them were dispersed duplication. It indicated that dispersed duplication was the sole driver of *NtDnaJAs* evolution. In contrast, other species, such as maize, have different types of gene expansion, with retrotransposons (50%) being the dominant

duplication events [32]. Further analysis of K_a and K_s values indicated that the duplicated gene pairs in tobacco have undergone purifying selection, as all the K_a/K_s ratios were below 1. Similar findings were also observed in soybean and *Arabidopsis* [33, 34]. Overall, these results suggested that the distribution among the *NtDnaJA* subfamilies and the influence of polyploidy were more complex and warranted further investigation.

By homolog search in STRING and PlantTFDB, 97 TF members from 10 different families were identified, which might play significant roles in regulating *NtDnaJAs*. These TF families were known to be involved in stress responses and plant development, including HSF, WRKY, bZIP, and bHLH. The interactions between DnaJ proteins and TFs have been widely investigated in plants. In soybean, HSP40s were targeted by four SNAP-WRKY TFs, which could influence the high nitrogen transcriptional response [35]. *AtERdj3A* might function as the downstream regulators of bZIP28 and bZIP60, and be involved in thermotolerance of plants [36]. The overexpression of *SIDnaJ20* contributed to higher expression levels of HsfA1 and HsfB1 under heat stress [37]. In our study, we also found *NtDnaJA1* and *NtDnaJA3* might interact with HSFs, suggested that they might also played a crucial role during heat stress response. Furthermore, identifying potential miRNA target sites offered valuable insights into the complexity of gene regulatory networks and how miRNAs modulate biological processes by regulating specific genes. In *Medicago sativa*, novel-miR205 could target chaperone DnaJ-like protein and abscisic acid receptor PYL4, which might participate in ABA signal transduction [38]. In *Brassica juncea*, miR2926 could target DnaJ, and involve in heat response [39]. In wheat, the interaction between ata-miR172c-3p and DnaJ chaperone binding proteins, might be crucial in both seed development and the plant's response to heat [40]. In this study, we also tried to predict miRNAs that might target tobacco DnaJA genes. Our analysis revealed 35 miRNAs from 17 miRNA families, which might involve in the regulation of *NtDnaJAs*. Previous studies [40] have demonstrated that *S. viridis* miR397 could increase salt stress sensitivity by repressing three *Arabidopsis* LAC genes. In our study, we found that miR397 might target *NtDnaJA15*, *NtDnaJA18*, and *NtDnaJA22*, indicating that they might also be significant regulators to tobacco's drought tolerance. As we all know, the miR167 family played crucial roles in the development of roots, stems, leaves, and flowers, as well as in flowering time, embryonic development, seed development, and stress responses, by regulating auxin response factors (*ARFs*) and *IAA-Ala resistant3 (IAR3)* genes [41, 42]. Our study revealed that miR167 might target *NtDnaJA19* and *NtDnaJA20*. Further tissue-specific expression patterns (Fig. 6A) indicated that these two *NtDnaJAs* were specifically expressed in roots. Hence, we speculated that these two *NtDnaJAs*, regulated by miR167, might play important roles during root development of tobacco.

By studying cis-regulatory elements, we could not only gain a deeper understanding of the complex networks governing gene regulation, but also uncovered how plants adapt to changing environments through modulation of gene expression during evolution. In this study, a large number of CREs associated with light response, plant growth and development, hormone response, stress response and transcription factor were identified for the promoter regions of *NtDnaJAs* in tobacco. Comparable results have been observed in other plant species, including soybean and *Arabidopsis* [18, 43]. Specifically, light responsive element constituted approximately 39.9% of the CREs in the *NtDnaJA* promoter regions. A

large number of *NtDnaJAs* in cluster II exhibited high expression in leaf tissues (Fig. 6A), suggesting their specific roles in leaf growth and development. Besides, stress responsive element constituted approximately 16.4% of the CREs in the *NtDnaJA* promoter regions, suggesting that *NtDnaJAs* might play crucial roles in tobacco's adaptation to various environmental stresses. Most genes such as *NtDnaJA10*, *NtDnaJA15*, *NtDnaJA18* in cluster II, showed a consistent response pattern under high-temperature (leaves and roots). The significant upregulation of these genes under high-temperature suggested a specialized role in heat stress tolerance, which were critical under thermal stress conditions. Besides, large number of genes in cluster I showed an upregulation pattern under various stresses, including cadmium stress (both in leaves and roots), drought stress in leaves, salt stress in leaves and stems, and PVY infection in leaves. However, an interesting exception was observed for the topping condition, which showed reverse patterns. Previous studies [44, 45] had shown that topping significantly affected the growth and development of the plant, including root growth and leaf yield production. Meanwhile, topping also altered many biological processes such as carbon and nitrogen metabolism, photosynthesis, and secondary metabolism [46–48], which might cause resident to different diseases. These unique physiological changes induced by topping might be attributed to different responses comparing with disease responses. It is noteworthy that *NtDnaJAs* response to drought treatment was specifically explored in this study. Qrt-PCR data indicated that several *NtDnaJAs* exhibited notable changes in expression levels under drought treatment. For instance, the expression level of *NtDnaJA3* was significantly increased under drought treatment in leaf, suggesting that it might play an important role in the plant's response to drought stress. To further investigate it roles during drought stress, VIGS was employed to silence it. Preliminary results indicated that silencing of *NtDnaJA3* led to a compromised drought tolerance phenotype. In summary, there were functional heterogeneity and diversity for the expression patterns of the *NtDnaJAs* under growth and stress treatment. Further studies were still needed to investigate the detail mechanisms of *NtDnaJA* candidates in various stresses.

Conclusions

In this study, we identified and systematically examined the *DnaJA* gene family in tobacco through analyses of phylogeny, gene structure, conserved domains, and motifs. Additionally, comprehensive evolutionary and collinearity analyses, along with the identification of duplicate genes, revealed that purifying selection was the primary factor driving the expansion of *NtDnaJAs*. Our regulatory network analysis indicated that *NtDnaJA* genes might play crucial roles in tobacco development and in responses to various abiotic/biotic stresses. Moreover, silencing the expression of *NtDnaJA3* led to a significant reduction in drought tolerance in tobacco. These findings were pivotal for boosting tobacco's resilience to stress and deserve deeper exploration in future research.

Methods

Identification and analysis of DnaJA family genes in tobacco

Two approaches were used to identify tobacco DnaJA genes. To identify all members of DnaJA family genes in tobacco, the protein sequences of DnaJA from *Arabidopsis* (TAIR: The Arabidopsis Information Resource) [49] were used as queries against the tobacco genome [27] using BLASTP. HMMER (v3.4) [50] was utilized to identify potential DnaJA family genes by employing hidden Markov models (HMM) based on the conserved domains of DnaJA: PF00684 (DnaJ central domain), PF01556 (DnaJ_C-terminal domain), and PF00226 (DnaJ domain), using the default settings. The molecular weight, isoelectric point, and subcellular localization of DnaJA proteins were determined using the SIB database and CELLO (v2.0), respectively [51, 52].

The phylogenetic classification, gene structures, three-dimensional structure and conserved motifs analysis

The DnaJA protein sequences from *N. tabacum*, and other plant species derived from previous studies including *A. thaliana*, *Z. mays*, *O. sativa*, *G. max*, and *S. bicolor*, were used for ClustalW alignment [53]. Using MEGA (v11.0), an unrooted neighbor-joining phylogenetic tree was generated with 1000 bootstrap replicates [54]. ITOL [55] was used to visualize phylogenetic tree. Conserved protein motifs in NtDnaJA proteins were identified with MEME [56]. The identified motifs were annotated using InterProScan [57]. NtDnaJAs tertiary structural characteristics were predicted using SWISS-MODEL [58].

Promoter analysis and interaction network prediction

The identification of CREs in the promoter regions of *NtDnaJAs* was performed using the PlantCARE database [59]. The regulatory relationships between miRNAs and *NtDnaJAs* were investigated using psRNATarget [60] with default settings, and the sequence for tobacco miRNAs were obtained from the miRBase database [61]. The regulatory interactions between *NtDnaJAs* and TFs were derived from the Plant Transcription Factor Database (PlantTFDB) and STRING (v12.0) [62, 63]. Cytoscape (v3.10.2) [64] was used to visualize the interaction networks.

Chromosomal localization, collinearity analysis, and identification of gene duplication events

Based on the annotation file, all *NtDnaJA* genes have been located on the chromosomes and visualized using Tbtools (v2) software [65]. The duplication events of regulatory *NtDnaJA* genes were identified using dupgen_finder (v1) [66], with *Arabidopsis* serving as an outgroup. Syntenic blocks between *N. tabacum*, *A. thaliana*, *O. sativa*, *G. max*, *Z. mays*, and *S. bicolor* DnaJA genes were identified with Tbtools software. *Ka/Ks* rates were calculated using KaKs_Calculator (v3.0) [67]. For the collinearity analysis of DnaJA genes across different species, TBtools and MCScanX [68] were utilized.

Expression profiles based on transcriptome data

Transcriptome data were collected from eight representative tobacco tissues, including roots, stems, young flowers, mature flowers, senescent flowers, young leaves, mature leaves, and senescent leaves. Additionally, transcriptome data under different stress conditions, including RS and PVY, along with abiotic stresses such as high/low temperature, cadmium, drought, salt, and topping, were also collected. All of these data were sourced from the NCBI database. The quantification of all clean reads was performed using Salmon software (v1.10.1) [69], and the R package 'tximport' was employed to calculate transcripts per kilobase million (TPM) values. All accession numbers for transcriptome samples involved in this study can be found in Table S11.

RNA extraction, isolation, reverse transcription, and qPCR analysis

Total RNA was extracted from the above-mentioned tobacco samples using the GenePure Plantpilot RNA Kit (Condox, Zhengzhou, China). DNase I Column Purification Kit (Condox) was used to eliminate the DNA contamination. TransScript One-Step gDNA Removal and cDNA Synthesis SuperMix (Transgen) was used to synthesize cDNA. PerfectStart® Green qPCR SuperMix (Transgen) was used for qPCR. Primers specific for the selected *NtDnaJA* genes were designed with Primer-BLAST tools in Primer premier 5. The transcriptional level of genes was calculated using the $2^{-\Delta\Delta C_t}$ method. The gene expression levels were normalized against the GAPDH gene. The primers utilized in this study are listed in Table S12.

Plant materials and treatments

Using the widely cultivated tobacco variety K326, we investigated the expression patterns of *NtDnaJAs* in different tissues and stress treatment. The seedlings were cultivated in plastic pots under a 16-hour light cycle, with daytime temperatures of 28°C and nighttime temperatures of 23°C. Root, pistil, sepal, stem, senescent leaf, young leaf, stamen, ovary, flower bud samples were collected as described in our previous study [70]. During the vigorous growth phase of tobacco, drought, cold, cadmium, salt, and *Ralstonia solanacearum* (RS) stress experiments were conducted. For drought treatment, 28-day-old tobacco plants were subjected to 5-day water deprivation. For cold treatment, five-leaf stage seedlings was placed in a growth chamber set to 4°C. During cadmium treatment, Cd^{2+} was added to the pots in the form of a $CdCl_2$ aqueous solution at a concentration of 200 μM . The salt treatment was carried out in a growth chamber with 150 mM NaCl for 7 days. A suspension of RS (OD = 0.5) was applied to the root system of tobacco plants. Observations were made one week after the treatment. Control plantlets remained untreated. Following the treatment, both treated and control groups were quickly frozen in liquid nitrogen, and subsequently stored at -80°C for future uses. All samples were subjected to three independent biological replicates.

Construction of subcellular localization vectors and virus-induced gene silencing vectors

To determine the subcellular localization of NtDnaJA3 protein, the full-length cDNA of *NtDnaJA3* was amplified using PCR with primers specific to the *NtDnaJA3* sequence. The amplified *NtDnaJA3*

sequence was inserted into the PC1300s-GFP vector through homologous recombination to create an NtDnaJA3-PC1300s-GFP fusion protein expression construct. The recombinant vector was introduced into *Agrobacterium tumefaciens* strain GV3101, and after verifying the positive colonies, they were cultured at 28°C until an OD600 of 1 was reached. These colonies were resuspended in MgCl₂ + As + MES buffer. After incubation in darkness for 3 hours, the suspension was injected into leaves of *N. benthamiana*. After the plants were injected, they were kept in the dark for one day and then returned to normal growing conditions. The GFP signal was then detected using a confocal microscopy system (Nikon C2-ER, Japan). The amplified *NtDnaJA3* gene fragment was inserted into the pYY13 vector to construct the pYY13-*NtDnaJA3* silencing vector. Simultaneously, TRV1 and TRV2 control vectors were also constructed. The correctness of the vector construction was verified by sequencing. The verified pYY13-*NtDnaJA3* vector was then transformed into *A. tumefaciens* GV3101. After the transformation, the bacteria were inoculated into LB medium containing selective antibiotics and cultured at 28°C for 2 days. The transformation efficiency was assessed by colony PCR, and positive clones were selected for subsequent experiments. The positive clones were cultured overnight in LB liquid medium containing antibiotics at 28°C with shaking. After overnight cultivation, the bacterial suspension was centrifuged at 4000 rpm to collect the cells. The supernatant was discarded, and the bacterial pellet was resuspended in MgCl₂+As + MES buffer. The resuspended *Agrobacterium* was incubated in the dark for 3 hours. The resuspended pYY13-*NtDnaJA3*, TRV2, and PDS vectors were each mixed with TRV1. The mixture was then injected into the lower leaf surfaces of *N. benthamiana* plants. After injection, the plants were kept in the dark for 24 hours, then transferred to a growth chamber set at 25°C, with 70% relative humidity, and a 16-hour light/8-hour dark cycle, for an additional 14 days. When the *PDS*-silenced plants exhibited a severe phenotype with complete bleaching of newly emerging leaves, uninfected leaves from each group were collected, and the expression levels of the *NtDnaJA3* gene were analyzed using qRT-PCR to confirm the silencing effect. GAPDH was used as an internal control for data normalization. Plants with successful silencing of *NtDnaJA3* and control plants were subjected to drought stress. After 5 days of withholding water, the phenotypic changes of each plant were observed and recorded.

Abbreviations

CAT-box meristem expression element

MSA-like cell cycle regulation element

GCN4_motif endosperm expression element

miRNA Micro RNA

HSF Heat Shock Factors

WRKY WRKY Transcription Factors

bZIP Basic Leucine Zipper

bHLH Basic Helix-Loop-Helix

CREs Cis-acting regulatory elements

qRT-PCR Quantitative real-time polymerase chain reaction

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

All data generated or analyzed in this study are included in the materials and methods section of this article.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the Natural Science Foundation of HeNan (232300420220); Beijing Life Science Academy [BLSA: 2023000CC0110, 2023200CB0090, 2023200CB0070]; Zhengzhou Tobacco Research Institute (CNTC: 110202201001(JY-01)); Chief scientist innovation project of State Tobacco Monopoly Administration/China National Tobacco Corporation [902023CK0880, 902023CK0890].

Authors' contributions

JJJ conceived and designed the experiments, revised and reviewed the manuscript. PJC conceived and designed the experiments, revised and reviewed the manuscript, and contributed to writing the manuscript. QW performed bioinformatics data analysis, conducted the qRT-PCR experiments, subcellular localization, and VIGS experiments, wrote and checked the manuscript, and created the figures. HS performed bioinformatics data analysis. LJM conducted the qRT-PCR experiments, subcellular localization, and VIGS experiments. ZCQ conducted the qRT-PCR experiments, subcellular localization, and VIGS experiments. JMT contributed to writing the manuscript. PL contributed to writing the manuscript. JFZ contributed to writing the manuscript.

Acknowledgments

Not applicable.

References

1. Ritossa FM. Experimental activation of specific loci in polytene chromosomes of *Drosophila*. *Experimental cell research*. 1964;35:601–7.
2. Tutar L, Tutar Y. Heat Shock Proteins; An Overview. *Current Pharmaceutical Biotechnology*. 2010;11:216–22.
3. Aolymat I, Hatmal MM, Olaimat AN. The Emerging Role of Heat Shock Factor 1 (HSF1) and Heat Shock Proteins (HSPs) in Ferroptosis. *Pathophysiology*. 2023;30:63–82.
4. The role of heat shock proteins (HSPs) in type 2 diabetes mellitus pathophysiology. *Journal of Diabetes and its Complications*. 2023;37:108564.
5. Cabaud-Gibouin V, Durand M, Quéré R, Girodon F, Garrido C, Jegou G. Heat-Shock Proteins in Leukemia and Lymphoma: Multitargets for Innovative Therapeutic Approaches. *Cancers*. 2023;15:984.
6. Wu S, Zhao Y, Wang D, Chen Z. Mode of Action of Heat Shock Protein (HSP) Inhibitors against Viruses through Host HSP and Virus Interactions. *Genes*. 2023;14:792.
7. Jeyachandran S, Chellapandian H, Park K, Kwak I-S. A Review on the Involvement of Heat Shock Proteins (Extrinsic Chaperones) in Response to Stress Conditions in Aquatic Organisms. *Antioxidants*. 2023;12:1444.
8. Heat shock protein TaHSP17.4, a TaHOP interactor in wheat, improves plant stress tolerance. *International Journal of Biological Macromolecules*. 2023;246:125694.
9. Janni M, Gullì M, Maestri E, Marmioli M, Valliyodan B, Nguyen HT, et al. Molecular and genetic bases of heat stress responses in crop plants and breeding for increased resilience and productivity. *Journal of Experimental Botany*. 2020;71:3780–802.
10. Pulido P, Leister D. Novel DNAJ-related proteins in *Arabidopsis thaliana*. *The New Phytologist*. 2018;217:480–90.
11. Matsuoka E, Kato N, Hara M. Induction of the heat shock response in *Arabidopsis* by heat shock protein 70 inhibitor. *Frontiers in Plant Science*. 4:273.
12. Kampinga HH, Craig EA. The HSP70 chaperone machinery: J proteins as drivers of functional specificity. *Nature Reviews Molecular Cell Biology*. 2010;11:579–92.
13. Kim J-S, Liu L, Vázquez-Torres A. The DnaK/DnaJ Chaperone System Enables RNA Polymerase-DksA Complex Formation in *Salmonella* Experiencing Oxidative Stress. *mBio*. 2021;12:e03443-20.
14. Feng H, Li C, Zhou J, Yuan Y, Feng Z, Shi Y, et al. A cotton WAKL protein interacted with a DnaJ protein and was involved in defense against *Verticillium dahliae*. *International Journal of Biological Macromolecules*. 2021;167:633–43.
15. Jacob P, Hirt H, Bendahmane A. The heat-shock protein/chaperone network and multiple stress resistance. *Plant Biotechnology Journal*. 2017;15:405–14.

16. Yang Y, Zhao L, Wang J, Lu N, Ma W, Ma J, et al. Genome-wide identification of DnaJ gene family in *Catalpa bungei* and functional analysis of CbuDnaJ49 in leaf color formation. *Frontiers in Plant Science*. 2023;14.
17. Liu T, Xu M, Gao S, Zhang Y, Hu Y, Jin P, et al. Genome-wide identification and analysis of the regulation wheat DnaJ family genes following wheat yellow mosaic virus infection. *Journal of Integrative Agriculture*. 2022;21:153–69.
18. Zhang B, Liu Z, Zhou R, Cheng P, Li H, Wang Z, et al. Genome-wide analysis of soybean *DnaJA*-family genes and functional characterization of *GmDnaJA6* responses to saline and alkaline stress. *The Crop Journal*. 2023;11:1230–41.
19. Ray D, Ghosh A, Mustafi SB, Raha S. Plant Stress Response: Hsp70 in the Spotlight. *Heat Shock Proteins and Plants*. Cham: Springer International Publishing; 2016. p. 123–47.
20. Jang J, Lee S-H, Kang D-H, Sim D-W, Ryu K-S, Jo K-S, et al. Structural resemblance of the DNAJA-family protein, Tid1, to the DNAJB-family Hsp40. *BMB Reports*. 2022;55:488–93.
21. Mayer MP, Bukau B. Hsp70 chaperones: Cellular functions and molecular mechanism. *Cellular and Molecular Life Sciences*. 2005;62:670.
22. Meta-analysis of heat- and chemically upregulated chaperone genes in plant and human cells. *Cell Stress and Chaperones*. 2011;16:15–31.
23. Yang C, Owen HA, Yang P. Dimeric heat shock protein 40 binds radial spokes for generating coupled power strokes and recovery strokes of 9 + 2 flagella. *The Journal of Cell Biology*. 2008;180:403–15.
24. Shen L, Kang YGG, Liu L, Yu H. The J-Domain Protein J3 Mediates the Integration of Flowering Signals in *Arabidopsis*. *The Plant Cell*. 2011;23:499–514.
25. Gautam R, Meena RK, Rampuria S, Shukla P, Kirti PB. Ectopic expression of DnaJ type-I protein homolog of *Vigna aconitifolia* (VaDJI) confers ABA insensitivity and multiple stress tolerance in transgenic tobacco plants. *Frontiers in Plant Science*. 2023;14:1135552.
26. Xia Z, Zhang X, Li J, Su X, Liu J. Overexpression of a tobacco J-domain protein enhances drought tolerance in transgenic *Arabidopsis*. *Plant Physiology and Biochemistry*. 2014;83:100–6.
27. Wang J, Zhang Q, Tung J, Zhang X, Liu D, Deng Y, et al. High-quality assembled and annotated genomes of *Nicotiana tabacum* and *Nicotiana benthamiana* reveal chromosome evolution and changes in defense arsenals. *Molecular Plant*. 2024;17:423–37.
28. Liu T, Xu M, Gao S, Zhang Y, Hu Y, Jin P, et al. Genome-wide identification and analysis of the regulation wheat DnaJ family genes following wheat yellow mosaic virus infection. *Journal of Integrative Agriculture*. 2022;21:153–69.
29. Nagaraju M, Kumar A, Rajasheker G, Manohar Rao D, Kavi Kishor PB. DnaJs, the critical drivers of Hsp70s: genome-wide screening, characterization and expression of DnaJ family genes in *Sorghum bicolor*. *Molecular Biology Reports*. 2020;47:7379–90.
30. Chen T, Xu T, Zhang T, Liu T, Shen L, Chen Z, et al. Genome-Wide Identification and Characterization of DnaJ Gene Family in Grape (*Vitis vinifera* L.). *Horticulturae*. 2021;7:589.

31. Fan F, Yang X, Cheng Y, Kang Y, Chai X. The DnaJ Gene Family in Pepper (*Capsicum annuum* L.): Comprehensive Identification, Characterization and Expression Profiles. *Frontiers in Plant Science*. 2017;8:689.
32. Cao L, Wang G, Fahim AM, Pang Y, Zhang Q, Zhang X, et al. Comprehensive Analysis of the DnaJ/HSP40 Gene Family in Maize (*Zea mays* L.) Reveals that ZmDnaJ96 Enhances Abiotic Stress Tolerance. *Journal of Plant Growth Regulation*. 2024;43:1548–69.
33. Li K, Wong C, Cheng C, Cheng S, Li M, Mansveld S, et al. GmDNJ1, a type-I heat shock protein 40 (HSP40), is responsible for both Growth and heat tolerance in soybean. *Plant Direct*. 2021;5:e00298.
34. Arsovski AA, Pradinuk J, Guo XQ, Wang S, Adams KL. Evolution of Cis-Regulatory Elements and Regulatory Networks in Duplicated Genes of Arabidopsis. *Plant Physiology*. 2015;169:2982–91.
35. Wang X, Qiu Z, Zhu W, Wang N, Bai M, Kuang H, et al. The NAC transcription factors SNAP1/2/3/4 are central regulators mediating high nitrogen responses in mature nodules of soybean. *Nature Communications*. 2023;14:4711.
36. Ma Z-X, Leng Y-J, Chen G-X, Zhou P-M, Ye D, Chen L-Q. The THERMOSENSITIVE MALE STERILE 1 Interacts with the BiPs via DnaJ Domain and Stimulates Their ATPase Enzyme Activities in Arabidopsis. *PLOS ONE*. 2015;10:e0132500.
37. Cai G, Xu Y, Zhang S, Chen T, Liu G, Li Z, et al. A tomato chloroplast-targeted DnaJ protein, SlDnaJ20 maintains the stability of photosystem I/II under chilling stress. *Plant Signaling & Behavior*. 2022;17:2139116.
38. Li Z, Tong Z, He F, Li X, Sun J. Integrated mRNA and microRNA expression analysis of root response to phosphate deficiency in *Medicago sativa*. *Frontiers in Plant Science*. 2022;13:989048.
39. Singh I, Smita S, Mishra DC, Kumar S, Singh BK, Rai A. Abiotic Stress Responsive miRNA-Target Network and Related Markers (SNP, SSR) in *Brassica juncea*. *Frontiers in Plant Science*. 2017;8:1943.
40. Hou G, Du C, Gao H, Liu S, Sun W, Lu H, et al. Identification of microRNAs in developing wheat grain that are potentially involved in regulating grain characteristics and the response to nitrogen levels. *BMC Plant Biology*. 2020;20:87.
41. Liu X, Huang S, Xie H. Advances in the regulation of plant development and stress response by miR167. *Frontiers in Bioscience-Landmark*. 2021;26:655–65.
42. Li S-B, Xie Z-Z, Hu C-G, Zhang J-Z. A Review of Auxin Response Factors (ARFs) in Plants. *Frontiers in Plant Science*. 2016;7:47.
43. Zhang B, Qiu H-L, Qu D-H, Ruan Y, Chen D-H. Phylogeny-dominant classification of J-proteins in *Arabidopsis thaliana* and *Brassica oleracea*. *Genome*. 2018;61:405–15.
44. Mesbah R, Mohsenzadeh R, Seraji MR. Effect of Topping Height and Timing on Quantity and Quality Influe-Cured Tobacco (Var.K326). *International Journal of Advanced Biological and Biomedical Research*. 2014;2:1388–95.
45. Qi Y, Guo H, Li K, Liu W. Comprehensive analysis of differential genes and miRNA profiles for discovery of topping-responsive genes in flue-cured tobacco roots. *The FEBS journal*.

2012;279:1054–70.

46. Yan S, Niu Z, Yan H, Zhang A, Liu G. Transcriptome sequencing reveals the effect of biochar improvement on the development of tobacco plants before and after topping. *PLOS ONE*. 2019;14:e0224556.
47. Zhao J, Li L, Zhao Y, Zhao C, Chen X, Liu P, et al. Metabolic changes in primary, secondary, and lipid metabolism in tobacco leaf in response to topping. *Analytical and Bioanalytical Chemistry*. 2018;410:839–51.
48. Lei B, Chang W, Zhao H, Zhang K, Yu J, Yu S, et al. Nitrogen application and differences in leaf number retained after topping affect the tobacco (*Nicotiana tabacum*) transcriptome and metabolome. *BMC Plant Biology*. 2022;22:38.
49. Rhee SY, Beavis W, Berardini TZ, Chen G, Dixon D, Doyle A, et al. The Arabidopsis Information Resource (TAIR): a model organism database providing a centralized, curated gateway to Arabidopsis biology, research materials and community. *Nucleic Acids Research*. 2003;31:224–8.
50. Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar GA, Sonnhammer ELL, et al. Pfam: The protein families database in 2021. *Nucleic Acids Research*. 2021;49:D412–9.
51. Yu C, Chen Y, Lu C, Hwang J. Prediction of protein subcellular localization. *Proteins: Structure, Function, and Bioinformatics*. 2006;64:643–51.
52. Artimo P, Jonnalagedda M, Arnold K, Baratin D, Csardi G, de Castro E, et al. ExPASy: SIB bioinformatics resource portal. *Nucleic Acids Research*. 2012;40:W597–603.
53. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and Clustal X version 2.0. *Bioinformatics*. 2007;23:2947–8.
54. Tamura K, Stecher G, Kumar S. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Molecular Biology and Evolution*. 2021;38:3022–7.
55. Letunic I, Bork P. Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Research*. 2024;52:W78–82.
56. Bailey TL, Johnson J, Grant CE, Noble WS. The MEME Suite. *Nucleic Acids Research*. 2015;43:W39–49.
57. Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, et al. InterProScan 5: genome-scale protein function classification. *Bioinformatics*. 2014;30:1236–40.
58. Guex N, Peitsch MC. SWISS-MODEL and the Swiss-Pdb Viewer: An environment for comparative protein modeling. *ELECTROPHORESIS*. 1997;18:2714–23.
59. Lescot M, Déhais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, et al. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Research*. 2002;30:325–7.
60. Dai X, Zhuang Z, Zhao PX. psRNATarget: a plant small RNA target analysis server (2017 release). *Nucleic Acids Research*. 2018;46:W49–54.

61. Griffiths-Jones S. miRBase: The MicroRNA Sequence Database. *MicroRNA Protocols*. New Jersey: Humana Press; 2006. p. 129–38.
62. Zhang H, Jin J, Tang L, Zhao Y, Gu X, Gao G, et al. PlantTFDB 2.0: update and improvement of the comprehensive plant transcription factor database. *Nucleic Acids Research*. 2011;39:D1114–7.
63. Szklarczyk D, Kirsch R, Koutrouli M, Nastou K, Mehryary F, Hachilif R, et al. The STRING database in 2023: protein–protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Research*. 2023;51:D638–46.
64. 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 Research*. 2003;13:2498–504.
65. Chen C, Wu Y, Li J, Wang X, Zeng Z, Xu J, et al. TBtools-II: A “one for all, all for one” bioinformatics platform for biological big-data mining. *Molecular Plant*. 2023;16:1733–42.
66. Qiao X, Li Q, Yin H, Qi K, Li L, Wang R, et al. Gene duplication and evolution in recurring polyploidization–diploidization cycles in plants. *Genome Biology*. 2019;20:38.
67. Zhang Z. KaKs_Calculator 3.0: Calculating Selective Pressure on Coding and Non-Coding Sequences. *Genomics, Proteomics & Bioinformatics*. 2022;20:536–40.
68. Wang Y, Tang H, DeBarry JD, Tan X, Li J, Wang X, et al. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Research*. 2012;40:e49–e49.
69. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. *Nature Methods*. 2017;14:417–9.
70. Su H, Meng L, Qu Z, Zhang W, Liu N, Cao P, et al. Genome-wide identification of the N6-methyladenosine regulatory genes reveals NtFIP37B increases drought resistance of tobacco (*Nicotiana tabacum* L.). *BMC Plant Biology*. 2024;24:134.

Figures

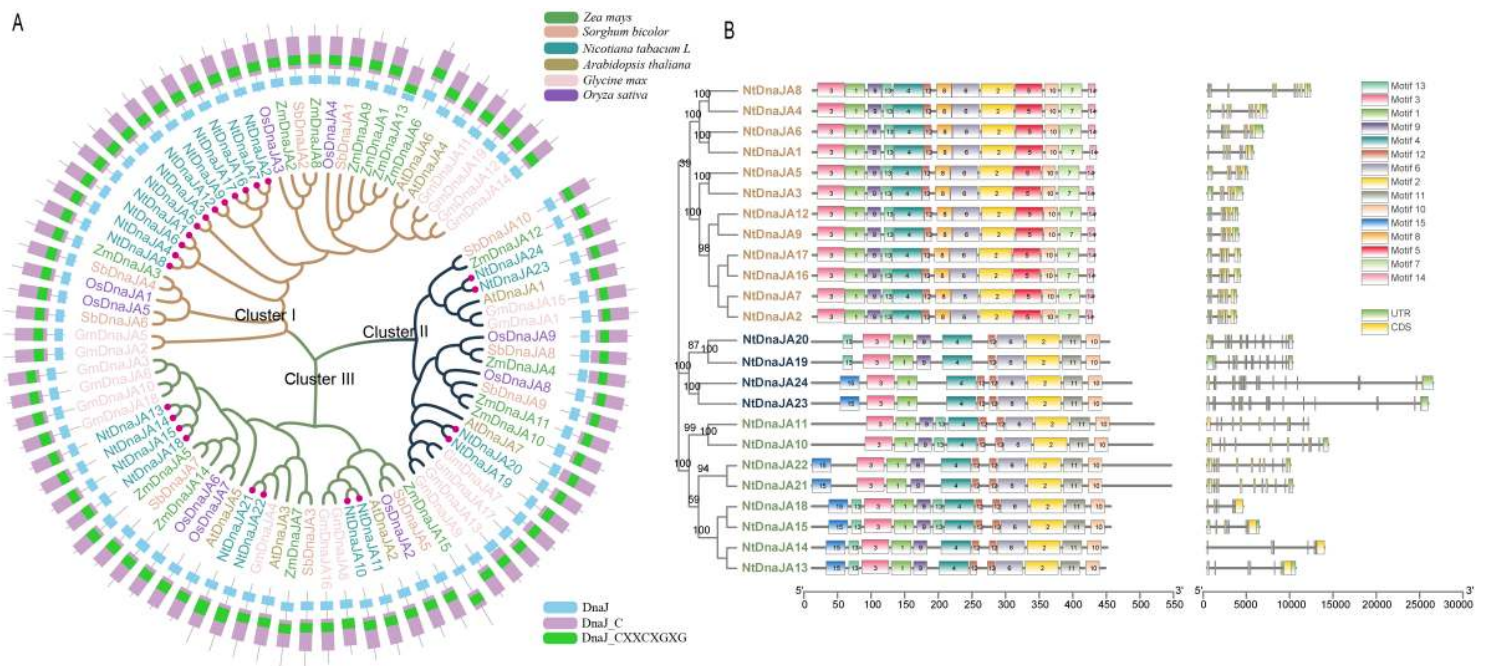


Figure 1

Phylogenetic, motif, and structural features for NtDnaJ family. **A** Phylogenetic tree was constructed using DnaJA proteins from *N. tabacum*, *G. max*, *A. thaliana*, *Z. mays*, *O. sativa*, and *S. bicolor*, with six species represented by different colors. Conserved protein domains were shown with blue (DnaJ), purple (DnaJ_C), and green (DnaJ_CXXCXGXG) boxes. **B** Structural features of NtDnaJAs. Conserved motifs within NtDnaJA proteins were highlighted in different colors to distinguish each motif. The exon-intron structures of the corresponding genes, with UTRs in green, CDS in yellow.

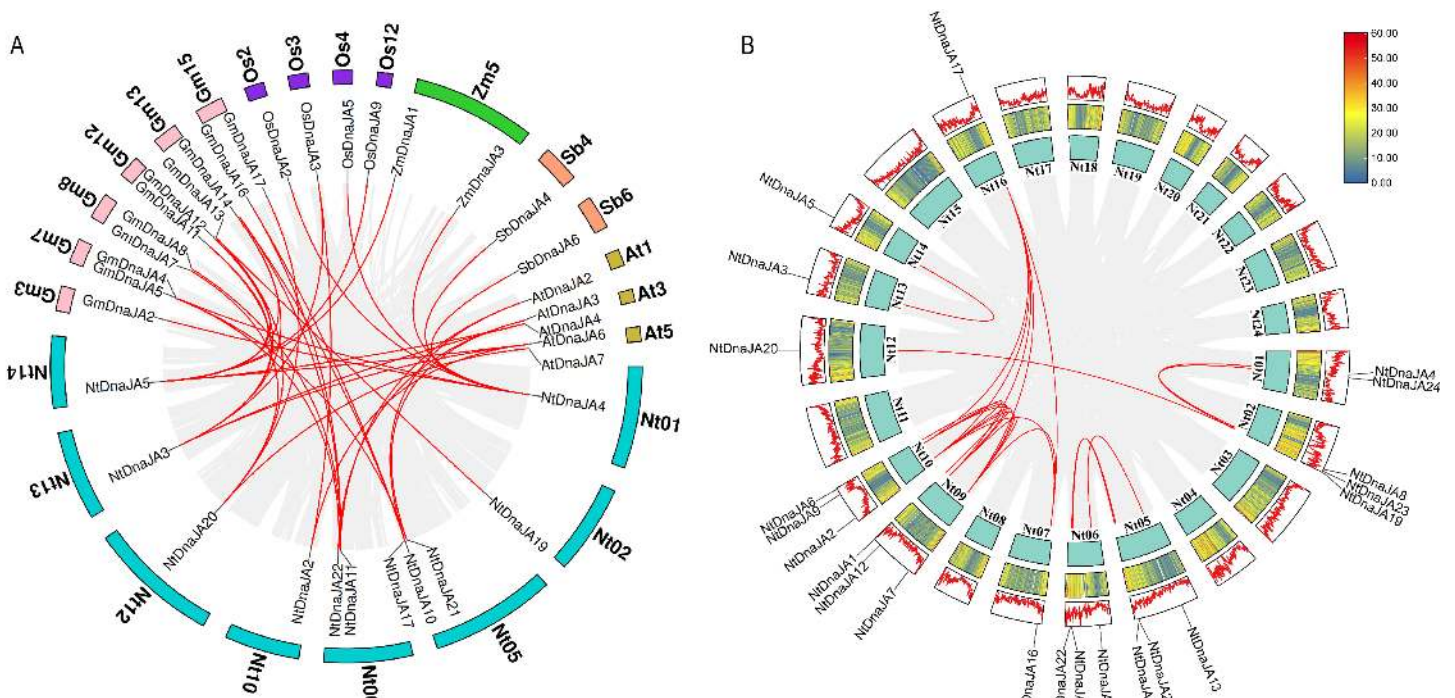


Figure 2

Collinearity analysis of *NtDnaJAs* between tobacco and other species. **A** Syntenic relationships between *NtDnaJAs* and DnaJA genes in representative plant species. The colored blocks represent different species, and the red lines represent syntenic gene pairs between *N. tabacum* and other species. **B** Intra-collinearity analysis for *NtDnaJAs*. Red lines are used to link repetitive gene pair fragments. Red lines and green boxes represent gene density, and blue boxes represent chromosomes.

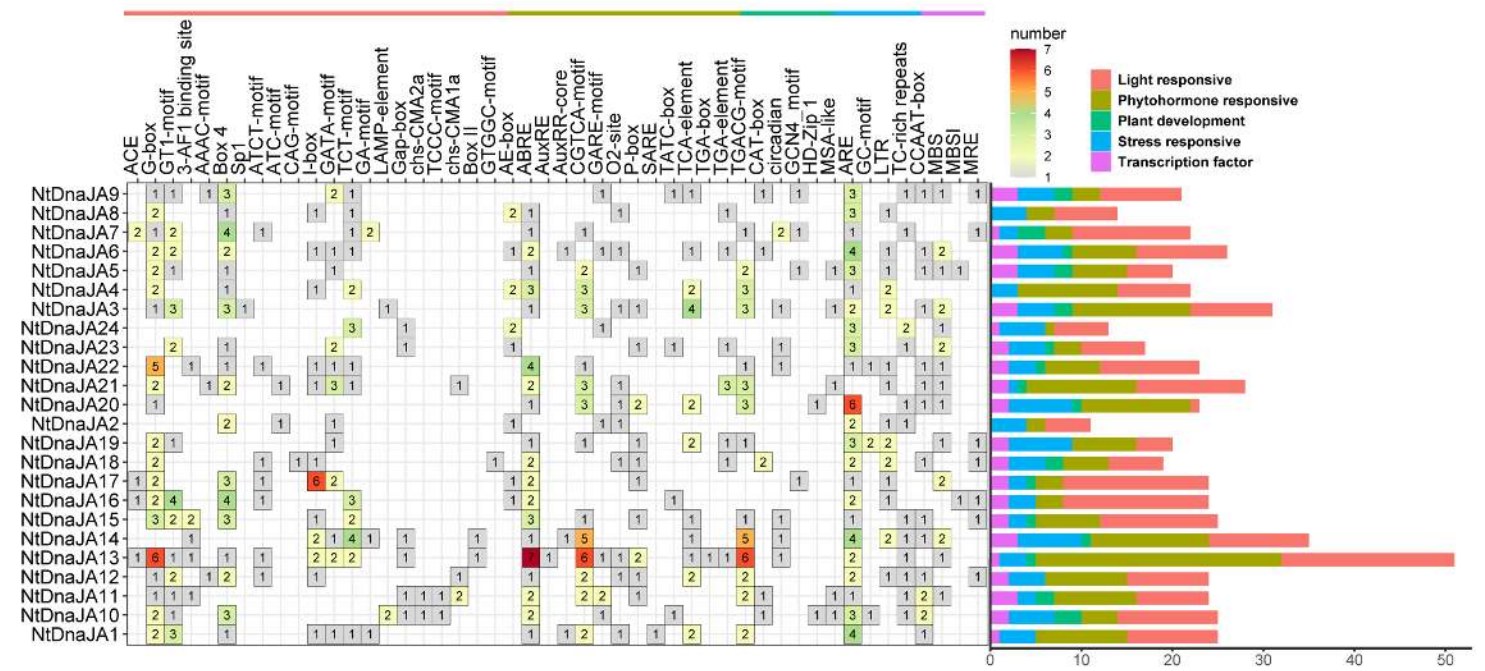


Figure 3

The distribution of CREs in the promoter regions of *NtDnaJA* genes. The heatmap shows the count of each element type within the promoter regions. The bar chart on the right shows the number of CREs in different categories, with different color representing different categories.

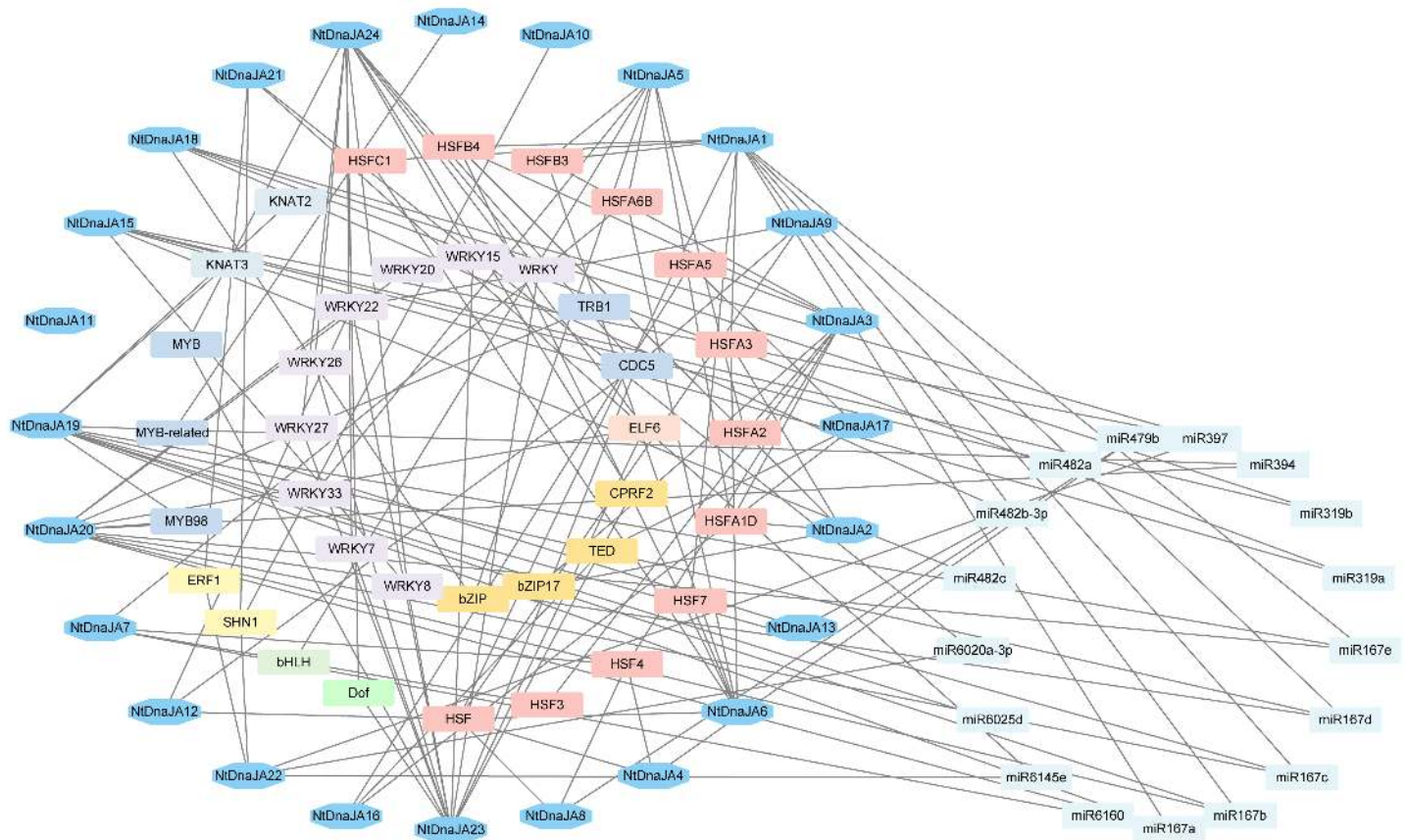


Figure 4

Regulation network of *NtDnaJA* in tobacco. The network illustrated the interactions between *NtDnaJA* genes (blue boxes), transcription factors (yellow, red, purple, sky blue and green shapes), and miRNAs (light blue ellipses). Transcription factors included various families such as ERF, HSF, bZIP, MYB, and WRKY, representing by different colors.

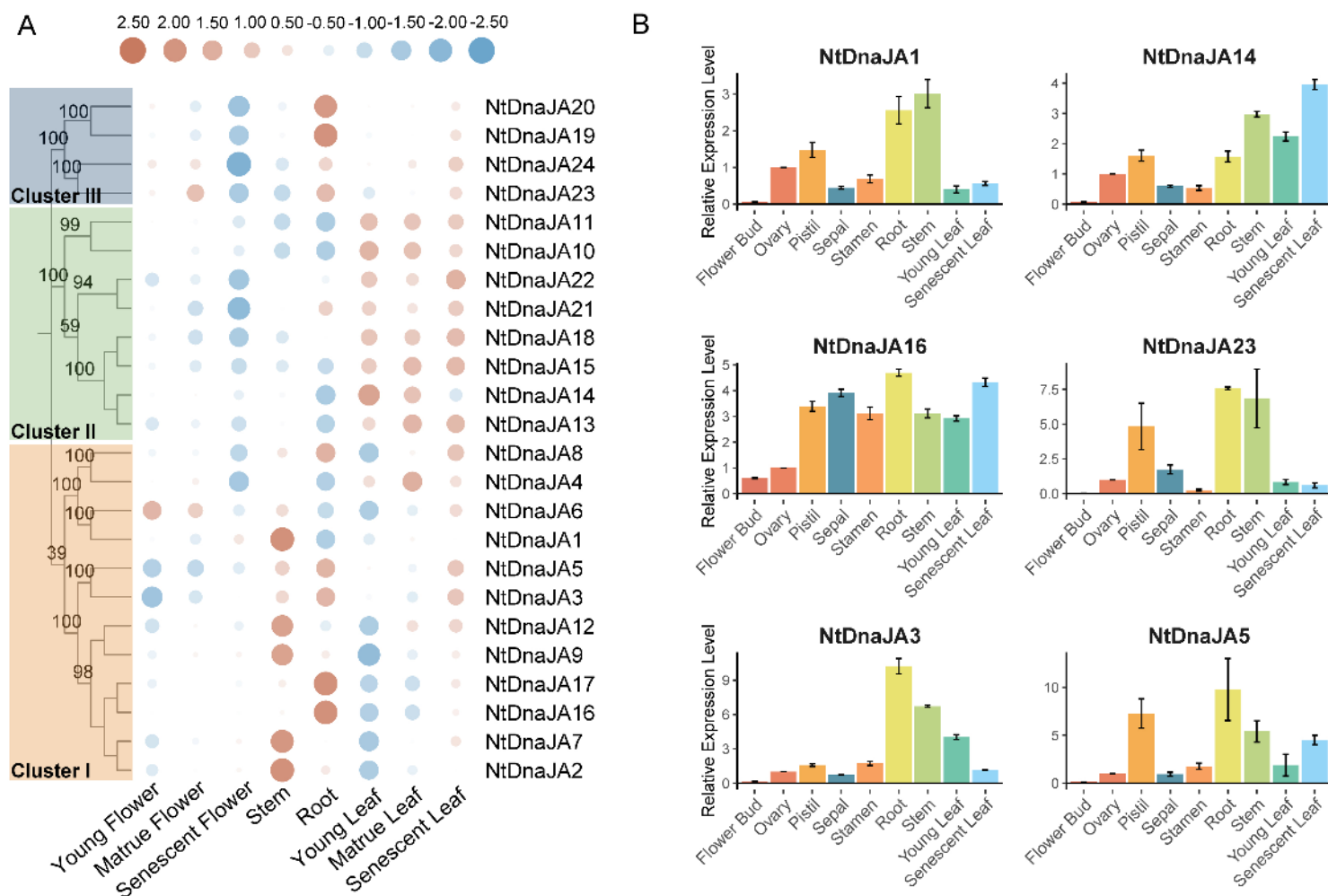


Figure 5

Expression profiles of tobacco *NtDnaJAs* in various tissues. **A** Heatmap depicting the expression levels of *NtDnaJA* genes across different tissues, with values scaled by rows. **B** qRT-PCR analysis of six randomly chosen *NtDnaJA* genes in various tissues at different developmental stages. Data are shown as mean $\pm$ standard error of the mean (SEM).

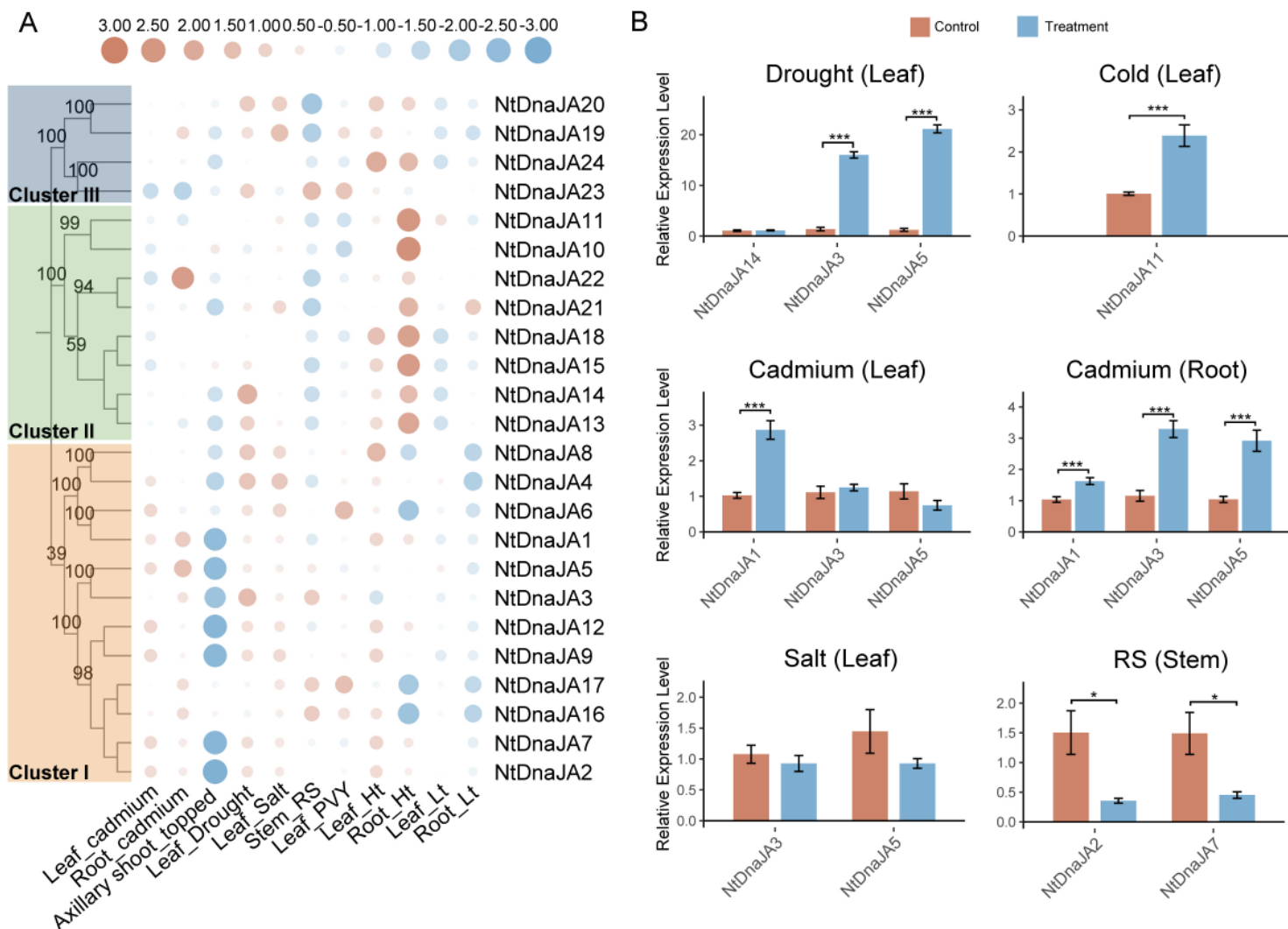


Figure 6

Expression profiles of tobacco *NtDnaJA* genes under various stress treatments. **A** Heatmap of the expression under cadmium, topping, drought, salt, PVY, RS, high/low temperature (Ht/Lt) treatment. **B** qRT-PCR analysis of six randomly selected *NtDnaJA* genes under drought, cold, cadmium, salt, RS treatment. Data are shown as mean $\pm$ SEM. Statistical significance is indicated by asterisks: one for $p < 0.05$, two for $p < 0.01$, and three for $p < 0.001$.

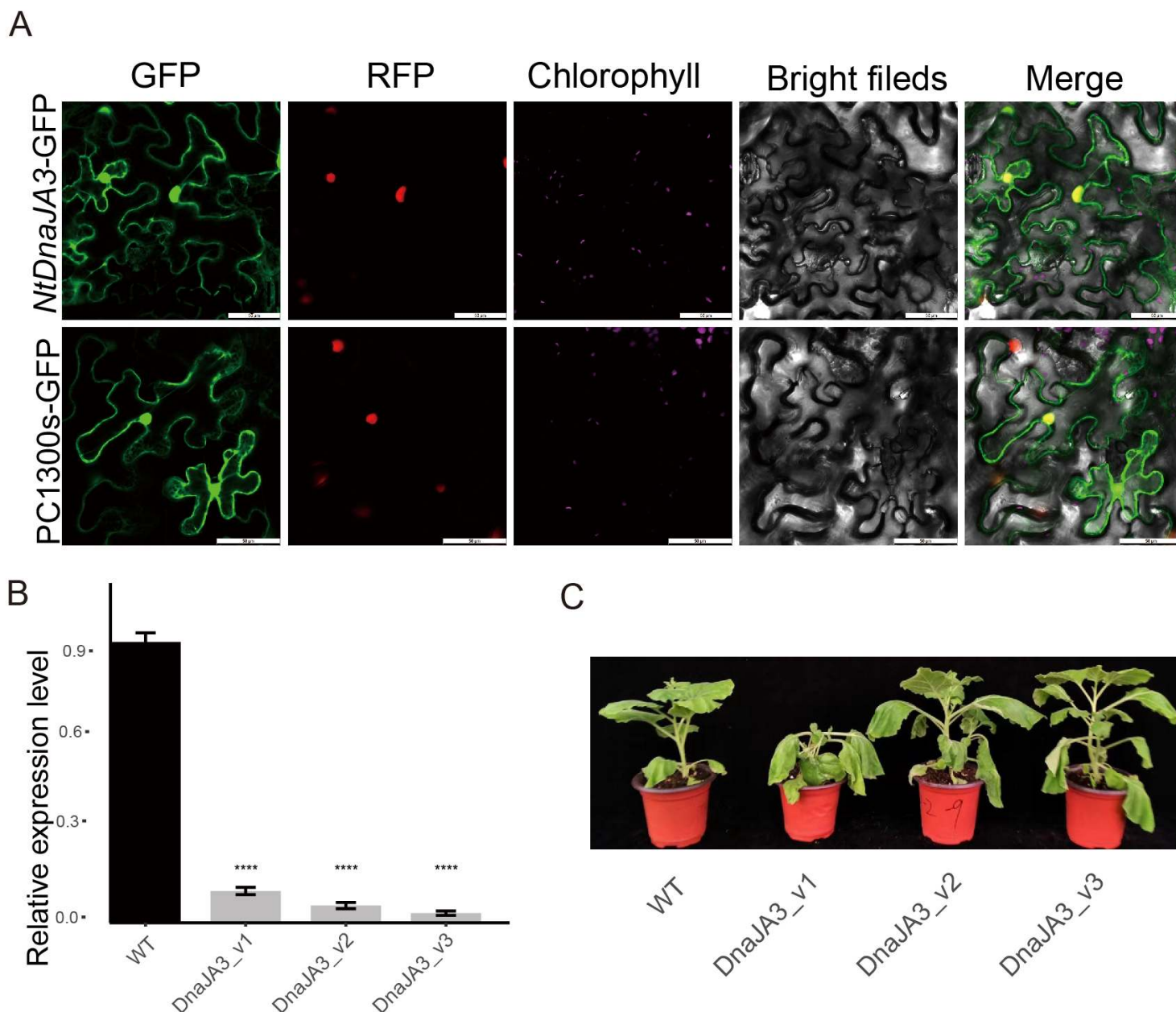


Figure 7

Subcellular localization of *NtDnaJA3* and phenotype for *NtDnaJA3*-silenced plants. **A** Subcellular localization of *NtDnaJA3*. **B** The relative expression of *NtDnaJA3* for *DnaJA3*-VIGS plants using qRT-PCR. **C** Phenotypic analysis of *DnaJA3*-silenced plants post five-day drought treatment.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [supplementarymaterials.docx](#)
- [supplementarytables.xlsx](#)
- [GraphicalAbstract.jpg](#)