Tropical Plant Biology (2021) 14:329–344
https://doi.org/10.1007/s12042-021-09295-2

Genome-wide Identification and Abiotic Stress Response Pattern Analysis of NF-Y Gene Family in Peanut (*Arachis Hypogaea L.*)

Qian Wan[1,2] · Lu Luo[1,2] · Xiurong Zhang[1,2] · Yuying Lv[1,2] · Suqing Zhu[1,2] · Lingrang Kong[1,2] · Yongshan Wan[1,2] · Fengzhen Liu[1,2] · Kun Zhang[1,2]

Received: 13 December 2020 / Accepted: 14 June 2021 / Published online: 30 June 2021

Abstract

The nuclear factor Y (NF-Y) transcription factor (TF) family consists of three subfamilies NF-YA, NF-YB and NF-YC. Many studies have proven that NF-Y complex plays multiple essential roles in stress response in *Arabidopsis* and other plant species. However, little attention has been given to these genes in peanut. In this study, thirty-three *AhNF-Y* genes were identified in cultivated peanut and they were distributed on 16 chromosomes. A phylogenetic analysis of the NF-Y amino acid sequences indicated that the peanut NF-Y proteins were clustered in pairs at the end of the branches and showed high conservation with previous reported plant NF-Ys. Evolutionary history analysis showed that only segmental duplication contributed to expansion of this gene family. Analysis of the 1500-bp regulatory regions upstream the start codon showed that, except for *AhNF-YB6*, peanut *NF-Y*s contained at least one abiotic stress response element in their regulatory region. Expression patterns of peanut *NF-Ys* in 22 tissues and developmental stages were analyzed. A few *NF-Y*s showed universal expression patterns, while most *NF-Y*s showed specific expression patterns. Through RNA-seq and qRT-PCR analyses, expression of six AhNF-Y genes was induced under salt stress in leaves or roots. In addition, *AhNF-YA4/8/11*, *NF-YB4* and *NF-YC2/8* also responded to osmotic stress, ABA (abscisic acid) and salicylic acid (SA) treatment.

Keywords Cultivated peanut (*Arachis hypogaea* L.) · NF-Y gene family · Phylogenetic analysis · Expression patterns · Abiotic stress

Introduction

Nuclear factor Y (NF-Y) transcription factor, also known as heme activator protein (HAP) or CCAAT-binding factor (CBF), is present in almost all higher eukaryotic genomes. Eukaryotic NF-Ys consist of three subfamilies (NF-YA, NF-YB and NF-YC), and each NF-Y subunit is encoded by a single gene with multiple splicing isoforms in animals and yeast (Li et al. 1992). In contrast, each subunit of NF-Y is encoded by a family of approximately 10 genes in plants (Petroni et al. 2012). Eukaryotic NF-Ys function as a heterotrimeric complex to regulate the expression of downstream target genes (Nardini et al. 2013). NF-YB and NF-YC first dimerize in cytoplasm to create a molecular scaffold and then move into the nucleus to recruit an NF-YA component and subsequently bind to the CCAAT-box specifically (Myers et al. 2018). The NF-YB-YC-YA trimer can further recruit other TFs, such as CONSTANS (CO) or bZIP28 (Cao et al. 2014; Liu et al. 2010). In addition, other TFs, such as bZIP67, CO2 and VRN2, can also bind to the NF-YB-YC heterodimer competitively with NF-YA subunits (Li et al. 2011; Yamamoto et al. 2009). The function of NF-Ys in *Arabidopsis thaliana*, wheat, rice, maize, legume plants and tomato have been reported to regulate a variety of growth and developmental processes including male gametogenesis, embryogenesis, seed development and germination (Huang et al. 2015; Mu et al. 2013),

Communicated by: Hongwei Cai

Qian Wan and Lu Luo contributed equally to this work

✉ Fengzhen Liu
liufz@sdau.edu.cn

Kun Zhang
kunzh@sadu.edu.cn

[1] State Key Laboratory of Crop Biology, Shandong Key Laboratory of Crop Biology, Shandong Agricultural University, Tai'an 271018, China

[2] College of Agronomy, Shandong Agricultural University, Tai'an 271018, China

photosynthesis and photomorphogenesis (Myers et al. 2016; Stephenson et al. 2011), root growth (Ballif et al. 2011; Sorin et al. 2014; Zanetti et al. 2017), nodule formation (Bu et al. 2020; Hossain et al. 2016; Ripodas et al. 2019), flowering and yield regulation (Hwang et al. 2019; Muday et al. 2016; Shen et al. 2020; Su et al. 2018; Zhang et al. 2019), and fruit maturation (Li et al. 2016).

Environmental stresses such as drought, salinity, and heat are the main challenges affecting the development, growth, and productivity of field crops, resulting in crop yield losses. A significant number of studies have revealed that NF-Ys are crucial regulators of plant stress response. *AtNF-YA5*, *AtNF-YB1*, *GmNF-YA3* and *PdNF-YB7* positively modulate drought tolerance of transgenic *Arabidopsis* (Chen et al. 2007; Han et al. 2013; Li et al. 2008; Ni et al. 2013), and *OsNF-YA7* and *ZmNF-YB2* confer drought tolerance to transgenic rice (Lee et al. 2015; Nelson et al. 2007). Overexpression of *AtNF-YA1* and *TaNF-YA10-1* significantly increased sensitivity to salt stress (Li et al. 2013; Ma et al. 2015a), and in constrast, *SiNF-YA5 and GhNF-YA10/23* enhance the salt tolerance of plants (Zhang et al. 2020). *AtNF-YA2*, *AtNF-YC10*, *SlNF-YA9/10* were reported to participate in heat tolerance regulation in *Arabidopsis* and tomato (Rao et al. 2021; Sato et al. 2014). Moreover, some NF-Y subunits are involved in multiple stress responses. *AtNF-YB2* and *AtNF-YB3* positively regulate the heat stress response in *Arabidopsis*, but play negative roles in drought tolerance and exhibit functional redundancy in both processes (Sato et al. 2019). Rice *OsHAP2E* confers salinity and drought tolerance and increases photosynthesis and tiller number (Alam et al. 2015). Although some aspects of plant NF-Y research have advanced sufficiently to provide a mechanistic understanding, the regulatory mechanisms of most plant NF-Ys in stress response remain less well understood.

Peanut (*Arachis hypogaea* L.) is an important oil and food crop worldwide (Zhang et al. 2018). More than half of the global peanut production comes from semiarid areas, where drought and soil salinity are the main limitations for peanut growth (Banavath et al. 2018). Cultivated peanut, which evolved from the hybridization and subsequent chromosome doubling of *Arachis duranensis* (A) and *Arachis ipaensis* (B), is an allotetraploid (AABB genome, 2n = 4x = 40) with a total genome size of approximately 2.7 Gb (Bertioli et al. 2016; Grabiele et al. 2012; Robledo et al. 2009). There is much less knowledge about the NF-Y TF involved in peanut stress response than the large genome. In the present study genome-wide identification and systematic analysis of the NF-Y gene family in cultivated peanut were performed. We identified the AhNF-Y gene family and analyzed their sequence features, phylogenetic relationships, chromosomal locations, gene duplication during the expansion. The expression profiles in 22 tissues and developmental stages and under salt stress were also investigated through reanalysis of published RNA-seq data. In addition, using quantitative real-time PCR (qRT-PCR), we analyzed the expression profiles of AhNF-Y genes under salt stress and identified several candidate genes responsive to abiotic stress and hormone treatment. The present results will facilitate future investigations on the functional characterization of NF-Y genes in peanut.

Materials and Methods

Identification of NF-Y Coding Genes in Peanut and Sequence Analysis

The genomic sequence of *A. hypogaea* cv. Tifrunner (Bertioli et al. 2019) and the annotated gene models were downloaded from PeanutBase (Version 1, https://www.peanutbase.org/peanut_genome). The NF-Y protein sequences of *A. thaliana* (Siefers et al. 2009) and *Glycine max* (Quach et al. 2015) used in this research (Additional file 4) were downloaded from the National Centre for Biotechnology Information (NCBI) database (http://www.ncbi.nlm.nih.gov/). The hidden Markov model (HMM) profile of CBFB_NFYA (PF02045) was obtained from the Pfam protein family database (http://pfam.xfam.org). The HMMER website server and BLASTP were used to search for NF-YA genes ($p < 0.01$) (Finn et al. 2015). The amino acid sequences of *Arabidopsis*, soybean and rice NF-Ys were used for BLAST of peanut NF-Y candidates. The putative AhNF-Y proteins were uploaded to the online Conserved Domain Database (https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi) (Marchler-Bauer et al. 2017), Pfam and SMART (http://smart.embl-heidelberg.de/) tools to verify the conserved NF-Y domain. DNAMAN software (LynnonBiosoft USA, version 6) was employed to perform multiple sequence alignments to determine whether the candidate genes could encode proteins carrying complete NF-Y subunit binding and DNA-binding domains. The ExPASy proteomics server (http://prosite.expasy.prg/) was used to acquire the sequence lengths, molecular weights and isoelectric points (Letunic et al. 2018; Robert et al. 2014).

Phylogenetic Analysis and Sequence Analysis

MEGA7 (Kumar et al. 2016) was used to perform sequence alignment and maximum-likelihood phylogenetic tree construction using the bootstrap method (number of bootstrap replications = 1000). iTOL (http://itol.embl.de/) was used for visualization and further editing of the phylogenetic tree. The MEME program (http://meme-suite.org/tools/meme) (Bailey et al. 2009) was employed to detect the conserved motifs in the full length of identified AhNF-Y proteins with the following parameters: the maximum number of motifs was 20 and the length range was 6—100 amino

acids. Exon–intron structure visualization was performed by comparing cDNA sequences with their corresponding full-length genomic DNA sequences using the online tool Gene Structure Display Server version 2.0 (gsds.cbi.pku.edu.cn/) (Hu et al. 2015).

Chromosomal Distribution and Gene Duplication Analysis

The Multiple Collinearity Scan toolkit (MCScanX)(Wang et al. 2012) was employed to detect gene duplication events with the default parameters. The visualization of chromosomal distribution and synteny analysis were performed by Circos and Mapchart (Krzywinski et al. 2009; Voorrips 2002). The amino acid sequences of duplicated gene pairs were first aligned and used to guide the alignment of cDNA sequences with in-house Perl-scripts. KaKs calculator was used to compute the nonsynonymous (Ka) and synonymous (Ks) substitution values of each duplicated gene pair using the YN model. The divergence time (million years ago, Mya) was calculated with the formula $T = Ks/2r$. The r (rate of divergence for nuclear genes) was taken to be 1.5×10^{-8} synonymous substitutions per site per year for dicotyledonous plants (Koch et al. 2000).

Analysis of Regulatory *cis*-elements Upstream of *AhNF-Y* Genes

The 1.5-kb upstream sequences of the initiation codon (ATG) of each *AhNF-Y* gene were submitted to the PlantCARE database (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) to identify six regulatory elements (Lescot et al. 2002).

Plant Materials and Growth Conditions

The plant material used in this study, *A. hypogea* L. cultivar Fenghua 2 (*Spanish type*), was bred and preserved by our group. Mature seeds were germinated on distilled water-wet degreasing cotton extended in seedling cultivation disks. These disks were placed at 26 °C in darkness for 3 days and then exposed to long-day conditions (LD; 16 h light and 8 h dark cycle). Two-functional-leaf seedlings were transplanted to a hydroponic box and cultured with 1/5 Hoagland's nutrient solution (Pan et al. 2017).

Stress Treatment, Total RNA Extraction and Reverse Transcription

To analyze the expression pattern of *AhNF-Y* genes, two-week-old seedlings were treated with nutrient solution containing 200 mM NaCl, 20% (w/v) mannitol, 100 mM abscisic acid (ABA) and 100 mM salicylic acid (SA), respectively. The leaves and roots of seedlings treated with NaCl were harvested at 0 and 16 h. Leaves of seedlings treated with mannitol were collected at 0, 2, 4, 6, 8, 12 and 24 h. For ABA and SA treatment, the leaves of seedlings were harvested at 0, 1, 2, 4, 6 and 8 h. All samples were frozen immediately in liquid nitrogen for RNA extraction. For each sample, leaves or roots from ten seedlings were harvested. Three independent biological replicates were performed for each treatment.

Total RNA was isolated with Quick RNA Isolation Kit (Waryong, Beijing, China) following the manufacturer's instructions. The concentration of the total RNA in each sample was quantified by a NanoDrop 2000 microvolume spectrophotometer (Thermo Fisher Scientific, Massachusetts, USA). Reverse transcription was performed using Advantage RT-for-PCR Kit (TaKaRa, Dalian, China) according to the manufacturer's instructions. Total RNA and cDNA samples were stored at -80°Cand -20 °C, respectively.

Expression Profile Analysis of *AhNF-Ys*

The expression atlas of 22 *A. hypogaea* tissues was downloaded from PeanutBase (https://www.peanutbase.org/gene_expression/atlas) (Clevenger et al. 2016; Dash et al. 2016). In these RNA-Seq data, the normalized reads were mapped to an in silico amphidiploid genome assembled from the genome of the diploid ancestor *A. duranensis* and *A. ipaensis* (Clevenger et al. 2016). BLAST was performed to identify homologous genes of *AhNF-Y* in the *A. duranensis* and *A. ipaensis* genomes. For each tetraploid peanut *AhNF-Y*, only when the *A. duranensis* or *A. ipaensis NF-Y* gene showed the highest similarity in amino acid sequence was it defined as the homologous gene. IDs of the homologous gene were used to extract the fragments per kilobase of transcript per million mapped reads (FPKM) values from the tissue expression atlas. Transcriptome data under salt stress were archived from the public repository of the NCBI (https://www.ncbi.nlm.nih.gov) under BioProject accession number PRJNA560660 (unpublished data from part of another work of our group, BioSample SAMN12594512-14, SAMN12594518-20, SAMN12594524-26, SAMN12594530-32). The heatmap was created using TBtools (Chen et al. 2020) with log2-transformed FPKM values, and row clustered.

SYBR Green real-time PCR was carried out using TB Green Premix Ex Taq (Tli RNaseH Plus, TaKaRa, Dalian, China) on a StepOne Plus system (Applied Biosystems, Waltham, USA) in a 20μL reaction volume according to the manufacturer's instructions. Three technical replicates were performed for each sample. The primers were designed using Beacon Designer 7.9. Actin was used as the internal reference gene. Sequences of the primers and actin are shown Table S1. The relative expression levels of the *AhNF-Y*s were evaluated by the $2^{-\Delta\Delta Ct}$ method. Statistical

differences were determined by Student's t-test (**P < 0.01, *P < 0.05, n = 3) using Excel.

Protein–protein Interaction Network Analysis

All differentially expressed genes (DEGs) encoding DNA-binding proteins detected under salt stress were screened from the online transcriptome data mentioned in "Expression profile analysis of *AhNF-Y*s". The interaction networks were created by the online software STRING (Version 11.0, https://string-db.org/) using the amino acid sequences of DEGs and visualized by Cytoscape 3.7.1. The interaction database of *A. thaliana* and *G. max* were selected as references.

Results

Identification and Phylogenetic Analysis of *AhNF-Y* Genes

In the allotetraploid peanut cultivar, 42 *AhNF-Y* candidates (19 *AhNF-YA*s, 14 *AhNF-YB*s and 9 *AhNF-YC*s) were identified. 33 of them (14 *NF-YA*s, 10 *NF-YB*s and 9 *NF-YC*s), which could encode NF-Y proteins with complete NF-Y subunit binding and DNA-binding domains, were defined as *AhNF-Y*s (Table 1, Fig. S1). All these genes were named according to chromosome name and chromosomal location (Fig. S2). Basic information on the AhNF-Y family members is listed in Table 1, including the IDs of the AhNF-Y-coding genes in *A. hypogaea* cv. Tifrunner genome (version 1), gene loci on the chromosome, length of the open reading frame (ORF), molecular weight (MW) and isoelectric point (pI). AhNF-YA12 was the largest protein with 492 amino acids and an MW of 55.45 kDa, while the smallest protein was AhNF-YB2 with 171 amino acids and an MW of 18.88 kDa. The pI ranged from 5.19 (AhNF-YC9) to 9.64 (AhNF-YA5).

To investigate the phylogenetic relationships of *NF-Y* genes in plants, phylogenetic tree was generated from 33 peanut NF-Ys (Table S2), 33 *A. thaliana* NF-Ys (AtNF-Ys) and 58 *G. max* NF-Ys (GmNF-Ys, Table S3). All the NF-Y proteins were clustered into three main clades, and each clade was equivalent to a single subfamily. AhNF-Y proteins tend to be present in pairs at the ends of branches, except AhNF-YC1, AhNF-YC2, and AhNF-YC8 (Fig. 1). Most of the pairs of peanut NF-Y genes were located at similar positions on the corresponding chromosome of the two subgenomes (Fig. S4). However, there were a few exceptions in the NF-YC subfamily, including *AhNF-YC2*, *AhNF-YC4*, *AhNF-YC5*, *AhNF-YC6*, and *AhNF-YC8*. Furthermore, as shown in Fig. 1, each subfamily consisted of a few secondary clades (from clade I to clade XII). NF-Ys from all three species were clustered in clades II, IV, VI, VII, IX, and XII. At the branch ends of clade I and XI, there were only NF-Ys from legume (soybean and peanut), while NF-Y proteins from peanut were not clustered in clades III, V, VIII, and X (gray strips in Fig. 1).

Chromosomal Distribution and Gene Duplication of *AhNF-Ys*

The chromosomes of the *A. hypogaea* genome are numbered Arahy.01-Arahy.20 (Bertioli et al. 2019). The 33 AhNF-Y genes were distributed on 16 chromosomes, that is on all chromosomes except Arahy.02, 05, 12 and 15 (Table 1, Fig. 2, Fig. S2). Arayh.01, Arayh.11, Arayh.14 and Arayh.18 carried 3 *AhNF-Y* genes, respectively, and only one *AhNF-Y* gene was located on Arahy.17, Arahy.19 and Arahy.20. Most of the *AhNF-Y*s were located near the end of the chromosome (Fig. S2).

Some studies indicated that both segmental duplication and tandem duplication played important roles in the generation of gene families during evolution, as well as crop domestication(Cannon et al. 2004; Kondrashov 2012; Salman-Minkov et al. 2016). We analyzed the gene duplication events of *AhNF-Y* genes in the *A. hypogaea* genome (Fig. 2). Duplication events occur only within subgenomes. Two duplicated pairs were detected in the A and B subgenomes, respectively (Table 2). The pairwise synonymous distances (Ks values) within the collinear blocks and divergence time of the segmental duplicated gene pairs were calculated (Table 2). The divergence of NF-YB6 and NF-YB7 occurred much earlier than that of the other pairs, and all the duplication events of peanut NF-Y occurred before speciation of the two wild diploids at approximately 2 Mya (Chen et al. 2019).

Gene Structure and Conserved Motif of AhNF-Ys

Phylogenetic analysis was performed for peanut NF-Y proteins, and most of the AhNF-Ys were clustered in pairs, except AhNF-YC4, AhNF-YC5 and AhNF-YC6 (Fig. 3A, Table S4). All the paired genes shared similar locations on the homologous chromosomes of the two subgenomes (Fig. S2). Furthermore, the gene structures of the paired genes showed a high degree of similarity, including the numbers and positions of both introns and exons (Fig. 3B). For example, there was an intron in the 5' untranslated region (UTR) of both *AhNF-YA7* and *AhNF-YA14*. The members of the *AhNF-YA* subfamily contain a larger number of introns (from 4 to 7), while there are only one or two introns in the *AhNF-YC* subfamily, except in *AhNF-YC1* (Fig. 3B).

The MEME analysis tool was used to predict the conserved motifs in *AhNF-Y* genes (Fig. 3C and Table S5). A total of 20 motifs were identified. Motifs 7 and 8 were identified only in the *AhNF-YA* subfamily, and they consist

Table 1 Identification of *AhNF-Ys* in peanut

Name	Gene ID	Chr. *	Genomic location (from)	Genomic location (to)	Strand	Length of ORF (bp)	Number of AAs	Number of exons	MW (kDa)	pI value
AhNF-YA1	arahy.P5AXXG	A.01	4,955,904	4,959,946	-	1020	339	5	36.81	6.4
AhNF-YA2	arahy.K8VWCH	A.01	94,322,773	94,326,097	+	975	224	4	24.60	9.6
AhNF-YA3	arahy.K4Y63P	A.03	45,004,146	45,007,537	-	639	212	4	23.34	9.45
AhNF-YA4	arahy.85VPHV	A.03	125,125,172	125,129,125	-	615	204	5	22.58	6.3
AhNF-YA5	arahy.VP74QW	A.04	9,362,311	9,370,508	-	1269	422	6	48.03	9.64
AhNF-YA6	arahy.TY3GKM	A.08	44,492,626	44,498,160	-	981	326	5	36.16	9.01
AhNF-YA7	arahy.HITK0G	A.10	111,351,581	111,356,009	+	978	325	5	35.49	9.03
AhNF-YA8	arahy.14IBSY	A.11	6,616,944	6,620,989	+	1248	415	7	45.95	6.76
AhNF-YA9	arahy.E0X5DU	A.11	149,088,288	149,091,730	-	975	224	4	24.63	9.6
AhNF-YA10	arahy.V1T15D	A.13	47,820,230	47,823,687	+	639	212	4	23.33	9.45
AhNF-YA11	arahy.40UHTZ	A.13	128,836,446	128,840,405	-	621	206	5	22.75	6.15
AhNF-YA12	arahy.D979GK	A.14	10,847,881	10,861,237	-	1479	492	7	55.45	9.34
AhNF-YA13	arahy.3C4V56	A.18	124,758,993	124,764,258	-	978	325	5	35.93	9.03
AhNF-YA14	arahy.RS70VX	A.20	137,968,920	137,973,224	+	978	325	5	35.51	9.12
AhNF-YB1	arahy.TAFY85	A.01	104,790,079	104,792,569	+	678	225	2	25.29	6.17
AhNF-YB2	arahy.M2R7LU	A.04	6,610,689	6,611,671	+	522	171	2	18.88	5.74
AhNF-YB3	arahy.S3VJK6	A.06	24,410,081	24,416,986	-	867	288	6	31.86	8.12
AhNF-YB4	arahy.BTE5MM	A.07	937,559	938,146	-	588	195	1	20.59	5.77
AhNF-YB5	arahy.SP9TFP	A.09	103,306,693	103,310,002	+	543	180	5	19.53	5.82
AhNF-YB6	arahy.CHFV0I	A.11	135,761,737	135,764,354	-	681	226	2	25.44	6.11
AhNF-YB7	arahy.D0EVJ3	A.14	8,058,879	8,060,003	+	525	174	2	19.15	5.83
AhNF-YB8	arahy.UWY0VP	A.16	36,381,530	36,388,018	-	699	232	7	25.56	8.7
AhNF-YB9	arahy.8XPX7R	A.17	2,248,428	2,249,015	-	588	195	1	20.61	5.77
AhNF-YB10	arahy.83ESWU	A.19	133,670,457	133,673,816	+	543	180	5	19.50	5.82
AhNF-YC1	arahy.0G5T8T	A.06	4,854,811	4,859,709	+	1122	373	7	41.49	8.72
AhNF-YC2	arahy.VV7AV3	A.07	33,598,286	33,601,060	+	1077	358	2	40.58	6.53
AhNF-YC3	arahy.DHT5D0	A.08	50,513,721	50,516,515	+	771	256	3	27.81	5.71
AhNF-YC4	arahy.QGX1M1	A.09	35,397,904	35,400,370	+	834	277	2	31.45	5.84
AhNF-YC5	arahy.Y83UC8	A.10	7,695,567	7,696,286	-	720	239	1	27.43	6.33
AhNF-YC6	arahy.FIIU1N	A.14	5,734,442	5,736,074	+	789	262	2	29.99	6.27
AhNF-YC7	arahy.WLF6LC	A.16	17,633,833	17,637,759	-	807	268	2	29.93	5.89
AhNF-YC8	arahy.R55U5A	A.18	94,597,401	94,600,161	+	1062	353	2	39.97	6.9
AhNF-YC9	arahy.HG1CGX	A.18	134,607,760	134,612,435	+	918	305	3	33.18	5.19

*A stand for Arahy

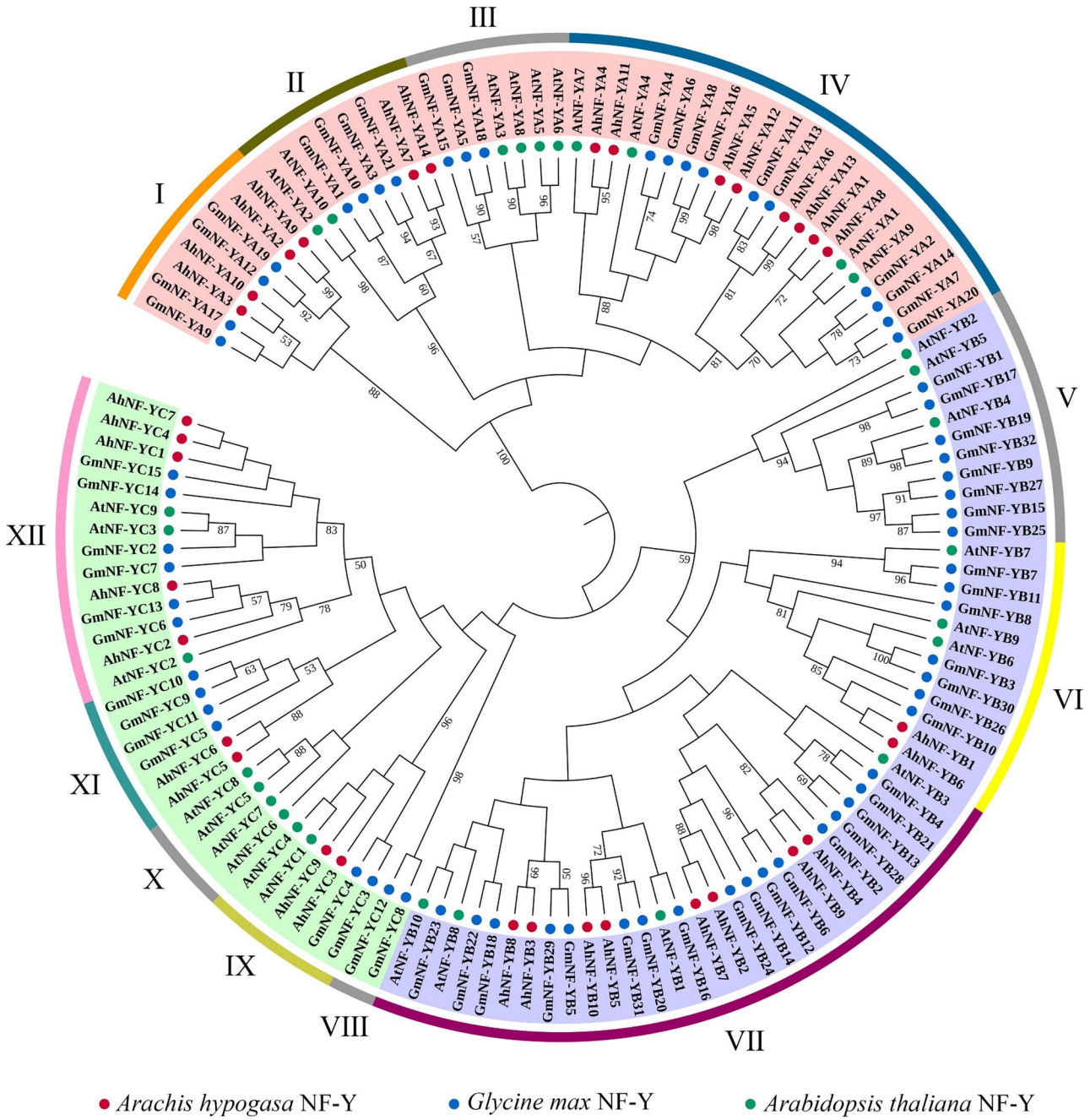

Fig. 1 Phylogenetic tree of NF-Y proteins from peanut, *Arabidopsis thaliana* and *Glycine max*. NF-Ys from the three subfamilies are indicated by different color ranges: NF-YA in pink, NF-YB in purple and NF-YC in light green, respectively. Genes from each individual species are denoted by circles with different colors. The colored strips on the outside represent the main clades of each subfamily. The bootstrap values are shown on the branches

the NF-YB/C interaction and DNA-binding domains of the NF-YA subunit (Fig. S1A) together with Motif 1. Motifs 2, 4 and 11 correspond to the NF-YA/NF-YC subunit interaction and DNA-binding domains of NF-YB protein (Fig. S1B), which were unique to the *AhNF-YB* subfamily. Motifs 3, 6 and 10 existed only in the AhNF-YC subfamily and corresponded to the NF-YA/NF-YB subunit interaction and DNA-binding domains in the NF-YC subunit

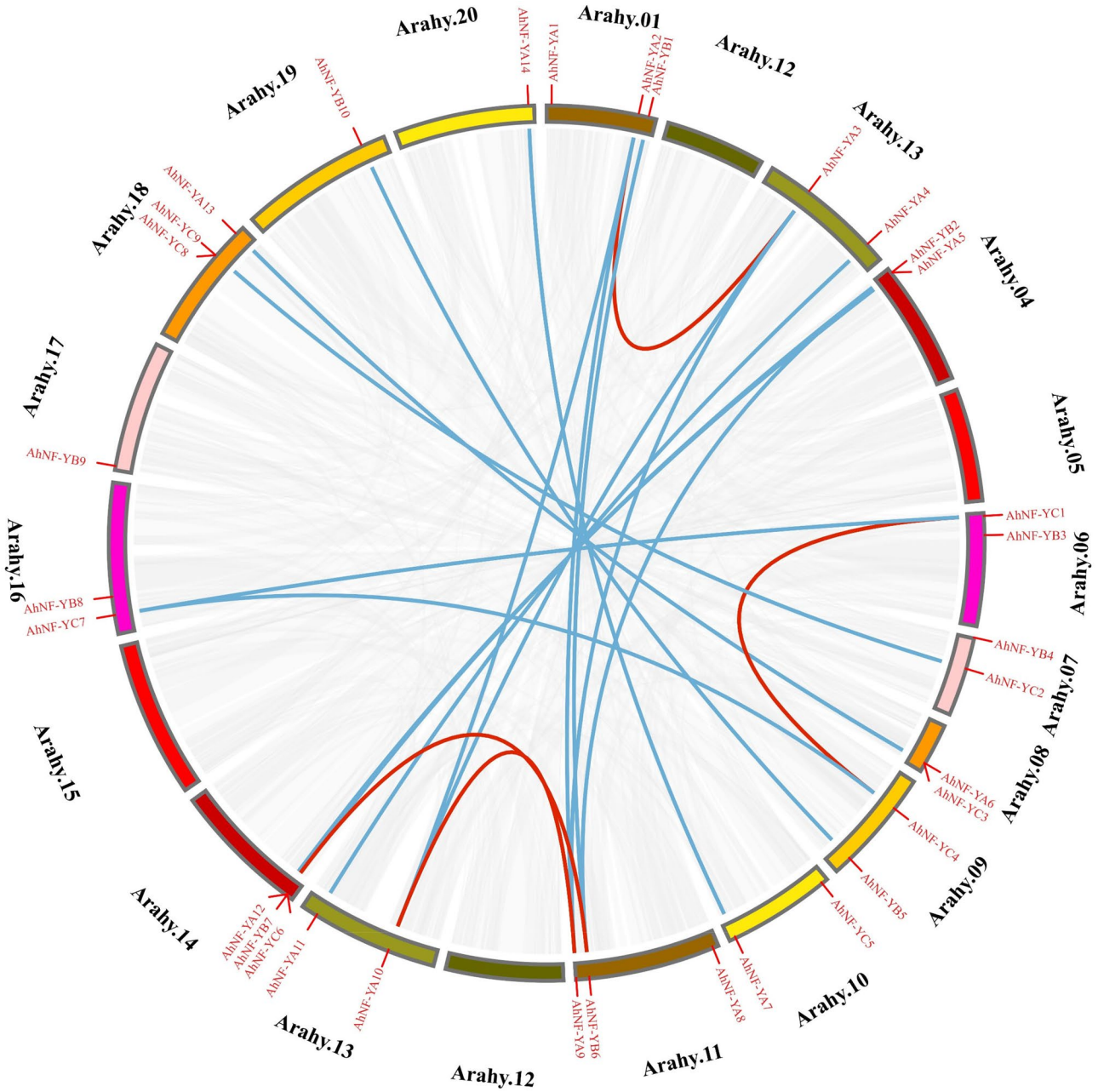

Fig. 2 Chromosome distribution and synteny analysis of *NF-Y* genes in peanut. Chromosomes are drawn in different colors. The approximate location of *AhNF-Y* genes is shown by the short red lines on the circle. The duplicated *AhNF-Y* paralogs were linked by red lines. The *NF-Y* orthologs between the A and B subgenomes are linked by blue lines. All synteny blocks are indicated by light grey lines

(Fig. S1C). These results indicated the high conversion of the protein–protein and protein-DNA interaction structures of the NF-Y family in peanut, which is consistent with previous reports (Myers et al. 2018). In addition, some motifs were identified only in a few proteins; for example, motif 15 was observed only in NF-YA4/11 and NF-YA5/12, and motif 16 was identified only in NF-YB3/8 (Fig. 3C). The specificity of these motifs may result in functional differences between NF-Y subunits within each subfamily.

Tissue/organ-specific Expression Analysis of *AhNF-Ys*

To determine the expression patterns of *AhNF-Y* genes, we used published RNA-seq data, which covered 22 tissues throughout the life cycle of peanut (Clevenger et al. 2016). Genes with the most similar expression patterns were clustered together by TBtools, and all the *AhNF-Ys* were classified into four main groups (Fig. 4, Table S6). Group 1 (G1) contained 3 genes (*AhNF-YA3*, *A10* and *C6*), which showed

Table 2 Ka, Ks, and Ka/Ks values for duplicated NF-Y gene pairs in *Arachis hypogaea* genome

Gene 1	Gene 2	Subgenome	Ka	Ks	Ka/Ks	P-value (Fisher)	Divergence-time (Mya)
AhNF-YA2	AhNF-YA3	A	0.96208	1.10194	0.87308	0.0119115	36.731
AhNF-YC1	AhNF-YC4	A	1.01304	0.95165	1.06451	0.18375	31.721
AhNF-YA9	AhNF-YA10	B	0.95219	1.13067	0.842148	0.0011769	37.689
AhNF-YB6	AhNF-YB7	B	0.41487	2.01636	0.20575	0	67.212

low expression levels in most tissues except seed. Group 2 (G2) comprised 8 genes. The expression of these genes was hardly detected in most tissues/organs except a few, such as AhNA-YB1/6 in seeds. The third group (G3) included only 2 genes (*AhNF-YC1* and *C8*), with extremely high expression levels throughout the life cycle and nearly all organs/tissues. Group 4 (G4) was composed of the other *AhNF-Ys*. These genes showed higher expression levels than genes belonging to group 1 and group 2 and displayed diverse expression patterns. For example, *AhNF-YB4*, *B9* and *C3* were expressed

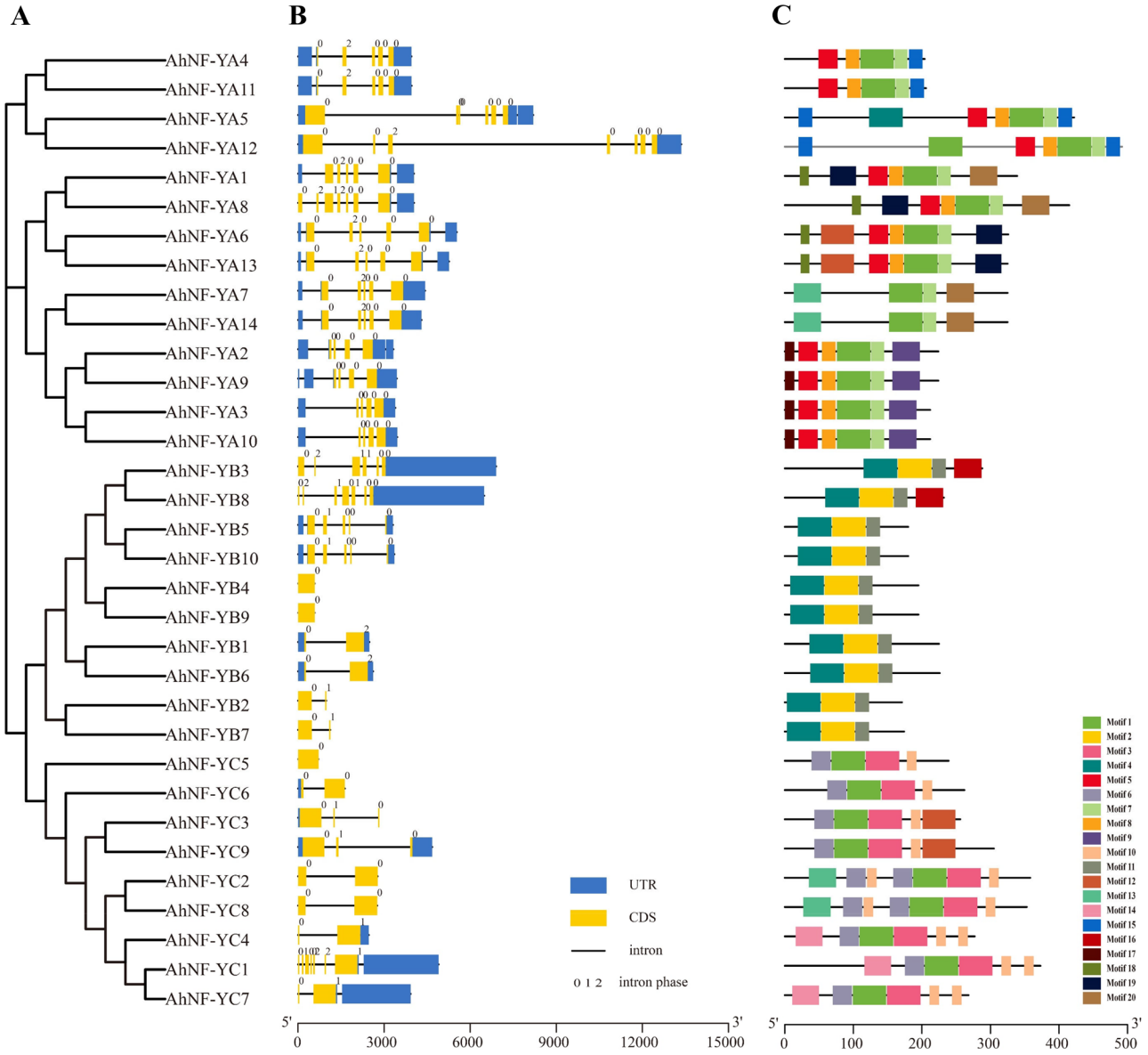

Fig. 3 Phylogenetic relationships, gene structures and motif compositions of *AhNF-Y* genes. **A** Unrooted maximum likelihood phylogenetic tree. **B** Exon/intron organization. Exons are shown as yellow boxes, and introns are shown as black lines. **C** Schematic representation of conserved MEME motifs of full-length NF-Y proteins. Colored boxes indicate different conserved motifs, while the black lines represent sequences no MEME motifs detected

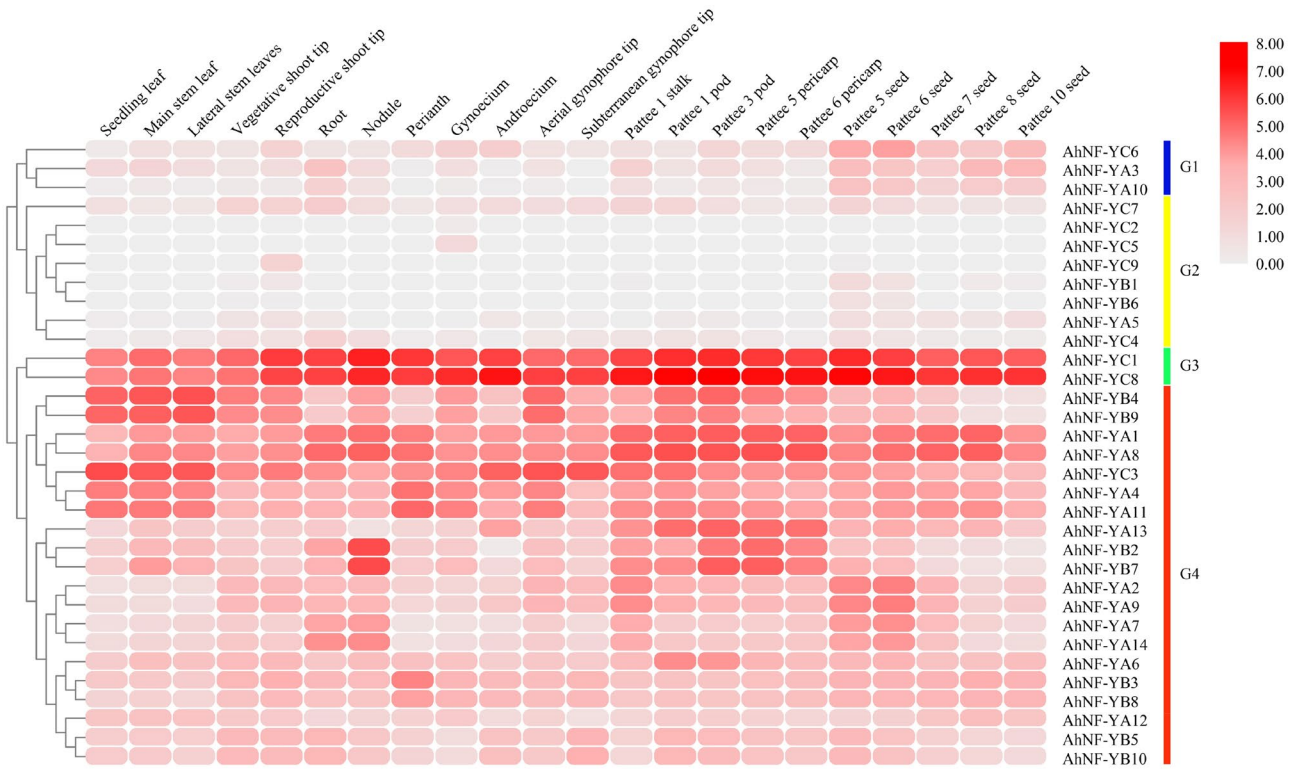

Fig. 4 Expression profiles of *AhNF-Y* genes in 22 peanut tissue types. The colored round rectangles indicate the log2-transformation of the transcripts. The visualization and clustering were performed by TBtools. Blue, yellow, green, and red indicate the different expression profile categories

in all the examined tissues and developmental stages, but the expression levels are quite different; *AhNF-YB2* and *B7* were highly expressed in nodules only. The above results indicated that *AhNF-Y* genes showed diverse expression patterns in peanut.

Expression and Protein–protein Interaction Analyses of *AhNF-Y* Genes under Salt Stress

Using the high-throughput RNA-seq data of peanut cultivar Fenghua 2, heatmap of the *AhNF-Y*s expression pattern under salt stress was established (Fig. 5A, Table S7). Cluster analysis showed that the expression of 11 *AhNF-Y* genes (*AhNF-YA1*, *A4*, *A6*, *A8*, *A9*, *A11*, *B7*, C1, *C2*, *C7* and *C8*) was upregulated in leaves, whereas the expression levels of 3 genes (*AhNF-B4*, *B9* and *C5*) were downregulated. Three *AhNF-Y* genes (*AhNF-YA3*, *A7* and *A14*) showed decreased expression in roots. The expression level of *AhNF-YB2* was upregulated in leaves but downregulated in roots. Furthermore, *AhNF-YB4* was upregulated in both leaves and roots. In contrast, 11 other *AhNF-Ys* (*AhNF-YA2*, *A5*, *A12*, *B3*, *B5*, *B8*, *B10*, *C3*, *C4*, *C6* and *C9*) could not be induced by salt stress. In addition, the expression of *AhNF-YA10* and *AhNF-YB1* was not detected. This distinction in the expression patterns indicated that members of the *AhNF-Y* family might have different responses and regulatory mechanisms under salt stress.

To further validate the RNA-seq data, the expression of 6 typical *AhNF-Y* genes (*AhNF-YA4*, *A8*, *A11*, *B4*, *C2* and *C8*) in both leaves and roots was analyzed by qRT-PCR (Fig. 5B and C). After 16 h of salt treatment (200 mM NaCl), the expression levels of *AhNF-YA4/A8/C8* were upregulated in both leaves and roots, while *AhNF-YC2* was induced only in leaves. These results are consistent with the RNA-seq data. In addition, the expression pattern of *AhNF-YA11* in leaves was not affected by salt stress, and the expression levels of *AhNF-YB4* in leaves and roots showed no significant change. The expression patterns of *AhNF-YA11* and *AhNF-YB4* were inconsistent with the RNA-seq data. The expression of the other 6 genes under salt stress was also detected by qRT-PCR (Fig. S4A). Correlation analysis between the RNA-seq and the qRT-PCR result of the 12 genes were performed, and the Pearson's coefficient was 0.8375 (Fig. S4B).

It has been reported that plant NF-Y proteins perform their functions by recruiting other TFs to specific promoter sequences. For example, *Arabidopsis* NF-YC3 could interact with ABF3/4 to promote flowering by inducing the transcription of SOC1 under drought stress (Hwang et al.

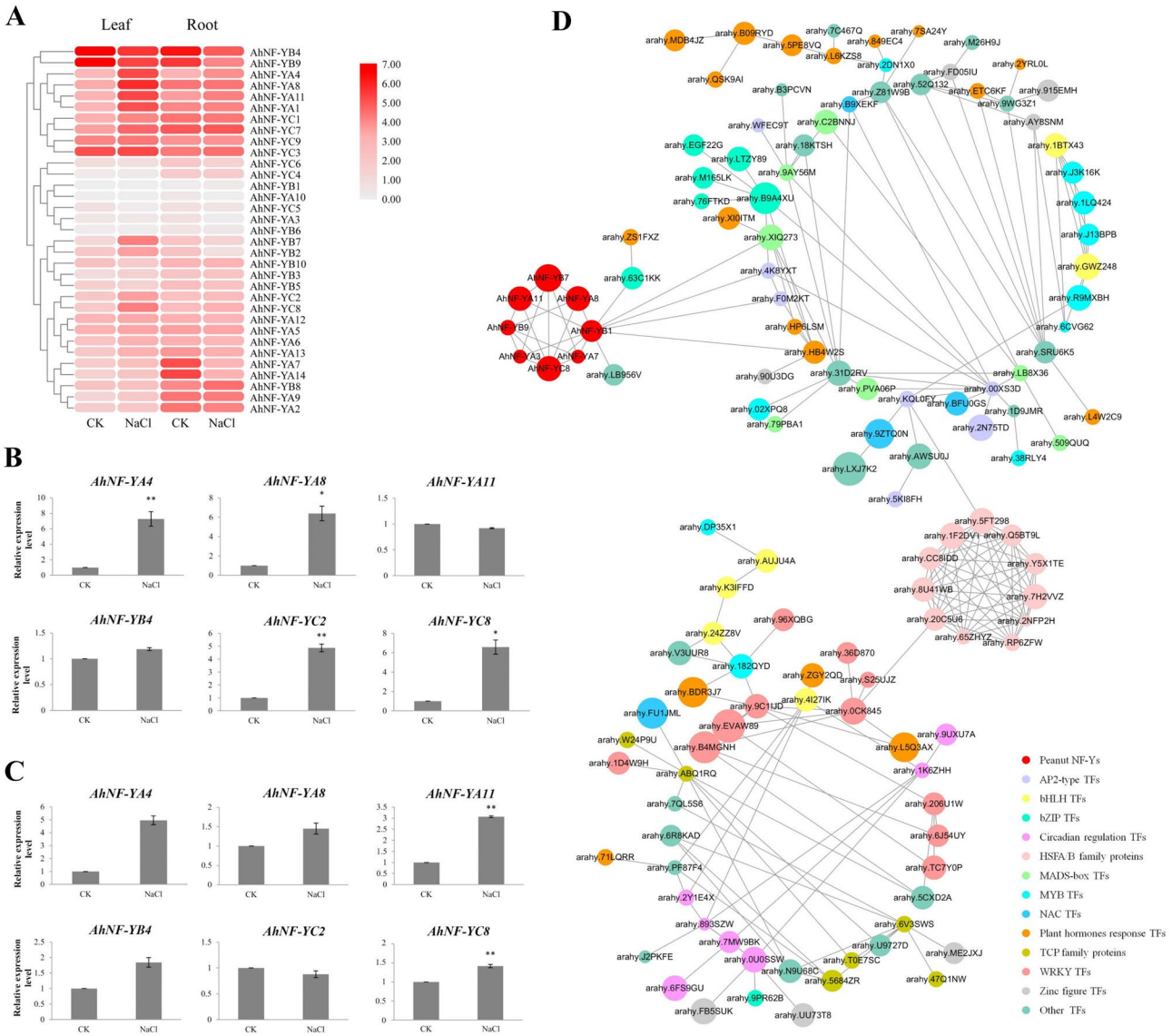

Fig. 5 Expression profile and protein–protein interaction analysis of *AhNF-Y* genes under salt stress. **A** Expression pattern of *AhNF-Y* genes in response to salt stress. The color scale indicates the log2-transformation of the transcript values. **B, C** qRT-PCR profiles of 6 *AhNF-Y* genes in leaves **B** and roots **C** under salt stress. Leaves of two-week-old seedlings were sampled at 16 h under a 16-h light/8-h dark cycle. Bars reflect the means ± SDs of three replicates. Asterisks indicate that the corresponding gene was significantly up- or downregulated compared with the untreated control (**P < 0.01 and *P < 0.05, Student's t-test). **D** Protein–protein interaction network of differentially expressed *AhNF-Y*s and other transcription factors. The size of the node indicates the value of change ratio under salt stress

2019). In addition, the CO protein could even interact with NF-YB2-YC9 dimer instead of the NF-YA subunit to binding the *CORE* element of the *FLOWERING LOCUS T* promoter (Gnesutta et al. 2017). To further obtain an overview of the transcriptional regulation function of AhNF-Ys in peanut under salt stress, we established a protein–protein interaction (PPI) network among the salt stress-induced DNA-binding protein encoding genes in peanut (Table S8). As shown in Fig. 5D, peanut NF-Ys could interact with NF-Y proteins belonging to other two subfamilies (red node). In addition to the AhNF-YA and AhNF-YC subunits, the predicted AhNF-YB1 could interact with other kinds of TFs directly, including a bZIP TF (arahy.63C1KK), AP2 TFs (arahy.4K8YXT, arahy.F0M2KT), a MADS-box TF (arahy.XIQ273), and plant hormone-responsive TFs (arahy.HP6LSM, arahy.HB4W2S). The above TFs further interact with TFs involve in plant hormone response, stress response, flowering regulation, circadian clock and so on (Fig. 5D, Table S9).

Abiotic Stress-related Response of *AhNF-Ys*

To investigate the potential regulatory mechanisms of *AhNF-Ys* in the abiotic stress response, the sequences 1.5 kb upstream from the initiation codon of the *AhNF-Y* genes were detected using the PlantCARE database to identify regulatory elements. Six stress-related regulatory elements, including ABA response element (ABRE), CGTCA-motif, MYB binding site (MBS), TCA-element, TC-rich repeat and TGACG-motif, are shown in Fig. S3. ABRE elements were detected in promoter regions of 18 *AhNF-Ys*. The methyl jasmonate (MeJA)-responsive elements CGTCA-motif and TGACG-motif were both detected in 22 *AhNF-Ys*. A total of 12 *AhNF-Ys* contained TCA-elements (SA responsiveness). In addition, MBS (drought responsiveness) and TC-rich repeats (defence and stress responsiveness) were found in 7 and 10 *AhNF-Ys*, respectively. At least one regulatory element was identified in the promoter region of the *NF-Y* genes, except in that of *AhNF-YB6*. These results suggested that *AhNF-Y* genes may be involved in many different abiotic stresses.

To further investigate whether the expression of these predicted *AhNF-Y* genes was influenced by other stress treatments (mannitol), ABA and SA, qRT-PCR was used to survey the transcript levels in leaves (Fig. 6). The results revealed that the transcript levels of *AhNF-YA4* and *AhNF-YA8* were downregulated, and both reached the lowest levels under osmotic stress at approximately 8 h. In contrast, *AhNF-YA11*, *AhNF-YC2* and *AhNF-YC8* had similar expression profiles and showed a trend toward upregulation. Under the same treatment, the expression level of *AhNF-YB4* was upregulated until approximately 12 h. The transcript levels of all 6 predicted *AhNF-Y* genes were increased under ABA treatment and peaked at approximately 2 h, except for those of *AhNF-YB4*. In addition, all these genes responded significantly to SA treatment. All the above results indicated that these 6 *AhNF-Y* genes responded to osmotic stress, ABA and SA with distinct expression patterns.

Discussion

The NF-Y TF family is widely distributed in eukaryotes and plays important roles in development and stress responses. Studies have shown that the protein–protein interaction domains of NF-Y proteins are highly similar within each subfamily (Laloum et al. 2013). Therefore, the NF-Y protein could theoretically interact with members from other subfamilies, and these interactions have been proven using yeast two-hybrid assay (Hackenberg et al. 2012). For example, there are at least 10 genes encoding each subunit type in *Arabidopsis*; therefore, up to 1690 theoretical NF-Y complexes can be formed (Siefers et al. 2009). This amplification creates a flexible formation, leading to new and divergent functions (Myers et al. 2018; Petroni et al. 2012). The complex interactions within NF-Y family, as well as between NF-Ys and other TFs, make functional analysis difficult. In addition, the overlapping functionality between NF-Y subunits resulted in inefficiency of the forward genetic mutant screen. Therefore, the function of more than one-third of NF-Y subunits has not been reported, and only a few NF-Y complexes have been described (Zhao et al. 2016). Before further study of the biological function of the NF-Y protein in peanut, it is necessary to obtain an overview of the NF-Y

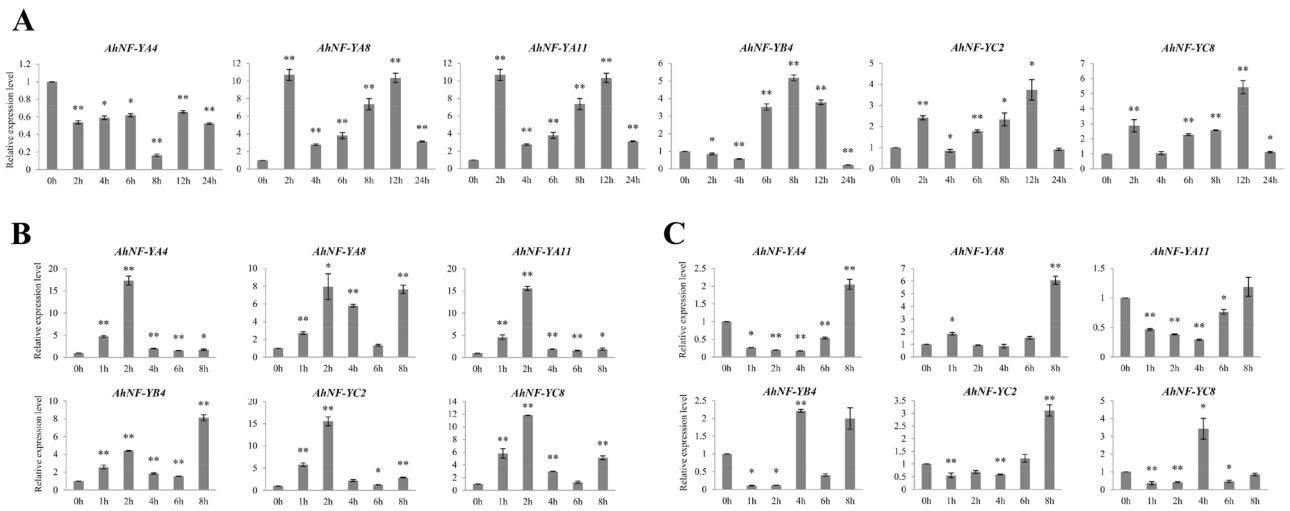

Fig. 6 Expression analysis of 6 *AhNF-Y* genes in response to mannitol **A**, ABA **B** and SA **C** in peanut leaves. Bars reflect the means ± SDs of three replicates. Asterisks indicate the corresponding gene significantly up- or downregulated compared with the untreated control (**P < 0.01 and *P < 0.05, Student's t-test)

family, including the gene structures, homology between members, and expression patterns in tissues and under stress.

In this study, we identified 33 NF-Y-coding genes throughout the peanut genome. The number of *AhNF-Y*s is similar to that in diploid plants such as *Arabidopsis* and rice (Siefers et al. 2009; Thirumurugan et al. 2008). Peanut is an allotetraploid crop with a genome size of approximately 2.7 Gb. Compared with crops, such as soybean (Shen et al. 2018), maize (Li et al. 2019) and cotton (Wang et al. 2019), peanut has fewer *NF-Y*s (Chen et al. 2018; Quach et al. 2015; Zhang et al. 2020, 2016). Similar situations have been reported in other gene families of peanut as well. For example, although cotton (*Gossypium hirsutum*, more than 60 K protein coding genes) and peanut are both tetraploid crops, there are 306 NAC TFs in cotton and only 162 in allotetraploid peanut (Mohanta et al. 2020; Wang et al. 2019). Furthermore, the numbers of *Hsf* genes in soybean and cotton were 38 and 40 (Li et al. 2014; Wang et al. 2014), which are further more than the total number of *Hsf* genes in the diploid ancestors of peanut (Wang et al. 2017). Gene duplication events are the main reason of gene family expansion. The genome of both diploid ancestors of the peanut cultivar underwent the early papilionoid whole-genome duplication (WGD) approximately 58 Mya (Bertioli et al. 2016), and WGD events occurred after tetraploidization have not been reported yet. Additionally, divergence time of the four duplication pairs identified in peanut NF-Y family were 36.731, 31.721, 37.689 and 67.689 Mya respectively (Table 2), which were long before speciation and hybridization of the two diploids ancestors reported (Chen et al. 2019; Zhuang et al. 2019). Therefore, number of NF-Y coding genes in allotetraploid peanut may mainly determine by those of the two diploid ancestors. BLASTP analysis of the NF-Y protein coding genes showed that, the total number of *NF-Y* in the two diploid ancestors is similar with that of allotetraploid peanut (Table S10), which is consistent with the above prediction. It is reported that, a new allele of the duplicated genes has a small probability to be fixed in a diploid population compared with pre-existing alleles (Cagliari et al. 2011). According to the phylogenetic analysis (Fig. 1), *AhNF-Y*s were missing on several branches and only *GmNF-Y* and *AtNF-Y* were presented. However, according to current knowledge about diploid and allotetraploid peanut genome, it is difficult to determine whether there were duplications not fixed by natural selection or massive gene lose in diploid ancestors. In conclusion, it can be inferred that the fewer number of peanut *NF-Y*s may due to the origin and evolution of *NF-Y*s in two diploid ancestors, and it will be an interesting area for further study.

There were relatively few NF-Y coding genes in peanut, and most of the *NF-Y*s existed as homologous gene pairs possibly indicating low functional redundancy among non-homologs within each subfamily. Homologs from different subgenomes usually share similar gene structures and

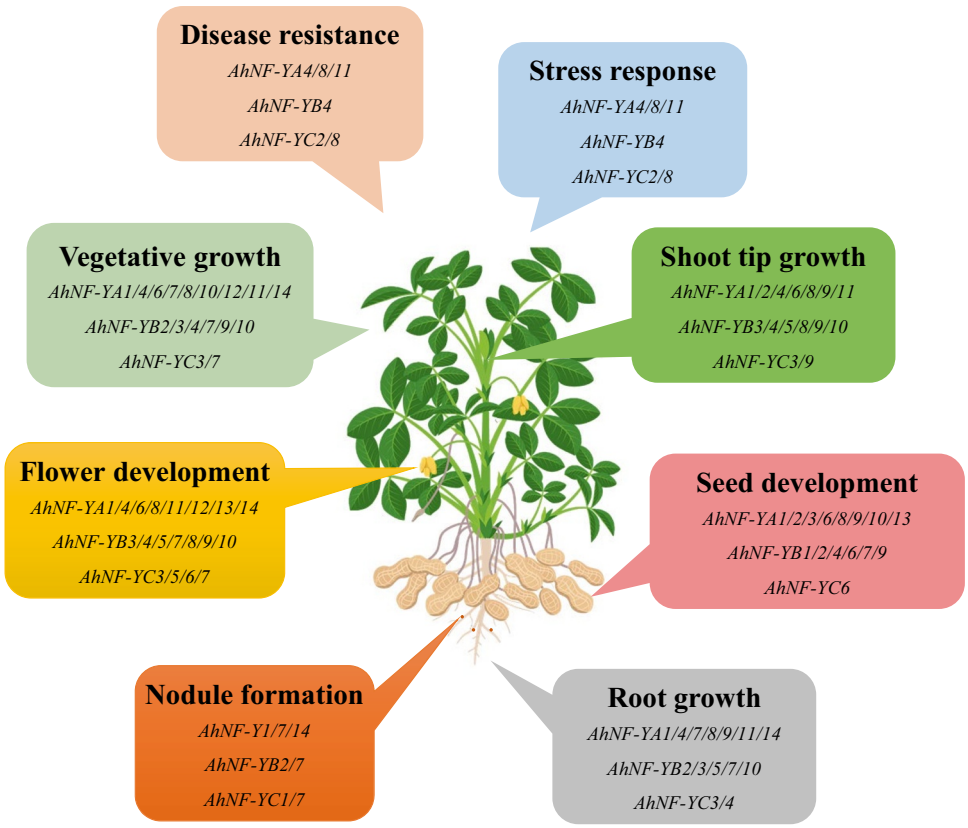

Fig. 7 Schematic diagram of the probable function of peanut *AhNF-Y*s. According to the tissue/organ expression patterns and the qRT-PCR data under stress, the probable biological processes peanut *AhNF-Y*s involved in were summarized. *AhNF-YC1* and *AhNF-YC8* showed relatively high expression level, but low differences between tissues/organs, therefore they are not included in this diagram

MEME motifs (Fig. 3). Some MEME motifs are conserved at the subfamily level, and they usually correspond to the DNA-binding and NF-Y interaction domains of NF-Y proteins. The MEME motifs that did not exhibit subfamily-level conservation will provide clues for NF-Y functional analysis. We performed row clustering for heatmap analysis, and the homologs showed high similarity in expression. There were 12 homologous pairs in the NF-YA and NF-YB subfamilies (Fig. 3A), seven of which (*NF-YA4/11*, *NF-YA1/8*, *NF-YA7/14*, *NF-YA2/9*, *NF-YB5/10*, *NF-YB4/9* and *NF-YB2/7*) showed the same tissue and salt stress-induced expression pattern (Figs. 4 and 5A) and the same MEME motifs within pairs (Fig. 3A). These pairs may play similar roles in most conditions; for example, *AhNF-YA4* and *AhNF-YA11* represent similar expression patterns under ABA and SA treatment (Fig. 6B, C). It has been reported that even if two *NF-Y* genes share very high sequence similarity and functional redundancy, some functional divergence remains (Sato et al. 2019). Under osmotic stress, the expression pattern of *AhNF-YA4* and *AhNF-YA11* were different (Fig. 6A). Although some homologous pairs carried the same MEME motifs, they presented different expression patterns, such as *NF-YA6/13*, *NF-YA3/10*, *NF-YB3/8* and *NF-YB1/6*. This difference may be due to the *cis*-elements in the promoter region (Fig. S3). *AhNF-YA5* and *AhNF-YA12* are homologs carrying different MEME motifs. There is a copy of motif 4 in NF-YA5, while motif 1 is present at a position similar to that in NF-YA12 (Fig. 3). These difference in promoter or function-unknown motifs may result in functional diversity of the *NF-YA* and *NF-YB* homologs. The NF-YC subfamily showed relatively low similarity in protein structure (Fig. 3) and expression in tissues (Fig. 4), which indicated that there may be low functional redundancy between NF-YCs and that NF-YC subunits may be the important determinants of NF-Y complex function in peanut.

According to the tissue/organ expression patterns and the qRT-PCR analysis under stress, the probable biological processes peanut *AhNF-Y*s involved in were summarized (Fig. 7). *AhNF-Y*s involve in both vegetative and reproductive, including shoot tip and root growth, flower and seed development. In addition, *AhAF-Y*s also play roles in abiotic stress response, disease resistance (SA-responsive genes) and nodule formation. Their functions in salt stress response were focused. Overexpression of *NF-YA* and *NF-YB* subunits could enhance the salt tolerance of *Arabidopsis* and *Paspalum vaginatum* O. Swartz (Li et al. 2013; Ma et al. 2015b; Wu et al. 2018). However, the regulatory mechanisms, especially the interaction proteins and downstream targets, have rarely been reported. In this study, we established a protein–protein interaction network including differentially expressed TFs under salt stress (Fig. 5D). In this network, AhNF-Ys could interact with multiple TFs and participate in several biological processes. These results will help in functional studies of plant NF-Ys under salt stress. *NF-Y*s in *Medicago truncatula*, *Phaseolus vulgaris*, *Lotus japonicus* and *Parasponia andersonii* play important roles in nodule formation, including *PvNF-YA1*, *LjNF-YA1*, *MtNF-YA1*, *PanNF-YA1*, *PvNF-YB7* and *PvNF-YC1* (Bu et al. 2020; Combier et al. 2006; Hossain et al. 2016; Laloum et al. 2014; Ripodas et al. 2019; Soyano et al. 2013; Zanetti et al. 2010). According to phylogenetic analysis (Fig. S5), their homologs genes in peanut, namely *AhNF-YA1/7/14*, *AhNF-YB2/7*, and *AhNF-YC1/7* showed relatively high expression levels in nodules, particularly *AhNF-YB2/7* (Fig. 4, Fig. 7). These proteins may participate in processes associated with peanut-rhizobia symbiosis in the form of NF-Y trimers or NF-Y-TF complexes.

In conclusion, here, we provide an overview of peanut NF-Y genes. The information described here will help in further investigation of the plant *NF-Y* gene family, especially in the context of salt stress response regulation and symbiosis with rhizobia.

Abbreviations AA: Amino acid; ABA: Abscisic acid; ABRE: Abscisic acid response element; bZIP: Basic region/leucine zipper motif; CBF: CCAAT binding factor; Co: CONSTANS; ER: Endoplasmic reticulum; HAP: Heme activator protein; HMMs: Hidden Markov models; MeJA: Methyl jasmonate; MW: Molecular weight; ORF: Open reading frame; pI: Isoelectric point; qRT-PCR: Real-time quantitative-polymerase chain reaction; SA: Salicylic acid

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12042-021-09295-2.

Authors' contributions QW, FZL and KZ conceived and designed this study. QW, LL, XRZ, YYL, and SQZ performed the research or analyzed the data. QW, LL and YSW wrote the manuscript. LRK and FZL were involved in revision of the manuscript. All authors read and approved the final manuscript.

Funding The authors acknowledge financial support from the earmarked fund for China Agriculture Research System (CARS-14), Integrated Research and Demonstration of Reduction Technology of Summer Peanut Fertilizer and Pesticide in the North (2018YFD0201008), the Peanut Seed Industry Project in Shandong Province of China (2020LZGC001), the earmarked fund for Agriculture Research System in Shandong Province of China (SDAIT-04–03).

Declarations

Conflict of interest The authors declare that they have no competing interests.

References

Alam MM, Tanaka T, Nakamura H, Ichikawa H, Kobayashi K, Yaeno T, Yamaoka N, Shimomoto K, Takayama K, Nishina H, Nishiguchi M (2015) Overexpression of a rice heme activator protein gene (OsHAP2E) confers resistance to pathogens, salinity and drought, and increases photosynthesis and tiller number. Plant Biotechnol J 13(1):85–96

Bailey T L, Boden M, Buske F A, Frith M, Grant C E, Clementi L, Ren J, Li W W, and Noble W S (2009) MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res 37(Web Server issue): W202–208

Ballif J, Endo S, Kotani M, MacAdam J, Wu Y (2011) Over-expression of HAP3b enhances primary root elongation in Arabidopsis. Plant Physiol Biochem 49(6):579–583

Banavath JN, Chakradhar T, Pandit V, Konduru S, Guduru KK, Akila C S, Podha S, and Puli C O R (2018) Stress inducible overexpression of AtHDG11 leads to improved drought and salt stress tolerance in peanut (*Arachis hypogaea* L.). Front Chem 6: 34

Bertioli DJ, Cannon SB, Froenicke L, Huang G, Farmer AD, Cannon EK, Liu X, Gao D, Clevenger J, Dash S, Ren L, Moretzsohn MC, Shirasawa K, Huang W, Vidigal B, Abernathy B, Chu Y, Niederhuth CE, Umale P, Araujo AC, Kozik A, Kim KD, Burow MD, Varshney RK, Wang X, Zhang X, Barkley N, Guimaraes PM, Isobe S, Guo B, Liao B, Stalker HT, Schmitz RJ, Scheffler BE, Leal-Bertioli SC, Xun X, Jackson SA, Michelmore R, Ozias-Akins P (2016) The genome sequences of Arachis duranensis and Arachis ipaensis, the diploid ancestors of cultivated peanut. Nat Genet 48(4):438–446

Bertioli DJ, Jenkins J, Clevenger J, Dudchenko O, Gao D, Seijo G, Leal-Bertioli SCM, Ren L, Farmer AD, Pandey MK, Samoluk SS, Abernathy B, Agarwal G, Ballen-Taborda C, Cameron C, Campbell J, Chavarro C, Chitikineni A, Chu Y, Dash S, El Baidouri M, Guo B, Huang W, Kim KD, Korani W, Lanciano S, Lui CG, Mirouze M, Moretzsohn MC, Pham M, Shin JH, Shirasawa K, Sinharoy S, Sreedasyam A, Weeks NT, Zhang X, Zheng Z, Sun Z, Froenicke L, Aiden EL, Michelmore R, Varshney RK, Holbrook CC, Cannon EKS, Scheffler BE, Grimwood J, Ozias-Akins P, Cannon SB, Jackson SA, Schmutz J (2019) The genome sequence of segmental allotetraploid peanut Arachis hypogaea. Nat Genet 51(5):877–884

Bu F, Rutten L, Roswanjaya YP, Kulikova O, Rodriguez-Franco M, Ott T, Bisseling T, van Zeijl A, Geurts R (2020) Mutant analysis in the nonlegume Parasponia andersonii identifies NIN and NF-YA1 transcription factors as a core genetic network in nitrogen-fixing nodule symbioses. New Phytol 226(2):541–554

Cagliari A, Carina A, Santos Maraschin F d, Loss G, Margis R, Margis-Pinheiro M (2011) The Evolutionary History of CBF Transcription Factors: Gene Duplication of CCAAT – Binding Factors NF-Y in Plants. Gene Duplication, vol 11. InTech, Rijeka, pp197-222

Cannon SB, Mitra A, Baumgarten A, Young ND, May G (2004) The roles of segmental and tandem gene duplication in the evolution of large gene families in Arabidopsis thaliana. BMC Plant Biol 4:10

Cao S, Kumimoto RW, Gnesutta N, Calogero AM, Mantovani R, Holt BF. Holt 3rd (2014) A distal CCAAT/NUCLEAR FACTOR Y complex promotes chromatin looping at the FLOWERING LOCUS T promoter and regulates the timing of flowering in Arabidopsis. Plant Cell 26(3):1009-1017

Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, Xia R (2020) TBtools: An integrative toolkit developed for interactive analyses of big biological data. Mol Plant 13(8):1194–1202

Chen NZ, Zhang XQ, Wei PC, Chen QJ, Ren F, Chen J, Wang XC (2007) AtHAP3b plays a crucial role in the regulation of flowering time in Arabidopsis during osmotic stress. J Biochem Mol Biol 40(6):1083–1089

Chen X, Lu Q, Liu H, Zhang J, Hong Y, Lan H, Li H, Wang J, Liu H, Li S, Pandey MK, Zhang Z, Zhou G, Yu J, Zhang G, Yuan J, Li X, Wen S, Meng F, Yu S, Wang X, Siddique KHM, Liu ZJ, Paterson AH, Varshney RK, Liang X (2019) Sequencing of cultivated peanut, *Arachis hypogaea*, yields insights into genome evolution and oil improvement. Mol Plant 12(7):920–934

Chen Y, Yang Z, Xiao Y, Wang P, Wang Y, Ge X, Zhang C, Zhang X, Li F (2018) Genome-wide analysis of the NF-YB gene family in *Gossypium hirsutum* L. and characterization of the role of GhDNF-YB22 in embryogenesis. Int J Mol Sc*i* 19(2): 483

Clevenger J, Chu Y, Scheffler B, Ozias-Akins P (2016) A developmental transcriptome map for allotetraploid *Arachis hypogaea*. Front Plant Sci 7:1446

Combier JP, Frugier F, de Billy F, Boualem A, El-Yahyaoui F, Moreau S, Vernie T, Ott T, Gamas P, Crespi M, Niebel A (2006) MtHAP2-1 is a key transcriptional regulator of symbiotic nodule development regulated by microRNA169 in Medicago truncatula. Genes Dev 20(22):3084–3088

Dash S, Cannon EKS, Kalberer SR, Farmer AD, Cannon S B (2016) PeanutBase and other bioinformatic resources for peanut. Peanut, vol 8. AOCS Press, pp241–252

Finn RD, Clements J, Arndt W, Miller BL, Wheeler TJ, Schreiber F, Bateman A, Eddy SR (2015) HMMER web server: 2015 update. Nucleic Acids Res 43(W1):W30–W38

Gnesutta N, Kumimoto RW, Swain S, Chiara M, Siriwardana C, Horner DS, Holt 3rd BF, Mantovani R (2017) CONSTANS Imparts DNA Sequence Specificity to the Histone Fold NF-YB/NF-YC Dimer. Plant Cell 29(6):1516-1532

Grabiele M, Chalup L, Robledo G, Seijo G (2012) Genetic and geographic origin of domesticated peanut as evidenced by 5S rDNA and chloroplast DNA sequences. Plant Syst Evol 298(6):1151–1165

Hackenberg D, Wu Y, Voigt A, Adams R, Schramm P, Grimm B (2012) Studies on differential nuclear translocation mechanism and assembly of the three subunits of the Arabidopsis thaliana transcription factor NF-Y. Mol Plant 5(4):876–888

Han X, Tang S, An Y, Zheng DC, Xia XL, Yin WL (2013) Overexpression of the poplar NF-YB7 transcription factor confers drought tolerance and improves water-use efficiency in Arabidopsis. J Exp Bot 64(14):4589–4601

Hossain MS, Shrestha A, Zhong S, Miri M, Austin RS, Sato S, Ross L, Huebert T, Tromas A, Torres-Jerez I, Tang Y, Udvardi M, Murray JD, Szczyglowski K (2016) Lotus japonicus NF-YA1 plays an essential role during nodule differentiation and targets members of the SHI/STY gene family. Mol Plant Microbe Interact 29(12):950–964

Hu B, Jin J, Guo AY, Zhang H, Luo J, Gao G (2015) GSDS 2.0: an upgraded gene feature visualization server. Bioinformatics 31(8):1296–1297

Huang M, Hu Y, Liu X, Li Y, Hou X (2015) Arabidopsis LEAFY COTYLEDON1 controls cell fate determination during post-embryonic development. Front Plant Sci 6:955

Hwang K, Susila H, Nasim Z, Jung JY, Ahn JH (2019) Arabidopsis ABF3 and ABF4 transcription factors act with the NF-YC complex to regulate SOC1 expression and mediate drought-accelerated flowering. Mol Plant 12(4):489–505

Koch MA, Haubold B, Mitchell-Olds T (2000) Comparative evolutionary analysis of chalcone synthase and alcohol dehydrogenase loci in Arabidopsis, Arabis, and related genera (Brassicaceae). Mol Biol Evol 17(10):1483–1498

Kondrashov FA (2012) Gene duplication as a mechanism of genomic adaptation to a changing environment. Proc Biol Sci 279(1749):5048–5057

Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA (2009) Circos: an information aesthetic for comparative genomics. Genome Res 19(9):1639–1645

Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol 33(7):1870–1874

Laloum T, De Mita S, Gamas P, Baudin M, Niebel A (2013) CCAAT-box binding transcription factors in plants: Y so many? Trends Plant Sci 18(3):157–166

Laloum T, Baudin M, Frances L, Lepage A, Billault-Penneteau B, Cerri MR, Ariel F, Jardinaud MF, Gamas P, de Carvalho-Niebel F, Niebel A (2014) Two CCAAT-box-binding transcription factors redundantly regulate early steps of the legume-rhizobia endosymbiosis. Plant J 79(5):757–768

Lee DK, Kim HI, Jang G, Chung PJ, Jeong JS, Kim YS, Bang SW, Jung H, Choi YD, Kim JK (2015) The NF-YA transcription factor OsNF-YA7 confers drought stress tolerance of rice in an abscisic acid independent manner. Plant Sci 241:199–210

Lescot M, Dehais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouze P, Rombauts S (2002) PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. Nucleic Acids Res 30(1):325–327

Letunic I, Bork P (2018) 20 years of the SMART protein domain annotation resource. Nucleic Acids Res 46(D1):D493–D496

Li C, Distelfeld A, Comis A, Dubcovsky J (2011) Wheat flowering repressor VRN2 and promoter CO2 compete for interactions with NUCLEAR FACTOR-Y complexes. Plant J 67(5):763–773

Li C, Song W, Luo Y, Gao S, Zhang R, Shi Z, Wang X, Wang R, Wang F, Wang J, Zhao Y, Su A, Wang S, Li X, Luo M, Wang S, Zhang Y, Ge J, Tan X, Yuan Y, Bi X, He H, Yan J, Wang Y, Hu S, Zhao J (2019) The HuangZaoSi maize genome provides insights into genomic variation and improvement history of maize. Mol Plant 12(3):402–409

Li P-S, Yu T-F, He G-H, Chen M, Zhou Y-B, Chai S-C, Xu Z-S, Ma Y-Z (2014) Genome-wide analysis of the Hsf family in soybean and functional identification of GmHsf-34 involvement in drought and heat stresses. BMC Genomics 15(1):1009

Li S, Li K, Ju Z, Cao D, Fu D, Zhu H, Zhu B, Luo Y (2016) Genome-wide analysis of tomato NF-Y factors and their role in fruit ripening. BMC Genomics 17:36

Li WX, Oono Y, Zhu J, He XJ, Wu JM, Iida K, Lu XY, Cui X, Jin H, Zhu JK (2008) The Arabidopsis NFYA5 transcription factor is regulated transcriptionally and posttranscriptionally to promote drought resistance. Plant Cell 20(8):2238–2251

Li XY, Hooft van Huijsduijnen R, Mantovani R, Benoist C, Mathis D (1992) Intron-exon organization of the NF-Y genes. Tissue-specific splicing modifies an activation domain. J Biol Chem 267(13): 8984–8990

Li YJ, Fang Y, Fu YR, Huang JG, Wu CA, and Zheng CC (2013) NFYA1 is involved in regulation of postgermination growth arrest under salt stress in Arabidopsis. PLoS One 8(4): e61289

Liu JX, Howell SH (2010) bZIP28 and NF-Y transcription factors are activated by ER stress and assemble into a transcriptional complex to regulate stress response genes in Arabidopsis. Plant Cell 22(3):782–796

Ma X, Li C, Wang M (2015a) Wheat NF-YA10 functions independently in salinity and drought stress. Bioengineered 6(4):245–247

Ma X, Zhu X, Li C, Song Y, Zhang W, Xia G, Wang M (2015b) Overexpression of wheat NF-YA10 gene regulates the salinity stress response in Arabidopsis thaliana. Plant Physiol Biochem 86:34–43

Marchler-Bauer A, Bo Y, Han L, He J, Lanczycki CJ, Lu S, Chitsaz F, Derbyshire MK, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Lu F, Marchler GH, Song JS, Thanki N, Wang Z, Yamashita RA, Zhang D, Zheng C, Geer LY, Bryant SH (2017) CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. Nucleic Acids Res 45(D1):D200–D203

Mohanta TK, Yadav D, Khan A, Hashem A, Tabassum B, Khan AL, Abd Allah EF, Al-Harrasi A (2020) Genomics, molecular and evolutionary perspective of NAC transcription factors. PLoS ONE 15(4):e0231425–e0231425

Mu J, Tan H, Hong S, Liang Y, Zuo J (2013) Arabidopsis transcription factor genes NF-YA1, 5, 6, and 9 play redundant roles in male gametogenesis, embryogenesis, and seed development. Mol Plant 6(1):188–201

Muday GK, Siriwardana CL, Gnesutta N, Kumimoto RW, Jones DS, Myers ZA, Mantovani R, Holt B F (2016) NUCLEAR FACTOR Y, subunit A (NF-YA) proteins positively regulate flowering and act through FLOWERING LOCUS T. PLOS Genetics 12(12): e1006496

Myers ZA, Holt BF 3rd (2018) NUCLEAR FACTOR-Y: still complex after all these years? Curr Opin Plant Biol 45 Pt A:96-102

Myers ZA, Kumimoto RW, Siriwardana CL, Gayler KK, Risinger JR, Pezzetta D, Holt Iii B F (2016) NUCLEAR FACTOR Y, subunit C (NF-YC) transcription factors are positive regulators of photomorphogenesis in *Arabidopsis thaliana.* PLoS Genet 12(9): e1006333

Nardini M, Gnesutta N, Donati G, Gatta R, Forni C, Fossati A, Vonrhein C, Moras D, Romier C, Bolognesi M, Mantovani R (2013) Sequence-specific transcription factor NF-Y displays histone-like DNA binding and H2B-like ubiquitination. Cell 152(1–2):132–143

Nelson DE, Repetti PP, Adams TR, Creelman RA, Wu J, Warner DC, Anstrom DC, Bensen RJ, Castiglioni PP, Donnarummo MG, Hinchey BS, Kumimoto RW, Maszle DR, Canales RD, Krolikowski KA, Dotson SB, Gutterson N, Ratcliffe OJ, Heard JE (2007) Plant nuclear factor Y (NF-Y) B subunits confer drought tolerance and lead to improved corn yields on water-limited acres. Proc Natl Acad Sci U S A 104(42):16450–16455

Ni Z, Hu Z, Jiang Q, Zhang H (2013) GmNFYA3, a target gene of miR169, is a positive regulator of plant tolerance to drought stress. Plant Mol Biol 82(1–2):113–129

Pan CL, Yao SC, Xiong WJ, Luo S Z, Wang YL, Wang A Q, Xiao D, Zhan J, He L F (2017) Nitric oxide inhibits Al-induced programmed cell death in root tips of peanut (*Arachis hypogaea* L.) by affecting physiological properties of antioxidants systems and cell wall. Front Physiol 8: 1037

Petroni K, Kumimoto RW, Gnesutta N, Calvenzani V, Fornari M, Tonelli C, Holt 3rd BF, Mantovani R (2012) The promiscuous life of plant NUCLEAR FACTOR Y transcription factors. Plant Cell 24(12):4777-4792

Quach TN, Nguyen HT, Valliyodan B, Joshi T, Xu D, Nguyen HT (2015) Genome-wide expression analysis of soybean NF-Y genes reveals potential function in development and drought response. Mol Genet Genomics 290(3):1095–1115

Rao S, Jha S, Bansal C, Gupta A, Sorin C, Crespi M, and Mathur S (2021) A conserved HSF:miR169:NF-YA loop involved in tomato and Arabidopsis heat stress tolerance. bioRxiv 2021.01.01.425064

Ripodas C, Castaingts M, Clua J, Villafane J, Blanco FA, Zanetti ME (2019) The PvNF-YA1 and PvNF-YB7 Subunits of the Heterotrimeric NF-Y Transcription Factor Influence Strain Preference in the Phaseolus vulgaris-Rhizobium etli Symbiosis. Front Plant Sci 10:221

Robledo G, Lavia GI, Seijo G (2009) Species relations among wild Arachis species with the A genome as revealed by FISH mapping of rDNA loci and heterochromatin detection. Theor Appl Genet 118(7):1295–1307

Salman-Minkov A, Sabath N, Mayrose I (2016) Whole-genome duplication as a key factor in crop domestication. Nat Plants 2:16115

Sato H, Mizoi J, Tanaka H, Maruyama K, Qin F, Osakabe Y, Morimoto K, Ohori T, Kusakabe K, Nagata M, Shinozaki K, Yamaguchi-Shinozaki K (2014) Arabidopsis DPB3-1, a DREB2A interactor, specifically enhances heat stress-induced gene expression by forming a heat stress-specific transcriptional complex with NF-Y subunits. Plant Cell 26(12):4954–4973

Shen C, Liu H, Guan Z, Yan J, Zheng T, Yan W, Wu C, Zhang Q, Yin P, Xing Y (2020) Structural insight into DNA recognition by CCT/

NF-YB/YC complexes in plant photoperiodic flowering. Plant Cell 32(11):3469–3484
Shen Y, Liu J, Geng H, Zhang J, Liu Y, Zhang H, Xing S, Du J, Ma S, Tian Z (2018) De novo assembly of a Chinese soybean genome. Sci China Life Sci 61(8):871–884
Siefers N, Dang KK, Kumimoto RW, Bynum WE, Tayrose G, Holt BF 3rd (2009) Tissue-specific expression patterns of Arabidopsis NF-Y transcription factors suggest potential for extensive combinatorial complexity. Plant Physiol 149(2):625-641
Sorin C, Declerck M, Christ A, Blein T, Ma L, Lelandais-Brière C, Njo MF, Beeckman T, Crespi M, Hartmann C (2014) A miR169 isoform regulates specific NF-YA targets and root architecture in Arabidopsis. New Phytol 202(4):1197–1211
Soyano T, Kouchi H, Hirota A, Hayashi M (2013) Nodule inception directly targets NF-Y subunit genes to regulate essential processes of root nodule development in Lotus japonicus. PLoS Genet 9(3): e1003352
Stephenson TJ, McIntyre CL, Collet C, Xue GP (2011) TaNF-YB3 is involved in the regulation of photosynthesis genes in Triticum aestivum. Funct Integr Genomics 11(2):327–340
Su H, Cao Y, Ku L, Yao W, Cao Y, Ren Z, Dou D, Wang H, Ren Z, Liu H, Tian L, Zheng Y, Chen C, Chen Y (2018) Dual functions of ZmNF-YA3 in photoperiod-dependent flowering and abiotic stress responses in maize. J Exp Bot 69(21):5177–5189
Thirumurugan T, Ito Y, Kubo T, Serizawa A, Kurata N (2008) Identification, characterization and interaction of HAP family genes in rice. Mol Genet Genomics 279(3):279–289
Voorrips RE (2002) MapChart: software for the graphical presentation of linkage maps and QTLs. J Hered 93(1):77–78
Wang J, Sun N, Deng T, Zhang L, Zuo K (2014) Genome-wide cloning, identification, classification and functional analysis of cotton heat shock transcription factors in cotton (Gossypium hirsutum). BMC Genomics 15(1):961
Wang M, Tu L, Yuan D, Zhu SC, Li J, Liu F, Pei L, Wang P, Zhao G, Ye Z, Huang H, Yan F, Ma Y, Zhang L, Liu M, You J, Yang Y, Liu Z, Huang F, Li B, Qiu P, Zhang Q, Zhu L, Jin S, Yang X, Min L, Li G, Chen LL, Zheng H, Lindsey K, Lin Z, Udall JA, Zhang X (2019) Reference genome sequences of two cultivated allotetraploid cottons, Gossypium hirsutum and Gossypium barbadense. Nat Genet 51(2):224–229
Wang Y, Tang H, Debarry JD, Tan X, Li J, Wang X, Lee TH, Jin H, Marler B, Guo H, Kissinger JC, Paterson AH (2012) MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res 40(7): e49
Wu X, Shi H, Guo Z (2018) Overexpression of a NF-YC gene results in enhanced drought and salt tolerance in transgenic Seashore Paspalum. Front Plant Sci 9:1355
Yamamoto A, Kagaya Y, Toyoshima R, Kagaya M, Takeda S, Hattori T (2009) Arabidopsis NF-YB subunits LEC1 and LEC1-LIKE activate transcription by interacting with seed-specific ABRE-binding factors. Plant J 58(5):843–856
Zanetti ME, Ripodas C, Niebel A (2017) Plant NF-Y transcription factors: Key players in plant-microbe interactions, root development and adaptation to stress. Biochim Biophys Acta Gene Regul Mech 1860(5):645–654
Zanetti ME, Blanco FA, Beker MP, Battaglia M, Aguilar OM (2010) A C subunit of the plant nuclear factor NF-Y required for rhizobial infection and nodule development affects partner selection in the common bean-Rhizobium etli symbiosis. Plant Cell 22(12):4142–4157
Zhang H, Zhu S, Liu T, Wang C, Cheng Z, Zhang X, Chen L, Sheng P, Cai M, Li C, Wang J, Zhang Z, Chai J, Zhou L, Lei C, Guo X, Wang J, Wang J, Jiang L, Wu C, Wan J (2019) DELAYED HEADING DATE1 interacts with OsHAP5C/D, delays flowering time and enhances yield in rice. Plant Biotechnol J 17(2):531–539
Zhang Q, Zhang J, Wei H, Fu X, Ma L, Lu J, Wang H, Yu S (2020) Genome-wide identification of NF-YA gene family in cotton and the positive role of GhNF-YA10 and GhNF-YA23 in salt tolerance. Int J Biol Macromol 165(Pt B):2103–2115
Zhang X, Zhu S, Zhang K, Wan Y, Liu F, Sun Q, Li Y (2018) Establishment and evaluation of a peanut association panel and analysis of key nutritional traits. J Integr Plant Biol 60(3):195–215
Zhang Z, Li X, Zhang C, Zou H, Wu Z (2016) Isolation, structural analysis, and expression characteristics of the maize nuclear factor Y gene families. Biochem Biophys Res Commun 478(2):752–758
Zhao H, Wu D, Kong F, Lin K, Zhang H, Li G (2016) The Arabidopsis thaliana Nuclear Factor Y Transcription Factors. Front Plant Sci 7:2045
Zhuang WJ, Chen H, Yang M, Wang JP, Pandey MK, Zhang C, Chang WC, Zhang LS, Zhang XT, Tang RH, Garg V, Wang XJ, Tang HB, Chow CN, Wang JP, Deng Y, Wang D P, Khan AW, Yang Q, Cai TC, Bajaj P, Wu KC, Guo BZ, Zhang XY, Li JJ, Liang F, Hu J, Liao BS, Liu SY, Chitikineni A, Yan HS, Zheng YX, Shan SH, Liu QZ, Xie DY, Wang ZY, Khan SA, Ali N, Zhao CZ, Li XG, Luo ZL, Zhang SB, Zhuang RR, Peng Z, Wang SY, Mamadou G, Zhuang YH, Zhao ZF, Yu WC, Xiong FQ, Quan WP, Yuan M, Li Y, Zou HS, Xia H, Zha L, Fan JP, Yu JG, Xie WP, Yuan JQ, Chen K, Zhao SS, Chu WT, Chen YT, Sun PC, Meng FB, Zhuo T, Zhao YH, Li CJ, He GH, Zhao YL, Wang CC, Kavikishor PB, Pan RL, Paterson AH, Wang XY, Ming R, Varshney RK (2019) The genome of cultivated peanut provides insight into legume karyotypes, polyploid evolution and crop domestication. Nature Genetics 51(5): 865-876.