Genome-wide identification and characterization of NF-Y gene family in peanut (Arachis hypogaea L.)

**CURRENT STATUS:** UNDER REVIEW

BMC Genomics

Qian Wan
Shandong Agricultural University

Lu Luo
Shandong Agricultural University

Xiurong Zhang
Shandong Agricultural University

Yuying Lv
Shandong Agricultural University

Suqing Zhu
Shandong Agricultural University

Lingrang Kong
Shandong Agricultural University

Yongshan Wan
Shandong Agricultural University

Fengzhen Liu liufz@sdau.edu.cn
Corresponding Author
ORCiD: 0000-0002-0537-8788

Kun Zhang
Shandong Agricultural University

**DOI:**
10.21203/rs.2.11389/v1

**SUBJECT AREAS**
- Plant Molecular Biology and Genetics

**KEYWORDS**
- Cultivated peanut (Arachis hypogaea L.), NF-Y gene family, Transcription factor, Phylogenetic analysis, Expression patterns, Salt stress
Abstract

Background

Nuclear factor Y (NF-Y) gene family consists of NF-YA, NF-YB and NF-YC subfamilies. Many members of NF-Y family have been involved in plant development processes, phytohormone signaling and tolerance to stresses in Arabidopsis and other plant species. However, little attention has been given in peanut.

Results

A total of 33 AhNF-Y genes (AhNF-Ys) were identified and distributed on 16 chromosomes. A phylogenetic analysis indicated that NF-Y genes possessed highly conservatism in different plants. Gene duplication analyze indicated that only segmental duplication were detected. The abiotic stress-related regulatory elements analysis showed that AhNF-Ys, except for AhNF-YB6, contained at least one abiotic stress response element. With RNA-seq data, the tissue/organ-specific expression and differential expression profiling under salt stress were analyzed, indicating that six selected AhNF-Y gene may play potential roles in the regulation of salt stress response. qRT-PCR results suggested that these AhNF-Y genes also responded to osmotic, ABA (Abscisic Acid) and SA (Salicylic acid) stresses.

Conclusions

In this study, thirty three AhNF-Y genes were identified in cultivated peanut and the phylogeny, gene structures, motif composition, chromosomal location, gene duplication, stress-related regulatory elements, and expression patterns were also examined. These results may contribute to functional characterization of AhNF-Y genes in further research.

Background

Nuclear factor Y (NF-Y), also known as the heme activator protein (HAP) or CCAAT-binding factor (CBF), is present in almost all higher eukaryotic genomes. NF-Y specifically binds
CCAAT boxes in promoters as a heterotrimeric complex [1]
consists of NF-YA (HAP2 or CBFB), NF-YB (HAP3 or CBFA), and NF-YC (HAP5 or CBFC) subunits [1].

A single subunit has no DNA-binding activity. Initially, NF-YB and NF-YC form a
tight dimer in the cytoplasm. This dimer subsequently interacts with NF-YA protein in the
nucleus to form a heterotrimeric complex, which binds to DNA non-specifically [1, 3].

The NF-Y complex can bind to different transcription factors or regulator to
active or repress transcription. In the past decade some researches indicate that the
CONSTANS (CO) can interact with the NF-YB-YC dimer to form a NF-YB-YC-TF trimer. This
complex can be substitute for the NF-YA-YB-YC to bind to the promoter region of the
downstream target gene, and regulate the expression [1, 5].

In addition, basic region/leucine zipper motif (bZIP) type transcription factors
bZIP28 and bZIP67 can play a similar role [1, 7].

In animals and yeast, each NF-Y subunit is encoded by a single gene [9].

However, in the plant lineage, each subunit type is encoded by a family of
around 10 genes. Thus, hundreds of unique combinations are possible in theory. This
amplification of NF-Y family in plants creates a flexible formation, leads to new and
divergent functions [10].

NF-Y genes have been reported to regulate many plant developmental processes, such as
gametogenesis [10].
early embryogenesis [11-15] seed development and germination [16-18] root growth [19-30] and fruit maturation [31-
They have been also found to be involved in physiological processes, for instance, the regulation of endoplasmic reticulum (ER) stress [8,32] photosynthesis [33-35] and photomorphogenesis. In addition, some NF-Y genes also involve in the response to abiotic stresses [18,30,36-47].

In leguminous plants, NF-Ys have been suggested as important regulators of organogenesis and development of symbiotic root nodules [18,48].

Peanut (Arachis hypogaea L.) is an important oil and food crop worldwide [49]. Cultivated peanut, which evolved from the hybridization and subsequent chromosome doubling A. duranensis (A) and A. ipaensis (B), is an allotetraploid (AABB genome, 2n = 4x = 40) with a total genome size of about 2.7 Gb [50-52].

]. The A subgenome is represented as Arah.01-Arah.10 and the B subgenome as Arah.11-Arah.20. However, the identification and function of AhNF-Y have been little
studied. Recently, the whole genome sequence of cultivated peanut was reported by two independent research groups, respectively [53, 54]. The genome-wide identification and systematic analysis of the NF-Y gene family in cultivated peanut genomes becomes fully feasible. In this study, we identified AhNF-Y gene family, analyzed the sequence features, phylogenetic relationship, chromosomal locations, gene duplication on the expansion, and abiotic stress/hormone-related regulatory elements in promoter of AhNF-Y gene family. The expression profiles in various tissues and organs of peanut during developmental stages were also investigated by RNA-seq method. In addition, using the RNA-seq and quantitative real-time PCR (qRT-PCR) methods, we analyzed the expression profiles of AhNF-Y genes under salt stress, and identified several candidate genes responsive to abiotic stress and hormonal treatment. The present results facilitate future investigations to elucidate the functional characterization of NF-Y genes in peanut.

Result

Identification and analysis of AhNF-Y genes

By using BLASTP and HMMER software, the amino acid sequences of Arabidopsis and rice NF-Ys were used to search the peanut genome database. Pfam, NCBI-CDD and SMART were used to confirm the NF-Y domain. Finally, 14 NF-YA genes, 10 NF-YB genes and 9 NF-YC genes were identified as peanut NF-Ys (AhNF-Ys). All these genes were named on the base
of the exact position on chromosome. The results of sequence alignment indicated that $AhNF-Y_A$ certain conserved domains genes. Gene characteristics, including gene names, gene IDs, chromosomal locations, open reading frame (ORF) lengths, exon numbers, amino acid (AA) numbers, molecular weights (MW) and the isoelectric points (pI) were showed in Table 1. $AhNF-Y$ proteins contained 171 ($AhNF-Y_B2$) to 492 ($AhNF-Y_A12$) amino acid. The MW of $AhNF-Y$s ranged from 18.88 kDa ($AhNF-Y_B2$) to 55.45 kDa ($AhNF-Y_A12$), and the pI values were between 5.19 ($AhNF-Y_C9$) and 9.64 ($AhNF-Y_A5$).

**Phylogenetic analysis of $AhNF-Y$ genes**

To investigate the evolutionary relationships of $NF-Y$ genes in plants, an un-rooted phylogenetic tree of $NF-Y$ complete protein sequences from *Arabidopsis*, rice, peanut and human was constructed using MEGA 7. The unrooted maximum likelihood phylogenetic tree was generated from 33 AhNF-Ys, 33AtNF-Ys, 28 OsNF-Ys and 3 NF-Y_HSs. The NF-Y proteins were divided into 3 subfamilies. As indicated in Fig. 1, the members of $NF-Y$ protein family exhibit high homology. In NF-YA subfamily, AhNF-YA1 and A8 exhibit highest homology with AtNF-YA1 and therefore they were orthologous proteins. In NF-YB subfamily, AhNF-YB4/9 and AtNF-YB2 belong to orthologous proteins. Besides, AhNF-YC3/9 share highest similarity with AtNF-YC1. This highly conservatism suggests that the $NF-Y$ proteins in different plant species might share similar function.

**Gene structure and motif analysis**

The cultivated peanut is an allotetraploid containing of A and B subgenome. With multiple sequences alignment (Figs. 2) and phylogeny evolution analysis (Fig. 2a), 7 pairs of homologous $AhNF-Y_A$ genes, 5 pairs of homologous $AhNF-Y_B$ genes and 2 pairs of homologous $AhNF-Y_C$ genes were identified, respectively (Table 2). Gene Structure Display Server 2.0 was used to analyze and visualize the structures of NF-
Y genes in peanut. As shown in Fig. 2a and 2b, AhNF-Ys share high similarity in number and length of introns when the amino acid sequence similarity of two genes was the highest. Comparing the members of the 3 subfamilies, the members of AhNF-YA subfamily contain a larger number of introns. Except for AhNF-YC1, the intron numbers of AhNF-YC subfamily were generally at a low level.

The MEME analysis tool was used to predict the conserved motifs in AhNF-Y genes (Fig. 2c and Additional file 6). Motif 7, 8 were only identified in AhNF-YA subfamily. Motif 5 was observed in each member of AhNF-YA subfamily except for AhNF-YA7 and A14. Motif 2, 4 and 11 were unique to AhNF-YB subfamily. Motif 3, 6 and 10 were existed only in AhNF-YC subfamily. These results indicated that NF-Y is a kind of highly conservative gene family in peanut. Moreover, motif 16 was found only in AhNF-YB3 and B8, and motif 9 and 17 were both identified in 4 AhNF-Y genes (AhNF-YA2, A3, A9 and A10). These specific motifs may result in some functional differences among AhNF-Y genes.

Chromosomal distribution and gene duplication of AhNF-Ys

The 33 AhNF-Y genes were unevenly distributed on 16 chromosomes except chr2, 5, 12 and 15 (Table 1, Fig. 3 and Figs. 1). Chr1, 11, 14 and 18 contained 3 AhNF-Y genes, respectively. Only one AhNF-Y gene was located on Chr17, 19 and 20. Significantly, all pairwise homologous genes of AhNF-YA and AhNF-YB subfamilies were located on the similar chromosomal positions of two subgenomes (Figs. 1).

Some studies indicated that both segmental duplication and tandem duplication played an important role to the generation for gene family during the evolution. [55]

Thus, we analyzed the duplication events of AhNF-Y genes (Fig. 3). Two gene pairs (AhNF-YA2/AhNF-YA3 and AhNF-YC1/AhNF-YC4) were detected as segmental duplications in A subgenome, and also two gene pairs (AhNF-YA9/AhNF-YA10 and AhNF-
*YB6/AhNF-YB7* in B subgenome. Segmental duplication accounted for around a quarter of the *AhNF-Y* genes. Moreover, tandem duplication of *AhNF-Ys* was not detected in the whole genome.

**Abiotic stress-related regulatory elements in *AhNF-Y* promoters**

To investigate the potential regulatory mechanisms of *AhNF-Ys* in the abiotic stress response, the 1.5kb upstream sequences from the initiation codon of *AhNF-Y* genes were detected using the PlantCARE database to identify regulatory elements. Six stress-related regulatory elements, including ABRE (abscisic acid responsiveness), CGTCA-motif, MBS, TCA-element, TC-rich and TGACG-motif, were showed in Fig. 4. The ABRE were located in 18 *AhNF-Ys*. The MeJA-responsive elements CGTCA-motif and TGACG-motif were both detected in 22 *AhNF-Ys*. A number of 12 *AhNF-Ys* contained TCA-element (salicylic acid responsiveness). In addition, MBS (drought responsiveness) and TC-rich repeats (defense and stress responsiveness) were found in 7 and 10 *AhNF-Ys*, respectively. At least one regulatory element was identified in the promoter region of *NF-Y* genes except *AhNF-YB6*. These results suggested that *AhNF-Y* genes may involve in many different abiotic stresses.

**Tissue/organ-specific expression analysis of *AhNF-Ys***

To determine the expression patterns of *AhNF-Y* genes, we used a published RNA-seq data, which covered gene expression profiles of 22 tissues throughout the entire life cycle of peanut.

Based on the expression characteristics, the tissue/organ expression profiles of *AhNF-Ys* were classified into four categories [Fig. 5]. Group 1 contained 3 genes (*AhNF-YA3*, *A10* and *C6*), which showed low expression level in most tissues except seed. Group 2 comprised 8 genes, which were hardly detected or only at low levels in all tissues and organs. The third group included only 2 genes (*AhNF-YC1* and *C8*), which expressed
extremely high, specifically in seed developmental stages. Group 4 was composed of the other AhNF-Ys. These genes represented higher expression levels compared with genes belong to Group 2 and display tissue/organ-specific expression patterns. For example, AhNF-YA1 and A8 exhibited root-specific expression; AhNF-YB4, B9 and C3 showed preferential expression in leaves; AhNF-YB2 and B7 were expressed highly only in nodule. The above results indicated that most AhNF-Y genes were constitutively expressed. Moreover, the tissues and organs-specific expression of some AhNF-Y genes suggests that they may play different roles in different tissues and growth stages peanut.

Expression analysis of AhNF-Y genes under salt treatment

Using the high through put RNA-seq data of AhNF-Y genes, a heatmap was established to analyze the response to salt stress. (Fig. 6a) Cluster analysis showed that 11 AhNF-Y genes (AhNF-YA1, A1, A4, A6, A8, A9, A11, B7, C1, C2, C7 and C8) indicated up-regulated expression in leaf, whereas the expression level of 3 genes (AhNF-YB6, B9 and C5) exhibited down-regulation. 3 AhNF-Y genes (AhNF-YA3, A7 and A14) showed decreased expression in root. The expression level of AhNF-YB2 was up-regulated in leaf, but down-regulated in root. Furthermore, AhNF-YB4 exhibited up-regulation in both leaf and root. In contrast, other 11 AhNF-Ys (AhNF-YA2, A5, A12, B3, B5, B8, B10, C3, C4, C6 and C9) were not induced by salt stress. In addition, the expression of AhNF-YA10 and AhNF-YB1 were not detected. This distinction of expression patterns indicated that the members of AhNF-Y family might have different response and regulatory mechanisms under salt stress.

To further validate RNA-seq data, 6 typical AhNF-Y genes (AhNF-YA4, A8, A11, B4, C2 and C8) were analyzed by qRT-PCR (Fig. 6b). After 16 h salt treatment (200mM), the expression levels of AhNF-YA4 and A11 up-regulated in both leaf and root. AhNF-YC2 and C8 indicated up-regulated expression only in leaf, but not induced in root. These results are consistent with RNA-seq data. Besides, AhNF-YA11 was not affected by salt stress in
leaf, and the expression level of AhNF-YB4 showed no significant change in both leaf and root. The expression patterns of AhNF-YA11 and AhNF-YB4 were inconsistent with RNA-seq data.

**Expression patterns of AhNF-Y genes in response to different treatments**

To further investigate whether the expression of these predicted AhNF-Y genes was influenced by other various stress treatments (mannitol, ABA and SA), qRT-PCR was used to survey the transcript levels in leaves (Fig.7 and 8). The result revealed that the transcript level of AhNF-YA4 and AhNF-YA8 were down-regulated, and both reached the nadir under osmotic stress at about 8 h. In contrast, AhNF-YA11, AhNF-YC2 and AhNF-YC8 had similar expression profiles, represented up-regulated trend. Under the same treatment, the expression level of AhNF-YB4 was up-regulated until about 12 h. The transcript levels of all 6 predicted AhNF-Y genes were increased under ABA treatment and reached peak at about 2 h except for AhNF-YB4. In addition, all these genes were significantly responded to SA tress. All above results indicated that these 6 AhNF-Y genes responded to osmotic, ABA and SA stress with distinctive expression patterns.

**Discussion**

NF-Y gene family, as a transcription factor family, is present in almost all plant species. However, the function in cultivated peanut is little known. In our study, 33 AhNF-Y genes were identified from the cultivated peanut genome database.

In our study, four pairs of AhNF-Y genes were detected as segment duplications, but no tandem duplication event was found, suggesting that only the segmental duplication contribute to the expansion of AhNF-Y gene family. And the remaining AhNF-Y genes may evolve in an early divergence time or be obtained from gene translocation.

To further investigate the putative function of AhNF-Ys, the RNA-seq data was analyzed.
As our results suggested that the majority of $AhNF$-Ys were expressed in all investigated tissues and organs. Individual members of $AhNF$-Ys exhibited tissue/organ-specific expression patterns. This phenomenon was also confirmed by previous studies in several other plants [29, 31, 47, 58, 59]. For instance, $AhNF$-YC1 and C8 exhibited preferential expression in seed developmental stages. Both of them may play important roles in seed development. Previous study of their ortholog $PvNF$-YB7 in Phaseolus vulgaris was consistent with our speculation [48].

The homologous genes $AhNF$-YB2 and $AhNF$-YB7 which had similar expression profiles were specifically expressed in nodule, thereby suggesting that their functions may be related to the formation of roots and nodules. Moreover, the expression levels of $AhNF$-YB4 and $AhNF$-YB9 were relatively higher in leaf and stem than root. These 2 genes may play similar roles with their orthologs $AtNF$-YB3 and $AtNF$-YB2, which were concern with the response to the endoplasmic reticulum stress and/or in the promotion of flowering in Arabidopsis [10, 60]. Additionally, some homologous genes shared similar transcript patterns, suggesting that they may have redundant functions. These hypotheses should be researched in future work.

$AhNF$-YA8 was orthologous to $AtNF$-YA1, which had been proved to be a salt stress responsive gene in Arabidopsis [11]
AhNF-YA8 had a similar expression pattern with AtNF-YA1 in response to salt stress, indicating their similar function in conferring salt tolerance. Our results showed that the expression level of AhNF-YA8 was up-regulated only in leaf under salt stress, but showed no significant changes in root, implying it may mainly function in leaf. In addition, AhNF-YA8 is also a positive regulator of ABA, which is consistent with the research of AtNF-YA1. Therefore, AhNF-YA8 should be a focus of the future study of peanut salt tolerance mechanism. The expression profiles of AhNF-YA4 and AhNF-YA8 were up-regulated under salt stress, but showed opposite expression features in response to osmotic stress. These results indicated that up-regulation of AhNF-YA4 and AhNF-YA8 may not due to the osmotic stress caused by salt stress.

Additionally, AhNF-YA11 and AhNF-YB4 showed different expression patterns in transcriptome data and real-time RT-PCR analysis. According to the principle of the two tests, the expression level detected by real-time RT-PCR cannot truly reflect level of complete mRNA, mRNAs partly degraded can also be amplified. On the other hand, assembling of transcripts from reads may also introduce errors. The SA responsive elements were not detected in the 1.5 kb upstream sequence of SA induced genes AhNF-YA8/B4/C2/C8 identified via qRT-PCR analysis in this study. Therefore, it can be inferred that the upstream regulators of the four genes may can be induced by SA, or unreported SA responsive elements exist in the promoter region, or the probable SA responsive elements located more than 1.5 kb far from the initiation codon.

Conclusions

A genome-wide analysis of peanut AhNF-Y family was performed, and a total of 33 AhNF-Y genes were identified. Analyses of AhNF-Y genes, including phylogeny, gene structures, motif composition, chromosomal location, gene duplication, stress-related regulatory elements, were conducted based on bioinformatics methods. Published RNA-seq data was
used to explore their tissue/organ-specific expression patterns. Combining our own results of RNA-seq data and qRT-PCR under several kinds of stresses, we also identified some candidate AhNF-Y genes that may have functions under these stresses. These results provide comprehensive information on the AhNF-Y gene family in peanut and could facilitate further research on their functional mechanisms.

Methods

Identification of NF-Y genes in peanut genomes

The genomic sequence of *Arachis hypogaea* cv. Tifrunner and the annotated gene models were downloaded from peanutbase (http://www.peanutbase.org/). The NF-Y protein sequences of *A. thaliana*, *Homo sapiens* and *Oryza sativa* used in this research (Additional file 2, 3) were downloaded from NCBI (http://www.ncbi.nlm.nih.gov/). The hidden Markov model (HMM) profile of CBFB_NFYA (PF02045) was obtained from Pfam protein family database (http://pfam.xfam.org/). HMMER was used as the query (p<0.01) to search the NF-YA genes [1]

[1]The known NF-Y proteins of Arabidopsis and rice were used for BLAST analyses. The supposed AhNF-Y proteins were uploaded to CDD (https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi/) [2]

[2]Pfam and SMART (http://smart.embl-heidelberg.de/) to verify the conserved NF-Y domain. ExPASy Proteomics Server (http://prosite.expasy.org/) was used to acquire the length of sequences, molecular weights and isoelectric points [3]

[3]

[4]

Phylogenetic analysis and sequence analysis

The phylogenetic trees were conducted by MEGA7, and maximum likelihood method was
used with 1000 bootstrap replications [§5]

iTOL (http://itol.embl.de/) was used for further editing. The MEME program (http://meme-suite.org/tools/meme) was used to predict the conserved motifs in the identified AhNF-Ys proteins, with the following parameters: maximum number of 20 motifs and the width between 6 and 100 amino acid residues [§6].

The exon-intron structures analysis was visualized by comparing cDNA sequences with their corresponding full-length DNA sequences by using the online tool Gene Structure Display Server version 2.0 (gsds.cbi.pku.edu.cn/) [§7].

Chromosomal distribution and gene duplication

MCScanX was used to analyze the gene duplication events, with the default parameters. The charts of chromosomal distribution and synteny analysis were produced by Circos and Mapchart [§8, §9].

Prediction of regulatory elements in AhNF-Y genes

The 1.5 kb upstream sequences of the initiation codon (ATG) of each AhNF-Y genes were submitted to PlantCARE database (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) to identify six regulatory elements [§10].

Plant Materials and Growth Conditions

Mature seeds of Arachis hypogasa L. cultivar Fenghua 2 (Spanish type) were used in this research. The seeds were germinated on distilled water-wet degreasing cotton in seedling-
raising disks. These disks were placed at 26°C in darkness for 3 days, and then the germinated seeds were exposed to long-day conditions (LD; 16h light and 8h dark cycle). Two-functional-leaves seedlings were transplant to hydroponic-box, and cultured with 1/5 Hoagland's nutrient solution.

Stress treatment, total RNA extraction and revers transcription
To analysis expression pattern of AhNF-Y genes, two-week old seedlings were treated with nutrient solution containing 200mM NaCl, 20% (w/v) mannitol, 100mM ABA and 100mM SA, respectively. Leaf and root of seedlings treated with NaCl were harvested at 0h and 16h. These samples were used for transcriptome and qRT-PCR analysis. Leaves of seedlings treated with mannitol were collected at 0h, 2h, 4h, 6h, 8h, 12h and 24h. For ABA and SA treatment, leaves of seedlings were harvested at 0h, 1h, 2h, 4h, 6h and 8h. All samples were frozen immediately in liquid nitrogen for RNA extraction. The experiments were conducted with three independent biological replicates.

Using Quick RNA Isolation Kit (Waryong, Beijing, China), total RNA was isolated following manufacturer’s instructions. Concentration of the total RNA in each sample was quantified via NanoDrop 2000 Microvolume Spectrophotometers (Thermo Fisher Scientific, Massachusetts, USA). For revers transcription, 1μg of total RNA was used according to the manufacturer’s instruction of Advantage RT-for-PCR Kit (TaKaRa, Dalian, China). Total RNA and cDNA were frozen at -80°C and -20°C, respectively. Using the NEBNext Ultra RNA Library Prep Kit for Illumina (New England Biolabs Inc., Ipswich, USA), the RNA libraries preparation was performed. Paired-end sequencing were performed by HiSeq X Ten System (Illumina, San Diego, USA). The high-throughput sequencing of transcriptome was carried out by Beijing ORI-GENE science and technology co., LTD.
Expression profile analysis of AhNF-Ys

A published transcriptome dataset of cultivated peanut was obtained from the website PeanutBase (https://www.peanutbase.org/).

The assembled transcripts of this study were annotated as either “A”- or “B”- derived.

Previous study showed that the genomic sequences of A and B subgenome in A. hypogaea share high similarity with A. duranensis and A. ipaensis, respectively.

BLAST was used to query transcripts for AhNF-Y genes. For each AhNF-Y gene, only the subject “A”- or “B”- derived gene with the highest identity and the least mismatches was retained for extraction of gene expression.

The heatmap was created by TBtools with FPKM values be standardized by log2.

The SYBR green real-time PCR was carried out using TB Green Premix EX Taq (Tli RNaseH Plus, TaKaRa, Dalian, China) and performed by StepOne Plus system (Applied Biosystems, Waltham, USA) in a 20 μL reaction volume according to the manuscript. Three biological
repeats were performed for each sample, with three technical replicates each. The primers were designed using the Beacon Designer 7.9. Actin was used as the internal reference gene. Sequences of the primers and actin were shown in Additional file 9. The relative expression levels of the AhNF-Ys genes were evaluated by the method of the \(2^{-\Delta\Delta Ct}\). Statistical difference was determined by Student’s t test \((**P < 0.01, *P < 0.05, n = 3)\) using Excel.

Abbreviations

AA: amino acid; ABA: Abscisic Acid; ABRE: abscisic acid responsiveness; 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 weights; ORF: open reading frame; pI: the isoelectric points; qRT-PCR: Real-time quantitative; SA: Salicylic acid

Declarations

Acknowledgments

The authors thank to lab members for assistance.

Funding

This work was supported by 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 Industy Project in Shandong Province of China (2017LZGC003), the earmarked fund for Agriculture Research System in Shandong province of China (SDAIT-04-03).

Availability of data and materials

Data generated or analyzed during this research are included in this article and its Additional files.
Authors’ contributions

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

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

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

2. Mantovani R: The molecular biology of the CCAAT-binding factor NF-Y. (0378-1119 (Print)).

3. Frontini M, Imbriano C, Manni I, Mantovani R: Cell cycle regulation of NF-YC nuclear localization. Cell cycle 2004, 3(2):217-222.

4. Kahle J, Baake M, Doenecke D, Albig W: Subunits of the Heterotrimeric Transcription Factor NF-Y Are Imported into the Nucleus by Distinct Pathways Involving Importin and Importin 13. Molecular and cellular biology 2005, 25(13):5339-5354.

5. Hou X, Zhou J, Liu C, Liu L, Shen L, Yu H: Nuclear factor Y-mediated H3K27me3 demethylation of the SOC1 locus orchestrates flowering responses of Arabidopsis. Nature communications 2014, 5:4601.
6. Xu F, Li T, Xu PB, Li L, Du SS, Lian HL, Yang HQ: DELLA proteins physically interact with CONSTANS to regulate flowering under long days in Arabidopsis. *FEBS letters* 2016, 590(4):541-549.

7. Yamamoto A, Kagaya Y, Toyoshima R, Kagaya M, Takeda S, Hattori T: Arabidopsis NF-YB subunits LEC1 and LEC1-LIKE activate transcription by interacting with seed-specific ABRE-binding factors. *The Plant journal : for cell and molecular biology* 2009, 58(5):843-856.

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

9. Li XY, Hooft van Huijsduijnen R, Mantovani R, Benoist C, Mathis D: Intron-exon organization of the NF-Y genes. Tissue-specific splicing modifies an activation domain. *The Journal of biological chemistry* 1992, 267(13):8984-8990.

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

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

12. Fornari M, Calvenzani V, Masiero S, Tonelli C, Petroni K: The Arabidopsis NF-YA3 and NF-YA8 genes are functionally redundant and are required in early embryogenesis. *PloS one* 2013, 8(11):e82043.

13. Kwong RW: LEAFY COTYLEDON1-LIKE Defines a Class of Regulators Essential for Embryo Development. *The Plant Cell Online* 2002, 15(1):5-18.

14. Lee H, Fischer RL, Goldberg RB, Harada JJ: Arabidopsis LEAFY COTYLEDON1 represents a functionally specialized subunit of the CCAAT binding transcription factor. *Proceedings
of the National Academy of Sciences 2003, 100(4):2152-2156.

15. Huang M, Hu Y, Liu X, Li Y, Hou X: Arabidopsis LEAFY COTYLEDON1 controls cell fate determination during post-embryonic development. Frontiers in plant science 2015, 6:955.

16. Sorin C, Declerck M, Christ A, Blein T, Ma L, Lelandais-Brière C, Njo MF, Beeckman T, Crespi M, Hartmann C: A miR169 isoform regulates specific NF-YA targets and root architecture in Arabidopsis. New Phytologist 2014, 202(4):1197-1211.

17. Ballif J, Endo S, Kotani M, MacAdam J, Wu Y: Over-expression of HAP3b enhances primary root elongation in Arabidopsis. Plant physiology and biochemistry : PPB 2011, 49(6):579-583.

18. Zanetti ME, Ripodas C, Niebel A: Plant NF-Y transcription factors: Key players in plant-microbe interactions, root development and adaptation to stress. Biochimica et biophysica acta Gene regulatory mechanisms 2017, 1860(5):645-654.

19. Muday GK, Siriwardana CL, Gnesutta N, Kumimoto RW, Jones DS, Myers ZA, Mantovani R, Holt BF: NUCLEAR FACTOR Y, Subunit A (NF-YA) Proteins Positively Regulate Flowering and Act Through FLOWERING LOCUS T. PLOS Genetics 2016, 12(12):e1006496.

20. Zhao H, Wu D, Kong F, Lin K, Zhang H, Li G: The Arabidopsis thaliana Nuclear Factor Y Transcription Factors. Frontiers in plant science 2016, 7:2045.

21. Siefers N, Dang KK, Kumimoto RW, Bynum WEt, Tayrose G, Holt BF, 3rd: Tissue-specific expression patterns of Arabidopsis NF-Y transcription factors suggest potential for extensive combinatorial complexity. Plant physiology 2009, 149(2):625-641.

22. Ben-Naim O, Eshed R, Parnis A, Teper-Bamnolker P, Shalit A, Coupland G, Samach A, Lifschitz E: The CCAAT binding factor can mediate interactions between CONSTANS-like proteins and DNA. The Plant journal : for cell and molecular biology 2006, 46(3):462-476.

23. Cai X, Ballif J, Endo S, Davis E, Liang M, Chen D, DeWald D, Kreps J, Zhu T, Wu Y: A putative CCAAT-binding transcription factor is a regulator of flowering timing in
21 Arabidopsis. *Plant physiology* 2007, 145(1):98-105.

24. Kumimoto RW, Adam L, Hymus GJ, Repetti PP, Reuber TL, Marion CM, Hempel FD, Ratcliffe OJ: The Nuclear Factor Y subunits NF-YB2 and NF-YB3 play additive roles in the promotion of flowering by inductive long-day photoperiods in Arabidopsis. *Planta* 2008, 228(5):709-723.

25. Kumimoto RW, Zhang Y, Siefers N, Holt BF, 3rd: NF-YC3, NF-YC4 and NF-YC9 are required for CONSTANS-mediated, photoperiod-dependent flowering in Arabidopsis thaliana. *The Plant journal : for cell and molecular biology* 2010, 63(3):379-391.

26. Wenkel S, Turck F, Singer K, Gissot L, Le Gourrierec J, Samach A, Coupland G: CONSTANS and the CCAAT Box Binding Complex Share a Functionally Important Domain and Interact to Regulate Flowering of Arabidopsis. *The Plant Cell Online* 2006, 18(11):2971-2984.

27. Hackenberg D, Keetman U, Grimm B: Homologous NF-YC2 Subunit from Arabidopsis and Tobacco Is Activated by Photooxidative Stress and Induces Flowering. *International Journal of Molecular Sciences* 2012, 13(3):3458-3477.

28. Brambilla V, Fornara F: Y flowering? Regulation and activity of CONSTANS and CCT-domain proteins in Arabidopsis and crop species. *Biochimica et biophysica acta Gene regulatory mechanisms* 2017, 1860(5):655-660.

29. Quan S, Niu J, Zhou L, Xu H, Ma L, Qin Y: Identification and characterization of NF-Y gene family in walnut (Juglans regia L.). *BMC Plant Biology* 2018, 18(1).

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

31. Li S, Li K, Ju Z, Cao D, Fu D, Zhu H, Zhu B, Luo Y: Genome-wide analysis of tomato NF-Y factors and their role in fruit ripening. *BMC genomics* 2016, 17:36.
32. Yoshida H, Okada T, Haze K, Yanagi H, Yura T, Negishi M, Mori K: Endoplasmic reticulum stress-induced formation of transcription factor complex ERSF including NF-Y (CBF) and activating transcription factors 6alpha and 6beta that activates the mammalian unfolded protein response. *Molecular and cellular biology* 2001, 21(4):1239-1248.

33. Stephenson TJ, McIntyre CL, Collet C, Xue GP: TaNF-YC11, one of the light-upregulated NF-YC members in Triticum aestivum, is co-regulated with photosynthesis-related genes. *Functional & integrative genomics* 2010, 10(2):265-276.

34. Kusnetsov V, Landsberger M Fau - Meurer J, Meurer J Fau - Oelmuller R, Oelmuller R: The assembly of the CAAT-box binding complex at a photosynthesis gene promoter is regulated by light, cytokinin, and the stage of the plastids. *The Journal of biological chemistry* 1999, 274:36009-36014.

35. Stephenson TJ, McIntyre CL, Collet C, Xue GP: TaNF-YB3 is involved in the regulation of photosynthesis genes in Triticum aestivum. *Functional & integrative genomics* 2011, 11(2):327-340.

36. Lee DK, Kim HI, Jang G, Chung PJ, Jeong JS, Kim YS, Bang SW, Jung H, Choi YD, Kim JK: The NF-YA transcription factor OsNF-YA7 confers drought stress tolerance of rice in an abscisic acid independent manner. *Plant science : an international journal of experimental plant biology* 2015, 241:199-210.

37. Ma X, Zhu X, Li C, Song Y, Zhang W, Xia G, Wang M: Overexpression of wheat NF-YA10 gene regulates the salinity stress response in Arabidopsis thaliana. *Plant physiology and biochemistry : PPB* 2015, 86:34-43.

38. Manimaran P, Venkata Reddy S, Moin M, Raghurami Reddy M, Yugandhar P, Mohanraj SS, Balachandran SM, Kirti PB: Activation-tagging in indica rice identifies a novel transcription factor subunit, NF-YC13 associated with salt tolerance. *Scientific reports* 2017, 7(1):9341.
39. Sato H, Mizoi J, Tanaka H, Maruyama K, Qin F, Osakabe Y, Morimoto K, Ohori T, Kusakabe K, Nagata M et al: 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. *The Plant cell* 2014, 26(12):4954-4973.

40. Filichkin SA, Ansariola M, Fraser VN, Megraw M: Identification of transcription factors from NF-Y, NAC, and SPL families responding to osmotic stress in multiple tomato varieties. *Plant science : an international journal of experimental plant biology* 2018, 274:441-450.

41. Chen M, Zhao Y, Zhuo C, Lu S, Guo Z: Overexpression of a NF-YC transcription factor from bermudagrass confers tolerance to drought and salinity in transgenic rice. *Plant biotechnology journal* 2015, 13(4):482-491.

42. Zhang T, Zhang D, Liu Y, Luo C, Zhou Y, Zhang L: Overexpression of a NF-YB3 transcription factor from *Picea wilsonii* confers tolerance to salinity and drought stress in transformed *Arabidopsis thaliana*. *Plant physiology and biochemistry : PPB* 2015, 94:153-164.

43. Wang Y, Xu W, Chen Z, Han B, Haque ME, Liu A: Gene structure, expression pattern and interaction of Nuclear Factor-Y family in castor bean (*Ricinus communis*). *Planta* 2018, 247(3):559-572.

44. Chu H, Nguyen K, Watanabe Y, Le D, Pham T, Mochida K, Tran L-S: Identification, Structural Characterization and Gene Expression Analysis of Members of the Nuclear Factor-Y Family in Chickpea (*Cicer arietinum L.*) under Dehydration and Abscisic Acid Treatments. *International Journal of Molecular Sciences* 2018, 19(11):3290.

45. Su H, Cao Y, Ku L, Yao W, Cao Y, Ren Z, Dou D, Wang H, Ren Z, Liu H et al: Dual functions of ZmNF-YA3 in photoperiod-dependent flowering and abiotic stress responses in maize. *Journal of experimental botany* 2018, 69(21):5177-5189.
46. Li YJ, Fang Y, Fu YR, Huang JG, Wu CA, Zheng CC: NFYA1 Is Involved in Regulation of Postgermination Growth Arrest Under Salt Stress in Arabidopsis. *PloS one* 2013, 8(4).

47. Ren C, Zhang Z, Wang Y, Li S, Liang Z: Genome-wide identification and characterization of the NF-Y gene family in grape (*vitis vinifera* L.). *BMC genomics* 2016, 17(1):605.

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

49. Zhang X, Zhu S, Zhang K, Wan Y, Liu F, Sun Q, Li Y: Establishment and evaluation of a peanut association panel and analysis of key nutritional traits. *Journal of integrative plant biology* 2018, 60(3):195-215.

50. Bertioli DJ, Cannon SB, Froenicke L, Huang G, Farmer AD, Cannon EK, Liu X, Gao D, Clevenger J, Dash S et al: The genome sequences of *Arachis duranensis* and *Arachis ipaensis*, the diploid ancestors of cultivated peanut. *Nat Genet* 2016, 48(4):438-446.

51. Robledo G, Lavia GI, Seijo G: Species relations among wild *Arachis* species with the A genome as revealed by FISH mapping of rDNA loci and heterochromatin detection. *TAG Theoretical and applied genetics Theoretische und angewandte Genetik* 2009, 118(7):1295-1307.

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

53. Zhuang W, Chen H, Yang M, Wang J, Pandey MK, Zhang C, Chang W-C, Zhang L, Zhang X, Tang R et al: The genome of cultivated peanut provides insight into legume karyotypes, polyploid evolution and crop domestication. *Nature Genetics* 2019, 51(5):865-876.
54. Bertioli DJ, Jenkins J, Clevenger J, Dudchenko O, Gao D, Seijo G, Leal-Bertioli SCM, Ren L, Farmer AD, Pandey MK et al: The genome sequence of segmental allotetraploid peanut Arachis hypogaea. *Nat Genet* 2019, 51(5):877-884.

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

56. Kondrashov FA: Gene duplication as a mechanism of genomic adaptation to a changing environment. *Proceedings Biological sciences* 2012, 279(1749):5048-5057.

57. Clevenger J, Chu Y, Scheffler B, Ozias-Akins P: A Developmental Transcriptome Map for Allotetraploid Arachis hypogaea. *Frontiers in plant science* 2016, 7:1446.

58. Liang M, Yin X, Lin Z, Zheng Q, Liu G, Zhao G: Identification and characterization of NF-Y transcription factor families in Canola (Brassica napus L.). *Planta* 2014, 239(1):107-126.

59. Stephenson TJ, McIntyre CL, Collet C, Xue GP: Genome-wide identification and expression analysis of the NF-Y family of transcription factors in Triticum aestivum. *Plant molecular biology* 2007, 65(1-2):77-92.

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

61. Finn RD, Clements J, Arndt W, Miller BL, Wheeler TJ, Schreiber F, Bateman A, Eddy SR: HMMER web server: 2015 update. *Nucleic acids research* 2015, 43(W1):W30-38.

62. Marchler-Bauer A, Bo Y, Han L, He J, Lanczycki CJ, Lu S, Chitsaz F, Derbyshire MK, Geer RC, Gonzales NR et al: CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic acids research* 2017, 45(D1):D200-D203.

63. Robert X, Gouet P: Deciphering key features in protein structures with the new ENDscript server. *Nucleic acids research* 2014, 42(Web Server issue):W320-324.
Letunic I, Bork P: 20 years of the SMART protein domain annotation resource. *Nucleic acids research* 2018, 46(D1):D493-D496.

Kumar S, Stecher G, Tamura K: MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular biology and evolution* 2016, 33(7):1870-1874.

Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS: MEME SUITE: tools for motif discovery and searching. *Nucleic acids research* 2009, 37(Web Server issue):W202-208.

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

Voorrips RE: MapChart: software for the graphical presentation of linkage maps and QTLs. *The Journal of heredity* 2002, 93(1):77-78.

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

Lescot M, Dehais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouze P, Rombauts S: 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(1):325-327.

Pan CL, Yao SC, Xiong WJ, Luo SZ, Wang YL, Wang AQ, Xiao D, Zhan J, He LF: 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. *Frontiers in physiology* 2017, 8:1037.

Dash S, Cannon EKS, Kalberer SR, Farmer AD, Cannon SB: PeanutBase and Other Bioinformatic Resources for Peanut. 2016:241-252.

Chen C, Xia R, Chen H, He Y: TBtools, a Toolkit for Biologists integrating various HTS-data handling tools with a user-friendly interface. *bioRxiv* 2018:289660.
Table 1. The identification of *AhNF*-Ys in peanut

| Name          | Gene ID      | CHR. | Genomic Location(from) | Genomic Location(to) | Strand | ORF | AA | MW(kDa) |
|---------------|--------------|------|------------------------|----------------------|--------|-----|----|---------|
| AhNF-YA1      | arahy.P5AXXG | 1    | 4955904                | 4959946              | -      | 1020| 33 | 36.81   |
| AhNF-YA2      | arahy.K8VWCH | 1    | 94322773               | 94326097             | +      | 975 | 27 | 24.60   |
| AhNF-YA3      | arahy.K4Y63P | 3    | 45004146               | 45007537             | -      | 639 | 21 | 23.34   |
| AhNF-YA4      | arahy.85VPHV | 3    | 125125172              | 125129125            | -      | 615 | 20 | 22.58   |
| AhNF-YA5      | arahy.VP74QW | 4    | 9362311                | 9370508              | -      | 1269| 42 | 48.03   |
| AhNF-YA6      | arahy.TY3GKM | 8    | 44492626               | 44498160             | -      | 981 | 32 | 36.16   |
| AhNF-YA7      | arahy.HITK0G | 10   | 111351581              | 111356009            | +      | 978 | 32 | 35.49   |
| AhNF-YA8      | arahy.14IBSY | 11   | 6616944                | 6620989             | +      | 1248| 41 | 45.95   |
| AhNF-YA9      | arahy.E0X5DU | 11   | 149088288              | 149091730            | -      | 975 | 22 | 24.63   |
| AhNF-YA10     | arahy.V1T15D | 13   | 47820230               | 47823687             | +      | 639 | 21 | 23.33   |
| AhNF-YA11     | arahy.40UHTZ | 13   | 128836446              | 128840405            | -      | 621 | 20 | 22.75   |
| AhNF-YA12     | arahy.D979GK | 14   | 10847881               | 10861237            | -      | 1479| 49 | 55.45   |
| AhNF-YA13     | arahy.3C4V56 | 18   | 124758993              | 124764258            | -      | 978 | 32 | 35.93   |
| AhNF-YA14     | arahy.RS70VX | 20   | 137968920              | 137973224            | +      | 978 | 32 | 35.51   |
| AhNF-YB1      | arahy.TAFY85 | 1    | 104790079              | 104792569            | +      | 678 | 22 | 25.29   |
| AhNF-YB2      | arahy.M2R7LU | 4    | 6610689                | 6611671             | +      | 522 | 17 | 18.88   |
| AhNF-YB3      | arahy.S3VJK6 | 6    | 24410081               | 24416986            | -      | 867 | 28 | 31.86   |
| AhNF-YB4      | arahy.BTE5MM | 7    | 937559                 | 938146              | -      | 588 | 19 | 20.59   |
| AhNF-YB5      | arahy.SP9TFP | 9    | 103306693              | 103310002           | +      | 543 | 18 | 19.53   |
| AhNF-YB6      | arahy.CHFVOI | 11   | 135761373              | 135764354            | -      | 681 | 22 | 25.44   |
| AhNF-YB7      | arahy.D0EVJ3 | 14   | 8058879                | 806003             | +      | 525 | 17 | 19.15   |
| AhNF-YB8      | arahy.UWY0VP | 16   | 36381530               | 36388018            | -      | 699 | 23 | 25.56   |
| AhNF-YB9      | arahy.8XPX7R | 17   | 2248428                | 2249015             | -      | 588 | 19 | 20.61   |
| AhNF-YB10     | arahy.83ESWU | 19   | 133670457              | 133673816           | +      | 543 | 18 | 19.50   |
| AhNF-YC1      | arahy.0G5T8T | 6    | 4854811                | 4859709             | +      | 1122| 37 | 41.49   |
| AhNF-YC2      | arahy.V7AV3  | 7    | 33598286               | 33601060            | +      | 1077| 35 | 40.58   |
| AhNF-YC3      | arahy.DHT5D0 | 8    | 50513721               | 50516515            | +      | 771 | 25 | 27.81   |
| AhNF-YC4      | arahy.QGX1M1 | 9    | 35397904               | 35400370           | +      | 834 | 27 | 31.45   |
| AhNF-YC5      | arahy.Y83UC8 | 10   | 7695567                | 7696286             | -      | 720 | 23 | 27.43   |
| AhNF-YC6      | arahy.FIU1N  | 14   | 5734442                | 5736074             | +      | 789 | 26 | 29.99   |
| AhNF-YC7      | arahy.WLF6LC | 16   | 17633833               | 17637759            | -      | 807 | 26 | 29.93   |
| AhNF-YC8      | arahy.R55U5A | 18   | 94597401               | 94600161           | +      | 1062| 35 | 39.97   |
| AhNF-YC9      | arahy.HG1CGX | 18   | 134607760              | 134612435        | +      | 918 | 30 | 33.18   |
Table 2. The similarity of amino acid and nucleotide sequences of pairwise homologous AhNF-Y genes.

| Pairwise homologous genes | Similarity (%) |
|---------------------------|----------------|
|                           | A genome       | B genome       | Amino Acid Sequence | Nucleotide Sequence |
| AhNF-YA1                  | AhNF-YA8       | 99.41          | 99.41               |
| AhNF-YA2                  | AhNF-YA9       | 99.55          | 99.11               |
| AhNF-YA3                  | AhNF-YA10      | 95.75          | 97.5                |
| AhNF-YA4                  | AhNF-YA11      | 97.09          | 98.07               |
| AhNF-YA5                  | AhNF-YA12      | 72.67          | 79.93               |
| AhNF-YA6                  | AhNF-YA13      | 97.55          | 98.47               |
| AhNF-YA7                  | AhNF-YA14      | 98.15          | 98.77               |
| AhNF-YB1                  | AhNF-YB6       | 95.13          | 95.89               |
| AhNF-YB2                  | AhNF-YB7       | 95.85          | 96.19               |
| AhNF-YB3                  | AhNF-YB8       | 68.4           | 74.86               |
| AhNF-YB4                  | AhNF-YB9       | 99.49          | 99.32               |
| AhNF-YB5                  | AhNF-YB10      | 99.44          | 99.82               |
| AhNF-YC1                  | AhNF-YC7       | 69.71          | 70.5                |
| AhNF-YC3                  | AhNF-YC9       | 81.97          | 81.7                |

Additional File Legend

Additional files 1 The CDS sequences of validated AhNF-Y genes in peanut. (XLSX 21kb)
Additional file 2 The protein sequences of validated AhNF-Y genes in peanut. (XLSX 14kb)
Additional file 3 Amino acid sequences of the Arabidopsis, rice and human NF-Ys. (DOC
Additional file 4 Fig S1 Gene location of AhNF-Y genes. (PDF 411kb)

Additional file 5 Fig S2 Multiple sequences alignment of NF-Y genes in peanut, rice and Arabidopsis. (PDF 2442kb)

Additional file 6 Analysis and distribution of conserved motifs in peanut NF-Y proteins. (XLSX 12kb)

Additional file 7 FPKM values of 33 AhNF-Y genes in 22 tissues/organisms of peanut from online database. (XLSX 17kb)

Additional file 8 FPKM value of 33 AhNF-Y genes under salt stress obtained from RNA-seq data. (XLSX 13kb)

Additional file 9 Sequences of the primers used in this study. (XLSX 11kb)

Figures
Figure 1

Phylogenetic tree of NF-Y proteins in peanut, rice, Arabidopsis and Homo sapiens. 

Phylogenetic tree was constructed by MEGA7 with maximum likelihood method and bootstrap of 1000 replications. Red, blue and orange indicate the NF-YA, NF-YB, and NF-YC subfamilies, respectively. The bootstrap values are shown on branches.
Figure 2

Phylogenetic relationships, gene structures and motif compositions of AhNF-Y genes. a Unrooted maximum likelihood phylogenetic tree. b Schematic representation of conserved motifs. Colored boxes indicate different conserved motifs. c Exon/intron organization. Exons are shown as yellow boxes, and introns are shown as black lines.
The chromosome distribution and synteny analysis of NF-Y genes in peanut. All chromosomes were drawn in different colours. The approximate location of AhNF-Y genes was shown by short black line on the circle. All synteny blocks were indicated by gray lines, and the segmental duplication events were represented by red lines in A subgenome, and green in B subgenome. The NF-Y orthologs between A and B subgenomes were linked by blue lines.
Prediction of abiotic stress-related regulatory elements in AhNF-Y promoters.

Regulatory elements are represented by round corner rectangles of different color. ABRE: cis-acting element involved in the abscisic acid responsiveness; CGTCA-motif: cis-acting regulatory element involved in the MeJA-responsiveness; MBS: MYB binding site involved in drought-inducibility; TCA-element: cis-acting element involved in salicylic acid responsiveness; TC-rich repeats: cis-acting element involved in defense and stress responsiveness; TGACG-motif: cis-acting regulatory element involved in the MeJA-responsiveness. The upstream length to the initiation codon can be estimated according to the scale per 200 bp at the
Figure 5

Expression profiles of AhNF-Y genes in 22 peanut tissue types. The color round rectangles indicates the log2 values of transcript per million. Blue, yellow, green and red indicated 4 expression profile categories.
Expression profile 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 values of transcript per million. b qRT-PCR profiles of 6 AhNF-Y genes in response to salt stress. Two-week old seedling leaves were sampled at 16 h under a 16-h light/8-h dark cycle. Bars reflect the means ± SD of three replicates. Asterisks indicate the corresponding gene significantly up- or down-regulated compared with untreated control (**P < 0.01 and *P < 0.05, Student’s t-test).
Figure 7

qRT-PCR profiles of 6 AhNF-Y genes in response to mannitol. 14 days-old seedling leaves were sampled at 0, 2, 4, 8, 12 and 24 h under a 16-h light/8-h dark cycle. Bars reflect the means ± SD of three replicates. Asterisks indicate the corresponding gene significantly up- or down-regulated compared with untreated control (**P < 0.01 and *P < 0.05, Student’s t-test).
Figure 8

qRT-PCR profiles of 6 AhNF-Y genes in response to ABA and SA. 14 days-old
seedling leaves were sampled at 0, 1, 2, 4, 6 and 8 h under a 16-h light/8-h dark cycle. Bars reflect the means ± SD of three replicates. Asterisks indicate the corresponding gene significantly up- or down-regulated compared with untreated control (**P < 0.01 and *P < 0.05, Student’s t-test).

**Supplementary Files**

This is a list of supplementary files associated with the primary manuscript. Click to download.

Additional file 1.xlsx
Additional file 2.xlsx
Additional file 3.docx
Additional file 4.pdf
Additional file 5.pdf
Additional file 6.xlsx
Additional file 7.xlsx
Additional file 8.xlsx
Additional file 9.xlsx