

RESEARCH ARTICLE

Open Access

Genome-wide analysis of the *NF-Y* gene family in peach (*Prunus persica* L.)



Miao Li, Guixiang Li, Wei Liu, Xiaomin Dong and Anning Zhang*

Abstract

Background: Nuclear Factor Y (NF-Y) is a heterotrimeric complex composed of three unique subunits: NF-YA, NF-YB, and NF-YC. The NF-Y transcription factor complex binds to the CCAAT box of eukaryotic promoters, playing a vital role in various biological processes in plants. However, the *NF-Y* gene family has not yet been reported from the peach genome. The current study identified and classified candidate peach *NF-Y* genes for further functional analysis of this family.

Results: The current study identified 24 Nuclear Factor Y (NF-Y) transcription factor subunits (6 NF-YA, 12 NF-YB, and 6 NF-YC subunits) in peach. These NF-Y subunits were described with respect to basic physicochemical characteristics, chromosome locations, gene structures, and conserved domains. Based on an analysis of the phylogenetic relationships among peach *NF-Ys*, six pairs of paralogous *NF-Ys* were detected. The expansion of the peach *NF-Y* family occurred by segmental and tandem duplication. Phylogenetic gene synteny of NF-Y proteins was observed between peach and *Arabidopsis*, and five pairs of paralogous NF-Y proteins from peach and *Arabidopsis* were identified. Twenty-four peach *NF-Ys* displayed a diversity of tissue expression patterns. In addition, drought-responsive *cis*-elements were observed in peach *NF-Y* promoters, and 9 peach *NF-Y* genes were shown to distinctly increase their transcript abundances under drought stress.

Conclusions: This study identified 24 *NF-Y* genes in the peach genome and analysed their properties at different levels, providing a foundation for researchers to understand this gene family in peach. The up-regulation of 9 *NF-Y* genes under drought stress indicates that they can serve as candidate functional genes to further study drought resistance in peach.

Keywords: Peach, Nuclear Factor Y, Genomics, Drought stress, Transcription factors

Background

Transcription factors (TFs) play an important role in regulating multiple physiological processes in all living organisms. In general, TFs can be divided into different families by their recognizable conserved domains, such as MYB [17], GRAS [38], NAC [51, 52] and BZIP [44, 45]. Similar protein sequences generally indicate that TF family members have similar biological functions, which facilitates the identification and analysis of unknown proteins in other species. For example, *AtNF-YB1* can enhance resistance in *Arabidopsis* under drought conditions and an orthologous maize *NF-Y* factor, *ZmNF-YB2*, was shown to have an equivalent activity in maize [28]. NUCLEAR FACTOR Y (NF-Y), also called heme-activated protein (HAP) or

CCAAT binding factor (CBF), is a heterotrimeric complex composed of three unique subunits: NF-YA, NF-YB, and NF-YC [25]. In the promoter regions, the NF-Y TF binds to *cis*-elements with the conserved core sequence CCAAT to activate or inhibit the expression of related functional genes in metabolism [25]. Three subfamilies (NF-YA, NF-YB, and NF-YC) can be characterized by their conserved domains and sequence lengths. In general, NF-YA sequences are longer than those of NF-YB and NF-YC. NF-YA has a core region composed of 53 amino acids, which contains two conserved domains, A1 and A2 [27]. NF-YB sequences are shorter than those of NF-YC. The protein structures of NF-YB and NF-YC are similar to those of H2B and H2A histones, respectively [27]. Initially, NF-YB and NF-YC form a dimer in the cytoplasm. This dimer is transferred into the nucleus, where it interacts with NF-YA to complete the assembly of the heterotrimeric complex [39].

* Correspondence: zan_hope@163.com
Shandong Institute of Pomology, Taian, Shandong, China



In recent years, the role of members of the *NF-Y* gene family under drought stress has been a popular research topic. In several species, *NF-Y* gene family members have been identified across their genomes and analysed under drought stress. In *Citrus*, a total of 22 *CsNF-Y* genes have been identified and the candidate gene *CsNF-YA5* was shown to exerted distinct effects on the dehydration tolerance of transgenic tobacco [33]. In castor bean, 25 *RcNF-Y* genes have been identified across the genome and their expression changes were investigated under four types of abiotic stresses (drought, cold, heat and salt stresses) [44, 45]. In chickpea, a total of 40 *CaNF-Y* genes have been identified and some *CaNF-Y* genes were found to be responsive to dehydration and abscisic acid treatments [9].

Peach (*Prunus persica* L.) is one of the most popular fruits worldwide. Because of its small genome size, economic and nutritional importance, and short reproductive cycle [7], peach has become a model tree species for plant physiology and genetics research. With the release of the whole genome sequence of peach [15], it has become convenient to analyse entire gene families. The *NF-Y* gene family plays vital roles in various physiological processes and regulatory networks. However, a detailed analysis of the peach *NF-Y* (designated *PpNF-Y*) gene family has not yet been performed. Considering all these points, the aim of this study was to understand the members of the *NF-Y* gene family in peach and explore the roles of these genes under drought stress. In this study, based on the sequence information derived from public databases, *PpNF-Y* members were identified and classified. Gene duplication was analysed to identify the evolutionary origins of *PpNF-Y* gene family. In addition, the current study analysed the phylogenetic relationships of *NF-Y* proteins in peach and *Arabidopsis* to further elucidate the biological functions of *PpNF-Y* genes. Tissue expression profiles were examined to further investigate the roles of *PpNF-Y* genes during the development of different organs. Moreover, expression analysis during drought stress treatment indicated that some *PpNF-Y* genes responded to drought stress and these genes served as candidate genes for drought resistance in peach.

Methods

Plant materials and drought treatments

Prunus davidiana seeds (Pd) were collected from the Jinniu Mountain Scientific Research base of Shandong Institute of Pomology, Shandong, China, located at 35°38'N and 116°20'E. The seeds were grown in plastic pots (upper diameter 7.0cm, lower diameter 5.0cm, height 7.8cm) filled with mixed soil (vermiculite:humus = 1:1) (without chemical fertilizer application) and were well watered for 30d under greenhouse conditions with

24/18°C day/night, 16/8 h light/dark and 70% relative humidity [32]. In the experiment, there were two groups, a control group and a treated group. Each of them contained 10 peach seedlings. The drought stress treatments were carried out on ten Pd seedlings with similar stem lengths and leaf areas for ten days by withholding water until distinct wrinkling was observed in the top three leaves [29]. The control pot contained another ten similar Pd seedlings that were slightly watered every three days to keep the soil moist. Leaf samples from the treated and control groups were collected, immediately frozen in liquid nitrogen and stored at -70°C [1]. In the tissue expression pattern experiment, samples of five tissues including roots, stems, leaves, flowers and fruits were collected from the cultivar Sunagowase, immediately frozen in liquid nitrogen and stored at -70°C [1].

Identification and analysis of *PpNF-Ys*

Siefers et al. [37] described the conserved regions of the three *NF-Y* subfamilies: the *NF-YA* conserved region contains the amino acid sequence f-V-N-A-K-Q-Y-h-x-I-l-r-R-R-q-x-R-A-k-l-E-a-x-x-K-l-i-k-x- R-K-P-Y-l-H-E-S-R-H-x-H-A-x-r-R-p-R-G-s-G-G-R-E, the *NF-YB* conserved region contains the amino acid sequence r-e-q-D-r-x-L-P-I-A- N-v-x-R-I-M-K-x-x-L-P-x-x-n-x-k-i-s-k-D-A-K-e-t-x-Q-E-C-v-s-E-F-I-S-F-v-T-s-E-A-s-d-k-C-q-x-E-k-R-K-T-I-n-g-d-D-x-L-w-A-m-x-t-L-G-F-x-d-Y-x-e-p-L-x-k-x-Y-x-L-x-k-y-R-e-x-x-e-g-e, and the *NF-YC* conserved region contains the amino acid sequence l-P-l-a-R-I-K-K-I-M- K-x-D-e-D-V-x-m-I-s-a-e-A-P-x-l-f-a-K-A-c-E-M-F-I-x-e-L-T-x-R-s-W-x-h-t-e-e-n-k-R-r-T-l-q-k-x-d-i-a-a-A-v-x-r-x-d-x-x-f-D-F-L-x-x-D-x-V-P, where the uppercase letters represent completely conserved sites, the lowercase letters except x represent relatively conserved sites, and the lowercase x represents non-conserved sites. This study characterized 24 *NF-Y* family members in peach by using the above three conserved regions to blast the peach protein database (version 2.0, <https://www.rosaceae.org/blast/protein/protein>) [16]. The candidate *NF-Y* genes were considered to be *PpNF-Ys*. The online program Conserved Domains (<https://www.ncbi.nlm.nih.gov/Structure/>) was used to ensure the conserved domains of the candidate *NF-Ys* [24]. *PpNF-Y* conserved motifs were identified using the MEME program (<http://www.meme.sdsc.edu/meme/meme.html>) [42].

PpNF-Y sequence structure and genome distribution

The distribution of gene exons and introns was analysed using Gene Structure Display Server 2.0 software (<http://gsds.cbi.pku.edu.cn/index.php>) [14]. The detailed structures of the genes were drawn with Illustrator for Biological Sequences software (Version 1.0.3; <http://ibs.biocuckoo.org/index.php>) [22]. The locations of the *PpNF-Ys* on the genome were collected from a database (*Prunus persica* v2.1)

in JGI (<https://phytozome.jgi.doe.gov/jbrowse/index.html>) [31]. Their genome distribution was displayed using the MapInspect tool (<http://mapinspect.software.informer.com>) [10].

Calculating the parameters of *NF-Y* gene duplication events

The paralogous nucleotide sequences were pairwise aligned by MEGA 5.0 [40]. The parameters K_s (synonymous substitution rate) and K_a (nonsynonymous substitution rate) were calculated by the program DNASP 6.0 [34]. The date (T) of the duplication events was estimated by the formula $T = K_s/2\lambda$, where λ represents the estimated clock-like rate of synonymous substitution, which was 1.5×10^{-8} substitutions/synonymous site/year in dicots [4, 6].

PpNF-Y sequence alignment and phylogenetic analysis

The PpNF-Y protein sequence alignments were analysed using MEGA 5.0 software (<https://www.megasoftware.net>) [40] and the conserved domains were marked using GeneDoc 2.7.000 software [30]. A phylogenetic tree of PpNF-Y family proteins was constructed using the neighbour-joining (NJ) method [35] of MEGA 5.0 software with a bootstrap test ($n=1000$).

NF-Y protein sequences of *Arabidopsis*

The *Arabidopsis* *NF-Y* gene family was identified and numbered in a previous study [37]. The gene sequences and the corresponding protein sequences were collected by searching the gene names such as *NF-YA1* from the *Arabidopsis* Information Resource (TAIR) database (<https://www.arabidopsis.org/>) [3]. *Arabidopsis* *NF-Y* proteins were designated AtNF-Ys.

Drought-responsive *cis*-elements of the *PpNF-Y* promoter

The promoter sequences (length, 1.5 kb) of *PpNF-Y*s were collected from the Genome Database for Rosaceae (<https://www.rosaceae.org/>). Drought-responsive *cis*-elements were analysed in the online software PlantCARE (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>).

The models of *cis*-elements in the promoters were made with Illustratior for Biological Sequences software [22].

RNA isolation and expression analysis

Total RNA from different tissues of *Prunus davidiana* (Pd) was extracted using TRIzol® reagent (Invitrogen Inc., USA). According to the supplier's protocol, the extracted RNA was used as a template with M-MLV reverse transcriptase (Toyobo, Japan) to synthesize complementary DNA (cDNA). Real-time PCR was performed using a 7500 Fast Real Time PCR System (Applied Biosystems, NY, USA). One microlitre cDNA template, 10 μ L SYBR Premix Ex

Taq (Takara, Kyoto, Japan), 0.8 μ L gene-specific primers, and 8.2 μ L ddH₂O were mixed to compose 20 μ L reaction systems. The PCR thermal cycle was as follows: 95°C for 30 s; 40 cycles of 95°C for 5 s, 52°C for 30 s, and 72°C for 20 s. *Translation elongation factor 2* (*TEF2*) was used as an internal control to normalize gene expression [2, 43]. Each sample was analysed with three biological replicates.

Statistical analysis

All data were subjected to analysis of variance according to Student's *t* test using SPSS statistical software 17.0 (SPSS Inc., USA) [21].

Result

Identification of PpNF-Y family members

After removing different transcripts of the same gene, a total of 24 non-redundant protein sequences (6 PpNF-YAs, 12 PpNF-YBs and 6 PpNF-YCs) representing the primary transcript were identified. All 24 PpNF-Ys contained one highly conserved domain of the three *NF-Y* subfamilies that were also confirmed by the online database of Conserved Domains. To distinguish these newly identified genes, the current study renamed these genes based on subfamily branch and chromosomal distribution (*PpNF-YA1-PpNF-YA6*, *PpNF-YB1-PpNF-YB12* and *PpNF-YC1-PpNF-YC6*). The protein lengths of the 24 PpNF-Ys ranged from 121 AA to 398 AA (Table 1), showing a wide distribution of PpNF-Y lengths. Among the three subfamilies of PpNF-Y, the PpNF-YA lengths (range from 202 AA to 398 AA, average length 322.5 AA) were generally longest, the PpNF-YC lengths (range from 121 AA to 280 AA, average length 228.5 AA) were shorter, and the PpNF-YB lengths (range from 167 AA to 254 AA, average length 203.2 AA) were shortest. The predicted molecular weights (M_w) of the 24 PpNF-Ys ranged from 13.26 (PpNF-YC4) to 44.05 (PpNF-YA4), and the predicted theoretical isoelectric points (pI) of the 24 PpNF-Ys ranged from 5.27 (PpNF-YB2) to 9.43 (PpNF-YA5).

Phylogenetic relationships, genome distribution and gene structure of PpNF-Ys

To investigate the phylogenetic relationships among the *PpNF-Y*s, three phylogenetic trees were constructed based on an alignment of the *PpNF-Y* nucleotide sequences (Additional file 1) using MEGA 5.0 (Fig. 1). The neighbour-joining (NJ) phylogenetic trees were constructed to show the structural classification of the *PpNF-Y*s. As shown in Fig. 1, the NJ tree was distinctly divided into three subgroups, marked by three different colours (*PpNF-Y*As, red; *PpNF-Y*Bs, green; *PpNF-Y*Cs, blue), which were consistent with our identification results. The *PpNF-YA* branch had a simple structure, and a pair of paralogous *PpNF-Y*As (*PpNF-YA2* and 6) was detected. In the *PpNF-YB* branch, there were four major groups. A group containing *PpNF-YB2* and 10

Table 1 Information of the *NF-Y* genes family in peach

| Name | Gene | Chromosome localization | CDS | Length (AA) | Mw (kD) | PI | Homologs in <i>Arabidopsis</i> | Functional annotation in <i>Arabidopsis</i> |
|------------------|-----------------------|--------------------------|------|-------------|---------|------|--------------------------------|---|
| <i>PpNF-YA1</i> | <i>Prupe.2G014500</i> | Pp02:1285333..1293724+ | 1035 | 344 | 37.66 | 8.32 | <i>AtNF-YA1,7</i> | Salt Stress [21] |
| <i>PpNF-YA2</i> | <i>Prupe.3G043900</i> | Pp03:3121596..3127378+ | 1032 | 343 | 38.00 | 9.20 | <i>AtNF-YA3,8</i> | Embryo development [48] |
| <i>PpNF-YA3</i> | <i>Prupe.4G090200</i> | Pp04:4494268..4499893- | 609 | 202 | 22.13 | 7.90 | <i>AtNF-YA4,7</i> | Flower [11] |
| <i>PpNF-YA4</i> | <i>Prupe.4G249900</i> | Pp04:16710014..16714553+ | 1197 | 398 | 44.05 | 5.89 | <i>AtNF-YA1,9</i> | Seed development [50] |
| <i>PpNF-YA5</i> | <i>Prupe.7G093100</i> | Pp07:12566137..12572248+ | 966 | 321 | 35.35 | 9.43 | <i>AtNF-YA2,10</i> | Seed germination [41] |
| <i>PpNF-YA6</i> | <i>Prupe.7G203500</i> | Pp07:18875767..18879196+ | 984 | 327 | 36.37 | 8.81 | <i>AtNF-YA3,8</i> | Embryo development [48] |
| <i>PpNF-YB1</i> | <i>Prupe.1G138500</i> | Pp01:10829440..10830284+ | 492 | 163 | 18.55 | 5.62 | <i>AtNF-YB5,7</i> | Unknown |
| <i>PpNF-YB2</i> | <i>Prupe.2G219500</i> | Pp02:24828695..24832311+ | 543 | 180 | 19.19 | 5.27 | <i>AtNF-YB1</i> | Drought stress [25] |
| <i>PpNF-YB3</i> | <i>Prupe.3G197300</i> | Pp03:20770866..20771306- | 546 | 181 | 19.11 | 5.38 | <i>AtNF-YB3</i> | Flowering, root growth [47] |
| <i>PpNF-YB4</i> | <i>Prupe.4G003800</i> | Pp04:258372..261155- | 738 | 245 | 26.87 | 6.29 | <i>AtNF-YB6</i> | Embryonic desiccation-intolerance [18] |
| <i>PpNF-YB5</i> | <i>Prupe.4G004200</i> | Pp04:271278..272705- | 765 | 254 | 29.78 | 9.00 | <i>AtNF-YB3</i> | Flowering, root growth [47] |
| <i>PpNF-YB6</i> | <i>Prupe.4G242700</i> | Pp04:16019863..16021262+ | 600 | 199 | 21.41 | 6.66 | <i>AtNF-YB3</i> | Flowering, root growth [47] |
| <i>PpNF-YB7</i> | <i>Prupe.4G243500</i> | Pp04:16087429..16090365- | 750 | 249 | 26.72 | 6.44 | <i>AtNF-YB9</i> | Embryonic desiccation-intolerance [18] |
| <i>PpNF-YB8</i> | <i>Prupe.4G273800</i> | Pp04:22774061..22774753- | 693 | 230 | 25.92 | 6.20 | <i>AtNF-YB6</i> | Embryonic desiccation-intolerance [18] |
| <i>PpNF-YB9</i> | <i>Prupe.6G175100</i> | Pp06:17959164..17964486- | 651 | 216 | 24.17 | 5.77 | <i>AtNF-YB3</i> | Flowering, root growth [47] |
| <i>PpNF-YB10</i> | <i>Prupe.6G241800</i> | Pp06:24020199..24022272- | 522 | 173 | 18.69 | 6.15 | <i>AtNF-YB8,10</i> | Unknown |
| <i>PpNF-YB11</i> | <i>Prupe.6G323400</i> | Pp06:28504283..28505052+ | 519 | 172 | 19.11 | 5.41 | <i>AtNF-YB4,5</i> | Unknown |
| <i>PpNF-YB12</i> | <i>Prupe.8G256000</i> | Pp08:21852595..21853514+ | 531 | 176 | 19.91 | 6.66 | <i>AtNF-YB5</i> | Unknown |
| <i>PpNF-YC1</i> | <i>Prupe.1G152500</i> | Pp01:12042456..12046186+ | 783 | 260 | 28.87 | 5.89 | <i>AtNF-YC9</i> | Flowering [37] |
| <i>PpNF-YC2</i> | <i>Prupe.3G149300</i> | Pp03:16308590..16312026+ | 636 | 211 | 24.53 | 5.37 | <i>AtNF-YC2,3</i> | Flowering [37] |
| <i>PpNF-YC3</i> | <i>Prupe.4G132200</i> | Pp04:7318206..7321121+ | 843 | 280 | 31.40 | 5.76 | <i>AtNF-YC2,3</i> | Flowering [37] |
| <i>PpNF-YC4</i> | <i>Prupe.5G096000</i> | Pp05:10523596..10524073+ | 366 | 121 | 13.26 | 7.82 | <i>AtNF-YC3,9</i> | Flowering [37] |
| <i>PpNF-YC5</i> | <i>Prupe.5G150400</i> | Pp05:13652222..13654960+ | 714 | 237 | 26.04 | 5.31 | <i>AtNF-YC4</i> | Floral induction [45] |
| <i>PpNF-YC6</i> | <i>Prupe.6G328500</i> | Pp06:28759728..28761956- | 789 | 262 | 29.16 | 5.96 | <i>AtNF-YC3,9</i> | Flowering [37] |

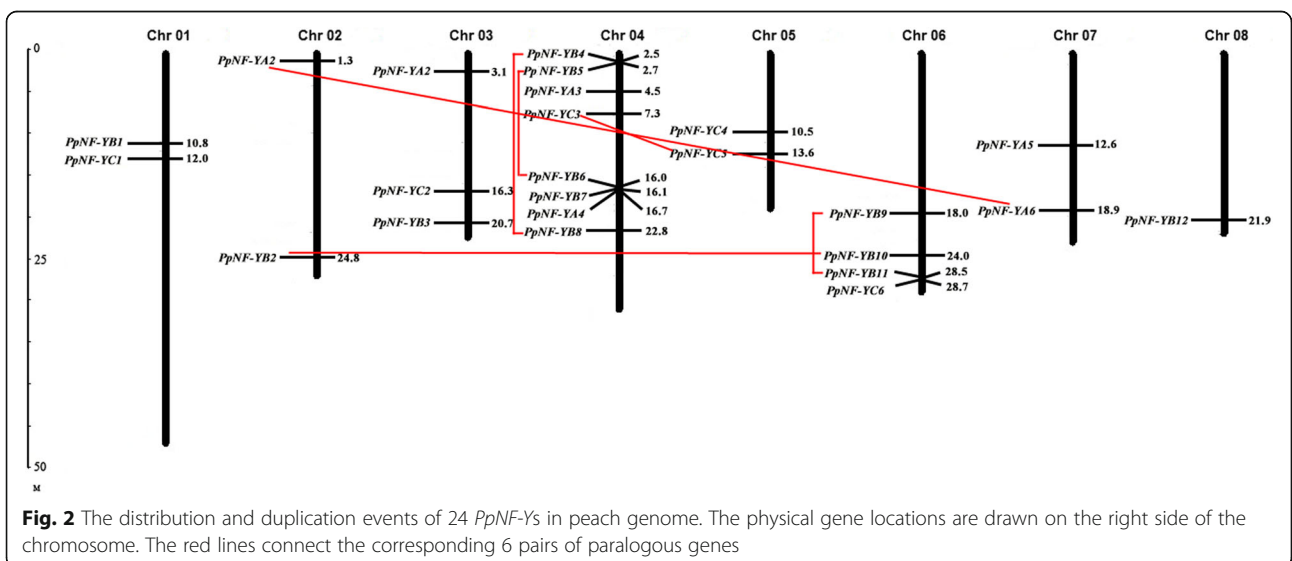
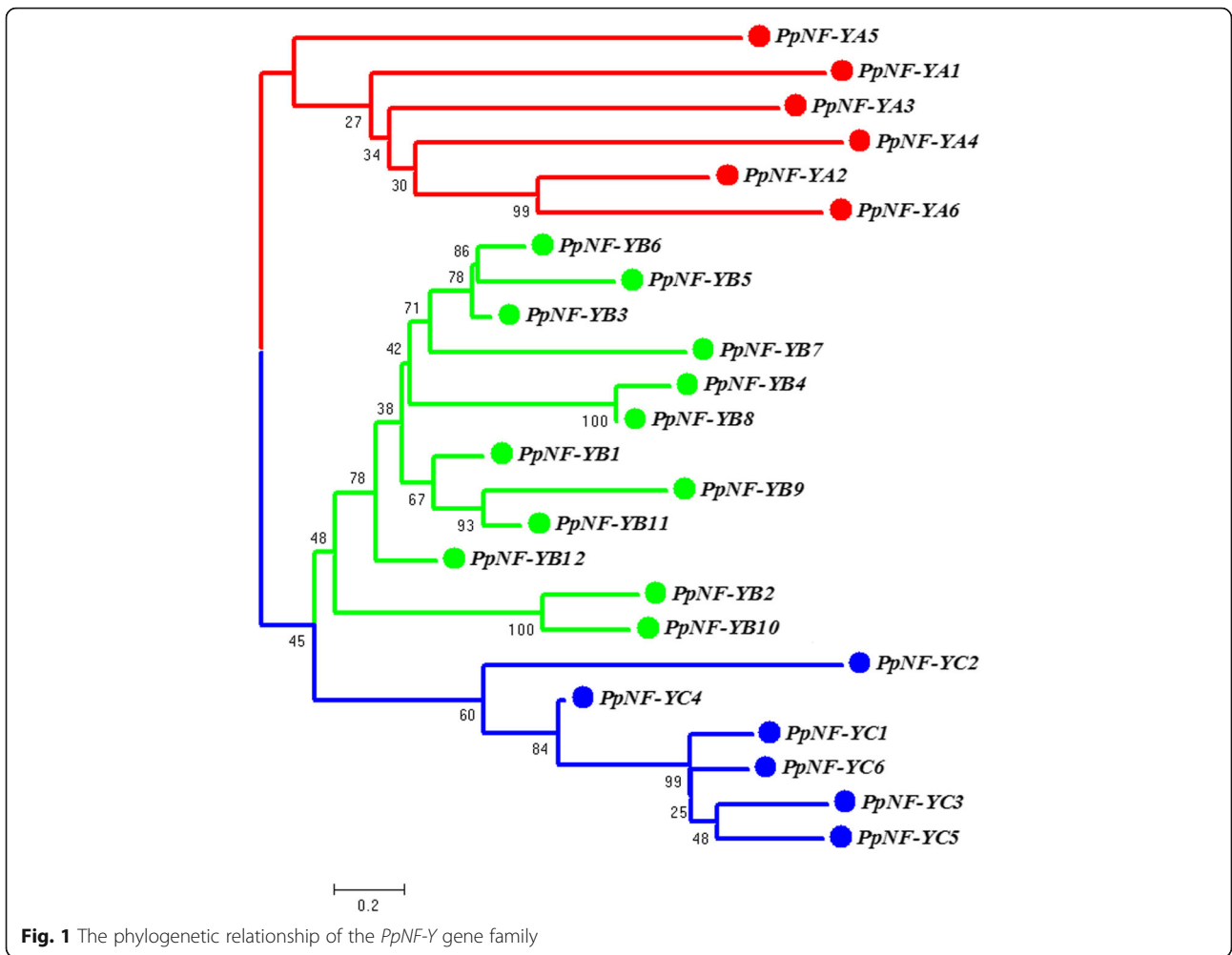
was located outside the three other groups. The presence of a bifurcation put *PpNF-YB12* outside a secondary subgroup formed by *PpNF-YB1*, 9 and 11. The third group was composed of *PpNF-YB4* and 8. In the fourth group, there were also two bifurcations separating *PpNF-YB3* and 7 from a second subgroup composed of *PpNF-YB5* and 6. The branch structure of the *PpNF-YCs* was similar to that of the *PpNF-YAs*. A pair of paralogous *PpNF-YCs* (*PpNF-YC3* and 5) was located on the innermost side, and the remaining *PpNF-YCs* were located around the outside.

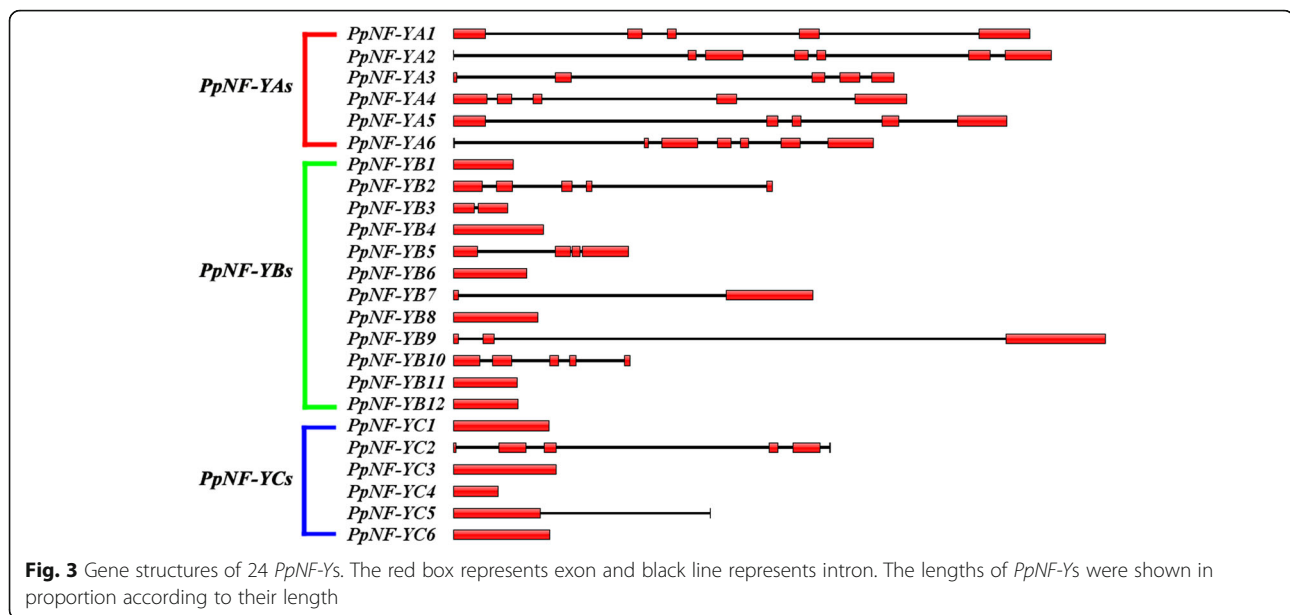
Twenty-four *PpNF-Ys* were distributed on all eight chromosomes (Fig. 2). Most *PpNF-Ys* with eight members were distributed on the 4th chromosome. The fewest of *PpNF-Ys* were located on the 8th chromosome, with only one member. The other six chromosomes contained 2-4 *PpNF-Ys*. The exons and introns of the *PpNF-Ys* were drawn with Illustrator for Biological Sequences (IBS, version 1.0.3) software using the genomic DNA sequences of the *PpNF-Ys* and the corresponding coding sequence (Fig. 3). All *PpNF-YAs* contained 4-6 introns. Some *PpNF-YBs* (*PpNF-YB1*, 4,

6, 8, 11, 12) and *PpNF-YCs* (*PpNF-YC1*, 3, 4, 6) were composed of only one exon.

Duplication events of the *NF-Y* genes in the peach genome

In the phylogenetic analysis of the *NF-Y* gene family, 6 pairs of paralogous *NF-Y* genes were detected (Fig. 2). Among them, 3 pairs (*PpNF-YA2* and *PpNF-YA6*, *PpNF-YB2* and *PpNF-YB10*, *PpNF-YC3* and *PpNF-YC5*) were randomly scattered on different chromosomes, and the other three pairs (*PpNF-YB5* and *PpNF-YB6*, *PpNF-YB4* and *PpNF-YB8*, *PpNF-YB9* and *PpNF-YB11*) were located on the same chromosomes. To further trace the dates of the duplication events, the parameters *Ks* and *Ka* and the *Ka/Ks* ratio were estimated using DNASP 6.0 software [34]. The *Ka/Ks* ratios of *PpNF-YB9* and *PpNF-YB11* were greater than 1 and those of the other five pairs of paralogous *PpNF-Ys* were less than 1. The approximate dates of the duplication events are shown in Table 2. The origin dates of the three pairs of paralogous *PpNF-Ys* on





the different chromosomes ranged from 20.6 to 59.67 million years ago. The dates of the other three pairs on the same chromosomes ranged from 13.33 to 68.33 million years ago.

Conserved regions of PpNF-Ys

To further investigate the conserved regions of the three subfamilies in peach, multiple protein sequence alignments of the PpNF-YAs, PpNF-YBs, and PpNF-YCs were analysed using MEGA 5.0. All three subfamilies contained conserved regions. Multiple alignment of the PpNF-YA proteins indicated that there was a conserved region composed of approximately 50 amino acids (Fig. 4a). As with the PpNF-YAs, two conserved regions of PpNF-YB (approximately 102 amino acids) and PpNF-YC (approximately 81 amino acids) were identified by protein sequence alignment (Fig. 4b and c).

Table 2 The parameters and dates of the duplication events in the paralogous *PpNF-Ys*.

| The paralogous gene | | KS | KA | Ka/Ks ratio | Date (million year ago) |
|---------------------|------------------|------|------|-------------|-------------------------|
| <i>PpNF-YA2</i> | <i>PpNF-YA6</i> | 0.66 | 0.55 | 0.83 | 22.00 |
| <i>PpNF-YB5</i> | <i>PpNF-YB6</i> | 1.08 | 0.21 | 0.19 | 36.00 |
| <i>PpNF-YB4</i> | <i>PpNF-YB8</i> | 2.05 | 0.43 | 0.21 | 68.33 |
| <i>PpNF-YB9</i> | <i>PpNF-YB11</i> | 0.4 | 0.55 | 1.38 | 13.33 |
| <i>PpNF-YB2</i> | <i>PpNF-YB10</i> | 0.62 | 0.28 | 0.45 | 20.67 |
| <i>PpNF-YC3</i> | <i>PpNF-YC5</i> | 1.79 | 0.22 | 0.12 | 59.67 |

Ks and Ka were calculated using the program DNASP 6.0 [34]. The date of the duplication events in the paralogous *PpNF-Ys* were analyzed by the formula $T = Ks/2\lambda$, λ representing the estimated clock-like rates of synonymous substitution, which was 1.5×10^{-8} substitutions/synonymous site/year in dicots [4]

To reveal the putative motifs of the *NF-Y* family in peach, 24 *PpNF-Ys* were analysed using the program MEME. All members contained three distinct motifs (Fig. 5), which was consistent with the fact that they belonged to the same gene family.

Expression patterns of PpNF-Ys in different peach organs

The expression patterns of *PpNF-Ys* in the five organs (roots, stems, leaves, flowers and fruits) were analysed (Fig. 6). The *PpNF-Ys* showed different expression patterns. Some were constitutively expressed in every organ such as *PpNF-YB5*, while others showed specific high expression in one or two peach organs, such as *PpNF-YA4* (leaves and fruits), *PpNF-YB6* (flowers) and *PpNF-YB4* (stems). This diversity of expression patterns of *PpNF-Ys* suggested a divergence in the biological functions of *PpNF-Ys* during peach growth and development.

Evolutionary relationships of the *NF-Y* family in peach and *Arabidopsis*

The functions of some *AtNF-Y* family members have been identified in *Arabidopsis*. However, the biological functions of PpNF-Y proteins are unknown. To predict the functions of PpNF-Y members, we explored the phylogenetic relationships of the *NF-Y* members from *Arabidopsis* and peach using 60 *NF-Y* protein sequences (Fig. 7; Additional files 2 and 3). The NJ phylogenetic tree showed that 58 *NF-Y* members (excluding *AtNF-YB11* and *AtNF-YC11*) could be classified into three main groups distinguished by three different colours (red, green and blue), which were consistent with the subfamily classifications of the PpNF-Ys. In each group, we found that some pairs of paralogous *NF-Y* proteins were composed of one PpNF-Y

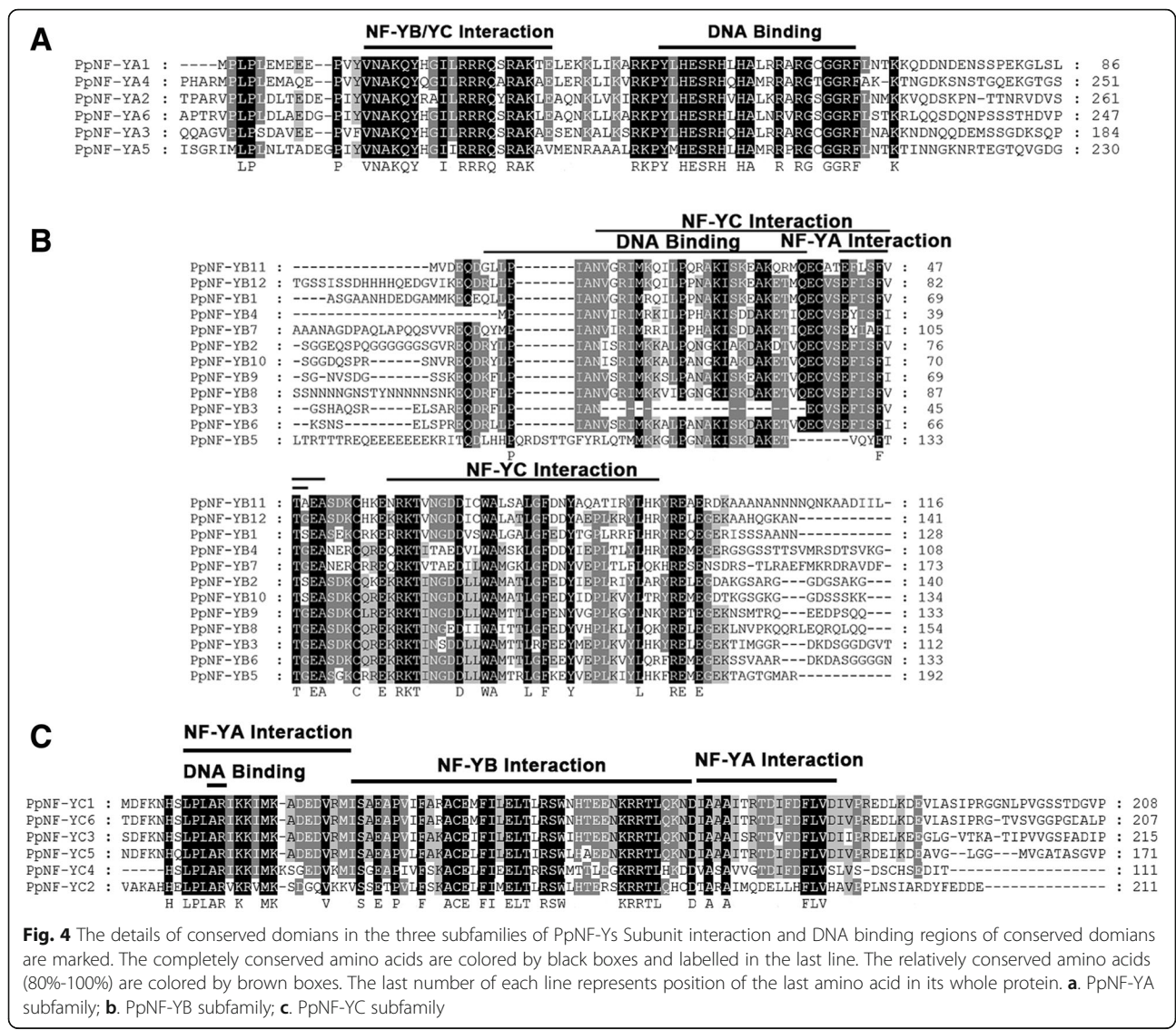


Fig. 4 The details of conserved domains in the three subfamilies of PpNF-Ys Subunit interaction and DNA binding regions of conserved domains are marked. The completely conserved amino acids are colored by black boxes and labelled in the last line. The relatively conserved amino acids (80%-100%) are colored by brown boxes. The last number of each line represents position of the last amino acid in its whole protein. **a.** PpNF-YA subfamily; **b.** PpNF-YB subfamily; **c.** PpNF-YC subfamily

and one AtNF-Y, such as PpNF-YA4 and AtNF-YA9, and this close evolutionary relationship generally suggested the similarity of their biological functions. Both the NF-YA and NF-YC groups contained one pair of paralogous NF-Y proteins, and the NF-YB group contained three pairs.

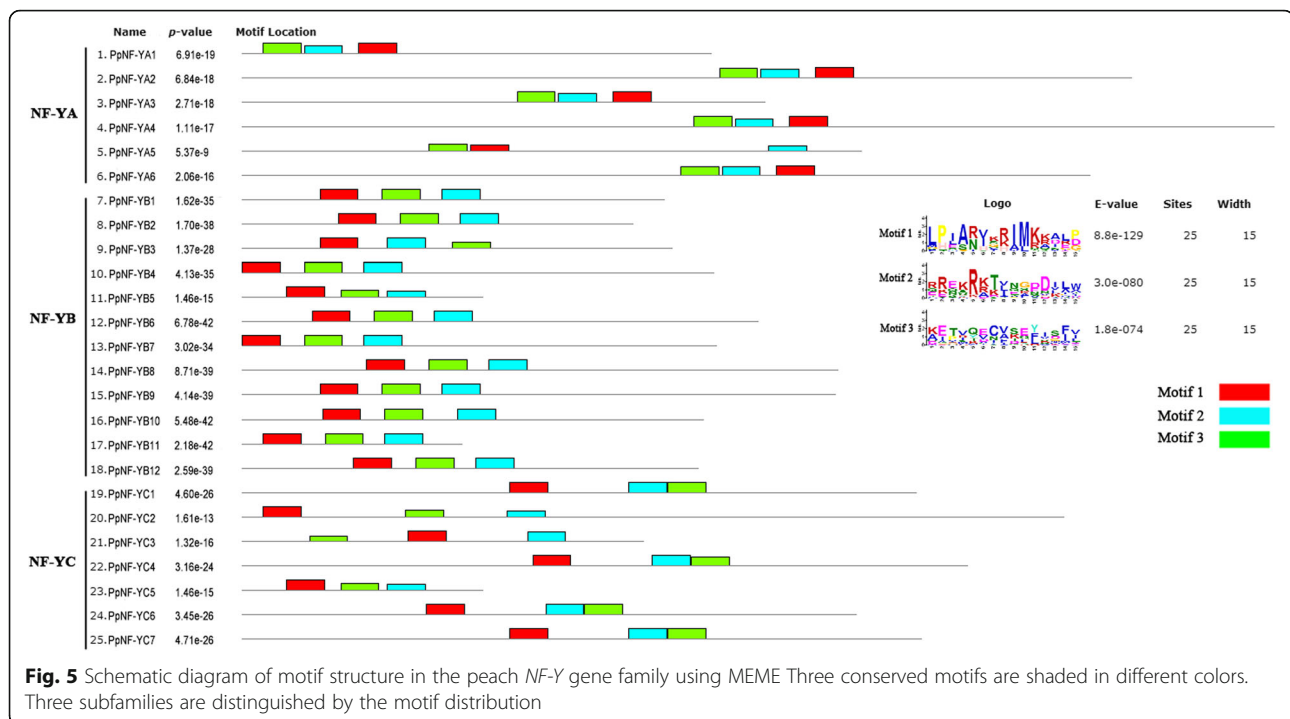
Analysis of five drought-responsive cis-elements in the promoter sequences of PpNF-Y genes

To explore the involvement of the PpNF-Y genes in drought tolerance, their promoter sequences were analysed using PlantCARE software [19]. It is generally known that five cis-elements, ABREs, MBSs, G-boxes, W-boxes and DREs, respond to drought-induced signalling and regulation of downstream gene expression [18, 26, 36]. All five cis-elements in the promoter regions of the PpNF-Y gene family members are shown in Fig. 8. The results indicated that all

24 PpNF-Y promoter regions contained one or more drought-responsive cis-elements. ABREs, MBSs, G-boxes, W-boxes and DREs were distributed within 17, 9, 8, 10 and 7 PpNF-Y promoter regions, respectively. The numbers of drought-responsive cis-elements in the 24 PpNF-Y promoter regions ranged from 1 (PpNF-YC4) to 8 (PpNF-YB1), and the types ranged from 1 (PpNF-YC4) to 4 (PpNF-YB1).

Expression analysis of PpNF-Ys under drought stress

Nelson et al. [28], Chen et al. [8] and Pereira et al. [33] have reported that the NF-Y gene family is closely related to drought tolerance in some species, including Arabidopsis, Citrus, maize, and Bermuda grass. In peach, it is possible that one or more PpNF-Ys are involved in tolerance to drought stress. To investigate drought-responsive PpNF-Ys, an expression analysis of the 24 PpNF-Ys from Pd seedlings under drought



stress was performed (Fig. 9; Additional file 4). The results showed that 4 *PpNF-YAs* (*PpNF-YA3*, 4, 5, and 6), 4 *PpNF-YBs* (*PpNF-YB2*, 6, 7, and 12) and *PpNF-YC4* were upregulated under drought stress. Among these members, the expression level of *PpNF-YA5* increased the most, to approximately fifteen times that of the control sample. The upregulated gene *PpNF-YA4* and *PpNF-YB7* showed the smallest increases in transcript abundance, less than two times that of the control. The increased expression of the other six upregulated *PpNF-Ys* ranged from two to eight times that of the control.

Discussion

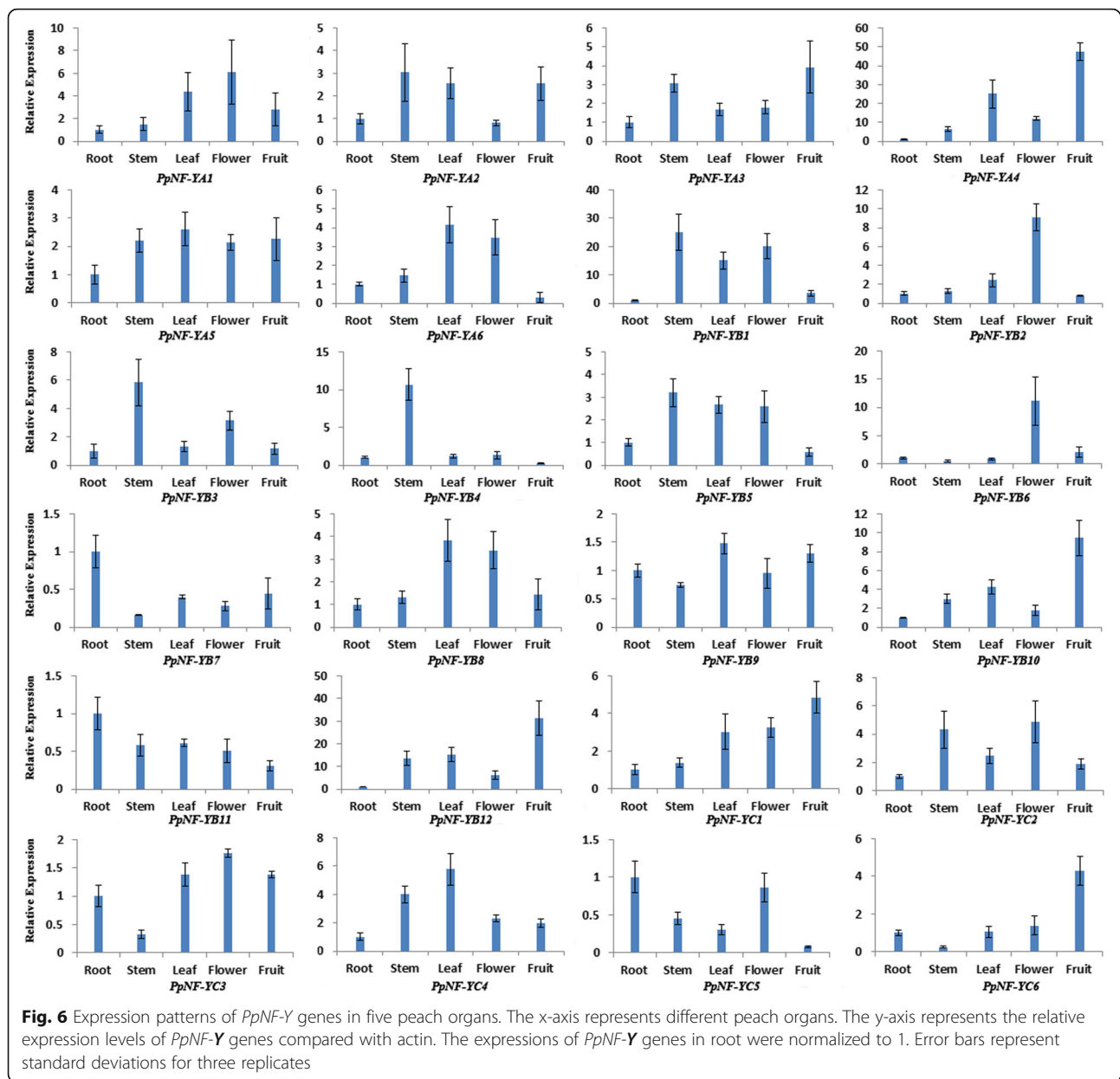
Peach contains only 24 diverse *NF-Ys*

The *NF-Y* TFs are found in all sequenced eukaryotes. This gene family has been studied in some higher plant species, such as maize, rice, tomato, and banana [20, 46, 47, 51, 52]. Mantovani [23] has revealed some properties of this family, including relatively conserved binding and interaction domains, three types of subunits, and diverse biological functions, which were beneficial to our research on the *NF-Y* gene family in peach. A total of 24 *NF-Y* members were identified in peach, fewer in *Arabidopsis* (36 *NF-Y* members), tomato (59 *NF-Y* members), rice (34 *NF-Y* members), and wheat (39 *NF-Y* members). The small number of peach *NF-Y* members may be associated with the small size of the peach genome. The 24 *PpNF-Ys* demonstrated diversity in several aspects, including protein length, gene structure, molecular

weight, theoretical isoelectric point and expression pattern, which suggested a diversity of biological functions among *PpNF-Ys*.

Putative segmental and tandem duplication events in the peach genome

The paralogous genes distributed on different chromosomes are generally designated as segmental duplication events, and those co-located on the same chromosome are considered tandem duplication events [6]. In this study, the locations of 6 paralogous *NF-Y* genes were either on the same chromosome or on different chromosomes, implying that both tandem and segmental duplication events contributed to the expansion of the *NF-Y* genes in peach. When the Ka/Ks ratios of the five paralogous *PpNF-Ys* were less than 1, purifying selection of duplication events occurred and the corresponding paralogous *PpNF-Y* proteins were considered to be unchanged. In contrast, the Ka/Ks ratios of *PpNF-YB9* and *PpNF-YB11* were greater than 1, indicating that distinct variation between the *PpNF-YB9* and *PpNF-YB11* proteins occurred during the duplication event. As the initial point of peach evolution was still unclear, we could not determine whether the duplication events in this gene family predated the formation of the peach species. However, the oldest date of the duplication events among the 6 pairs of paralogous *PpNF-Ys* could be 68.33 million years ago, suggesting that this is an ancient gene family. At a minimum, it could indicate traces of peach evolution.

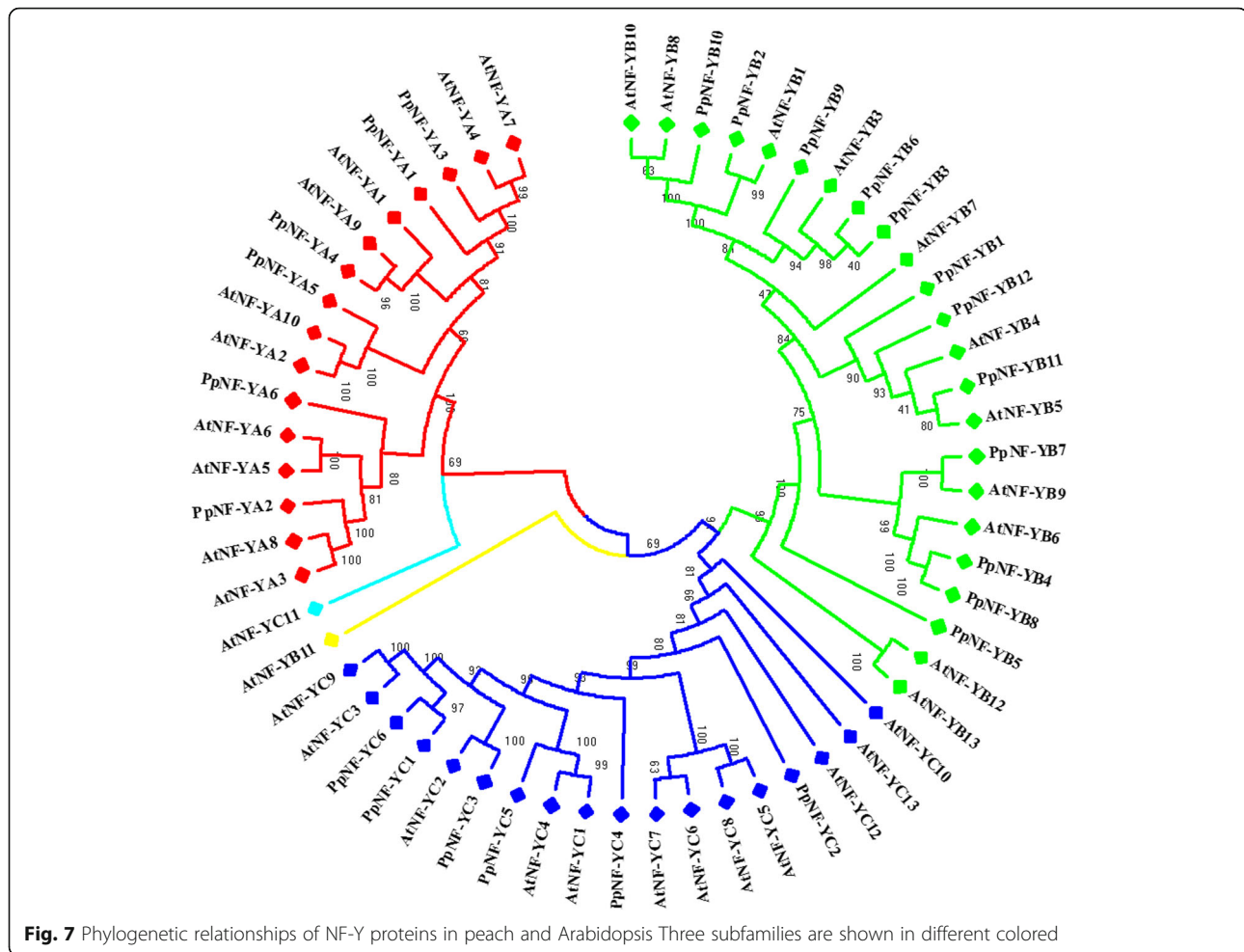


Potential conserved domains in three NF-Y subfamilies

In general, proteins can be characterized and classified by their conserved regions, which play vital roles in heterodimerization, heterotrimerization, and DNA interactions at *CCAAT* sites (Zambelli and Pavesi 2017). Previous studies have shown that all three subfamilies (NF-YA, NF-YB and NF-YC) are recognized by interaction domains that interact with NF-Y members and DNA-binding domains that bind to downstream targeted *CCAAT* sites [25]. Based on the recognized DNA-binding domain of NF-YA in plants, mammals and yeasts, the similar 17-amino-acid sequence Y-L-H-E-... -G-G-R-F in the C-terminus of the PpNF-YA conserved region was considered to interact with DNA at *CCAAT* sites

(Fig. 4a). It was inferred that the 21-amino-acid sequence Y-V-N-A- ... -A-K-L-E in the N-terminus of the PpNF-YA conserved region interacted with the other two subunits (PpNF-YB and PpNF-YC) (Fig. 4a).

The structure and amino acid composition of the NF-YB conserved regions were similar to those of H2B histone fold motifs [12]. Based on the conserved regions of NF-YB members in *Arabidopsis*, a 31-amino-acid sequence R-x-L-P-... -E-T-x-Q was considered to be the DNA-binding domain of PpNF-YB. Similarly, two core regions, the 40-amino-acid sequence A-N-V-x-... -T-x-E-A and the 32-amino acid sequence x-R-K-T-... Y-L-x-x, were considered to interact with the other two subunits (PpNF-YA and PpNF-YC) (Fig. 4b).



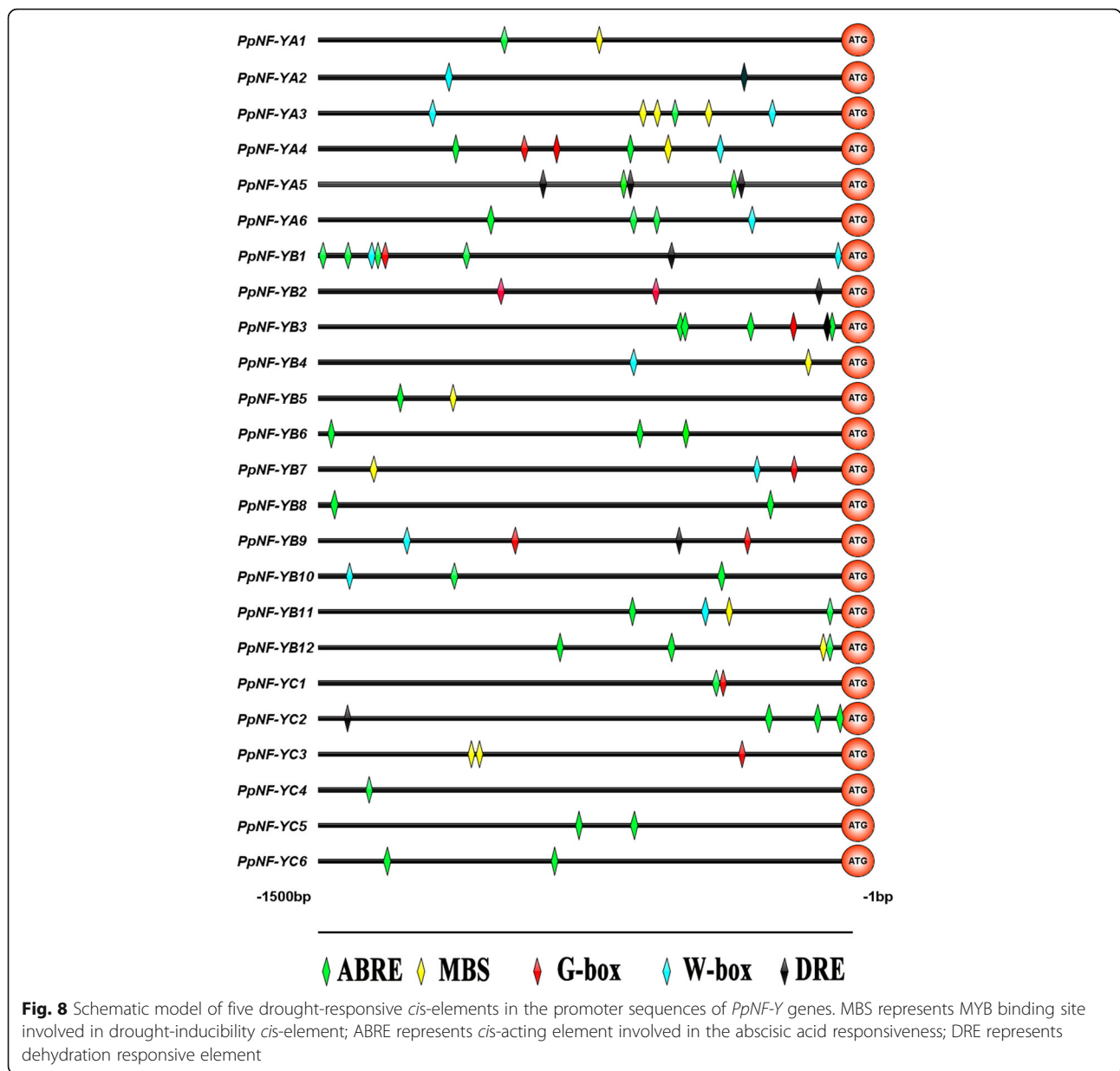
In the alignment analysis of 7 PpNF-YC members, the conserved 74-amino-acid sequence L-P-L-A... -D-F-L-V had similarities to an *Arabidopsis* subunit interaction domain (Fig. 4c). Thus, we inferred that this fragment was the core region by which PpNF-YC associated PpNF-YA/PpNF-YB. In addition, the DNA-binding domain of PpNF-YC was composed of only two residues in series “A” and “R”, which were necessary for the formation of a complex between heterotrimeric NF-Y and DNA.

Similar core protein regions between PpNF-Ys and AtNF-Ys imply similar biological functions

Based on the analysis of the protein alignment, the conserved core amino acids of the PpNF-YA interaction and DNA-binding domains were almost highly consistent with those of AtNF-YAs, RnCFB-B (Rn, *Rattus norvegicus*), SchAP2 (Sc, *Saccharomyces cerevisiae*), indicating that there may be similar biological functions of several NF-YA members among plants, mammals and yeasts. For example, most AtNF-YAs increased their expression under drought stress in *Arabidopsis* [13], and two-thirds

of PpNF-YAs were also upregulated under drought stress in this study.

Boulard et al. [5] showed that the substitution of a required amino acid residue (from Lys to Asp) caused LEC1 (AtNF-YB9) and LEC1-like (AtNF-YB6) to differ from plant and animal NF-YB proteins, and this variation caused LEC1 to fail to rescue the *lec1* embryonic desiccation-intolerant phenotype. The alternating distribution of PpNF-Y and AtNF-Y proteins in the NJ tree implied that some NF-Y members from these two species might originate from common ancestors, which might be reflected in their protein sequences. For example, we found that two PpNF-YB proteins, PpNF-YB4 and PpNF-YB7, were closely related to AtNF-YB9 and AtNF-YB6, and these four members together formed a subgroup in the NF-YB branch, which might suggest a similar biological function for them. Specifically, a substitution of required amino-acid residues (from Lys to Asp; the column marked by *) emerged in PpNF-YB4 and PpNF-YB7 (Fig. 4b), indicating that these two PpNF-YB members may have similar biological functions to those of AtNF-YB9 and AtNF-YB6. In *Arabidopsis*,

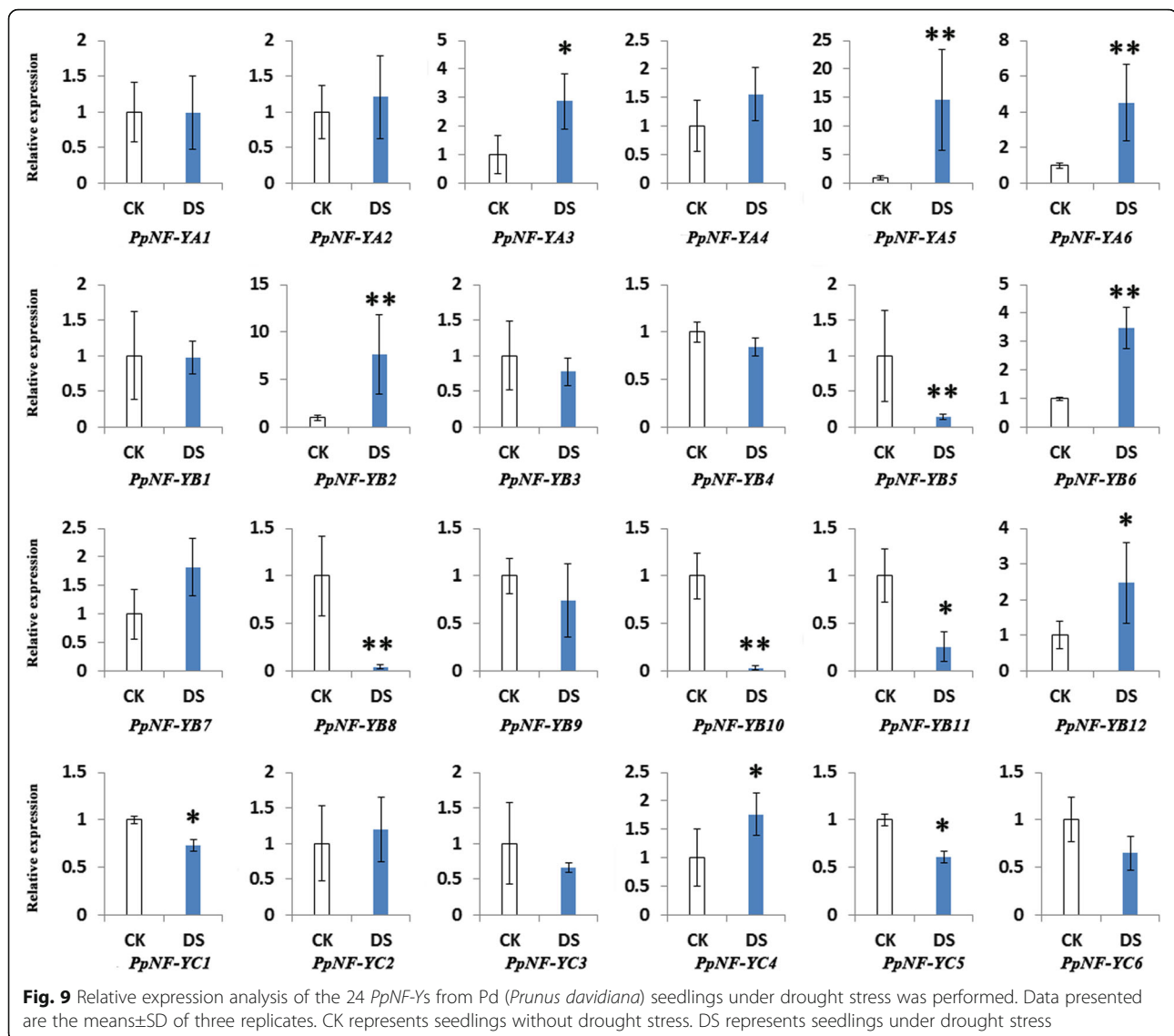


AtNF-YB members have been shown to have biological functions in various developmental processes, such as stimulating cell division and expansion, promoting flowering, and synthesizing chloroplasts [39]. Based on the phylogenetic relationships of NF-Y proteins between *Arabidopsis* and peach, more biological functions of PpNF-YB members could be identified in various developmental processes.

Diverse expression patterns of *PpNF-Ys* indicated a diversity of biological functions

As gene expression patterns can provide important clues for gene function, the current study analysed the expression patterns of *PpNF-Ys* in five peach organs. Some

PpNF-Ys such as *PpNF-YB* and *PpNF-YB7*, were highly expressed in vegetative tissues, indicating that they might be involved in vegetative growth. Some *PpNF-Ys* such as *PpNF-YC1*, were highly expressed in both vegetative tissues and reproductive tissues, showing that they might play multiple roles in the developmental process. However, some *PpNF-Ys* such as *PpNF-YB10* and *PpNF-YC6*, were specifically expressed in reproductive tissues, implying that they might participate in the development of floral organs and fruits. The expression patterns of *PpNF-Ys* can reveal the behaviours of *PpNF-Ys* during peach growth and development, providing useful information for further identification of the biological functions of *PpNF-Ys*.



*PpNF-Y*s as candidate drought-tolerance genes

The *NF-Y* TFs are closely related to drought stress tolerance. *NF-Y* members, including *Arabidopsis NF-YA5* and *NF-YB1*, maize *NF-YB2*, soybean *NF-YA3*, poplar *NF-YB7*, and bermudagrass *NF-YC1*, have been identified to participate in tolerance to drought stress [8, 29, 49]. Moreover, the transcriptional expression of drought-tolerance genes is generally induced by drought stress. This character could be used for identifying unknown drought-tolerance genes in peach.

The analysis of drought-responsive *cis*-elements in the *PpNF-Y* promoter sequences implied that some members might be involved in the drought-responsive pathway. The *PpNF-Y* promoter regions containing multiple types of drought-responsive *cis*-elements suggested that *PpNF-Y*s might be involved in different drought-responsive pathways. Based on an analysis of gene expression under drought

stress, 9 upregulated *PpNF-Y*s could serve as candidate genes to analyse drought tolerance in peach. The NJ tree of *PpNF-Y* and *AtNF-Y* proteins showed a close relationship between *AtNF-YB1* and *PpNF-YB2*, which was consistent with our inference that *PpNF-YB2* might be a drought-resistance gene, similar to *AtNF-YB1*. *PpNF-YA5*, another candidate drought-resistance gene, also showed a close relationship with the drought-resistance gene *AtNF-YA5*, supporting our analysis and prediction. This study provided the possibility to further research the novel drought-tolerance pathway related to the *NF-Y* genes in peach.

The *NF-Y* gene family is a popular research topic

The strong correlation between the *NF-Y* gene family and drought resistance has become a popular research topic and has been demonstrated in many species. Recently, in *Citrus*, castor bean and chickpea, *NF-Y* gene

family has been identified and candidate drought resistance genes members were analysed. Compared with recent published studies, this study contained some unique insights. First, the analysis of the *PpNF-Y* gene duplication indicated that most gene duplication events occurred long ago and generated similar duplicate genes, providing important clues into the evolutionary origins of the *PpNF-Y* gene family. Second, the details of drought-responsive *cis*-elements in the *PpNF-Y* promoters were shown, which is useful for exploring the upstream elements of the *PpNF-Y* genes involved in the drought resistance pathway. Third, the results of the phylogenetic relationships of *NF-Y* proteins between *Arabidopsis* and peach suggested that a continued targeted functional analysis of the *PpNF-Y* genes could be performed, such as for *PpNF-YB2-AtNF-YB1* (drought resistance) and *PpNF-YB4-AtNF-YB9* (desiccation-intolerant phenotype).

Conclusions

Since the *NF-Y* gene family was first recognized and classified, many studies have focused on this gene family. Recently, published studies have identified the *NF-Y* gene family in various plant species including maize, tomato, rice, banana and analysed the functions of this gene family members in some processes such as fruit ripening and abiotic stress. The current study found that *NF-Y* gene family members were frequently identified in field crop and vegetable crop species, and this research trend has been extended to fruit trees. Our study, carried out in peach, closely followed the current research focus. Twenty-four *PpNF-Ys* were first identified in the peach genome in the current study. Each subfamily of the *PpNF-Ys* contained typical characteristics. The analysis of the duplication events clearly displayed the expansion of *PpNF-Ys* across the genome. The current study indicated that there were some structural similarities of *NF-Ys* between peach and *Arabidopsis*. The reported *AtNF-Ys* were useful to predict and analyse the biological functions of the peach *NF-Y* gene family. In particular, the current study predicted the functions of some *PpNF-Ys* such as *PpNF-YB4-AtNF-YB9*, *PpNF-YB2-AtNF-YB1*, with the help of similar conserved domains and close evolutionary relationships to the reported *AtNF-Ys*, which allowed detailed research into those *PpNF-Ys*. In addition, the identification of drought-responsive *cis*-elements in the promoter regions of *PpNF-Ys* was useful to analyse and extend the drought-resistance pathway in peach. This study explored the performance of *NF-Ys* under drought stress in a notable current research area. Moreover, the current study identified two *PpNF-Ys* (*PpNF-YB2* and *PpNF-YA5*) as candidate genes for drought resistance, providing a foundation for further investigating the functions of *PpNF-Ys* and the molecular mechanism of peach drought-resistance.

Additional files

- Additional file 1:** Peach *NF-Ys* gene sequence. (DOCX 32 kb)
- Additional file 2:** Peach *NF-Ys* protein sequence. (DOCX 18 kb)
- Additional file 3:** *Arabidopsis* protein sequence. (XLSX 11 kb)
- Additional file 4:** Primer of Peach *NF-Y* family. (DOCX 19 kb)

Abbreviations

μL: Microlitre; ABRE: Abscisic Acid Responsive Element; *AtNF-Y*: *Arabidopsis NF-Y*; *CaNF-Y*: *Cicer arietinum NF-Y*; cDNA: Complementary DNA; *CsNF-Y*: *Citrus sinensis NF-Y*; DNA: Deoxyribonucleic acid; IBS: Illustrator for Biological Sequences; Kb: Kilobase; MBS: MYB binding site involved in drought-inducibility; mRNA: Messenger RNA; Mw: Molecular weight; *NF-Y*: Nuclear Factor Y; NJ: Neighbour-joining; Pd: *Prunus davidiana*; Pi: Isoelectric point; *PpNF-Y*: Peach *NF-Y*; *RcNF-Y*: *Ricinus communis NF-Y*; RNA: Ribonucleic acid

Acknowledgement

We sincerely thank Bo Li (Shandong Institute of Pomology, Taian, Shandong, China) for insightful discussions.

Authors' contributions

ANZ planned and designed the research. ML, GXL, WL and XMD performed experiments, analyzed data and made conclusions based on results. ANZ and ML wrote the manuscript. All authors read and approved the final manuscript.

Funding

This study is supported by China Agriculture Research System (CARS-30-Z-08). The funding body had no contribution in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

Availability of data and materials

- Additional file 1. Peach *NF-Ys* gene sequence
- Additional file 2. Peach *NF-Ys* protein sequence
- Additional file 3. *Arabidopsis* protein sequence
- Additional file 4. Primer of Peach *NF-Y* family

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Received: 17 January 2019 Accepted: 11 July 2019

Published online: 26 July 2019

References

1. Alves PC, Hartmann DO, Nunez O, Martins I, Gomes TL, Garcia H, Galceran MT, Hampson R, Becker JD, Pereira CS. Transcriptomic and Metabolomic Profiling of Ionic Liquid Stimuli Unveils Enhanced Secondary Metabolism in *Aspergillus nidulans*. *BMC Genomics*. 2016;17(1):284.
2. Ben S, Guasmi F, Mohamed MB, Benhaj K, Boussora F, Triki T, Kammoun NG. Identification of Internal Control Genes for Gene Expression Studies in Olive Mesocarp Tissue during Fruit Ripening. *S Arf J Bo*. 2018;117:11–6.
3. Berardini TZ, Reiser L, Li D, Mezheritsky Y, Muller R, Strait E, Huala E. The *Arabidopsis* information resource: Making and mining the “gold standard” annotated reference plant genome. *Genesis*. 2015;53(8):474–85.
4. Blanc G, Wolfe KH. Widespread paleopolyploidy in model plant species inferred from age distributions of duplicate genes. *Plant Cell*. 2004;16(7):1667–1678.
5. Boulard C, Thévenin J, Tranquet O, Laporte V, Lepiniec L, Dubreucq B. LEC1 (*NF-YB9*) directly interacts with LEC2 to control gene expression in seed. *Biochim Biophys Acta - Gene Regul Mech*. 2018;1861(5):443–50.

6. Cai X, Zhang Y, Zhang C, Zhang T, Hu T, Ye J, Zhang J, Wang T, Li H, Ye Z. Genome-wide analysis of plant-specific Dof transcription factor family in tomato. *J Integr Plant Biol.* 2013;55(6):552–66.
7. Cao K, Zhou Z, Wang Q, Guo J, Zhao P, Zhu G, Fang W, Chen C, Wang X, Wang X, Tian Z, Wang L. Genome-Wide Association Study of 12 Agronomic Traits in Peach. *Nat Commun.* 2016;8(7):13246.
8. 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 Biotechnol J.* 2014;13(4):482–91.
9. Chu HD, Nguyen KH, Watanabe Y, Le DH, Pham TLT, Mochida K, Tran LSP. 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. *Int J Mol Sci.* 2018;19:E3290.
10. Garcia-Albornoz M, Thankaswamy-Kosalai S, Nilsson A, Varembo L, Nookaew I, Nielsen J. BioMet Toolbox 2.0: genome-wide analysis of metabolism and omics data. *Nucleic Acids Res.* 2014;42(1):175–81.
11. Giovannoni JJ. Genetic regulation of fruit development and ripening. *Plant Cell.* 2004;16:170–80.
12. Gnesutta N, Kumimoto RW, Swain S, Chiara M, Siriwardana C, Horner DS, Holt BF, Mantovani R. CONSTANS Imparts DNA Sequence Specificity to the Histone Fold NF-YB/NF-YC Dimer. *Plant Cell.* 2017;29(6):1516–32.
13. Hackenberg D, Keetman U, Grimm B. Homologous NF-YC2 Subunit from Arabidopsis and Tobacco Is Activated by Photooxidative Stress and Induces Flowering. *Int J Mol Sci.* 2012;13(3):3458–77.
14. Hu B, Jin J, Guo A, Zhang H, Luo J, Gao G. GSDS 2.0: an upgraded gene feature visualization server. *Bioinformatics.* 2015;31(8):1296–7.
15. Initiative, I.P.G. Verde I, Abbott AG, Scalabrini S, Jung S, Shu S, Marroni F, Zhebentyayeva T, Dettori MT, Grimwood J, Cattonaro F, Zuccolo A, Rossini L, Jenkins J, Vendramin E, Meisel LA, Decroocq V, Sosinski B, Prochnik S, Mitros T, Policriti A, Cipriani G, Dondini L, Ficklin S, Goodstein DM, Xuan P, Fabbro CD, Aramini V, Copetti D, Gonzalez S, Horner DS, Falchi R, Lucas S, Mica E, Maldonado J, Lazzari B, Bielenberg D, Pirona R, Miculan M, Barakat A, Testolin R, Stella A, Tartarini S, Tonutti P, Arus P, Orellana A, Wells C, Main D, Vizzotto G, Silva H, Salamini F, Schmutz J, Morgante M, Rokhsar DS. The High-Quality Draft Genome of Peach (*Prunus Persica*) Identifies Unique Patterns of Genetic Diversity, Domestication and Genome Evolution. *Nat Genet.* 2013;45(5):487–94.
16. Jung S, Lee T, Cheng C, Buble K, Zheng P, Yu J, Humann J, Ficklin SP, Gasic K, Scott K, Frank M, Ru S, Hough H, Evans K, Peace C, Olmstead M, DeVetter LW, McFerson J, Coe M, Wegrzyn JL, Staton ME, Abbott AG, Main D. 15 Years of GDR: New Data and Functionality in the Genome Database for Rosaceae. *Nucleic Acids Res.* 2018;47(1):137–45.
17. Katiyar A, Smita S, Lenka SK, Rajwanshi R, Chinnusamy V, Bansal KC. Genome-wide classification and expression analysis of MYB transcription factor families in rice and Arabidopsis. *BMC Genomics.* 2012;13:544.
18. Lee SC, Kim SH, Kim SR. Drought inducible *OsDhn1* promoter is activated by *OsDREB1A* and *OsDREB1D*. *J Plant Biol.* 2013;56(2):115–21.
19. Lescot M, Dehais P, Thijs G, Marchal K, Moreau Y, Peer VY, 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 Res.* 2002;30:325–7.
20. 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.
21. Link J, Pachaly J. Intranarcotic infusion therapy: A computer interpretation using the program package SPSS (Statistical package for the social sciences). *Infusionsther Klin Ernahr.* 1975;2(4):255–9.
22. Liu W, Xie Y, Ma J, Luo X, Nie P, Zuo Z, Lahrmann U, Zhao Q, Zheng Y, Zhao Y, Xue Y, Ren J. BS: an illustrator for the presentation and visualization of biological sequences. *Bioinformatics.* 2015;31(20):3359–61.
23. Mantovani R. The molecular biology of the CCAAT-binding factor NF-Y. *Gene.* 1999;239(1):15–27.
24. Marchler-Bauer A, Zheng C, Chitsaz F, Derbyshire MK, Geer LY, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Lanczycki CJ, Lu F, Lu S, Marchler GH, Song JS, Thanki N, Yamashita RA, Zhang D, Bryant SH. CDD: Conserved domains and protein three-dimensional structure. *Nucleic Acids Res.* 2012;41(1):348–52.
25. Myers ZA, Holt BF. NUCLEAR FACTOR-Y: Still Complex after All These Years? *Curr Opin Plant Biol.* 2018;45:96–102.
26. Nakashima K, Jan A, Todaka D, Maruyama K, Goto S, Shinozaki K, Yamaguchi-Shinozaki K. Comparative functional analysis of six drought-responsive promoters in transgenic rice. *Planta.* 2014;239(1):47–60.
27. Nardone V, Chaves-Sanjuan A, Nardini M. Structural determinants for NF-Y/DNA interaction at the CCAAT box. *Biochim Biophys Acta - Gene Regul Mech.* 2017;1860(5):571–80.
28. 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. 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.* 2007;104(42):16450–5.
29. Ni Z, Zheng H, Jiang Q, Zhang H. GmNFYA3, a target gene of miR169, is a positive regulator of plant tolerance to drought stress. *Plant Mol Biol.* 2013;82(1–2):113–29.
30. Nicholas KB, Nicholas HB, Deerfield DW, Nicholas HBJ, Nicholas K, Nicholas RW, Maszle DR, Nicholas HBJ, Nicholas A, Deerfield DW, Nicholas H, Gauch H. GeneDoc: a tool for editing and annotating multiple sequence alignments. *Embnet.news.* 1997.
31. Nordberg H, Cantor M, Dusheyko S, Hua S, Poliakov A, Shabalov I, Smirnova T, Grigoriev IV, Dubchak I. The genome portal of the Department of Energy Joint Genome Institute: 2014 updates. *Nucleic Acids Res.* 2014;42:26–31.
32. Olias R, Eljakaoui Z, Li J, De M, Alvarez P, Marin-Manzano MC, Pardo JM, Belver A. The plasma membrane Na⁺/H⁺ antiporter SOS1 is essential for salt tolerance in tomato and affects the partitioning of Na⁺ between plant organs. *Plant Cell Environ.* 2009;32(7):904–16.
33. Pereira SLS, Martins CPS, Sousa AO, Camillo LR, Araujo CP, Alcantara GM, Camargo DS, Cidade LC, de Almeida AF, Costa MGC. Genome-wide characterization and expression analysis of citrus NUCLEAR FACTOR-Y (NF-Y) transcription factors identified a novel NF-YA gene involved in drought-stress response and tolerance. *PLoS One.* 2018;13(6):e0199187.
34. Rozas J, Ferrer-Mata A, Sanchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, Sanchez-Gracia A. DnaSP 6: DNA Sequence Polymorphism Analysis of Large Datasets. *Mol Biol Evol.* 2017;34(12):3299–302.
35. Saitou N, Nei M. The Neighbor-Joining Method—a New Method for Reconstructing Phylogenetic Trees. *Mol Biol Evol.* 1987;4(4):406–25.
36. Shi J, Zhang L, An H, Wu C, Guo X. GhMPK16, a novel stress-responsive group DMAPK gene from cotton, is involved in disease resistance and drought sensitivity. *BMC Mol Biol.* 2011;12(1):22.
37. Siefers N, Dang K, Kumimoto R, Bynum W, Tayrose G, Holt B. Tissue-Specific Expression Patterns of Arabidopsis NF-Y Transcription Factors Suggest Potential for Extensive Combinatorial Complexity. *Plant Physiol.* 2009;149(2):625–41.
38. Song XM, Liu TK, Duan WK, Ma QH, Ren J, Wang Z, Li Y, Hou XL. Genome-Wide Analysis of the GRAS Gene Family in Chinese Cabbage (*Brassica Rapa* Ssp. *Pekinensis*). *Genomics.* 2014;103(1):135–46.
39. Swain S, Myers Z, Chamindika S, Holt B. The Multifaceted Roles of NUCLEAR FACTOR-Y in Arabidopsis Thaliana Development and Stress Responses. *Biochim Biophys Acta - Gene Regul Mech.* 2016;1860(5):636–44.
40. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: Molecular Evolutionary Genetic Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Mol Biol Evol.* 2011;28(10):2731–9.
41. Thon M, Al Abdallah Q, Hortschansky P, Scharf DH, Eisendle M, Haas H, Brakhage AA. The CCAAT-binding complex coordinates the oxidative stress response in eukaryotes. *Nucleic Acids Res.* 2010;38(4):1098–113.
42. Timothy B, Elkan C. Fitting a mixture model by expectation maximization to discover motifs in bipolymers. *Proc Int Conf Intell Syst Mol Biol.* 1994;2:28–36.
43. Tong Z, Gao Z, Wang F, Zhou J, Zhang Z. Selection of reliable reference genes for gene expression studies in peach using real-time PCR. *BMC Mol Biol.* 2009;10:71.
44. Wang L, Zhu J, Li X, Wang S, Wu J. Salt and Drought Stress and ABA Responses Related to BZIP Genes from *V. Radiata* and *V. Angularis*. *Gene.* 2018a;651(20):152–60.
45. 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.* 2018b;247:559–72.
46. Yan H, Wu F, Jiang G, Xiao L, Li Z, Duan X, Jiang Y. Genome-wide identification, characterization and expression analysis of NFY gene

- family in relation to fruit ripening in banana. *Postharvest Biol Technol.* 2019;151:98–110.
47. Yang W, Lu Z, Xiong Y, Yao J. Genome-wide identification and co-expression network analysis of the OsNF-Y gene family in rice. *Crop J.* 2017;5(1):21–31.
 48. Zambelli F, Pavesi G. Genome wide features, distribution and correlations of NF-Y binding sites. *Biochim. Biophys. Acta - Gene Regul. Mech.* 2017;1860(5):581–9.
 49. Zanetti ME, Ripodas C, Niebel A. Plant NF-Y transcription factors: Key players in plant-microbe interactions, root development and adaptation to stress. *Biochim Biophys Acta - Gene Regul Mech.* 2017;1860(5):645–54.
 50. Zhang H, Gao S, Lercher MJ, Hu S, Chen WH. EvolView, an online tool for visualizing, annotating and managing phylogenetic trees. *Nucleic Acids Res.* 2012;40(1):569–72.
 51. Zhang H, Kang H, Su C, Qi Y, Liu X, Pu J. Genome-Wide Identification and Expression Analysis of the NAC Transcription Factor Family in Cassava. *PLoS One.* 2015a;10(6):e0136993.
 52. Zhang Z, Li X, Yu R, Han M, Wu Z. Isolation, structural analysis, and expression characteristics of the maize TIFY gene family. *Mol Gen Genomics.* 2015b;290(5):1849–58.

Publisher's Note

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

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

