



# Genome-Wide Identification and Analysis of the NF-Y Gene Family in Potato (*Solanum tuberosum* L.)

Zhen Liu<sup>1</sup>, Yuanming Li<sup>2</sup>, Jinyong Zhu<sup>3</sup>, Wenjing Ma<sup>3</sup>, Zhitao Li<sup>3</sup>, Zhenzhen Bi<sup>3</sup>, Chao Sun<sup>3</sup>, Jiangping Bai<sup>3</sup>, Junlian Zhang<sup>2</sup> and Yuhui Liu<sup>1\*</sup>

<sup>1</sup> State Key Laboratory of Aridland Crop Science, Gansu Agricultural University, Lanzhou, China, <sup>2</sup> College of Horticulture, Gansu Agricultural University, Lanzhou, China, <sup>3</sup> College of Agronomy, Gansu Agricultural University, Lanzhou, China

## OPEN ACCESS

### Edited by:

Rocio Deanna,  
University of Colorado Boulder,  
United States

### Reviewed by:

Jingwei Yu,  
Southern University of Science  
and Technology, China  
Tao Xu,  
Jiangsu Normal University, China  
Fuyou Fu,  
Agriculture and Agri-Food Canada  
(AAFC), Canada

### \*Correspondence:

Yuhui Liu  
yuhui2021@163.com

### Specialty section:

This article was submitted to  
Plant Genomics,  
a section of the journal  
Frontiers in Genetics

**Received:** 12 July 2021

**Accepted:** 20 August 2021

**Published:** 17 September 2021

### Citation:

Liu Z, Li Y, Zhu J, Ma W, Li Z, Bi Z,  
Sun C, Bai J, Zhang J and Liu Y  
(2021) Genome-Wide Identification  
and Analysis of the NF-Y Gene Family  
in Potato (*Solanum tuberosum* L.).  
*Front. Genet.* 12:739989.  
doi: 10.3389/fgene.2021.739989

Nuclear factor Y (NF-Y) is a ubiquitous transcription factor in eukaryotes, which is composed of three subunits (NF-YA, NF-YB, and NF-YC). NF-Y has been identified as a key regulator of multiple pathways in plants. Although the NF-Y gene family has been identified in many plants, it has not been reported in potato (*Solanum tuberosum*). In the present study, a total of 41 NF-Y proteins in potato (StNF-Ys) were identified, including 10 StNF-YA, 22 StNF-YB, and nine StNF-YC subunits, and their distribution on chromosomes, gene structure, and conserved motif was analyzed. A synteny analysis indicated that 14 and 38 pairs of StNF-Y genes were orthologous to *Arabidopsis* and tomato (*Solanum lycopersicum*), respectively, and these gene pairs evolved under strong purifying selection. In addition, we analyzed the expression profiles of NF-Y genes in different tissues of double haploid (DM) potato, as well as under abiotic stresses and hormone treatments by RNA-seq downloaded from the Potato Genome Sequencing Consortium (PGSC) database. Furthermore, we performed RNA-seq on white, red, and purple tuber skin and flesh of three potato cultivars at the tuber maturation stage to identify genes that might be involved in anthocyanin biosynthesis. These results provide valuable information for improved understanding of StNF-Y gene family and further functional analysis of StNF-Y genes in fruit development, abiotic stress tolerance, and anthocyanin biosynthesis in potato.

**Keywords:** potato (*Solanum tuberosum*), nuclear factor Y transcription factor, expression profiles, abiotic stress, anthocyanin biosynthesis

## INTRODUCTION

Nuclear factor Y (NF-Y) transcription factor, also known as CCAAT-box binding factor (CBF) or heme-associated protein (HAP) (Mantovani, 1999), is a ubiquitous transcription factor in eukaryotes (Zhao et al., 2017). NF-Y is a trimeric transcription factor composed of three distinct subunits, namely, NF-YA, NF-YB, and NF-YC (Nardini et al., 2013). To form the NF-Y complex, the histone folding motifs of NF-YB and NF-YC interact in the cytoplasm to form heterodimer,

**Abbreviations:** StNF-Y, nuclear factor Y transcription factor in potato; DM, double monoploid potato; HMM, hidden Markov model; Ka, non-synonymous substitution rate; Ks, synonymous substitution rate; PG, PGSC0003DMG40; qPCR, quantitative real-time PCR; CC, correlation coefficient.

which is transferred to the nucleus. Upon arrival in the nucleus, the NF-YA subunit is recruited and combines with NF-YB/NF-YC heterodimer to produce mature NF-Y complex (Kahle et al., 2005; Laloum et al., 2013). In mature NF-Y complex, the NF-YA subunit provides a distinctive sequence that specifically binds to the CCAAT *cis*-element and regulates the expression of target genes (Hackenberg et al., 2012; Yan et al., 2013; Cao et al., 2014; Xu et al., 2016). In yeast and mammals, each NF-Y subunit is encoded by a single gene; however, three NF-Y subunits are encoded by multiple genes in plants (Li et al., 1992; Petroni et al., 2012).

In recent years, many NF-Ys were identified in *Arabidopsis* (Riechmann et al., 2000), rice (*Oryza sativa*) (Yang W. et al., 2017), tomato (*Solanum lycopersicum*) (Li et al., 2016), banana (*Musa acuminata*) (Yan et al., 2019), sorghum (*Sorghum bicolor*) (Malviya et al., 2016), watermelon (*Citrullus lanatus*) (Yang J. et al., 2017), and canola (*Brassica napus*) (Liang et al., 2014). And, the functions of many NF-Y genes have been characterized to play multiple roles in plants, including embryogenesis (Kwong et al., 2003), seed germination (Liu et al., 2016), root growth (Sorin et al., 2014), flowering (Cao et al., 2014), fruit ripening (Li et al., 2016), as well as flavonoid biosynthesis (Wang et al., 2020) and response to various abiotic stresses (Li et al., 2008; Ni et al., 2013; Sato et al., 2014; Wang et al., 2018). For example, *AtLEC1* (*AtNF-YB9*) is a key regulator of late embryogenesis and seed development in *Arabidopsis*, and the ectopic expression of *AtLEC1* caused abnormality in tobacco transgenic seedling (Guo et al., 2013). The overexpression of *AtNF-YA1* and *AtNF-YA9* significantly affects male gametophyte development, embryo development, seed morphology, and seed germination (Mu et al., 2013). The overexpression of *AtNF-YA5*, which is strongly induced under drought, increases drought tolerance of *Arabidopsis* plants (Li et al., 2008). The constitutive expression of *TaNF-YA10* significantly increases the sensitivity of *Arabidopsis* plants to salinity (Ma et al., 2015). A recent report suggested that NF-Y complexes composed of NF-YB8, NF-YC1/9, and NF-YA1/9 play an important role in tomato flavonoid biosynthesis, and these complexes are involved in transcriptional regulation and H3K27me3 marking of the *chalcone synthase* (*CHS1*) locus (Wang et al., 2020). Although the NF-Y gene family has been extensively studied in plants, little report is yet available for potato.

Potato, which originated from the Andean regions of Peru and Bolivia (Spooner et al., 2005), is an important food crop worldwide. As a new kind of natural pigment and antioxidant resource, the pigmented potato has attracted more and more attention (Fossen and Andersen, 2000). Flavonoids not only attract insects for pollination but also protect plants against UV-induced damage (He and Giusti, 2010), as well as play a key role in cold and drought stress tolerance (Castellarin et al., 2007; Kim et al., 2017). Furthermore, many flavonoids have strong antioxidant activity and free radical-scavenging ability, which are beneficial to human health, such as preventing cardiovascular disease, controlling obesity, alleviating diabetes, anti-cancer, and so on (Hollman and Katan, 1997; Castañeda-Ovando et al., 2009). However, abiotic stresses can severely restrict the growth and productivity of potato, such as high-temperature stress, cold

stress, salinity stress, and drought stress. Given the important role that NF-Y genes play in plants, it is essential to identify and study the NF-Y gene family in the potato genome.

In the present study, we identified a NF-Y gene family with 41 members in potato (*StNFYs*), and comprehensive analyses of the phylogenetic relationships, chromosome distribution, gene structure, sequence features, and gene duplications were further performed. In addition, we analyzed the expression profiles of NF-Y genes in different tissues of double haploid (DM) potato, as well as under abiotic stresses and hormone treatments by RNA-seq downloaded from the Potato Genome Sequencing Consortium (PGSC) database. Furthermore, we performed RNA-seq on white, red, and purple tuber skin and flesh of three potato cultivars at the tuber maturation stage to identify genes that might be involved in anthocyanin biosynthesis. The results could provide a basis for further study on the functional characterization of the NF-Y gene family.

## MATERIALS AND METHODS

### Identification of StNF-Ys in Potato

The amino acid and nucleotide sequences of potato were downloaded from the PGSC<sup>1</sup>. To identify the potato NF-Y members, two methods were used individually and then combined: (1) the 36 reported *Arabidopsis* NF-Y (*AtNF-Y*) protein sequences (Riechmann et al., 2000) were downloaded from the *Arabidopsis* information resource (TAIR;<sup>2</sup>). The BLASTP (Altschul et al., 1997) was used to search for *StNF-Y* members based on 36 *AtNF-Y* amino acid sequences with an *E*-value  $\leq 1e-5$ . (2) The hidden Markov model (HMM) profiles of the NF-Y domain (PF00808 and PF02045) were downloaded from Pfam<sup>3</sup>. The *StNF-Y* proteins were identified by HMMER3.1 software<sup>4</sup>. After removing all redundant sequences, these candidate members were submitted to SMART<sup>5</sup> and NCBI Conserved Domain Data (CDD) to manually screen *StNF-Y* members for further analysis.

### Sequence Analysis and Structural Characterization

The number of amino acids, theoretical isoelectric point (pI), and molecular weight (MW) were calculated using the ExPasy site<sup>6</sup> (Gasteiger et al., 2005). The MEME program<sup>7</sup> was used to identify motifs in *StNF-Y* sequence with the following parameters: 20 motifs with an optimal motif width of 6–50 amino acid residues and any number of repeats (Bailey et al., 2009). The gene structures of the *StNF-Ys* were drawn using the Gene Structure Display Server (GSDS 2.0, <sup>8</sup>) (Guo et al., 2007).

<sup>1</sup>[http://solanaceae.plantbiology.msu.edu/pgsc\\_download.shtml](http://solanaceae.plantbiology.msu.edu/pgsc_download.shtml)

<sup>2</sup><http://www.Arabidopsis.org>

<sup>3</sup><http://pfam.xfam.org>

<sup>4</sup><http://hmmer.org/download.html>

<sup>5</sup><http://smart.embl-heidelberg.de>

<sup>6</sup><http://web.expasy.org/protparam>

<sup>7</sup><http://alternate.meme-suite.org/tools/meme>

<sup>8</sup><http://gsds.cbi.pku.edu.cn>

## Chromosomal Localization and Gene Duplication

All *StNF-Y* genes were mapped to 12 potato chromosomes based on physical location information from PGSC using MapChart software (Voorrips, 2002). MCScanX (Wang et al., 2012) was used to analyze the duplication events of the *StNF-Y* genes and identified the synteny of *NF-Y* genes between potato and *Arabidopsis*. Ka (non-synonymous) and Ks (synonymous) of each pair of duplicated *NF-Y* genes were calculated using KaKs Calculator 2.0 (Wang et al., 2010).

## Multiple Alignments, Phylogenetic Analysis, and Classification of NF-Ys

The multiple alignments were performed using ClustalW on the 41 *StNF-Ys* amino acid under default parameters. The full-length amino acid sequences of 41 *StNF-Ys*, 54 *SINF-Ys* (tomato *NF-Y*), and 36 *AtNF-Ys* were used for phylogenetic analysis. The Molecular Evolutionary Genetics Analysis (MEGA) 7.0 software (Kumar et al., 2016) was used to construct an unrooted phylogenetic tree with the following parameters: Poisson model and 1,000 bootstrap replications.

## Plant Materials and Treatments

Three potato cultivars, including “Xindaping” (XD—white skin and white flesh), “Lingtianhongmei” (LT—red skin and red flesh), and “Heimeiren” (HM—purple skin and purple flesh), were grown in a greenhouse at Gansu Agricultural University in Lanzhou, Gansu Province, China (Figure 1). Six fresh tubers (diameter = 4–5 cm) were collected from each cultivar for harvesting the skin tissue and flesh tissue. Skin tissues were carefully obtained from the cortical tissue using a scalpel. Flesh tissues were isolated at least 5-mm distance from the skin. The samples were immediately frozen in liquid nitrogen and stored in  $-80^{\circ}\text{C}$  refrigerator for further use.

## RNA-Seq Data Analysis

Total RNA of the aforementioned samples was chosen for further RNA-seq library construction. Each sample had three biological replications. Next-generation Illumina sequencing was

performed by Sagene Biotech Corporation (Guangzhou, China). The raw RNA-seq data were uploaded to the NCBI (Project ID PRJNA541919). After RNA sequencing, the clean reads were obtained by trimming the raw reads and filtering out contaminants, adapters, Phred scores less than 20, and uncertain bases. The cleaned data were aligned to PGSC\_DM\_v3.4 gene models downloaded from Solanaceae Genomics Resource at Michigan State University<sup>1</sup> by Bowtie2 (v2.2.9). Only reads with a perfect match or one mismatch were further analyzed.

## Differential Expression Genes Analysis

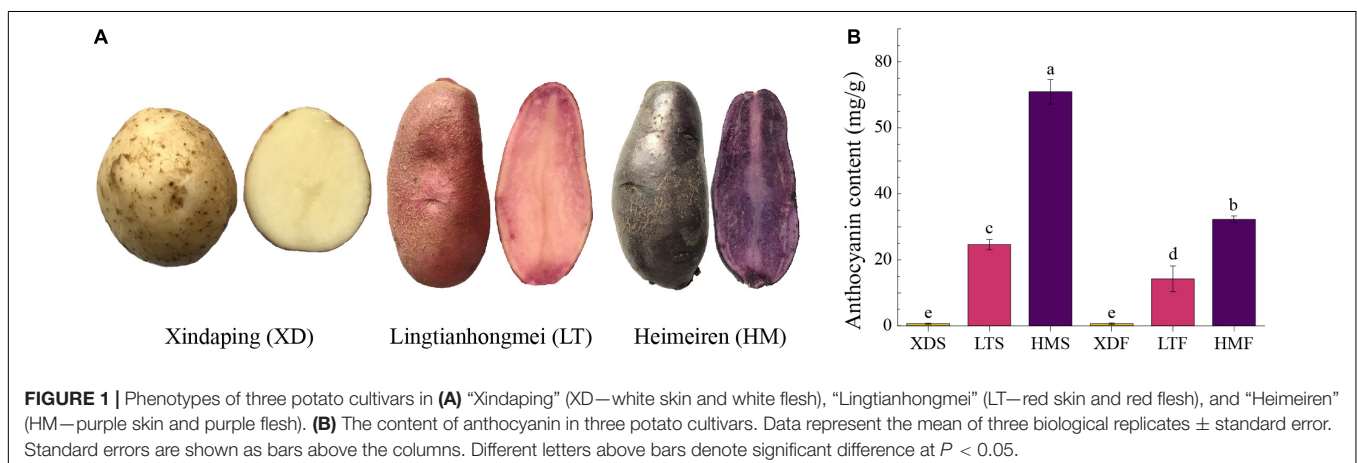
Differential genes were identified using fragments per kilobase of exon per million fragments mapped (FPKM). Genes with the absolute value of  $\log_2$  fold change (FC)  $\geq 1$  and a false discovery rate (FDR)  $< 0.05$  were considered significant differentially expressed genes (DEGs).

The DEGs were annotated against non-redundant database (Nr), SwissProt/UniProt Plant Proteins, Kyoto Encyclopedia of Genes and Genomes (KEGG), Cluster of Orthologous Groups of proteins (COG/KOG), and the potato protein database<sup>9</sup> by BLASTX, with a cut-off *E*-value of  $1e-5$ . Then, DEGs were subjected to enrichment analysis of GO functions and KEGG pathways.

## RNA Isolation and Quantitative Real-Time PCR

Total RNA extractions were isolated with the RNA extraction kit (Tiangen DP419, Beijing, China). The integrity of RNA was monitored by agarose gel electrophoresis, and the concentration was detected using a NanoDrop ND-2000 spectrophotometer (NanoDrop Technologies, United States). The cDNAs were synthesized using a FastKing RT kit with gDNase (Tiangen KR116, Beijing, China). Quantitative real-time PCR (qPCR) was performed with SuperReal PreMix Plus kit (SYBR Green FP205, Tiangen, Beijing, China) in Bio-Rad CFX96 (Bio-Rad, Hercules, CA, United States). The PCR program was as follows: 30 s at  $95^{\circ}\text{C}$ , followed by 40 cycles of 5 s at  $95^{\circ}\text{C}$  and 30 s at  $60^{\circ}\text{C}$ , followed by  $65-95^{\circ}\text{C}$  melting curve detection. Three independent

<sup>9</sup><http://ftp.jgi-psf.org/pub/compngen/phytozome/v9.0/Stuberosum>



biological replications were performed. We used the  $2^{-\Delta\Delta Ct}$  method (Livak and Schmittgen, 2001) to determine the relative expression levels of genes. The *StEF-1 $\alpha$*  (AB061263) was taken as the internal control (Tang et al., 2017). The primers are listed in **Supplementary Table 1**.

## Expression Pattern Analysis of StNF-Ys in Potato

The expressions of *StNF-Y* genes in 13 tissues (leaves, stamens, shoots, stolons, roots, tubers, carpels, petals, petioles, sepals, flowers, immature fruit, and mature fruit) in DM potato and the whole plant *in vitro* that was treated for abiotic stress (salt treatment: 150 mM NaCl, 24 h; mannitol-induced drought stress: 260  $\mu$ M mannitol, 24 h; heat treatment: 35°C, 24 h) and hormone treatments [benzylaminopurine (BAP): 10  $\mu$ M, 24 h; abscisic acid (ABA): 50  $\mu$ M, 24 h; indole acetic acid (IAA): 10  $\mu$ M, 24 h; gibberellic acid (GA3): 50  $\mu$ M, 24 h] were analyzed based on the Illumina RNA-seq data that was downloaded from the PGSC (The Potato Genome Sequencing Consortium et al., 2011). TBtools software was used to draw the heat map (Chen et al., 2018).

The expression patterns of *StNF-Y* genes in three pigmented potato cultivars (XD, LT, and HM) were analyzed based on RNA-seq. The Illumina sequencing was performed by Sagene Biotech Corporation (Guangzhou, China). The raw data were uploaded on NCBI (Project ID PRJNA541919).

## The Interaction Network of NF-Y Proteins

The Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) is a database of known and predicted protein-protein interactions. Based on the RNA-seq dataset of pigmented potato cultivars, we selected *StNF-Y*s with FPKM > 5 to predict the protein interaction network. The protein interactions were constructed using online STRING v11.0 software<sup>10</sup> with a combined score > 400 (medium confidence) (Szklarczyk et al., 2016). As active interaction sources, text mining, experiments, databases, co-expression, neighborhood, gene fusion, and co-occurrence were selected. The interaction network was visualized by Cytoscape v3.7.1 (Shannon et al., 2003).

## Statistical Analysis

For qPCR analyses, data were presented as means ( $\pm$ SE) of three biological replicates. Statistical significance was determined by one-way ANOVA followed by the least significant difference (LSD) computed at  $P < 0.05$ .

## RESULTS

### Identification and Chromosomal Distribution of StNF-Ys

To obtain the NF-Y family members in potato, the HMM and BLASTP algorithm search were combined to analyze the

potato genome. A total of 41 *StNF-Y* genes were identified and confirmed in the potato genome, including 10 *StNF-YA*, 22 *StNF-YB*, and 9 *StNF-YC* (**Supplementary Table 2**). The transcript IDs were converted to corresponding gene IDs and the “SC0003DMG40” in the gene ID was omitted for brevity. The bioinformatics data of *StNF-Y* members were analyzed, including the number of amino acid residues, theoretical molecular weight (MM), and theoretical isoelectric point (pI). The *StNF-Y* proteins varied in length and physicochemical properties, the number of amino acid residues of *StNF-Y* proteins ranged from 120 (PG0013302) to 311 (PG0021365), the molecular weights were between 13.82 kDa (PG0013302) and 33.97 kDa (PG0021365), and the pI values ranged from 4.64 (PG0023065) to 9.65 (PG0002484). The details are listed in **Supplementary Table 2**.

The identified 41 *StNF-Y* genes were distributed unevenly on 12 chromosomes (**Figure 2A**). Seven *StNF-Y* genes were located on chromosomes 1 and 5, which had the largest number of *StNF-Y* genes. In contrast, only one *StNF-Y* was distributed on chromosome 8, which had the least number of *StNF-Y* genes (**Figure 2B**). In addition, the high densities of *StNF-Y* genes were distributed at the proximal end of chromosome 1 and the distal end of chromosome 5.

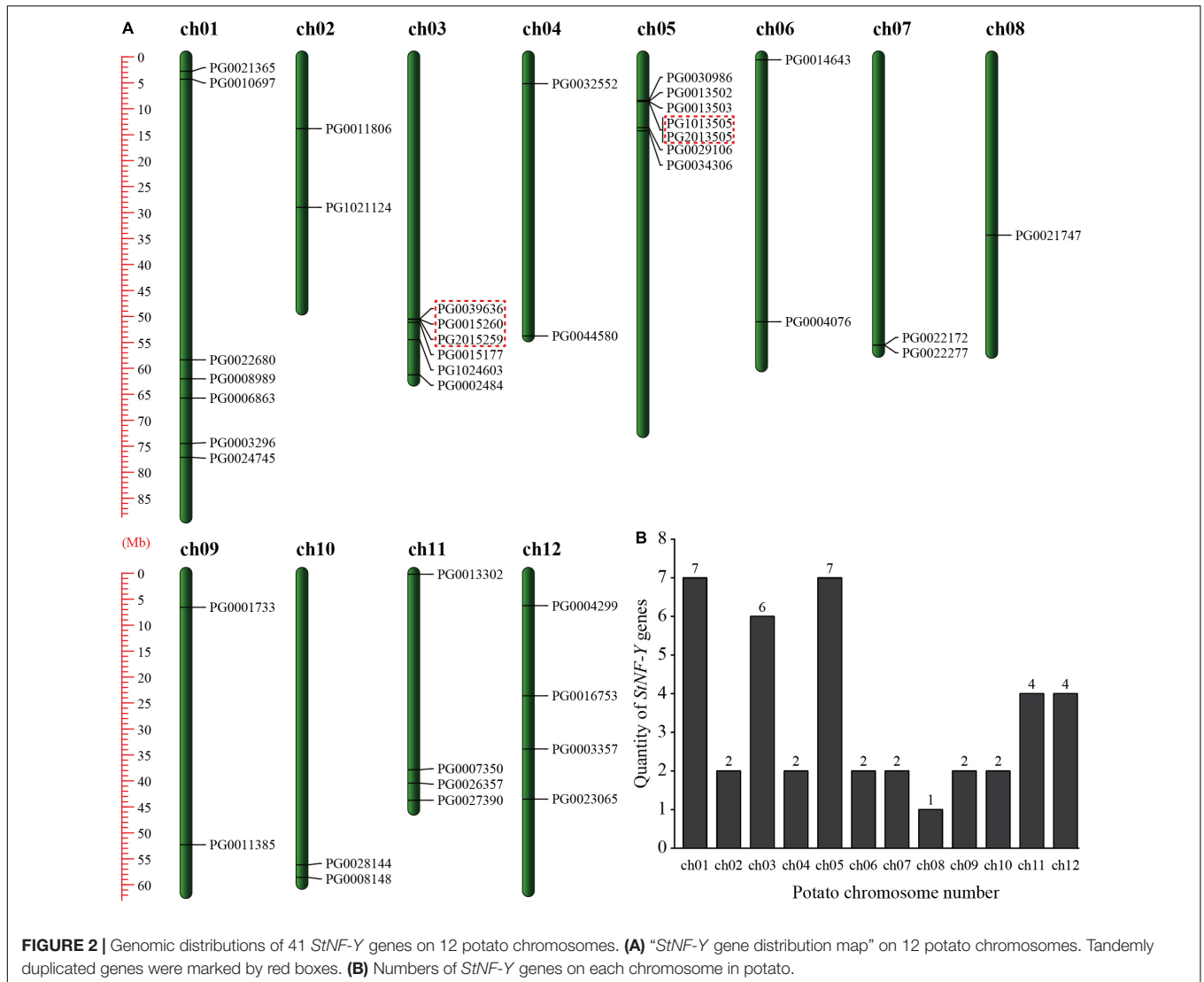
### Multiple Alignments, Phylogenetic Analysis, and Classification of StNF-Ys

Multiple sequence alignments showed that *StNF-Y* family proteins had conserved region domains as shown in **Figure 3**. The results showed that each member of *StNF-Y* contains a heterodimerization domain and a DNA-binding domain that recognizes the CCAAT site. The *StNF-YA* subunit contained two conserved domains (**Figure 3A**); one is the NF-YB and NF-YC interaction domain and the other is the DNA-binding domain. In this DNA-binding domain, three histidine (H) and three arginine (R) residues are absolutely conserved and essential (Xing et al., 1993). The conserved domain of *StNF-YB* subunit was about 93 AAs in length, which contained a NF-YA interaction region, a NF-YC interaction region, and a DNA-binding domain (**Figure 3B**). Previous studies have shown that the NF-YB subunits could be divided into LEC1 type and non-LEC1 type according to sequence similarity; the LEC1 type is composed of LEC1 and LEC1-like (L1L) (Lee et al., 2003; Siefers et al., 2009).

As shown in **Figure 3B**, the key amino acids that distinguish Lec1 from non-LEC1 were represented in the red box, and aspartate (D) at this site was considered to be specific for LEC1 type. The conserved domains of *StNF-YC* subunits contained three heterodimerization regions and one DNA-binding domain, which was about 79 AAs in length (**Figure 3C**).

To further investigate the phylogenetic relationships of NF-Ys among potato, tomato, and *Arabidopsis*, an unrooted phylogenetic tree of 41 *StNF-Y*s, 54 *SINF-Y*, and 36 *AtNF-Y* protein sequences was constructed by MEGA 7.0 software, as shown in **Figure 4**. The 131 NF-Ys were divided into three distinct subunits (NF-YA, NF-YB, and NF-YC). Ten *StNF-Y*s, 10 *SINF-Y*s, and 10 *AtNF-Y*s belonged to the NF-YA subunit. Twenty-two *StNF-Y*s, 27 *SINF-Y*s, and 13 *AtNF-Y*s were assigned to the NF-YB subunit. Nine *StNF-Y*s, 17

<sup>10</sup><http://string-db.org/>



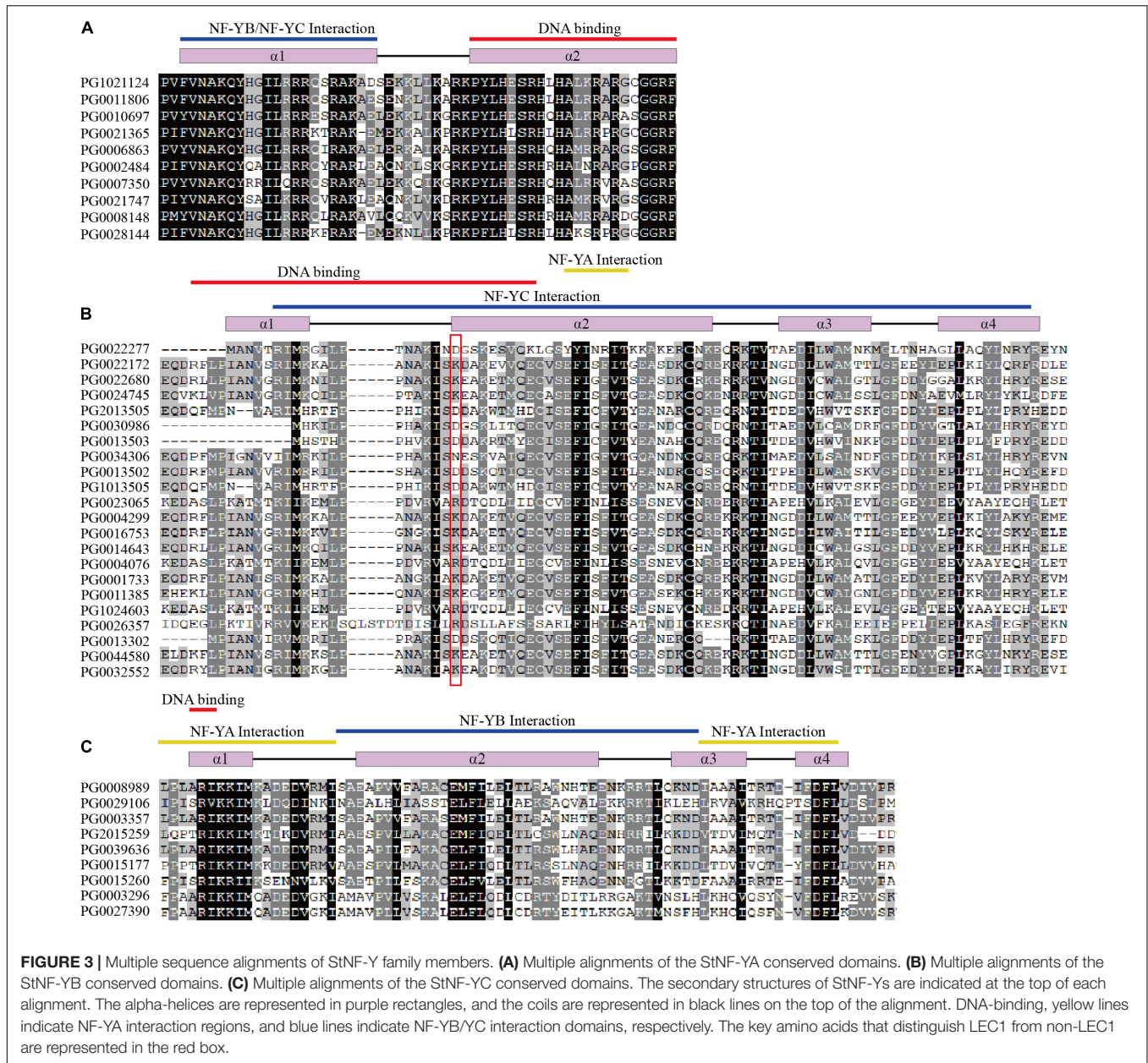
*StNF-Y*s, and 13 *AtNF-Y*s belonged to the *NF-YC* subunit. Notably, a distinct cluster was identified as *LEC1*-type group in *NF-YB*, including *AtNF-YB6* (*LEC1*-like), *AtNF-YB9* (*LEC1*), 13 *SlNF-Y*s, and eight *StNF-Y*Bs (PG1013505, PG2013505, PG0013503, PG0013302, PG0030986, PG0022277, and PG0034306).

## Gene Structure and Conserved Motifs of *StNF-Y*s

The full-length amino acid sequences of 41 *StNF-Y*s were used to construct an unrooted phylogenetic tree by MEGA 7.0 software. Based on phylogenetic analysis of potato, tomato, and *Arabidopsis*, 41 *StNF-Y*s were divided into three subunits (Figure 5A). A gene structure analysis can provide insight into the evolution of gene family. Therefore, we analyzed the organization of exons and introns (Figure 5B). The results showed that all of *StNF-YA* genes were separated by introns, of which eight *StNF-YA*s

contained four introns, one *StNF-YA* contained three introns, and one *StNF-YA* contained two introns. Most of *StNF-YB*s and *StNF-YC*s did not have introns, including 14 *StNF-YB*s and six *StNF-YC*s; the rest of *StNF-YB*s and *StNF-YC*s contained 1–4 introns. In general, the gene structure of *StNF-YA* members was more conserved than that of *StNF-YB*s and *StNF-YC*s.

In order to study the diversification of *StNF-Y*s, we searched for conserved motifs of these proteins using the online MEME program. The details of the 20 motifs are in **Supplementary Table 3**. The results showed that the three *StNF-Y* subunits have a unique motif distribution (Figure 5C). Motifs 12, 17, and 18 were specifically present in the *StNF-YA* members. Motifs 1, 8, 10, 11, 16, and 19 were unique to the *StNF-YB* members. Motifs 4, 9, 13, 15, and 20 were present only in *StNF-YC* members. Motif 2 was distributed in *StNF-YA* and *StNF-YB* subunits. Motif 3 was present in *StNF-YB* and *StNF-YC* subunits. Motif 6 was widely distributed in three *StNF-Y* subunits. Overall, each subunit had a conserved motif composition.



**FIGURE 3 |** Multiple sequence alignments of StNF-Y family members. **(A)** Multiple alignments of the StNF-YA conserved domains. **(B)** Multiple alignments of the StNF-YB conserved domains. **(C)** Multiple alignments of the StNF-YC conserved domains. The secondary structures of StNF-Ys are indicated at the top of each alignment. The alpha-helices are represented in purple rectangles, and the coils are represented in black lines on the top of the alignment. DNA-binding, yellow lines indicate NF-YA interaction regions, and blue lines indicate NF-YB/YC interaction domains, respectively. The key amino acids that distinguish LEC1 from non-LEC1 are represented in the red box.

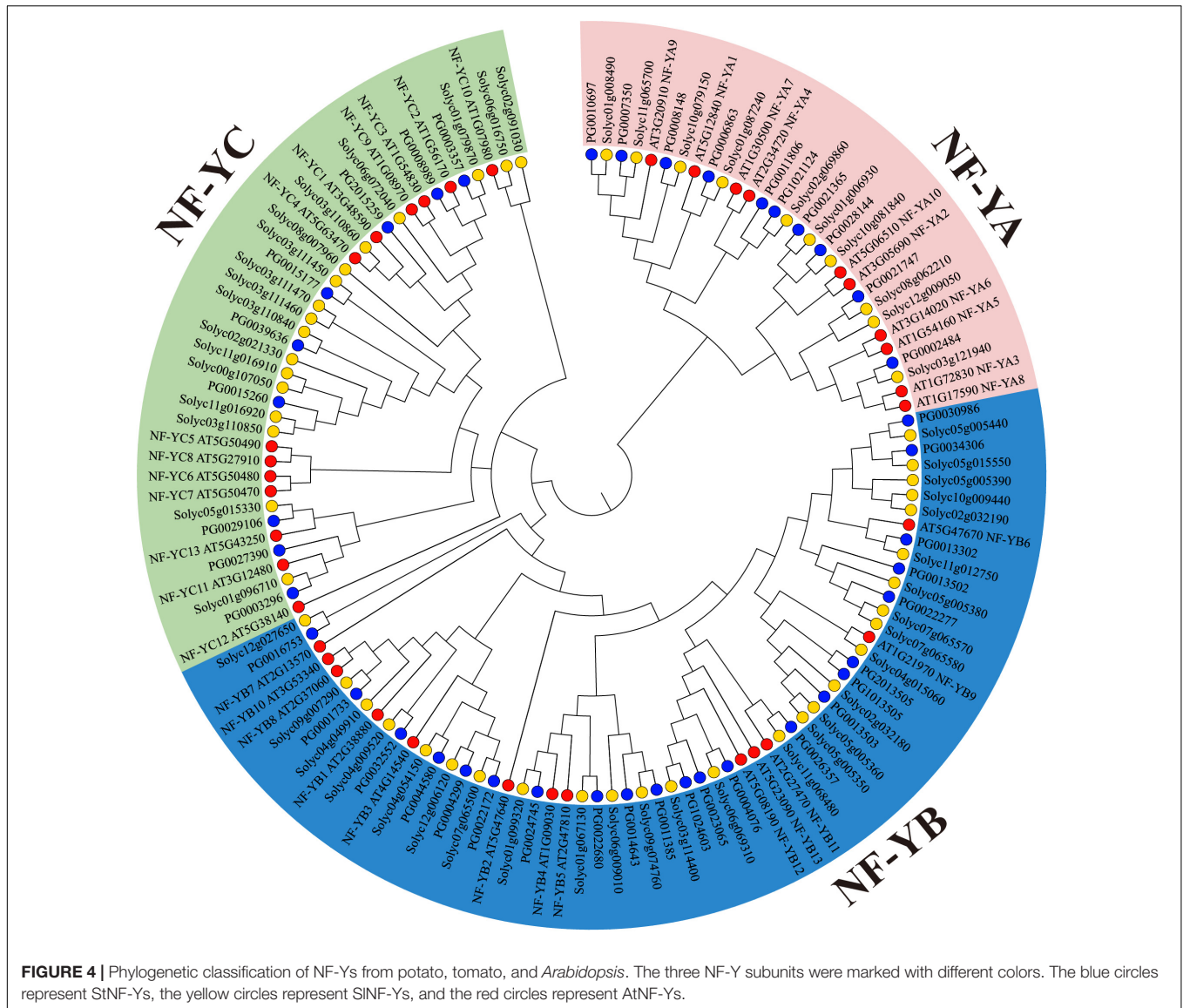
### Gene Duplication and Genome Synteny

Tandem and segmental duplications are crucial for the extension of gene family members and the realization of new functions (Cannon et al., 2004). In the present study, three pairs of *StNF-Y* genes (5/41, 12.20%) were identified as tandemly duplicated genes, of which chromosome 3 had two pairs and chromosome 5 had one pair (Figure 2A and Supplementary Table 4). In addition to tandem duplication, six pairs (10/41, 24.39%) of segmental duplication events were identified (Figure 6 and Supplementary Table 4).

In order to further investigate the potential evolution processes of the NF-Y gene family, we analyzed the synteny relationship of *NF-Ys* among potato, *Arabidopsis*, and tomato. A total of 14 pairs of orthologs were identified between potato and

*Arabidopsis*, and 38 pairs of orthologs between potato and tomato were identified (Figure 6 and Supplementary Tables 5, 6).

The substitution rate of non-synonymous ( $K_a$ ) and synonymous ( $K_s$ ) is the basis for evaluating the positive selection pressure of duplication events.  $K_a/K_s = 1$  indicates neutral selection,  $K_a/K_s < 1$  denotes purification selection, and  $K_a/K_s > 1$  signifies positive selection. The  $K_a/K_s$  of duplication *NF-Y* genes was calculated using KaKs Calculator 2.0. The results showed that the  $K_a/K_s$  of tandem and segmental duplications ranged from 0.0787 to 0.9482, with a mean value of 0.3569 (Supplementary Table 4). The  $K_a/K_s$  of the orthologous relationship between potato and *Arabidopsis* ranged from 0.0305 to 0.3507, with a mean value of 0.1736 (Supplementary Table 5). And, the  $K_a/K_s$  of the orthologous relationship between potato



and tomato ranged from 0.001 to 0.9260, with a mean value of 0.2863 (**Supplementary Table 6**). The Ka/Ks values were  $< 1$ , suggesting that these genes had evolved under the effect of purifying selection.

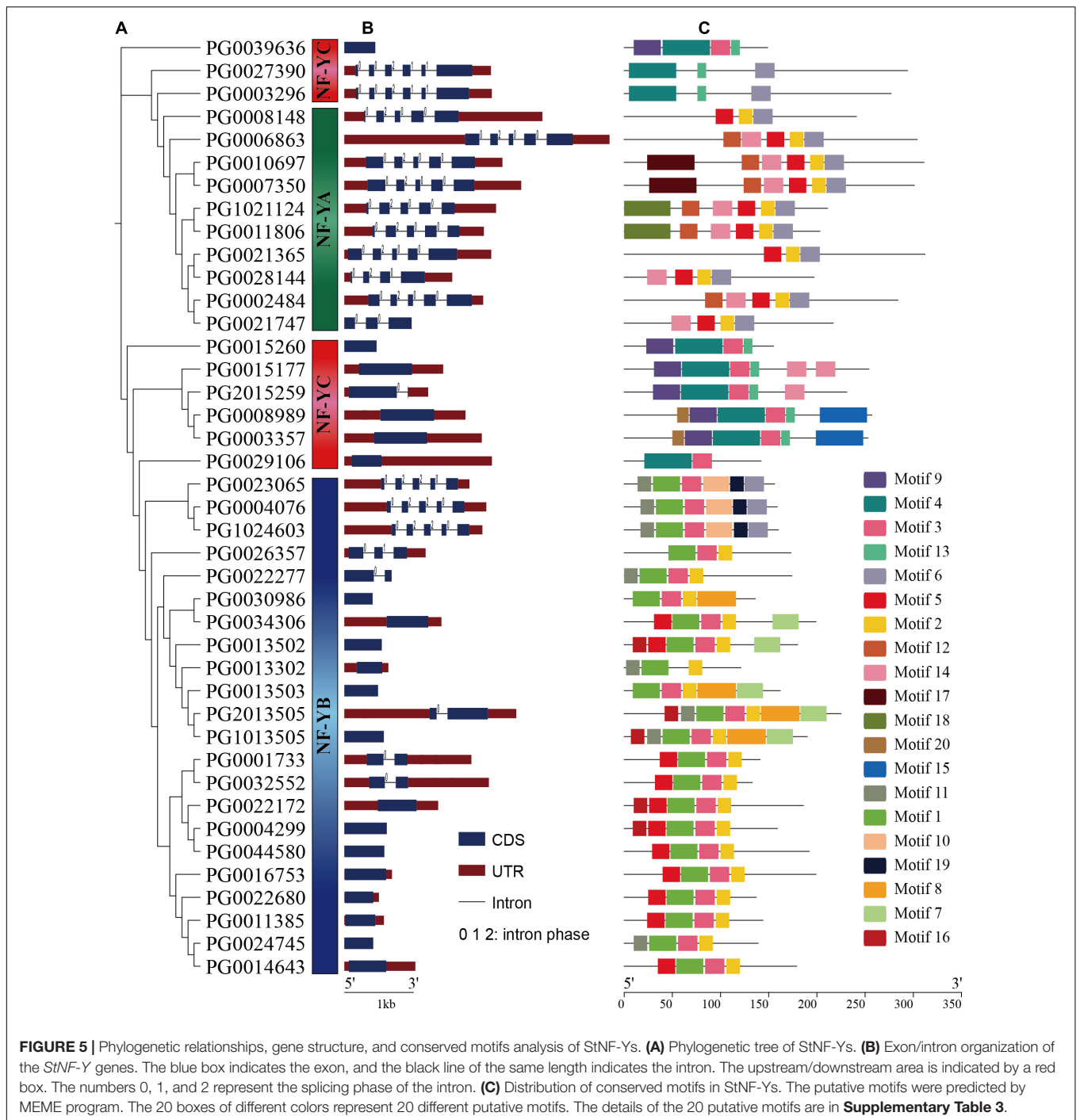
### Expression Profiles of StNF-Y Genes in Different Tissues

Based on transcriptome data downloaded from PGSC, the expression patterns of *StNF-Y* genes in 13 tissues (leaves, roots, shoots, tubers, sepals, stamens, stolons, mature flowers, petioles, petals, carpels, immature fruit, and mature fruit) of DM potato were analyzed (**Figure 7** and **Supplementary Table 7**). The results showed that 14 *StNF-Y* genes (*PG0001733*, *PG0003296*, *PG0003357*, *PG0004076*, *PG0004299*, *PG0006863*, *PG0007350*, *PG0008989*, *PG0010697*, *PG2015259*, *PG1021124*, *PG1024603*, *PG0026357*, and *PG0027390*) were highly expressed in all tissues with FPKM  $> 5$ . Four *StNF-Y* genes (*PG00011385*,

*PG1013505*, *PG0039636*, and *PG0024745*) were not expressed in all tissues (FPKM = 0). In addition, some *StNF-Y* genes showed tissue-specific expression patterns, e.g., *PG0021365* was highly expressed in stolon and *PG0044580* was highly expressed in the tuber. It is noteworthy that six *StNF-YB* genes (*PG0013502*, *PG0013503*, *PG0015177*, *PG0022277*, *PG0030986*, and *PG0034306*) were specifically expressed in immature fruit with FPKM  $> 5$  and  $\log_2FC > 1$ , which were in the same cluster as *AtNF-YB9* (*LEC1*) and *AtNF-YB6* (*LEC1-like*).

### Expression Profiles of StNF-Y Genes Under Abiotic Stresses and Hormone Treatments

The RNA-seq data was downloaded from the PGSC database to analyze the expression profiles of *StNF-Y* genes in DM potato under abiotic stresses. The results showed that six, nine, and 11 *StNF-Y* genes were differentially expressed

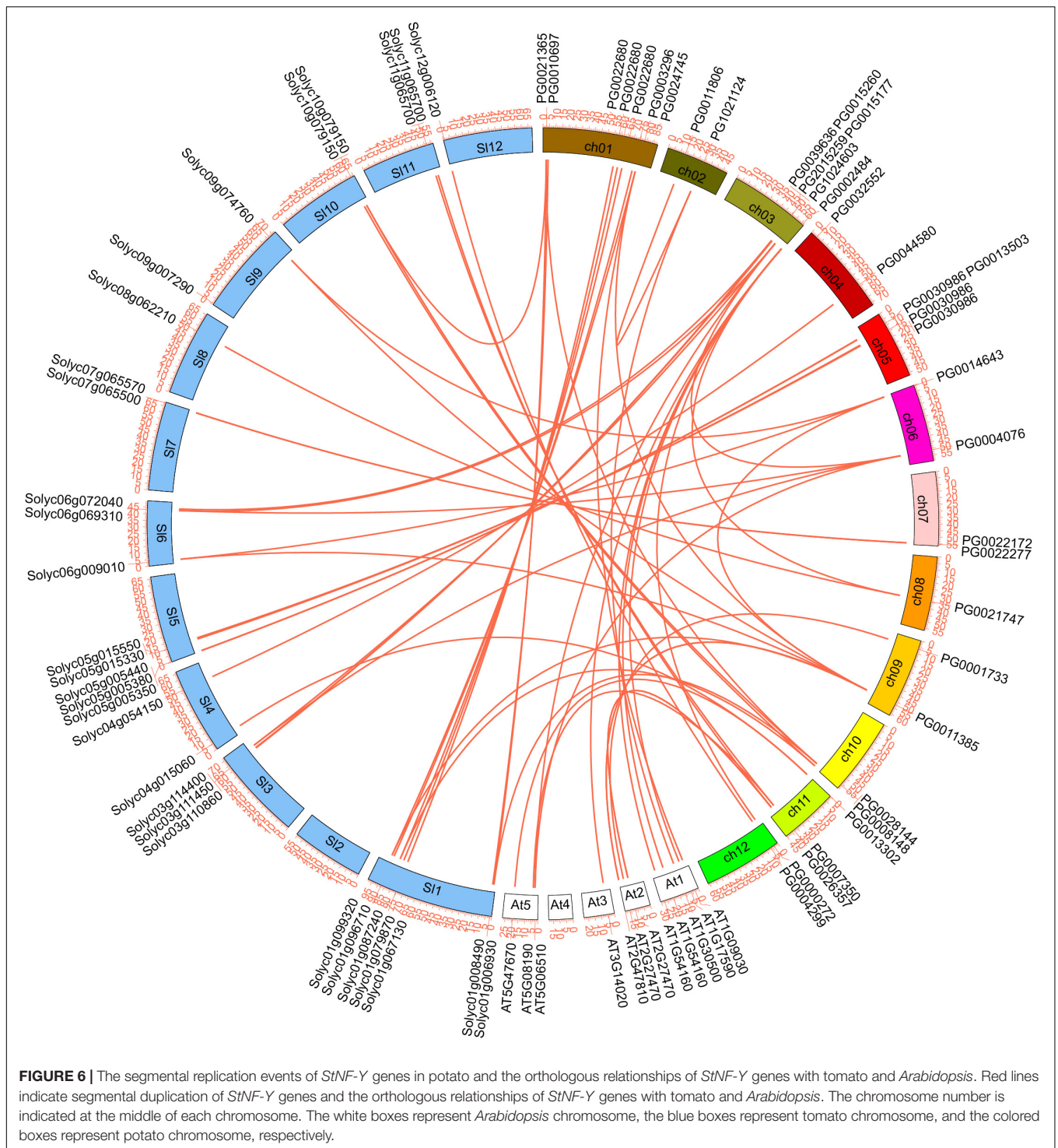


(FPKM > 1 and  $|\log_2FC| > 1$ ) under salt, mannitol, and heat treatments, respectively. Among them, four, seven, and nine *StNF-Y* genes were upregulated under salt, mannitol, and heat treatments, respectively. In addition, two *StNF-Y* genes (*PG2015259* and *PG0021365*) were differentially expressed under three stresses, five *StNF-Y* genes (*PG0002484*, *PG0004299*, *PG0022172*, *PG0028144*, and *PG0044580*) were differentially expressed under two stresses, and 10 genes (*PG0004076*, *PG0008148*, *PG0008989*, *PG0011385*, *PG0014643*, *PG1021124*,

*PG0021747*, *PG0022680*, *PG0023065*, and *PG0032552*) were only differentially expressed under single abiotic stress (**Figure 8A** and **Supplementary Table 8**).

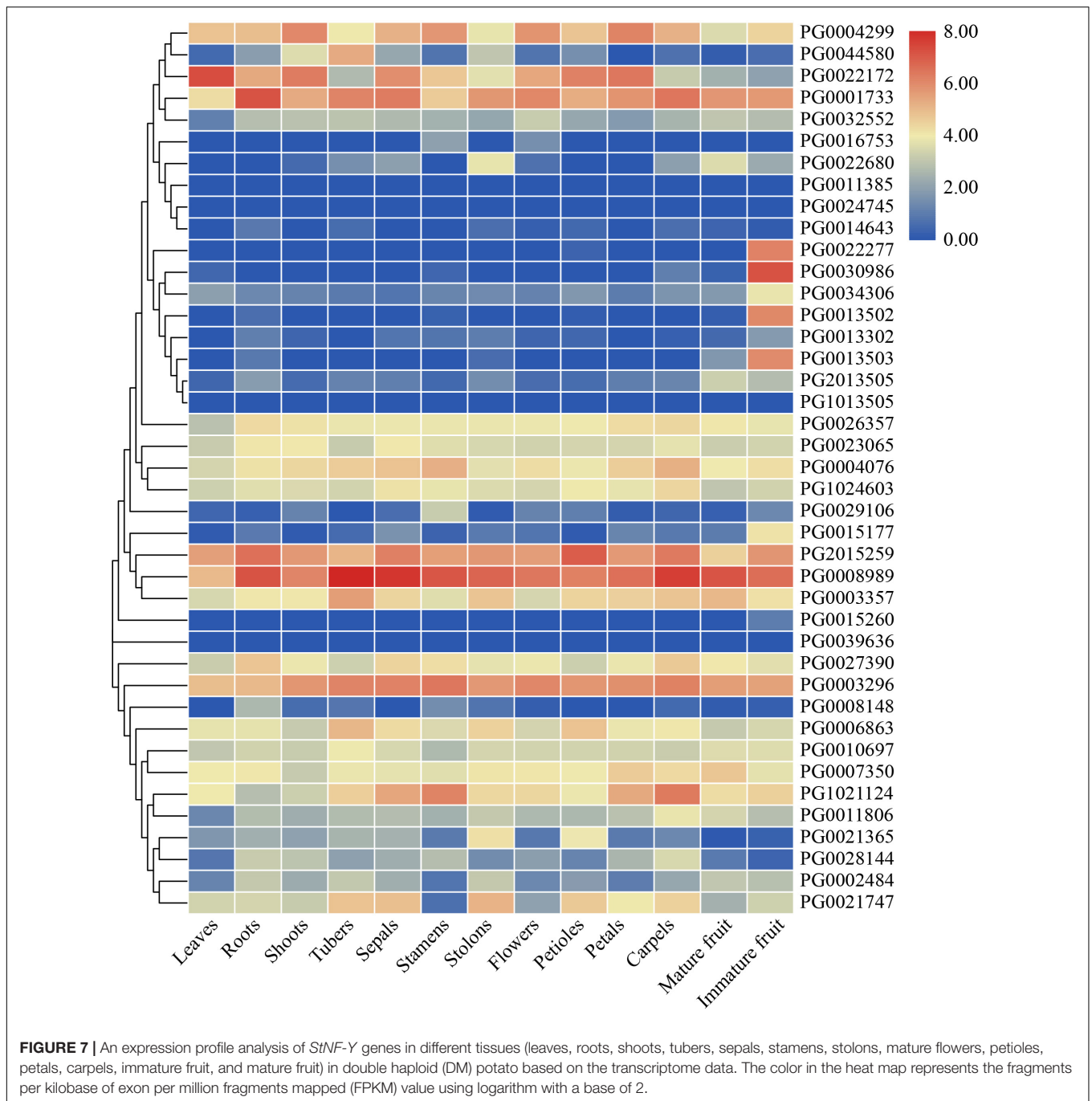
To further explore the expression changes in the *StNF-Y* genes under various hormone treatments (BAP, ABA, IAA, and GA3), the expression of 41 *StNF-Y* genes was analyzed using the RNA-seq data downloaded from the PGSC database. The results indicated that 22 and 10 *StNF-Y* genes responded to BAP and ABA treatments (FPKM > 1 and  $|\log_2FC| > 1$ ),





respectively. All of 22 *StNF-Y* genes were downregulated under BAP treatments, and all of 10 *StNF-Y* genes were upregulated under ABA treatments (FPKM > 1 and  $|\log_2FC| > 1$ ). However, all of *StNF-Y* genes were not differentially expressed under IAA and GA3 treatments, implying that *StNF-Y* genes did not respond to IAA and GA3 treatments (Figure 8A and Supplementary Table 8).

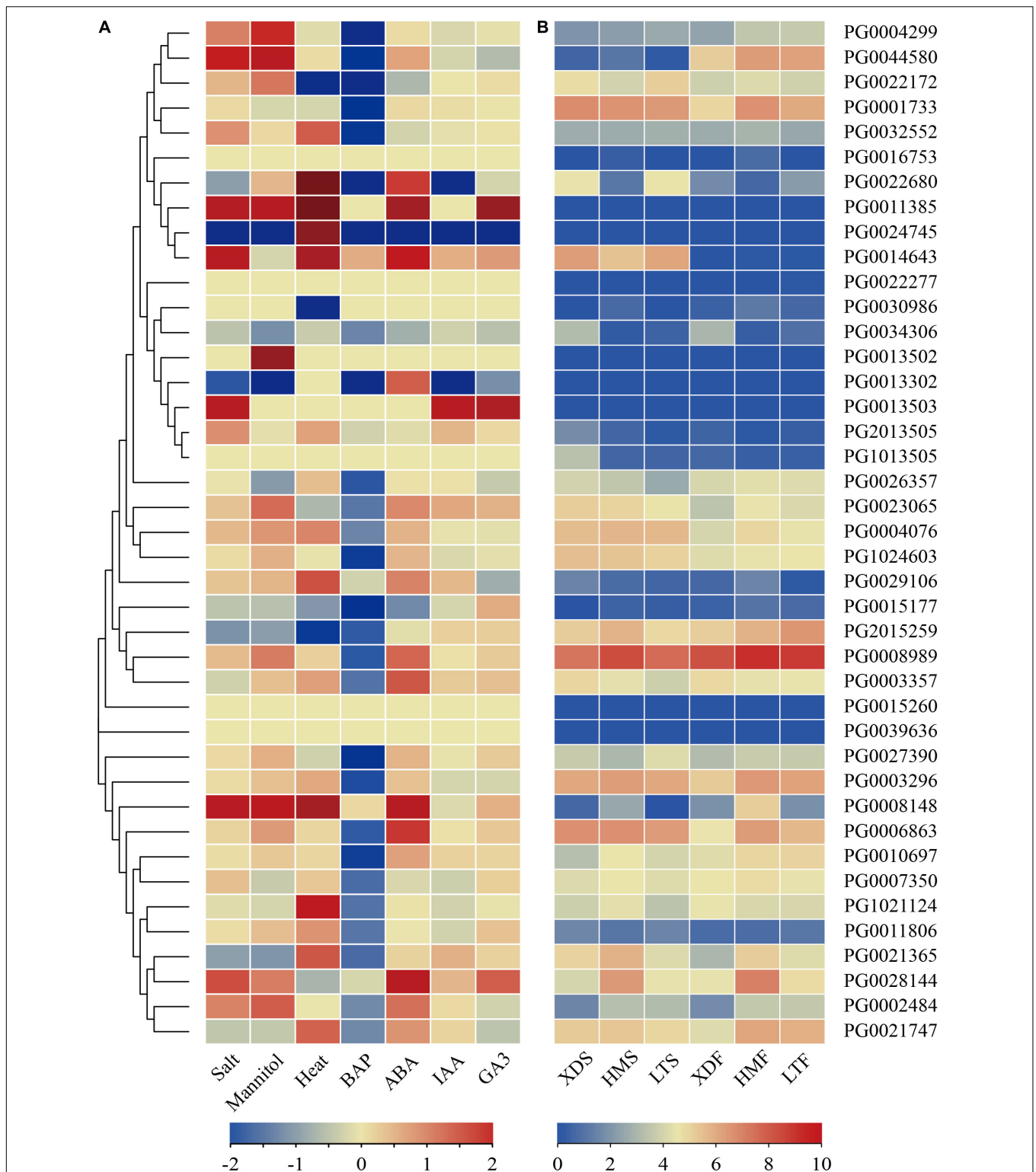
Combined with RNA-seq data analysis under abiotic stress and hormone treatment, four *StNF-Y* genes (*PG0002484*, *PG0008989*, *PG0023065*, and *PG0028144*) were differentially expressed (FPKM > 1 and  $|\log_2FC| > 1$ ) under mannitol and ABA treatments in DM potato. The results indicate that these four *StNF-Y* genes may be involved in drought stress response through an ABA-dependent pathway.



## Expression Profiles of *StNF-Y* Genes in Pigmented Potato Cultivars

To identify the potential functions of *StNF-Y* genes in flavonoid biosynthesis of potato, the expression patterns of the *StNF-Y* genes in tuber tissues (skin and flesh) of tetraploid pigmented potato cultivars were analyzed according to RNA-seq dataset. The results showed that eight *StNF-Y* genes (*PG0011385*, *PG0013502*, *PG0013503*, *PG0013302*, *PG0015260*, *PG0024745*, *PG0039636*, and *PG0022277*) were not expressed (FPKM = 0) in the skin, and 12 *StNF-Y* genes

were differentially expressed in pigmented skin compared with the white skin of XD (FPKM > 1 and  $|\log_2FC| > 1$ ). Among them, six genes (*PG0002484*, *PG0008148*, *PG0008989*, *PG0010697*, *PG0028144*, and *PG0044580*) were upregulated and four genes (*PG0022680*, *PG0034306*, *PG1013505*, and *PG2013505*) were downregulated in LTS and one gene (*PG0002484*) was upregulated and five genes (*PG0003357*, *PG0026357*, *PG0034306*, *PG1013505*, and *PG2013505*) were downregulated in HMS. In pigmented skin, one gene (*PG0002484*) was upregulated both in the skin of HM and LT,



**FIGURE 8 |** The expression profiles of *StNF-Y* genes. **(A)** The expression profiles of *StNF-Y* genes under abiotic stresses (salt, mannitol, and heat stress) and hormone treatments (BAP, ABA, IAA, and GA3) in DM potato. The color scale was plotted using the  $\log_2FC$  of each gene. **(B)** The expression profiles of *StNF-Y* genes in white and pigmented potato tuber skin and flesh. XDS, LTS, HMS represent the white skin of white potato cultivar (Xindaping, XD), the red skin of red potato cultivar (Lingtianhongmei, LT), and the purple skin of purple potato cultivar (Heimeiren, HM), respectively. XDF, LTF, and HMF represent the white flesh of XD, the red flesh of LT, and the purple flesh of HM, respectively. The color in the heat map represents the FPKM value using logarithm with a base of 2. The heatmap was constructed by TBtools software. BAP, benzylaminopurine; ABA, abscisic acid; IAA, indole acetic acid; GA3, gibberellic acid.

and three genes (*PG0034306*, *PG1013505*, and *PG2013505*) were downregulated both in the skin HM and LT (**Figure 8B** and **Supplementary Table 9**).

In tuber flesh, eight *StNF-Y* genes were not expressed (FPKM = 0), and these eight genes also were not expressed in the skin. Compared with the white flesh of XD, a total of 17 *StNF-Y* genes were differentially expressed (FPKM > 1 and  $|\log_2FC| > 1$ ) in pigmented flesh, of which eight genes (*PG0001733*, *PG0002484*, *PG0003296*, *PG0004299*, *PG0006863*, *PG0021365*, *PG0021747*, and *PG0044580*) were upregulated in both LTF and HMF and one gene (*PG0034306*) was downregulated in pigmented flesh (**Figure 8B** and **Supplementary Table 9**). It is noteworthy that one gene (*PG0002484*) was upregulated, and one gene (*PG0034306*) was downregulated in pigmented tissues (skin and flesh).

Eight *StNF-Y* genes with relatively high expression levels in pigmented tissues were selected as targets, and the reliability of RNA-seq dataset was verified by qPCR. The results confirmed that the qPCR expression patterns were in agreement with the RNA-seq dataset (**Figure 9A**), showing a high correlation ( $R^2 = 0.8393$ ) between the RNA-seq dataset and the simple linear regression equation of qPCR ( $y = 0.6430x + 0.0924$ ), indicating the consistency of the two analysis methods (**Figure 9B**).

## The Interaction Network of StNF-Y Proteins

Previous studies have shown that NF-Y subunits are involved in flavonoid synthesis, and they play a role through interaction with each other (Hackenberg et al., 2012; Mu et al., 2013; Wang et al., 2020). We further investigated the interaction between *StNF-Y* proteins that might be involved in flavonoid biosynthesis. Based on the RNA-seq dataset of pigmented potato cultivars, we selected highly expressed *StNF-Y* (FPKM > 5) to predict the protein interaction network with median confidence (combination score > 400). The results showed that a total of 89 pairs of interacting proteins were predicted. Twenty-one *StNF-Y* proteins (eight *StNF-Y*As, 10 *StNF-Y*Bs, and three *StNF-Y*Cs) were included in the protein interaction network, of which, eight *StNF-Y* genes (*PG0003296*, *PG0004299*, *PG0021365*, *PG0002484*, *PG0006863*, *PG0021747*, *PG0044580*, and *PG0034306*) were differentially expressed in pigmented tissues compared to that in white tissues. These eight *StNF-Y*s contained four *StNF-Y*A subunits, three *StNF-Y*Bs, and one *StNF-Y*C. In addition, we found that *StNF-Y*A subunits did not interact with each other at high confidence (combined score > 400), whereas they showed a direct interaction with *StNF-Y*Bs or *StNF-Y*Cs (**Figure 10** and **Supplementary Tables 10, 11**).

## DISCUSSION

NF-Ys are important transcription factors in plants, which play critical roles in plant growth, development, and stress responses (Yan et al., 2013; Liu et al., 2016; Xu et al., 2016; Yu et al., 2020). However, little is known about NF-Ys in potato. In the present study, a total of 41 *StNF-Y* genes were identified, including 10 *StNF-Y*As, 22 *StNF-Y*Bs, and nine *StNF-Y*Cs. The

bioinformatics analysis included a study of the phylogenetic relationships between potato and *Arabidopsis* NF-Y proteins and an analysis of chromosomal location, gene structures, encoded protein conserved motifs, and gene duplication events of *StNF-Y* genes. In addition, the expression pattern of *StNF-Y* genes in different tissues and their response to various abiotic stresses and hormone treatments were analyzed. This study provided comprehensive information for further investigation of their biological functions and the evolution of *StNF-Y* gene family.

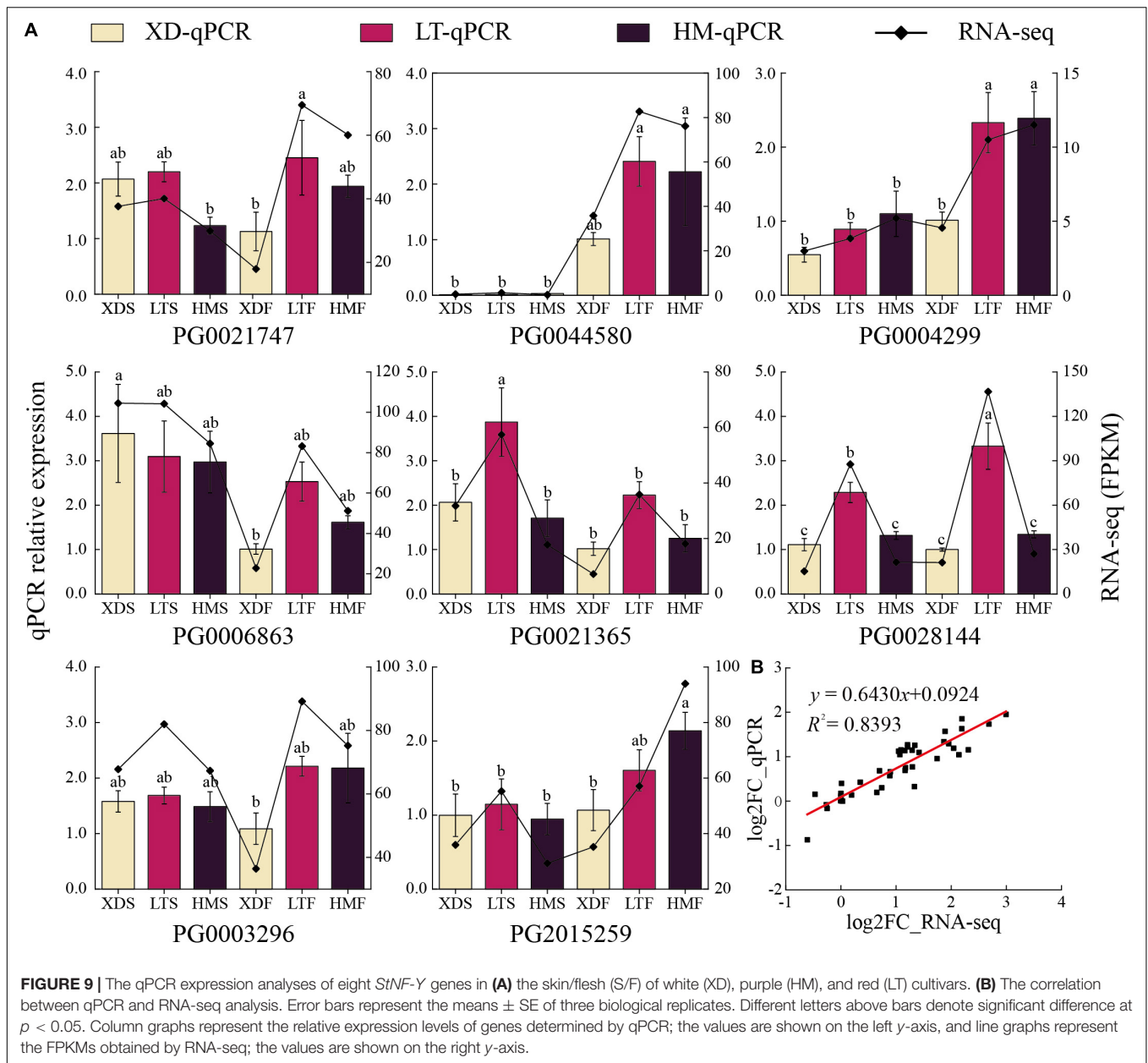
## Expansions of StNF-Y Gene Family in Potato

Gene duplication plays a major role in the expansion of gene families and may be an important driver of evolutionary diversification (Lawton-Rauh, 2003). In the present study, nine pairs of duplication genes were identified in potato, of which the number of *StNF-Y* genes arranged in tandem duplications contributed to 12.20% (5/41) and segmental duplication events contributed to 24.39% (10/41). These results suggested that tandem and segmental duplication events were the main mechanisms of *StNF-Y* gene family evolution and that some of the *StNF-Y* genes were possibly generated by gene duplication events.

Duplication events often lead to increased functional diversity between duplicated gene pairs, and some new members may acquire new functions or lose their original functions (Dias et al., 2003; Singh et al., 2012). In our study, three *StNF-Y* genes (*PG0039636*, *PG0015260*, and *PG2015259*) were tandem duplication, of which *PG2015259* was downregulated under various abiotic stresses (salt, mannitol, and heat) and highly expressed in 13 tissues (FPKM > 5), while the *PG0039636* and *PG0015260* were not expressed (FPKM = 0), indicating that *PG0039636* and *PG0015260* may become pseudogenes during duplication events and evolution. In addition, some duplication gene pairs were expressed consistently, but some were expressed inconsistently, such as *PG0021747/PG0002484* and *PG0021365/PG0028144* that were segmental duplication gene pairs. Among them, *PG0021747* and *PG0002484* were upregulated in pigmented flesh; however, *PG0021365* was downregulated and *PG28144* was upregulated under salt and mannitol stresses. In conclusion, gene duplication provided a large number of functional or non-functional genes for the potato genome.

## Phylogenetic Analysis and Evolution of StNF-Ys

To investigate the relationship of NF-Y proteins among potato, tomato, and *Arabidopsis*, an unroot phylogenetic tree was constructed using complete NF-Y protein sequences. The results showed that there were 10 NF-YA in all three plants (potato, tomato, and *Arabidopsis*). However, the amounts of NF-YB and NF-YC varied widely; there are 13 YBs and 13 YCs in *Arabidopsis*, 22 YBs and 9 YCs in potato, and 27 YBs and 17 YCs in tomato, respectively, suggesting that NF-YB and NF-YC play an important role in the expansion of NF-Y gene family.

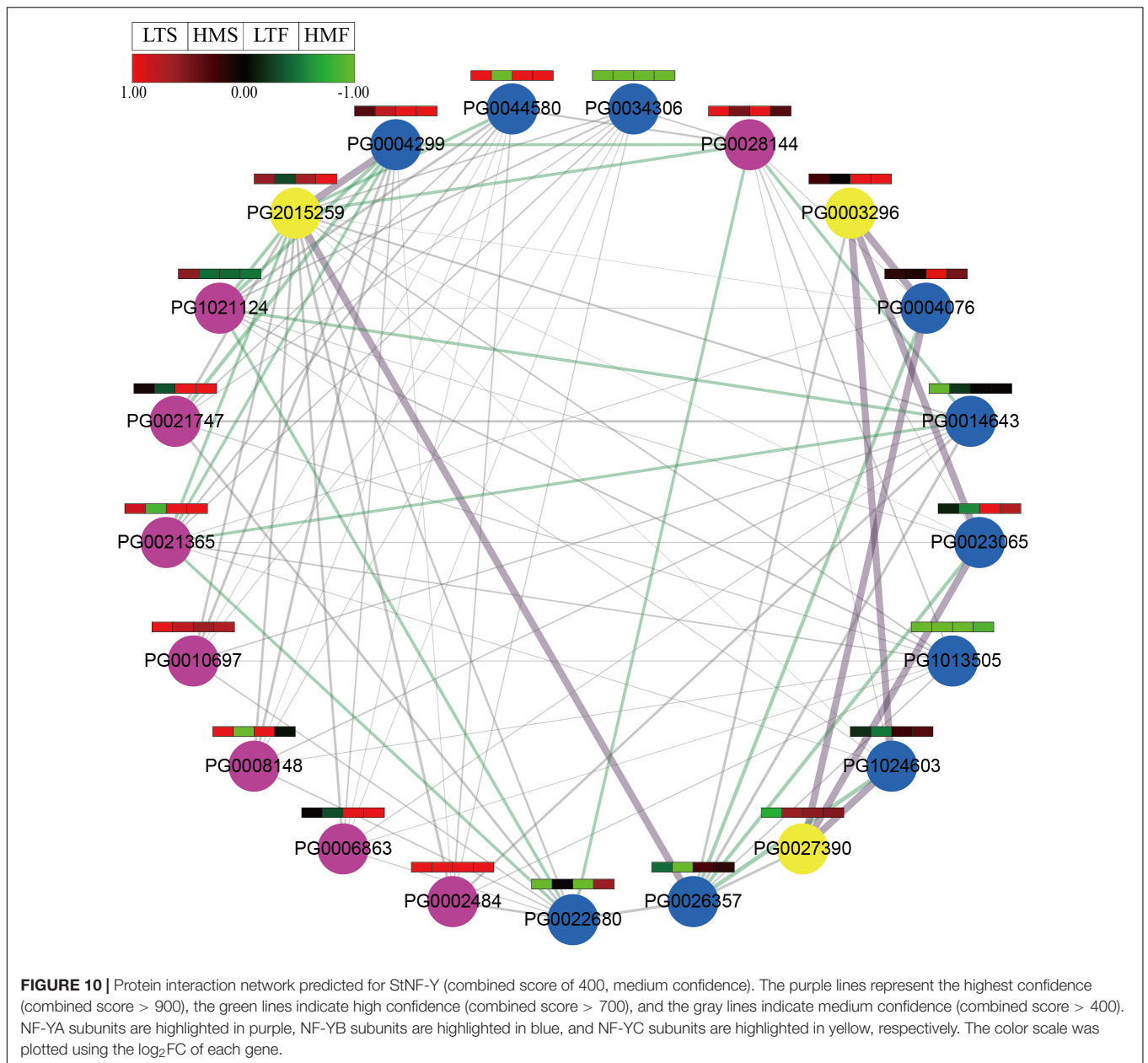


Studies have shown that *AtNF-YC1/2/3/4/9* controls flowering time (Siefers et al., 2009; Kumimoto et al., 2010; Hackenberg et al., 2012). Three *StNF-YCs* (PG0003357, PG0008989, and PG2015259) were clustered with *AtNF-YC1/2/3/4/9*, indicating that these *StNF-Ys* may have a similar function as *AtNF-YC1/2/3/4/9*. Previous studies have reported that the *AtLEC1* and *AtLEC1*-like play an important role in embryo development and fruit ripening (Kwong et al., 2003; Li et al., 2016; Gago et al., 2017). It is noteworthy that *AtLEC1* (*AtNF-YB9*) and *AtLEC1*-like (*AtNF-YB6*) formed a distinct cluster in *NF-YB* (Figure 4). Our study found that eight *StNF-YB* genes were clustered with *AtLEC1* and *AtLEC1*-like in one cluster, of which six *StNF-YB* genes were specifically expressed in immature fruits. It is suggesting that *LEC1* and *LEC1*-like genes were extensively

extended during potato evolution and that these genes may be associated with embryo development and fruit ripening.

### Expression Profile Analysis of *StNF-Y* Genes Under Abiotic Stress and Hormone Treatment

Previous studies have shown that the *NF-Y* family is involved in various abiotic stresses (Li L. et al., 2013; Li Y. J. et al., 2013; Alam et al., 2015; Feng et al., 2015). In *Arabidopsis*, *AtNF-YA1* (*At5G12840*) is significantly induced by NaCl, mannitol, PEG, and ABA, and it is involved in the regulation of post-germination growth arrest under salt stress (Li Y. J. et al., 2013). *PG0006863* was homologous to *AtNF-YA1*, and it was highly



expressed under mannitol treatment and upregulated under ABA treatment in DM potato, indicating that *PG0006863* may function in a similar way as *AtNF-YA1* under mannitol treatment. Li et al. (2008) found that overexpression of *AtNF-YA5* (*AT1G54160*) in *Arabidopsis* can increase the tolerance of plants to drought stress through the ABA-dependent pathway. *PG0002484* was homologous to *AT1G54160*, which was upregulated under salt, mannitol, and ABA treatments in DM potato, suggesting that *PG0002484* may have similar functions to *AtNF-YA5*. Overexpression of *AtNF-YA7* (*AT1G30500*) enhanced tolerance to various abiotic stresses in *Arabidopsis* (Leyva-González et al., 2012). *PG1021124*, homologous to *AtNF-YA7*, was upregulated under heat stress and differentially expressed under various hormone treatments in DM potato, suggesting that *PG1021124*

may have similar functions to *AtNF-YA7* under abiotic stresses. Overexpression of *NF-YA10* can improve tolerance to drought and salt stress in *Arabidopsis* (Ma et al., 2015; Yu et al., 2020). *PG0028144* and *PG0021365* were a pair of segmental duplication genes, which were clustered with *AtNF-YA10* (*AT6G06510*). *PG0021365* was differentially expressed under mannitol and heat treatments, while *PG0028144* was upregulated under salt, mannitol, and ABA treatments. The results indicated that *PG0028144* may be induced by drought stress in an ABA-dependent manner, while *PG0021365* was not. Sato et al. (2019) found that overexpression of *AtNF-YB2* (*AT5G47640*) and *AtNF-YB3* (*AT4G14540*) specifically enhances drought and heat stress tolerance, respectively. *PG0004299*, *PG0044580*, and *PG0022172* were clustered with *AtNF-YB2* and *AtNF-YB3* in

one cluster. Among them, *PG0004299* and *PG0044580* had the same expression pattern, which were upregulated under salt and mannitol treatments, indicating that *PG0004299* and *PG0044580* may have similar functions to *AtNF-YB2*. The expression of *PG0022172* was upregulated under mannitol treatment, and downregulated under heat treatment, suggesting that *PG0022172* may have functions of both *AtNF-YB2* and *AtNF-YB3*. In addition, *PG0008989* and *PG0023065* were upregulated under mannitol and ABA treatments, while *PG2015259* was downregulated under three abiotic stresses (salt, mannitol, and heat), suggesting that these *StNF-Y* genes might be putative regulators of response to abiotic stress.

## Expression of *StNF-Y* Genes in Pigmented Tuber Tissue

The expression profiles of *StNF-Y* genes in different tuber tissues (skin and flesh) of tetraploid pigmented potato cultivars were analyzed based on the RNA-seq dataset. Studies have shown that the NF-Y genes play crucial roles in flavonoid biosynthesis, e.g., NF-YB8 interacts with NF-YC1 and then recruits NF-YA1/9 to form NF-Y complexes that bind the CCAAT element in the CHS1 promoter to regulate the flavonoid biosynthesis in tomato (Wang et al., 2020). *PG0001733* was homologous to *AtNF-YB8* (*AT2G37060*) and clustered with *SINF-YB8c* in one cluster, which was upregulated in pigmented flesh (LTF and HMF). *PG2015259* was clustered with *AtNF-YC1* (*AT3G48590*) and *SINF-YC1a* (*Solyc03g110860*) in one cluster and was highly expressed in LTF and upregulated in HMF. *PG0006863* and *PG0008148* were homologous to *SINF-YA9* (*Solyc01g087240*), of which *PG0006863* was upregulated in LTF and HMF and *PG0008148* was upregulated in LTS and LTF. In addition, *PG0010697* was homologous to *SINF-YA1a* (*Solyc01g008490*) and was highly expressed in pigmented tissues and upregulated in LTS. These results indicated that these five genes may have similar functions to *SINF-YA1/9*, *SINF-YB8c*, and *SINF-YC1a* in flavonoid biosynthesis. Furthermore, six genes were upregulated in pigmented tissues (LTF and HMF), of which three genes (*PG0002484*, *PG0021365*, and *PG0021747*) belonged to *StNF-YA* subunit, two genes (*PG0004299* and *PG0044580*) were assigned to *StNF-YB* subunit, and one gene (*PG0003296*) was grouped into *StNF-YC* subunit.

In addition, we analyzed the correlation coefficient (CC) between differentially expressed *StNF-Ys* of potato tuber skin and flesh in RNA-seq data and anthocyanin content. The results showed that the expression of some *StNF-Y* genes was highly correlated to anthocyanin content in tuber flesh, such as *PG0044580* and *PG0021747* with  $CC > 0.7$ ; *PG0004299*, *PG0002484*, and *PG0010697* with  $CC > 0.8$ ; and *PG2015259* ( $CC = 0.8$ ) (Supplementary Table 12), suggesting that these *StNF-Y* genes may be involved in the anthocyanin biosynthesis in potato tuber flesh, and their functions require further investigation.

Through the prediction of *StNF-Y* protein interaction network, we found that there were interaction relationships (high confidence with combined score  $> 700$ ) between these differentially expressed genes (Figure 10 and

Supplementary Table 10), i.e., two *StNF-YAs* (*PG0021747* and *PG0021365*) directly interacted with one *StNF-YB* (*PG0004299*), and one *StNF-YB* (*PG0004299*) and one *StNF-YA* (*PG0021365*) interacted with one *StNF-YC* (*PG201259*). These *StNF-Ys* may be involved in the biosynthesis of anthocyanins by interacting with each other to form the completed NF-Y complex, and their functions need to be further studied.

## CONCLUSION

In this study, we identified a total of 41 *StNF-Y* genes from the potato genome, including 10 *StNF-YA*, 22 *StNF-YB*, and nine *StNF-YC*, and their distribution on chromosomes, multiple alignments, gene structure, conserved motifs, and phylogenetic relationships were analyzed. Synteny analysis showed that nine pairs of duplication genes were identified in the potato genome, which played an important role in the expansion of the *StNF-Y* gene family. Synteny analysis indicated that 14 and 38 pairs of *StNF-Y* genes were orthologous to *Arabidopsis* and tomato, respectively. Six *StNF-YB* genes were specifically expressed in immature fruit, which may be involved in fruit ripening. In addition, the *StNF-Y* genes displayed different expression patterns under the salt, mannitol, heat, and hormone treatments in DM potato. Furthermore, we analyzed the expression of *StNF-Y* genes in tetraploid pigmented potato cultivars and the protein interaction networks that were highly expressed in anthocyanin biosynthesis. The results of this study provide valuable information for further studies on the functions of the *StNF-Y* gene family.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author/s.

## AUTHOR CONTRIBUTIONS

YHL, JB, and JLZ conceptualized the study. ZL, YML, JYZ, WM, ZB, and ZTL curated the data. ZL and CS conducted the formal analysis. ZL, YHL, and JLZ acquired funding. ZL contributed the software and wrote the original draft. YHL wrote, reviewed, and edited the manuscript. All authors contributed to the article and approved the submitted version.

## FUNDING

This research was funded by the Scientific Research Start-up Funds for Openly-recruited Doctors of Gansu Agricultural University (GAU-KYQD-2020-11), the Program for Innovation Ability Improvement of the Higher Education Institutions of Gansu Province (2020A-056), the Program for Longyuan Youth Innovation Team of Gansu Province (LRYCT-2020-1), the National Natural Science Foundation of China (31601356 and 31860398), the Gansu Science Foundation for Distinguished

Young Scholars (17JR5RA138), the Fuxi Talent Project of Gansu Agricultural University (Ganfx-02Y04), and the China Agriculture Potato Research System of MOF and MARA (CARS-09-P14).

## ACKNOWLEDGMENTS

We would like to thank Yuanyuan Pu of Gansu Agricultural University for her help in the improvement of this manuscript.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2021.739989/full#supplementary-material>

**Supplementary Table 1** | Primers used for qPCR.

**Supplementary Table 2** | Characteristic features of 41 StNF-Y genes.

## REFERENCES

- Alam, M. M., Tanaka, T., Nakamura, H., Ichikawa, H., Kobayashi, K., Yaeno, T., et al. (2015). Overexpression of a rice heme activator protein gene (OsHAP2E) confers resistance to pathogens, salinity and drought, and increases photosynthesis and tiller number. *Plant Biotechnol. J.* 13, 85–96. doi: 10.1111/pbi.12239
- Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402. doi: 10.1093/nar/25.17.3389
- Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., et al. (2009). MEME Suite: tools for motif discovery and searching. *Nucleic Acids Res.* 37, W202–W208.
- Cannon, S. B., Mitra, A., Baumgarten, A., Young, N. D., and May, G. (2004). The roles of segmental and tandem gene duplication in the evolution of large gene families in *Arabidopsis thaliana*. *BMC Plant Biol.* 4:10. doi: 10.1186/1471-2229-4-10
- Cao, S., Kumimoto, R. W., Gnesutta, N., Calogero, A. M., Mantovani, R., and Holt, B. F. III (2014). A distal CCAAT/NUCLEAR FACTOR Y complex promotes chromatin looping at the FLOWERING LOCUS T promoter and regulates the timing of flowering in *Arabidopsis*. *Plant Cell* 26, 1009–1017. doi: 10.1105/tpc.113.120352
- Castañeda-Ovando, A., Pacheco-Hernández, M. D. L., Páez-Hernández, M. E., Rodríguez, J. A., and Galán-Vidal, C. A. (2009). Chemical studies of anthocyanins: a review. *Food Chem.* 113, 859–871. doi: 10.1016/j.foodchem.2008.09.001
- Castellarin, S. D., Pfeiffer, A., Sivilotti, M. D., Peterlunger, E., and Di Gasparo, G. (2007). Transcriptional regulation of anthocyanin biosynthesis in ripening fruits of grapevine under seasonal water deficit. *Plant Cell Environ.* 30, 1381–1399. doi: 10.1111/j.1365-3040.2007.01716.x
- Chen, C., Xia, R., Chen, H., and He, Y. (2018). TBtools, a Toolkit for Biologists integrating various HTS-data handling tools with a user-friendly interface. *BioRxiv* [Preprint] doi: 10.1101/289660v2
- Dias, A. P., Braun, E. L., McMullen, M. D., and Grotewold, E. (2003). Recently duplicated maize *R2R3 Myb* genes provide evidence for distinct mechanisms of evolutionary divergence after duplication. *Plant Physiol.* 131:610. doi: 10.1104/pp.012047
- Feng, Z.-J., He, G.-H., Zheng, W.-J., Lu, P., Chen, M., Gong, Y.-M., et al. (2015). Foxtail millet NF-Y families: genome-wide survey and evolution analyses identified two functional genes important in abiotic stresses. *Front. Plant Sci.* 6:1142. doi: 10.3389/fpls.2015.01142
- Supplementary Table 3** | Conserved motifs identified by MEME among StNF-Ys.
- Supplementary Table 4** | Ka/Ks ratios of tandemly and segmentally duplicated *StNF-Y* genes.
- Supplementary Table 5** | Ka/Ks ratios of the syntenic relationships for NF-Y genes between potato and *Arabidopsis*.
- Supplementary Table 6** | Ka/Ks ratios of the syntenic relationships for NF-Y genes between potato and tomato.
- Supplementary Table 7** | RNA-seq data of StNF-Y genes in different tissues.
- Supplementary Table 8** | RNA-seq data of StNF-Y genes under abiotic stress and hormone treatment.
- Supplementary Table 9** | The FPKM value of StNF-Y genes in white and pigmented tissues.
- Supplementary Table 10** | Protein interaction network of StNF-Ys.
- Supplementary Table 11** | The FPKM value of genes in protein interaction networks.
- Supplementary Table 12** | The correlation between the expression of StNF-Ys and anthocyanin content in potato tuber (skin and flesh).
- Fossen, T., and Andersen, ØM. (2000). Anthocyanins from tubers and shoots of the purple potato, *Solanum tuberosum*. *J. Horticult. Sci. Biotechnol.* 75, 360–363. doi: 10.1080/14620316.2000.11511251
- Gago, C. D., Paschalidis, K., Guerreiro, A., Miguel, G., Antunes, D., and Hilioti, Z. (2017). Targeted gene disruption coupled with metabolic screen approach to uncover the LEAFY COTYLEDON1-LIKE4 (L1L4) function in tomato fruit metabolism. *Plant Cell Rep.* 36, 1065–1082. doi: 10.1007/s00299-017-2137-9
- Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S. E., Wilkins, M. R., Appel, R. D., et al. (2005). “Protein identification and analysis tools on the ExPASy server,” in *The Proteomics Protocols Handbook*, ed. J. M. Walker (Totowa, NJ: Humana Press), 571–607. doi: 10.1385/1-59259-890-0:571
- Guo, A. Y., Zhu, Q. H., and Xin, C. (2007). GSDS: a gene structure display server. *Hereditas* 29, 1023–1026. doi: 10.1360/yc-007-1023
- Guo, F., Liu, C., Xia, H., Bi, Y., Zhao, C., Zhao, S., et al. (2013). Induced expression of AtLEC1 and AtLEC2 differentially promotes somatic embryogenesis in transgenic tobacco plants. *PLoS One* 8:e71714. doi: 10.1371/journal.pone.0071714
- Hackenberg, D., Wu, Y., Voigt, A., Adams, R., Schramm, P., and Grimm, B. (2012). Studies on differential nuclear translocation mechanism and assembly of the three subunits of the *Arabidopsis thaliana* transcription factor NF-Y. *Mol. Plant* 5, 876–888. doi: 10.1093/mp/ssp107
- He, J., and Giusti, M. M. (2010). Anthocyanins: natural colorants with health-promoting properties. *Annu. Rev. Food Sci. Technol.* 1, 163–187. doi: 10.1146/annurev.food.080708.100754
- Hollman, P., and Katan, M. B. (1997). Absorption, metabolism and health effects of dietary flavonoids in man. *Biomed. Pharmacother.* 51, 305–310. doi: 10.1016/s0753-3322(97)88045-6
- Kahle, J., Baake, M., Doenecke, D., and Albig, W. (2005). Subunits of the heterotrimeric transcription factor NF-Y are imported into the nucleus by distinct pathways involving importin beta and importin 13. *Mol. Cell Biol.* 25, 5339–5354. doi: 10.1128/mcb.25.13.5339-5354.2005
- Kim, J., Lee, W. J., Vu, T. T., Jeong, C. Y., Hong, S. W., and Lee, H. (2017). High accumulation of anthocyanins via the ectopic expression of AtDFR confers significant salt stress tolerance in *Brassica napus* L. *Plant Cell Rep.* 36, 1215–1224. doi: 10.1007/s00299-017-2147-7
- Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874.
- Kumimoto, R. W., Zhang, Y., Siefers, N., and Holt, B. F. III (2010). NF-YC3, NF-YC4 and NF-YC9 are required for CONSTANS-mediated, photoperiod-dependent flowering in *Arabidopsis thaliana*. *Plant J.* 63, 379–391. doi: 10.1111/j.1365-313x.2010.04247.x



- Kwong, R. W., Bui, A. Q., Lee, H., Kwong, L. W., Fischer, R. L., Goldberg, R. B., et al. (2003). LEAFY COTYLEDON1-LIKE defines a class of regulators essential for embryo development. *Plant Cell* 15, 5–18. doi: 10.1105/tpc.006973
- Laloum, T., De Mita, S., Gamas, M. B., and Niebel, A. (2013). CCAAT-box binding transcription factors in plants: y so many? *Trends Plant Sci.* 18, 157–166. doi: 10.1016/j.tplants.2012.07.004
- Lawton-Rauh, A. (2003). Evolutionary dynamics of duplicated genes in plants. *Mol. Phylogenet. Evol.* 29, 396–409. doi: 10.1016/j.ympev.2003.07.004
- Lee, H., Fischer, R. L., Goldberg, R. B., and Harada, J. J. (2003). *Arabidopsis* LEAFY COTYLEDON1 represents a functionally specialized subunit of the CCAAT binding transcription factor. *Proc. Natl. Acad. Sci. U.S.A.* 100, 2152–2156. doi: 10.1073/pnas.0437909100
- Leyva-González, M. A., Ibarra-Laclette, E., Cruz-Ramírez, A., and Herrera-Estrella, L. (2012). Functional and transcriptome analysis reveals an acclimatization strategy for abiotic stress tolerance mediated by *Arabidopsis* NF-YA family members. *PLoS One* 7:e48138. doi: 10.1371/journal.pone.0048138
- Li, L., Yu, Y., Wei, J., Huang, G., Zhang, D., Liu, Y., et al. (2013). Homologous HAP5 subunit from *Picea wilsonii* improved tolerance to salt and decreased sensitivity to ABA in transformed *Arabidopsis*. *Planta* 238, 345–356. doi: 10.1007/s00425-013-1894-0
- Li, S., Li, K., Ju, Z., Cao, D., Fu, D., Zhu, H., et al. (2016). Genome-wide analysis of tomato NF-Y factors and their role in fruit ripening. *BMC Genom.* 17:36. doi: 10.1186/s12864-015-2334-2
- Li, W. X., Oono, Y., Zhu, J., He, X. J., Wu, J. M., Iida, K., et al. (2008). The *Arabidopsis* NFYA5 transcription factor is regulated transcriptionally and posttranscriptionally to promote drought resistance. *Plant Cell* 20, 2238–2251. doi: 10.1105/tpc.108.059444
- Li, X. Y., Hoof van Huijsdijnen, R., Mantovani, R., Benoist, C., and Mathis, D. (1992). Intron-exon organization of the NF-Y genes. Tissue-specific splicing modifies an activation domain. *J. Biol. Chem.* 267, 8984–8990. doi: 10.1016/s0021-9258(19)50377-5
- Li, Y. J., Fang, Y., Fu, Y. R., Huang, J. G., Wu, C. A., and Zheng, C. C. (2013). NFYA1 is involved in regulation of postgermination growth arrest under salt stress in *Arabidopsis*. *PLoS One* 8:e61289. doi: 10.1371/journal.pone.0061289
- Liang, M., Yin, X., Lin, Z., Zheng, Q., Liu, G., and Zhao, G. (2014). Identification and characterization of NF-Y transcription factor families in canola (*Brassica napus* L.). *Planta* 239, 107–126. doi: 10.1007/s00425-013-1964-3
- Liu, X., Hu, M., Huang, Tang, Y., Li, Y., Li, L., and Hou, X. (2016). The NF-YC-RGL2 module integrates GA and ABA signalling to regulate seed germination in *Arabidopsis*. *Nat. Commun.* 7:12768.
- Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the  $2^{-\Delta\Delta CT}$  method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262
- Ma, X., Zhu, X., Li, C., Song, Y., Zhang, W., Xia, G., et al. (2015). Overexpression of wheat NF-YA10 gene regulates the salinity stress response in *Arabidopsis thaliana*. *Plant Physiol. Biochem.* 86, 34–43. doi: 10.1016/j.plaphy.2014.11.011
- Malviya, N., Jaiswal, P., and Yadav, D. (2016). Genome-wide characterization of Nuclear Factor Y (NF-Y) gene family of sorghum [*Sorghum bicolor* (L.) Moench]: a bioinformatics approach. *Physiol. Mol. Biol. Plants* 22, 33–49. doi: 10.1007/s12298-016-0349-z
- Mantovani, R. (1999). The molecular biology of the CCAAT-binding factor NF-Y. *Gene* 239, 15–27. doi: 10.1016/s0378-1119(99)00368-6
- Mu, J., Tan, H., Hong, S., Liang, Y., and Zuo, J. (2013). *Arabidopsis* transcription factor genes NF-YA1, 5, 6, and 9 play redundant roles in male gametogenesis, embryogenesis, and seed development. *Mol. Plant* 6, 188–201. doi: 10.1093/mp/sss061
- Nardini, M., Gnesutta, N., Donati, G., Gatta, R., Forni, C., Fossati, A., et al. (2013). Sequence-specific transcription factor NF-Y displays histone-like DNA binding and H2B-like ubiquitination. *Cell* 152, 132–143.
- Ni, Z., Hu, Z., Jiang, Q., and Zhang, H. (2013). GmNFYA3, a target gene of miR169, is a positive regulator of plant tolerance to drought stress. *Plant Mol. Biol.* 82, 113–129. doi: 10.1007/s11103-013-0040-5
- Petroni, K., Kumimoto, R. W., Gnesutta Calvenzani, N., Fornari, M., Tonelli, C., Holt, B. F., et al. (2012). The promiscuous life of plant nuclear factor Y transcription factors. *Plant Cell* 24:4777. doi: 10.1105/tpc.112.10.5734
- Riechmann, J. L., Heard, J., Martin, G., Reuber, L., Jiang, C., Keddie, J., et al. (2000). *Arabidopsis* transcription factors: genome-wide comparative analysis among eukaryotes. *Science* 290, 2105–2110. doi: 10.1126/science.290.5499.2105
- Sato, H., Mizoi, J., Tanaka, H., Maruyama, K., Qin, F., Osakabe, Y., et al. (2014). *Arabidopsis* DPB3-1, a DREB2A interactor, specifically enhances heat stress-induced gene expression by forming a heat stress-specific transcriptional complex with NF-Y subunits. *Plant Cell* 26, 4954–4973. doi: 10.1105/tpc.114.132928
- Sato, H., Suzuki, T., Takahashi, F., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2019). NF-YB2 and NF-YB3 have functionally diverged and differentially induce drought and heat stress-specific genes. *Plant Physiol.* 180, 1677–1690. doi: 10.1104/pp.19.00391
- Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504.
- Siefers, N., Dang, K. K., Kumimoto, R. W., Bynum, W. E. T., Tayrose, G., and Holt, B. F. III (2009). Tissue-specific expression patterns of *Arabidopsis* NF-Y transcription factors suggest potential for extensive combinatorial complexity. *Plant Physiol.* 149, 625–641. doi: 10.1104/pp.108.130591
- Singh, A., Pandey Baranwal, A., Kapoor, S., and Pandey, G. K. (2012). Comprehensive expression analysis of rice phospholipase D gene family during abiotic stresses and development. *Plant Signal. Behav.* 7, 847–855. doi: 10.4161/psb.20385
- Sorin, C., Declerck, M., Christ, A., Blein, T., Ma, L., Lelandais-Brière, C., et al. (2014). A miR169 isoform regulates specific NF-YA targets and root architecture in *Arabidopsis*. *New Phytol.* 202, 1197–1211. doi: 10.1111/nph.12735
- Spooner, D. M., McLean, K., Ramsay, G., Waugh, R., and Bryan, G. J. (2005). A single domestication for potato based on multilocus amplified fragment length polymorphism genotyping. *Proc. Natl. Acad. Sci. U.S.A.* 102, 14694–14699. doi: 10.1073/pnas.0507400102
- Szklarczyk, D., Morris, J. H., Cook, H., Kuhn, M., Wyder, S., Simonovic, M., et al. (2016). The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res.* 45, D362–D368.
- Tang, X., Zhang, N., Si, H., and Calderón-Urrea, A. (2017). Selection and validation of reference genes for RT-qPCR analysis in potato under abiotic stress. *Plant Methods* 13:85.
- The Potato Genome Sequencing Consortium, Xu X., Pan, S., Cheng, S., Zhang, B., Mu, D., et al. (2011). Genome sequence and analysis of the tuber crop potato. *Nature* 475:189.
- Voorrips, R. E. (2002). MapChart: software for the graphical presentation of linkage maps and QTLs. *J. Heredity* 93, 77–78. doi: 10.1093/jhered/93.1.77
- Wang, D., Zhang, Y., Zhang, Z., Zhu, J., and Yu, J. (2010). KaKs\_Calculator 2.0: a toolkit incorporating gamma-series methods and sliding window strategies. *Genom. Proteom. Bioinform.* 8, 77–80. doi: 10.1016/s1672-0229(10)60008-3
- Wang, J., Li, G., Li, C., Zhang, C., Cui, L., Ai, G., et al. (2020). NF-Y plays essential roles in flavonoid biosynthesis by modulating histone modifications in tomato. *New Phytol.* 229, 3237–3252. doi: 10.1111/nph.17112
- Wang, Y., Tang, H., DeBarry, J. D., Tan, X., Li, J., Wang, X., et al. (2012). MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* 40:e49. doi: 10.1093/nar/gkr1293
- Wang, Y., Xu, W., Chen, Z., Han, B., Haque, M. E., and Liu, A. (2018). Gene structure, expression pattern and interaction of Nuclear Factor-Y family in castor bean (*Ricinus communis*). *Planta* 247, 559–572. doi: 10.1007/s00425-017-2809-2
- Xing, Y., Fikes, J. D., and Guarente, L. (1993). Mutations in yeast HAP2/HAP3 define a hybrid CCAAT box binding domain. *EMBO J.* 12, 4647–4655. doi: 10.1002/j.1460-2075.1993.tb06153.x
- Xu, J. J., Zhang, X. F., and Xue, H. W. (2016). Rice aleurone layer specific OsNF-YB1 regulates grain filling and endosperm development by interacting with an ERF transcription factor. *J. Exp. Bot.* 67, 6399–6411. doi: 10.1093/jxb/erw409
- Yan, D. H., Xia, X., and Yin, W. (2013). NF-YB family genes identified in a poplar genome-wide analysis and expressed in populus euphratica are responsive to drought stress. *Plant Mol. Biol. Rep.* 31, 363–370. doi: 10.1007/s11105-012-0508-5

- Yan, H., Wu, F., Jiang, G., Xiao, L., Li, Z., Duan, X., et al. (2019). Genome-wide identification, characterization and expression analysis of NF-Y gene family in relation to fruit ripening in banana. *Postharvest Biol. Technol.* 151, 98–110. doi: 10.1016/j.postharvbio.2019.02.002
- Yang, J., Zhu, J., and Yang, Y. (2017). Genome-wide identification and expression analysis of NF-Y transcription factor families in watermelon (*Citrullus lanatus*). *J. Plant Growth Regul.* 36, 590–607. doi: 10.1007/s00344-017-9670-1
- Yang, W., Lu, Z., Xiong, Y., and Yao, J. (2017). Genome-wide identification and co-expression network analysis of the OsNF-Y gene family in rice. *Crop J.* 5, 21–31. doi: 10.1016/j.cj.2016.06.014
- Yu, Y., Bai, Y., Wang, Y., Wang, H., Liu, C., and Ni, Z. (2020). Soybean nuclear factor YA10 positively regulates drought resistance in transgenic *Arabidopsis thaliana*. *Environ. Exp. Bot.* 180:104249. doi: 10.1016/j.envexpbot.2020.104249
- Zhao, H., Wu, D., Kong, F., Lin, K., Zhang, H., and Li, G. (2017). The *Arabidopsis thaliana* nuclear factor Y transcription factors. *Front. Plant Sci.* 7:2045. doi: 10.3389/fpls.2016.02045

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Liu, Li, Zhu, Ma, Li, Bi, Sun, Bai, Zhang and Liu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.