Genome-wide identification and co-expression network analysis of the OsNF-Y gene family in rice

Wenjie Yang\textsuperscript{a}, Zhanhua Lu\textsuperscript{b}, Yufei Xiong\textsuperscript{a}, Jialing Yao\textsuperscript{a,}\textsuperscript{*}

\textsuperscript{a}College of Life Science and Technology, Huazhong Agricultural University, Wuhan, 430070, China
\textsuperscript{b}College of Plant Science and Technology, Huazhong Agricultural University, Wuhan, 430070, China

\textbf{A R T I C L E  I N F O}

Article history:
Received 24 March 2016
Received in revised form
10 June 2016
Accepted 11 July 2016
Available online 26 July 2016

\textbf{Keywords:}
Abiotic stress
Gene duplication
Nuclear factor Y
NF-Y
Oryza sativa

\begin{abstract}
Nuclear factor Y (NF-Y) is a ubiquitous transcription factor that regulates important physiological and developmental processes. In this study, we identified 34 OsNF-Y genes in rice, including 6 newly identified genes. Expression profile analysis covering the whole life cycle revealed that transcripts of OsNF-Y differentially accumulated in a tissue-specific, preferential or constitutive manner. In addition, gene duplication studies and expression analyses were performed to determine the evolutionary origins of the OsNF-Y gene family. Nine OsNF-Y genes were differentially expressed after treatment of seedlings with one or more abiotic stresses such as drought, salt and cold. Analysis of expression correlation and Gene Ontology annotation suggested that OsNF-Y genes were co-expressed with genes that participated in stress, accumulation of seed storage reserves, and plant development. Co-expression analysis also revealed that OsNF-Y genes might interact with each other, suggesting that NF-Y subunits formed complexes that take part in transcriptional regulation. These results provide useful information for further elucidating the function of the NF-Y family and their regulatory pathways.
\end{abstract}

© 2016 Crop Science Society of China and Institute of Crop Science, CAAS. Production and hosting by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

\section{Introduction}

Nuclear factor Y (NF-Y) transcription factor, which is also named CCAAT box binding factor (CBF) [1] and heme-associated protein (HAP), is a ubiquitous transcription factor in eukaryotes. NF-Y is a heterotrimeric complex composed of three subunits: NF-YA (CBF-B/HAP2), NF-YB (CBF-A/HAP3), and NF-YC (CBF-C/HAP5) [2]. NF-Y complexes are CCAAT-specific transcription factors that bind to CCAAT sites in DNA to control the expression of target genes [3]. Each NF-Y subunit in yeast and animals is encoded by a single gene, but in plants there is a multi-gene family that encodes each subunit [3–5]. In Arabidopsis, 10 NF-YA, 10 NF-YB, and 10 NF-YC were identified [6], and 7 NF-YA, 15 NF-YB, and 9 NF-YC were characterized in Brachypodium distachyon [5].

The core domain of NF-YA consists of two parts: one responsible for protein-protein interactions with NF-YB or NF-YC, and the other responsible for DNA binding [7]. Protein sequence analysis suggested that NF-YB and NF-YC contain a 65 amino acid histone-like fold motif (HFM) [8]. A comparison with core histones suggested that NF-YB was part of H2B and that NF-YC was a H2A family

\* Corresponding author. Fax: +86 27 87282866.
E-mail address: yaojilmy@mail.hzau.edu.cn (J. Yao).

Peer review under responsibility of Crop Science Society of China and Institute of Crop Science, CAAS.

\textcolor{red}{http://dx.doi.org/10.1016/j.cj.2016.06.014}

2214-5141/© 2016 Crop Science Society of China and Institute of Crop Science, CAAS. Production and hosting by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
histone [1]. Accordingly, misunderstandings have arisen among NF-Y, negative cofactors 2α/β (NC2), and DNA POLYMERASE II SUBUNIT B3/B4 (DPB3/4). NF-Ys are H2A/H2B-like histone fold domain (HFD)-containing proteins [9], whereas NC2 and DPB3/4 do not functionally overlap with NF-Y [6]. NC2 associates with the TATA box binding protein [10]. DPB3 and DPB4 are components of DNA polymerase II and are associated with the histone acetylating complex and the chromatin remodeling complex [9,11].

In plants, the roles of NF-Y were identified in regulation of embryogenesis, flowering time, chloroplast biogenesis, seed germination and stress tolerance [12]. LEAFY COTYLEDON 1 (LEC1, At-NF-YB9) is a key regulator of late embryogenesis and seed development in Arabidopsis [13,14]. LEC1 was also reported to play a role in hormone signaling and fatty acid biosynthesis [15]. Similar to LEC1, LEC1-LIKE (LL1, At-NF-YB6) is also an essential regulator in embryonic development [16]. NF-YB2 (HAP3b) promotes flowering by enhancing the expression of key flowering time genes such as FT and SOC1 [17]. A recent report suggested that NF-Y interacts with CONSTANS to modulate trimethylated H3K27 levels of SOC1 in the photoperiod pathway and DELLAs in the gibberellin pathway [18]. In addition, NF-Y also functions in response to abiotic stress. At-NF-YA1 is involved in regulation of postgermination growth under salt stress [19]. The expression of At-NF-YA5 is strongly induced by drought, and overexpression of At-NF-YA5 results in enhanced drought tolerance through activation of stress-responsive genes [20]. ATHAP5A binds to the CCAAT motif of AXTTH21 to modulate the freezing stress resistance in Arabidopsis [21]. OsHAP3A (NF-YB2), OsHAP3B (NF-YB3), and OsHAP3C (NF-YB4) were shown to regulate chloroplast biogenesis in rice, [22]. Overexpression of OsHAP3E (NF-YB7) resulted in various abnormal morphologies during both the vegetative and reproductive phases [23]. Ghδ8 (NF-YB11) has pleiotropic effects on grain yield, heading date and plant height in rice [24]. OsNF-YB1, an endosperm-specific gene, plays a role in maintenance of endosperm cell proliferation [25]. Recently, some reports indicated that NF-YB1 regulated the expression of sucrose transporters in aleurone to facilitate sugar loading to the endosperm [26]. OsHAP2E (NF-YA2) conferred resistance to biotic and abiotic stresses and increased photosynthesis and tiller number [27].

Twenty-eight NF-Y genes were already identified in rice [3]; however, with improvements in the rice genome data, new members of the NF-Y family need to be explored. In this study, based on the information derived from the public plant database, members and numbers of the NF-Y family in rice were reclassified. Co-expression analysis was performed, and four co-expression networks among NF-Y, negative cofactors 2α (NC2), and DNA POLYMERASE II SUBUNIT B3/B4 (DPB3/4) were employed to identify putative OsNF-Y genes [28]. A BlastP search was carried out using the HMM profile on NCBI Blast, the Pfam, and SMART (http://smart.embl-heidelberg.de/smart/batch.pl) were used to confirm each predicted OsNF-Y protein.

2. Materials and methods

2.1. Identification of OsNF-Y genes in rice

Hidden Markov Model (HMM) profiles of NF-YA (PF02045) and NF-YB/C (PF00808) downloaded from Pfam (http://pfam.xfam.org/) were employed to identify putative OsNF-Y genes [28]. A BlastP search was carried out using the HMM profile on NCBI Blast, the Pfam, and SMART (http://smart.embl-heidelberg.de/smart/batch.pl) were used to confirm each predicted OsNF-Y protein.

2.2. Phylogenetic analysis of the OsNF-Y family

The protein sequences of NF-Y in rice, Arabidopsis and mice were aligned using the ClustalX (version 1.83) program. Neighbor-joining phylogenetic trees were constructed using MEGA (version 5.1) based on the full-length protein sequences with default parameters. Bootstrap analysis was performed using 1000 replicates.

2.3. Chromosomal localization and gene duplication

The OsNF-Y genes were mapped on rice chromosomes according to their positions available in the Rice Genome Annotation Project (RGAP) database (http://rice.plantbiology.msu.edu/index.shtml). The distribution of OsNF-Y genes was drawn by MapInspect [29]. The duplicated genes were determined from the rice RGAP segmental duplication database, with a maximal length distance permitted between collinear gene pairs set at 500 kb. Tandem duplicates were defined as genes separated by five or fewer genes. The percentage of sequence similarities between the proteins encoded by these genes was determined by MEGA (version 5.1).

2.4. Expression analysis of OsNF-Y genes in rice

Expression profile data of the OsNF-Y gene family in 24 tissues of cultivar (cv.) Minghui 63 was extracted from the CREP database (http://crep.ncpgr.cn/, accession number: GSE19024) [30]. Expression values of each gene were logarithm transformed in Microsoft Excel, and cluster analyses were carried out using J-Express 2011 with Euclidean distances and the hierarchical cluster method of “complete linkage”. When more than one probe set was available for one gene, the probe set with the higher signal value was used. For stress treatments, the microarray data were downloaded from the NCBI Gene Expression Omnibus (GEO) under accession number GSE6901. Seven-day-old seedlings were treated under drought (dried between folds of tissue paper for 3 h), salt (kept in 200 mmol L⁻¹ NaCl solution for 3 h,) and cold (4 ± 1 °C for 3 h) stress conditions. Expression levels of treated seedlings were compared to expression in untreated seedlings using Student’s t-tests to determine statistically significant differences. Genes up- or down-regulated by more than two-fold and with P < 0.05 were considered to be differentially expressed.

2.5. Identification of correlated genes and network construction

The co-expression data for each OsNF-Y gene were downloaded from the CREP database. First, we ranked the correlated genes...
with PCCs higher than 0.75. Then, we recalculated the PCCs of the genes in which we were interested with the R project (version 3.1.3). The optimal threshold of the PCC was set at 0.8 [28]. Cytoscape (version 2.8.2) software was used to construct co-expression networks between OsNF-Y genes and co-expression genes. Go enrichment was performed by the Singular Enrichment Analysis (SEA) tool in the AgriGO database (http://bioinfo.cau.edu.cn/agriGO/analysis.php) with default parameters using the rice MSU 7.0 genome annotation as background.

2.6. RNA extraction and qRT-PCR

All samples were collected and stored at −80 °C until processed. Total RNA was isolated using TRIzol (TransGen) according to the manufacturer’s instructions. First-strand cDNA was reverse-transcribed from total RNA using M-MLV reverse transcriptase (Promega) with oligo (dT) as the primer. PCR was performed in a total volume of 10 μL containing 5 μL of 2× SYBR Premix Ex Taq (TaKaRa), 2 μmol L−1 of each gene-specific primer (Table S1), 0.5 μL of the cDNA sample, 0.2 μL of Rox Reference Dyell (TaKaRa) on an ABI StepOne Real-time PCR instrument (Applied Biosystems). The reactions were carried out using the following program: 95 °C for 30 s, 45 cycles of 95 °C for 5 s, and 60 °C for 34 s. The rice ubiquitin gene (LOC_Os03g13170) was used as an internal reference. Each experiment was performed with three technical replicates. Student’s t-tests were used to determine significant differences.

3. Results

3.1. Identification and phylogenetic analysis of the OsNF-Y gene family

Previous reports showed that there are 28 NF-Y genes in the rice genome [3,6]. According to the update of the Rice Genome Annotation Project (RGAP) data, some genes were verified as new OsNF-Y genes. Therefore, to identify the members of the OsNF-Y gene family, we sought proteins based on the amino acid sequences of conserved domains [6]. As a result, 40 putative NF-Y genes in rice were obtained, including 11 NF-YA, 13 NF-YB, and 16 NF-YC.

According to previous reports, three NC2 and three DPB3/4 in Arabidopsis were removed from the NF-Y family due to a lack of NC2 and DPB3/4 functional overlap [6]. Therefore, based on the sequence alignment of NF-Y proteins in Arabidopsis and rice, we constructed a phylogenetic tree (Fig. S1). Some candidates in rice were more closely related to the NC2 and DPB3/4 of Arabidopsis according to the phylogenetic tree. Thus, two genes in the NF-YB subunit family and four genes in the NF-YC family were removed, and 34 genes were identified as members of the NF-Y family in rice. To avoid confusion, we named the new members following the previous nomenclature of genes in the OsNF-Y family [31] (Table 1). All of the 11 NF-YA subunits contain a CBFB_NFYA domain (Fig. 1-D). As a new member, OsNF-YA11 was close to OsNF-YA2 and OsNF-YA6 in the phylogenetic tree (Fig. 1-A). The NF-YB and NF-YC family members contained 11 and 12 members, respectively. They both shared a CBFD_NFYB_HMF domain (Fig. 1-D). Five new members of NF-YS (OsNF-YC8, OsNF-YC9, OsNF-YC10, OsNF-YC11, and OsNF-YC12) were clustered in the same branch and were adjacent to each other (Fig. 1-C).

3.2. Chromosomal localization and gene duplication of OsNF-Y genes

The genomic distribution of OsNF-Y genes was determined by their positions in the rice chromosomes. The 34 OsNF-Y genes are unevenly dispersed on 11 of the 12 chromosomes (Fig. 2, the exact positions on rice chromosome pseudomolecules are given in Table S2).

Tandem and segmental duplications are common in the generation and maintenance of gene families in genomes [32]. In order to evaluate the likely mechanism of evolution of the OsNF-Y gene family, possible segmental duplication and tandem duplication events were examined (Table S3). Analysis of the rice segmental duplication database from RGAP revealed that 15 (7 groups) OsNF-Y genes could be assigned to RGAP segmental duplication blocks (Fig. 2). The overall identities of cDNA sequences of segmentally duplicated genes ranged from 18.9% to 60.8% (Table S3-b). In addition, 4 OsNF-Y genes (2 pairs) were probably produced from tandem duplication according to

---

Table 1 – Identification and classification of OsNF-Y genes in rice.

<table>
<thead>
<tr>
<th>NF-Y name</th>
<th>Gene locus</th>
<th>Other names</th>
</tr>
</thead>
<tbody>
<tr>
<td>NF-YA family</td>
<td>LOC_Os03g07880</td>
<td>HAP2C</td>
</tr>
<tr>
<td>OsNF-YA1</td>
<td>LOC_Os03g29760</td>
<td>HAP2E</td>
</tr>
<tr>
<td>OsNF-YA2</td>
<td>LOC_Os03g45450</td>
<td>HAP2H</td>
</tr>
<tr>
<td>OsNF-YA3</td>
<td>LOC_Os03g68970</td>
<td>HAP2D</td>
</tr>
<tr>
<td>OsNF-YA4</td>
<td>LOC_Os07g16720</td>
<td>HAP2G</td>
</tr>
<tr>
<td>OsNF-YA5</td>
<td>LOC_Os08g06410</td>
<td>HAP2A</td>
</tr>
<tr>
<td>OsNF-YA6</td>
<td>LOC_Os10g25850</td>
<td>HAP2F</td>
</tr>
<tr>
<td>OsNF-YA7</td>
<td>LOC_Os12g41880</td>
<td>HAP2B</td>
</tr>
<tr>
<td>OsNF-YA8</td>
<td>LOC_Os12g42400</td>
<td>HAP2F</td>
</tr>
<tr>
<td>OsNF-YA9</td>
<td>LOC_Os2g53620</td>
<td>HAP2F</td>
</tr>
<tr>
<td>OsNF-YA10</td>
<td>LOC_Os01g61810</td>
<td>HAP3A</td>
</tr>
<tr>
<td>OsNF-YA11</td>
<td>LOC_Os05g38820</td>
<td>HAP3B</td>
</tr>
<tr>
<td>OsNF-YB1</td>
<td>LOC_Os01g70880</td>
<td>HAP3J</td>
</tr>
<tr>
<td>OsNF-YB2</td>
<td>LOC_Os01g70890</td>
<td>HAP3E</td>
</tr>
<tr>
<td>OsNF-YB3</td>
<td>LOC_Os05g69780</td>
<td>HAP3C</td>
</tr>
<tr>
<td>OsNF-YB4</td>
<td>LOC_Os06g59370</td>
<td>HAP3I</td>
</tr>
<tr>
<td>OsNF-YB5</td>
<td>LOC_Os06g17480</td>
<td>HAP3D</td>
</tr>
<tr>
<td>OsNF-YB6</td>
<td>LOC_Os06g41580</td>
<td>HAP3F</td>
</tr>
<tr>
<td>OsNF-YB7</td>
<td>LOC_Os07g07740</td>
<td>HAP3H</td>
</tr>
<tr>
<td>OsNF-YB8</td>
<td>LOC_Os07g07740</td>
<td>HAP3H</td>
</tr>
<tr>
<td>OsNF-YB9</td>
<td>LOC_Os08g10560</td>
<td>HAP3S</td>
</tr>
<tr>
<td>OsNF-YB10</td>
<td>LOC_Os09g03010</td>
<td>HAP5E</td>
</tr>
<tr>
<td>OsNF-YB11</td>
<td>LOC_Os10g03010</td>
<td>HAP5E</td>
</tr>
<tr>
<td>OsNF-YC family</td>
<td>LOC_Os04g53680</td>
<td>HAP5G</td>
</tr>
<tr>
<td>OsNF-YC1</td>
<td>LOC_Os05g45640</td>
<td>HAP5B</td>
</tr>
<tr>
<td>OsNF-YC2</td>
<td>LOC_Os08g10160</td>
<td>HAP5F</td>
</tr>
<tr>
<td>OsNF-YC3</td>
<td>LOC_Os08g38780</td>
<td>HAP5D</td>
</tr>
<tr>
<td>OsNF-YC4</td>
<td>LOC_Os09g03010</td>
<td>HAP5E</td>
</tr>
<tr>
<td>OsNF-YC5</td>
<td>LOC_Os10g03010</td>
<td>HAP5E</td>
</tr>
<tr>
<td>OsNF-YC6</td>
<td>LOC_Os11g11580</td>
<td>HAP5F</td>
</tr>
<tr>
<td>OsNF-YC7</td>
<td>LOC_Os11g11580</td>
<td>HAP5F</td>
</tr>
<tr>
<td>OsNF-YC8</td>
<td>LOC_Os12g42400</td>
<td>HAP5F</td>
</tr>
<tr>
<td>OsNF-YC9</td>
<td>LOC_Os12g42400</td>
<td>HAP5F</td>
</tr>
<tr>
<td>OsNF-YC10</td>
<td>LOC_Os12g42400</td>
<td>HAP5F</td>
</tr>
<tr>
<td>OsNF-YC11</td>
<td>LOC_Os12g42400</td>
<td>HAP5F</td>
</tr>
<tr>
<td>OsNF-YC12</td>
<td>LOC_Os12g42400</td>
<td>HAP5F</td>
</tr>
</tbody>
</table>

Please cite this article as: W. Yang, et al., Genome-wide identification and co-expression network analysis of the OsNF-Y gene family in rice, The Crop Journal (2016), http://dx.doi.org/10.1016/j.cj.2016.06.014
the criteria used in the analysis. Both pairs produced from tandem duplication were NF-YB subunits. OsNF-YB1 and OsNF-YB7 were separated by 3 intervening genes, and the degree of homology between them was 69.4%. No intervening genes were present between OsNF-YB5 and OsNF-YB6; the homology of this gene pair was 47.2% (Table S3-a). These results indicated that the expansion of the OsNF-Y gene family is due to both tandem and segmental duplication events.

3.3. Expression pattern of OsNF-Y genes in rice

To examine the transcript accumulation of OsNF-Y genes in the entire rice life cycle the expression profiles at 24 developmental stages (Table S4) in cv. Minghui 63 (MH63) were analyzed by Affymetrix rice microarray data in the CREP database [30]. Probes for 33 of the 34 OsNF-Y genes could be used in the Affymetrix microarray; only the OsNF-YC3 probe is not in CREP.

---

Please cite this article as: W. Yang, et al., Genome-wide identification and co-expression network analysis of the OsNF-Y gene family in rice, The Crop Journal (2016), http://dx.doi.org/10.1016/j.cj.2016.06.014
OsNF-YA1 and OsNF-YB5 had two probe sets, and the probe set with higher signal value was used for analysis.

A hierarchical cluster was generated to display the signal values of the 33 OsNF-Y genes (Fig. 3). The expression patterns of the OsNF-Y genes were classified into two major groups. Group A contained 11 genes, most of which showed high expression in all tested tissues. These 11 genes were further divided into three subgroups (A1–A3). Subgroup A1 contained 5 genes (OsNF-YA1, OsNF-YA2, OsNF-YA4, OsNF-YB2, and OsNF-YC1), which displayed higher expression in most of the tested tissues, especially the vegetative tissues. Subgroup A2 consisted of 3 genes (OsNF-YA7, OsNF-YB4, and OsNF-YC4), with high expression in the panicles. Subgroup A3 has 3 genes (OsNF-YB3, OsNF-YC5, and OsNF-YC7) showing relatively high expression levels in all tested tissues.

Group B included 22 genes with relatively low expression in most tissues or tissue-specific/preferential expression patterns. OsNF-YA8, OsNF-YB1, OsNF-YB9, OsNF-YC8, OsNF-YC9, OsNF-YC10, OsNF-YC11, and OsNF-YC12 are endosperm-specific genes. OsNF-YA3 displayed higher expression in flag leaves, leaf sheaths and stems. OsNF-YA10 showed higher expression levels in flag leaves and panicles. OsNF-YA6 was predominantly expressed in endosperm, and OsNF-YB8 showed high expression in seedlings, shoots and flag leaves.

The expression patterns of 15 OsNF-Y genes were further confirmed by real-time PCR. RNA was extracted from various tissues of cv. Zhonghua 11 (ZH11). The expression level of OsNF-YC3, which had no probe in CREP, was examined at different stages. OsNF-YC3 was relatively highly expressed at the
reproductive stage, especially in the endosperm at 5 days post pollination (DAP) (Fig. 4). The expression levels of OsNF-YA1, OsNF-YA3, OsNF-YA8, OsNF-YA11, OsNF-YB1, OsNF-YB8, OsNF-YB9, OsNF-YC8, OsNF-YC9, OsNF-YC10, OsNF-YC11, and OsNF-YC12 were in accordance with the microarray data (Fig. 4). OsNF-YB2 was highly expressed in the microarray analysis; however, our results show a relatively higher expression level in the reproductive stage than in vegetative stage. Interestingly, OsNF-YC2 had low expression at different stages according to the microarray data, whereas our quantitative PCR displays indicated a relatively higher expression pattern in the examined tissues.

The expression patterns of segmentally and tandemly duplicated OsNF-Y genes were analyzed using Affymetrix microarray data. The six pairs of segmentally duplicated genes shared similar expression patterns (Fig. 5-A). Two pairs of tandemly duplicated genes had different expression patterns (Fig. 5-B). OsNF-YB5 and OsNF-YB6 had similar expression levels; the duplicated genes appeared to maintain their primary functions during evolution. OsNF-YB1 is an endosperm-specific gene, but OsNF-YB7 was not expressed at significant levels in any tissue, which might mean that it had degenerated into a pseudogene after duplication.

### 3.4. Responses of OsNF-Y genes to abiotic stress

Microarray data were analyzed to investigate whether the OsNF-Y genes are responsive to abiotic stress. The expression level of the corresponding OsNF-Y genes in 7-day-old seedlings without treatment was used as a control (Fig. 6). Expression fold changes of OsNF-Y genes under drought, salt and cold treatments are presented in Table S5. Nine OsNF-Y genes were differentially expressed under one or more of the abiotic stresses. OsNF-YA10 and OsNF-YA11 were down-regulated under both drought and salt stresses. OsNF-YA1 was down-regulated under both drought and cold stress. Of the 9 OsNF-Y genes, 5 showed differential expression under drought stress. Among them, OsNF-YA4 and OsNF-YB3 were up-regulated, whereas the others (OsNF-YB8, OsNF-YB9, and OsNF-YB11) were down-regulated. OsNF-YA5 was down-regulated in response to cold treatment. Expression of the other OsNF-Y genes was not significantly affected by drought, salt or cold treatments.

![Fig. 4](image-url) - Relative expression levels of 15 OsNF-Y genes in ZH11. The rice ubiquitin gene (LOC_Os03g13170) was used as an internal reference. Each experiment was performed with three biological replicates. Student’s t-tests were used for determination of significant differences. Root and leaf, roots and leaves at the trefoil stage; Stem, stems at heading stages; P4, panicles at the pistil and stamen primordial differentiation stage; P7, 4–5 cm young panicles; S1 and S3, seeds at 1 and 3 days post pollination; En, endosperm; the following number is days post pollination.
The functions participating in specific biological processes \cite{33}. To explore Co-expression analysis was used to identify functional factors of the genes were co-expressed with 9 starch synthesis genes; and OsNF-YB9 was co-expressed with 1 starch synthesis gene. Interestingly, these 9 OsNF-Y genes were also co-expressed with 35 storage protein synthesis related genes (including \textit{Ci}uA, \textit{Ci}uB1, and \textit{Ci}uD) (Fig. 7-B). Moreover, all of these 9 OsNF-Y genes displayed relatively high expression in the endosperm. The results of the expression correlation and GO annotation suggested that these OsNF-Ys might have roles in synthesis and accumulation of seed storage reserves in rice.

Fig. 7-C shows 12 OsNF-Y genes and 30 co-expressed genes that were associated with stress-related proteins based on their GO annotation. OsNF-YB1, OsNF-YB9, OsNF-YC11, and OsNF-YC12 were co-expressed with rice sucrose synthase (RSUS3). OsNF-YA8, OsNF-YB1, OsNF-YB9, OsNF-YC9, OsNF-YC10, OsNF-YC11, and OsNF-YC12 were co-expressed with trehalose-6-phosphate synthase (OsTPS9). OsNF-YA6, OsNF-YA8, OsNF-YB1, OsNF-YC8, OsNF-YC9, OsNF-YC10, OsNF-YC11, and OsNF-YC12 were co-expressed with mitogen-activated protein (OsMPK2). In addition, OsNF-YA6,

3.5. Gene ontology (GO) analyses and functional classification of the OsNF-Y family

Co-expression analysis was used to identify functional factors participating in specific biological processes \cite{33}. To explore the functions 33 OsNF-Y genes were selected as "guide genes" to seek co-expressed genes using expression data from the CREP database. Using a Pearson’s correlation coefficient (PCC) threshold of 0.75 \cite{34}, 1364 genes whose expression patterns closely correlated with 22 OsNF-Y genes were identified. Thereafter, we analyzed the GO annotations of the co-expressed genes using agrigo tools. The enriched annotations were mainly concentrated in metabolic processes, catalytic activity, transcriptional regulation and binding. Ten of 22 genes tightly correlated with transcription factors, material synthesis and stress related genes. Four networks were constructed with Cytoscape to display the relationship between OsNF-Y genes and their co-expressed genes (Fig. 7).

Nine OsNF-Y genes were co-expressed with 10 starch synthesis genes (Fig. 7-A). In this network, OsNF-YC8 and OsNF-YC9 were co-expressed with 10 starch synthesis genes; OsNF-YC10 was co-expressed with 9 starch synthesis genes; and OsNF-YA6, OsNF-YA8, and OsNF-YC11 were co-expressed with 8 starch synthesis genes. OsNF-YB1 was co-expressed with 5 starch synthesis genes; OsNF-YC12 was co-expressed with 2 starch synthesis genes; and OsNF-YB9 is co-expressed with 1 starch synthesis gene. Interestingly, these 9 OsNF-Y genes were also co-expressed with 35 storage protein synthesis related genes (including \textit{Ci}uA, \textit{Ci}uB1, and \textit{Ci}uD) (Fig. 7-B). Moreover, all of these 9 OsNF-Y genes displayed relatively high expression in the endosperm. The results of the expression correlation and GO annotation suggested that these OsNF-Ys might have roles in synthesis and accumulation of seed storage reserves in rice.

Fig. 7-C shows 12 OsNF-Y genes and 30 co-expressed genes that were associated with stress-related proteins based on their GO annotation. OsNF-YB1, OsNF-YB9, OsNF-YC11, and OsNF-YC12 were co-expressed with rice sucrose synthase (RSUS3). OsNF-YA8, OsNF-YB1, OsNF-YB9, OsNF-YC9, OsNF-YC10, OsNF-YC11, and OsNF-YC12 were co-expressed with trehalose-6-phosphate synthase (OsTPS9). OsNF-YA6, OsNF-YA8, OsNF-YB1, OsNF-YC8, OsNF-YC9, OsNF-YC10, OsNF-YC11, and OsNF-YC12 were co-expressed with mitogen-activated protein (OsMPK2). In addition, OsNF-YA6,
Fig. 6 – Differential expression of OsNF-Y genes following treatment of 7-day-old seedlings with drought (A), salt (B), and cold (C). “Control” represents expression levels of corresponding seedlings without stress treatment. Scores are average expression values from the microarray (GSE6901). Error bars represent standard errors for data obtained from three biological replicates.

Fig. 7 – Co-expression networks of OsNF-Y genes. Co-expression networks between OsNF-Y and (A) starch biosynthesis genes, (B) seed storage protein genes and (C) stress response genes. (D) Co-expression network of OsNF-Y genes.
OsNF-YA8, OsNF-YB1, OsNF-YC8, OsNF-YC9, OsNF-YC10, and OsNF-YC11 were co-expressed with 10 or more stress-related genes. Among these 12 OsNF-Y genes, two (OsNF-YB8 and OsNF-YB9) were down-regulated under drought treatment. These results indicated that these OsNF-Ys may participate in both abiotic and biotic stress responses.

Our analysis also indicated that the some OsNF-Y genes are co-expressed with other OsNF-Y genes. Two OsNF-YA genes, two OsNF-YB genes and five OsNF-YC genes co-expressed with each other (Fig. 7-D). OsNF-YA6 was co-expressed with OsNF-YC8, OsNF-YC9, and OsNF-YC10. OsNF-YA8 and OsNF-YB1 co-expressed and are both co-expressed with all OsNF-YC genes. OsNF-YB9 co-expressed with OsNF-YB1 and OsNF-YC12. Among the OsNF-YC genes, OsNF-YC10 and OsNF-YC11 co-expressed with all other OsNF-YCs; OsNF-YC8 and OsNF-YC9 co-expressed with 3 OsNF-YC genes; and OsNF-YC12 co-expressed with OsNF-YC10 and OsNF-YC11. All of these genes displayed similar expression patterns and the results suggested the possibility that OsNF-Y subunit interact with each other.

4. Discussion

4.1. OsNF-Y evolution and classification

Based on the amino acid sequence of the conserved domain 34 NF-Y genes were verified in rice; six of them were newly identified. Gene families were enlarged by three major mechanisms: segmental duplication, tandem duplication and retroposition [35]. Here, we showed that the number of OsNF-Y genes arranged in segmental duplications contributed to 44.1% of new births, and tandem duplication events contributed to 11.7%. It is suggested that segmental and tandem duplication events are the main mechanisms of OsNF-Y family evolution.

Duplication events often lead to increased functional diversity between duplicated gene pairs [36]. In this study, the fate of duplicated OsNF-Y genes could be described as non-functionalization and neo-functionalization. The expression patterns of six pairs of segmentally duplicated genes overlapped, suggesting that these gene pairs might undergo neo-functionalization. In the other two pairs of segmental duplications, one of the duplicated pair was not expressed at significant levels in all tissues. We infer that one copy lost function during non-functionalization. This was also seen in two pairs of tandemly duplicated genes in the OsNF-Y family. During the process of evolution of OsNF-Ys, the genes appear to lose function or gain new functions.

4.2. OsNF-Y might be involved in accumulation of seed reserves and interact with starch- and protein synthesis-related genes

In cereal crops, the endosperm is the major storage medium in which most seed nutrients are stored. Starch is a major component of rice endosperm. OsAGPL2, a large subunit of ADP-glucose pyrophosphorylase (AGPase), regulates starch contents through catalysis and allosteric regulation [37]. Starch synthases (SS) have been identified and are thought to act on amylopectin synthesis [38]. A mutant of OsSSI exhibited a change in amylopectin chain-length distribution [39]. OsBEI and OsBEIIb encode starch branch enzyme (BE); and changes in expression of OsBEIIb affect the structure of starch [40].

Our co-expression analyses revealed relationships between OsNF-Y genes and starch-related genes (Fig. 7-A). Five OsNF-Ys were co-expressed with OsAGPL2, OsAGPS2b (a small subunit of AGPase), OsSSI, SSIIa and OsBEIIb; OsNF-YA6 was co-expressed with OsAGPL2, OsAGPS2b, OsSSI, SSIIa, and OsBEIIb. OsNF-YB1 was co-expressed with OsAGPL2, OsSSI, SSIIa, and OsBEIIb. The expression patterns likely suggest the possible functions and metabolic pathways of the corresponding genes. Thus, co-expression networks of OsNF-Ys and starch-related genes suggest that these OsNF-Ys may take part in starch biosynthesis for accumulation of seed reserves.

Seed storage proteins are synthesized and accumulate in protein bodies, which account for 8–10% of the dry weight of seeds [41]. Prolamin and glutelin are major components of seed storage proteins. The basic leucine zipper (RISBZ1) and prolamin box binding (RPBF) factors are transcriptional activators of rice seed storage protein (SSP) genes; a double knock-down mutant (KD-RISBZ1/KD-RPBF) resulted in a significant reduction of most SSP gene expression (GluA, GluB1, GluD, Gib1, and RAG1) [42,43]. In our co-expression network (Fig. 7-B) seven OsNF-Y genes co-expressed with RISBZ1 and RPBF. Six OsNF-Y genes co-expressed with GluA, GluB1, GluD, Gib1, and RAG1. We speculate that these OsNF-Y genes are involved in the accumulation of seed storage protein.

### 4.3. OsNF-Ys respond to abiotic stress and may take part in stress resistance involving stress-related genes

It has been reported that the NF-Y family is involved in drought tolerance. Our analysis of abiotic stress treatments (drought, salt and cold) indicated that the expression levels of 9 OsNF-Y genes were significantly changed following stress. Among them, OsNF-YA10 was repressed by drought and salt, and a mutant of OsNF-YA10 enhanced drought tolerance in rice [44] (Table 2). Over-expression of OsNF-YA2 conferred resistance to salinity and drought [27]; however, OsNF-YA2 was not included in the above 9 OsNF-Y genes; possibly due to the different treatment conditions. Thus, it is likely possible that these 9 OsNF-Y genes respond to abiotic stress and might be candidate genes for stress resistance in rice.

Trehalose-6-phosphate synthase (TPS), an enzyme that catalyzes trehalose, has an important role in abiotic stress

| Table 2 – Summary of functions of OsNF-Y genes. |
|----------------|----------------|
| **Gene** | **Other name** | **Function** | **Reference** |
| OsNF-YA2 | HAP2E | Salinity and drought tolerance | [27] |
| OsNF-YA10 | HAP2F | Drought tolerance | [44] |
| OsNF-YB1 | HAP3K | Endosperm development | [25] |
| OsNF-YB2 | HAP3A | Chloroplast biogenesis | [22] |
| OsNF-YB3 | HAP3B | | |
| OsNF-YB4 | HAP3C | | |
| OsNF-YB7 | HAP3E | Vegetative and floral meristem development | [23] |
| OsNF-YB11 | GhB8, DTH8 | Flowering repressor | [24] |

Please cite this article as: W. Yang, et al., Genome-wide identification and co-expression network analysis of the OsNF-Y gene family in rice, The Crop Journal (2016), http://dx.doi.org/10.1016/j.cj.2016.06.014.
response in plants. Overexpression of OsTPS1 improved the tolerances of rice seedlings to cold, salinity and drought [45]. In this work, seven OsNF-Y genes co-expressed with OsTPS, suggesting that these OsNF-Y genes might be involved in abiotic stress responses mediated by OsTPS. One report indicated that mitogen-activated protein kinase (MAPK) cascades play crucial roles in biotic and abiotic stress responses [46]. Co-expression networks identified in the present work indicated that seven OsNF-Ys co-expressed with OsMPK2, implying that these OsNF-Ys might partner with OsMPK during stress response. It was reported that rice sucrose synthase (RSUS3) has a role in response to cold stress, light and sucrose concentration [47]. Germin-like protein (GLP) markers were associated with QTL for resistance to the rice blast pathogen [48]. Our analysis suggested that OsNF-YB1, OsNF-YB9, OsNF- YC11, and OsNF-YC12 were co-expressed with RSUS3. OsNF-YA6, OsNF-YA8, OsNF-YB1, OsNF-YC8, OsNF-YC9, OsNF-YC10, and OsNF-YC11 were co-expressed with OsGLP8. Hence, we infer that these OsNF-Ys are involved in stress resistance mediated by RSUS and OsGLP8.

Acknowledgments

We gratefully acknowledge Prof. Chungen Hu (HuaZhong Agricultural University, Wuhan, China) for helpful discussions. This research was supported by the National Natural Science Foundation of China (Nos. 31570321, 30971551).

Supplementary data

Supplementary data for this article can be found online at http://dx.doi.org/10.1016/j.cj.2016.06.014.

REFERENCES