frontiers in Plant Science

ORIGINAL RESEARCH
published: 07 November 2016
doi: 10.3389/fpls.2016.01638

Three Rice NAC Transcription Factors Heteromerize and Are Associated with Seed Size

*Iny Elizebeth Mathew, Sweta Das, Arunima Mahto and Pinky Agarwal**

National Institute of Plant Genome Research, New Delhi, India

NACs are plant-specific transcription factors (TFs) involved in multiple aspects of development and stress. In rice, three NAC TF encoding genes, namely *ONAC020*, *ONAC026,* and *ONAC023* express specifically during seed development, at extremely high levels. They exhibit significantly strong association with seed size/weight with the sequence variations located in the upstream regulatory region. Concomitantly, their expression pattern/levels during seed development vary amongst different accessions with variation in seed size. The alterations in the promoter sequences of the three genes, amongst the five rice accessions, correlate with the expression levels to a certain extent only. In terms of transcriptional properties, the three NAC TFs can activate and/or suppress downstream genes, though to different extents. Only ONAC026 is localized to the nucleus while ONAC020 and ONAC023 are targeted to the ER and cytoplasm, respectively. Interestingly, these two proteins interact with ONAC026 and the dimers localize in the nucleus. *Trans*-splicing between *ONAC020* and *ONAC026* results in three additional forms of *ONAC020*. The transcriptional properties including activation, repression, subcellular localization and heterodimerization of *trans*-spliced forms of ONAC020 and ONAC026 are different, indicating toward their role as competitors. The analysis presented in this paper helps to conclude that the three *NAC* genes, which are associated with seed size, have independent as well as overlapping roles during the process and can be exploited as potential targets for crop improvement.

Keywords: association analysis, NAC, rice, seed development, transcriptional properties

Edited by:
Niranjan Baisakh, Louisiana State University, USA

Reviewed by:
Sitakanta Pattanaik, University of Kentucky, USA
Surekha Katiyar-Agarwal, University of Delhi, India

****Correspondence:***
Pinky Agarwal
pinky.agarwal@nipgr.ac.in;
pinky.agarwal@gmail.com

Specialty section:
This article was submitted to Crop Science and Horticulture, a section of the journal Frontiers in Plant Science

Received: *25 July 2016*
Accepted: *17 October 2016*
Published: *07 November 2016*

Citation:
Mathew IE, Das S, Mahto A and Agarwal P (2016) Three Rice NAC Transcription Factors Heteromerize and Are Associated with Seed Size. Front. Plant Sci. 7:1638. doi: 10.3389/fpls.2016.01638

INTRODUCTION

Transcription factors (TFs) regulate the expression of the downstream target genes, in response to various external and/or internal stimuli, by binding to their upstream *cis* elements either as a monomer or a homodimer, or by interacting with other TFs or regulators. This specific binding is responsible for the spatial and temporal expression of the regulated genes, ultimately leading to a particular response. The functional specificity of these proteins is maintained by the presence of characteristic functional domains (Olsen et al., 2005; Agarwal et al., 2011). Rice seed development is regulated at the transcriptional level by a diverse group of TFs. Starch biosynthesis is controlled by OsBP-5, a MYC TF and OsEBP-89, an EREBP TF which act synergistically to regulate *Wx*, a starch synthase gene (Zhu et al., 2003). Another EREBP TF, rice starch regulator1 (RSR1), negatively regulates the expression of starch biosynthesis genes (Fu and Xue, 2010). A bZIP protein RISBZ1, and a DOF protein RPBF act synergistically in the regulation of storage protein synthesis and also affect starch biosynthesis in rice seeds (Kawakatsu et al., 2009). *OsNF-YB1*, an endosperm specific gene, is essential for proper endosperm development by regulating the cell proliferation genes (Sun et al., 2014). Similarly, *OsMADS6*, highly expressed in the flowers and endosperm, has been shown

to be essential for normal endosperm development in rice (Zhang J. et al., 2010). OsMADS29 is a key regulator of early grain development in rice. The gene, expressing preferentially in the nucellus and nucellar projection, promotes programmed cell death of the maternal tissues (Yin and Xue, 2012; Nayar et al., 2013). OsWRKY78 has been shown to act as a seed development regulator in rice (Zhang et al., 2011). OsWRKY24, OsWRKY53, and OsWRKY70 act in a partially redundant manner in regulating GA and ABA signaling pathways in aleurone cells (Zhang L. et al., 2015). Grain width 8 (GW8)/OsSPL16 is a positive regulator of cell proliferation and controls grain width and yield in rice (Wang et al., 2012). Another QTL, *grain length and width2* (*GLW2*) encodes a growth regulating factor 4 (OsGRF4), which regulates grain weight and interacts with OsGIF1 (GRF interacting factor 1), another positive regulator of grain size in rice (Li et al., 2016).

NAC is one of the largest group of plant-specific TFs, named after the first three reported members of the family, NO APICAL MERISTEM (NAM), *Arabidopsis thaliana* Activation Factor1 and 2 (ATAF1/2) and CUP-SHAPED COTYLEDON 2 (CUC2) (Souer et al., 1996; Aida et al., 1997; Christianson et al., 2010). Each plant genome has multiple encoding members, with 117 and 151 genes in *Arabidopsis* and rice, respectively (Nuruzzaman et al., 2010), and similar numbers in other species as well (Pereira-Santana et al., 2015). They have a conserved NAC domain of about 150 amino acids followed by a diversified transcriptional regulatory region (TRR). The NAC domain is further categorized into five subdomains designated A–E, of which A, C and D are highly conserved. The N-terminal regions of these TFs hold a large number of charged amino acid residues. Subdomains C–E exhibit a net positive charge, while the remaining two are negatively charged. The high conservation of the C and D subdomains and the richness in basic amino acids suggests the involvement of these two regions in conferring the DNA binding property. TRR determines the activation or repression property of the protein and may also influence oligomerization property (Kikuchi et al., 2000; Ooka et al., 2003; Fang et al., 2008).

NAC TFs have been found to regulate a wide array of plant functions. During development, the redundant maternal proteins, *Arabidopsis* NARS1 and NARS2 control PCD of the inner integument of the ovule (Kunieda et al., 2008), and *SlNAC1*, *SlNAC4*, *SlNAC48,* and *SlNAC19* control tomato fruit ripening (Kou et al., 2016). *OsNAP* overexpression delays leaf senescence causing an increase in seed yield (Liang et al., 2014), while *OsNAC5* senses the senescence signal and is responsible for iron remobilization to the seeds (Ricachenevsky et al., 2013). In wheat, single nucleotide polymorphism (SNPs) and differential expression of *NAM-G1*, in 12 accessions, causes variation in grain protein content (Hu et al., 2013). Hence, NACs are an important class of TFs controlling various aspects of seed development (Agarwal et al., 2011) and their functions and sequences need to be explored further. A number of studies have successfully demonstrated the efficiency of a combinatorial strategy which combines high-resolution SNP-based and candidate gene-based association analysis, traditional genetic/QTL mapping, differential expression profiling and molecular haplotyping, for quantitative dissection of complex yield component traits in diverse crop plants, including rice (Zhao et al., 2011; Zuo and Li, 2014; Kujur et al., 2015a; Agarwal et al., 2016). It would be interesting to utilize this integrated approach to identify the functional alleles of *NAC* genes regulating grain size/weight variation, in cultivars adapted to diverse natural agro-climatic conditions. This can aid in genomics-assisted rice crop improvement.

In accordance with biological diversification, similar proteins are known to exhibit variation in properties as well as interact amongst themselves. Two closely related *NAC* TFs, with 66% identity in the protein sequences, control stomata differently. JA2 and JA2L cause stomatal closure and opening, respectively (Du et al., 2014) while *OsSWN1* and *2* control different aspects of secondary wall formation (Yoshida et al., 2013). Similarly, four phytoene synthase (*PSY*) genes of loquat function differently (Fu et al., 2014) though they cause carotenoid production. Penetration-resistance genes *PEN1*, *PEN2,* and *PEN3* function in overlapping as well as distinct manners in cell wall-based defense (Johansson et al., 2014). Closely related amino acid transporter genes, *AtCAT2*, *3,* and *4* show overlap as well as distinctness in their subcellular localization as well as expression patterns (Yang et al., 2014). *SQUAMOSA promoter binding protein-like* genes, *GhSPL3* and *18* control flowering, second shoots and leaf development with *GhSOC1* binding to the promoter of *GhSPL3* but not *GhSPL18* (Zhang X. et al., 2015). Basic helix-loop-helix (bHLH) TF LONESOME HIGHWAY (LHW) interacts with two other members of the same TF family, *TARGET OF MONOPTEROS5* (*TMO5*) and *TMO5-LIKE1* (*T5L1*) to control various aspects of vasculature development (Ohashi-Ito et al., 2014). All the above examples stand testimony to the fact that genes from the same family, with overlapping expression and/or close phylogenetic relation, can control the same biological process both through independent and related pathways.

In order to deepen the knowledge about the role of NAC TFs in rice seed development, three seed-specific *NAC* genes have been chosen, on the basis of their expression pattern in diverse accessions and their phylogeny. Their promoter sequences have been compared in five accessions. Trait association mapping and association SNP analysis has been done for these genes to identify grain size/weight related allelic variants. Further, their transcriptional properties, including activation, repression, sub-cellular localization and heterodimerization have been examined. Transcriptional properties of multiple forms arising due to transcript fusion or *trans*-splicing among two of these genes have also been analyzed. In short, we have been able to assess the transcriptional properties of three NAC encoding genes, and their association with rice seed size/weight, and hence, put forth their method of regulation.

MATERIALS AND METHODS

Cloning and Expression Profiling of Rice NAC Genes

Rice accessions with variable seed weights namely, *indica*/aromatics cv. Sonasal (SN), Pusa Basmati1 (PB1), *indica* Rice 64 (IR64) and Long Grain Rice (LGR), and a *japonica* cv.

Nipponbare (NB) were grown in the field during crop growing season at NIPGR, New Delhi. Tissues were collected from five seed developmental stages, namely S1 [0–2 DAP (days after pollination)], S2 (3–4 DAP), S3 (5–10 DAP), S4 (11–20 DAP), and S5 (21–29 DAP) of these accessions (following Agarwal et al., 2007, 2011). Total RNA was isolated from these stages and was checked for quality as described previously (Agarwal et al., 2007; Sharma et al., 2012). To remove any contaminating DNA, RNA sample was treated with RNase-free DNase (QIAGEN) and further purified by RNeasy® MinElute Cleanup Kit (QIAGEN) according to the manufacturer's protocol. The purity and concentration of RNA samples were checked by Nanodrop 2000c Spectrophotometer (Thermo Scientific). Total cDNA was prepared from the RNA of different seed stages/tissues from IR64, for amplifying three rice *NAC* genes (*ONAC020*, *ONAC026,* and *ONAC023*) using the Oligo(dT) primers of Superscript® III First-Strand synthesis kit (Invitrogen™). The amplified PCR products generated by Phusion® High-Fidelity Polymerase (New England Biolabs®*Inc.*) were confirmed by sequencing of at least three positive colonies per gene and extended to a maximum of 19 positive colonies for *ONAC020*. GENERUNNER V3.05[1] was used for analyzing the various sequences. Further, a semi-quantitative PCR amplification using Phusion® High fidelity Polymerase was employed for estimating the expression levels of the *trans*-spliced forms, as per the manufacturer's protocol. Primers flanking the region of variation were used for this and the fragments were confirmed by sequencing. The amplicons were later separated on 2.5% MetaPhor™ agarose (Lonza) gel and were subsequently quantified in ChemiDoc™ MP imaging system by Image Lab v5.2.1 (BIORAD). In order to analyze the transcript abundance of the selected *NAC* genes, quantitative real-time PCR (Q-PCR) assay was carried out with gene-specific primers as previously described (Agarwal et al., 2007), using the Real-Time High-Capacity cDNA Reverse transcription Kit (Applied Biosystems™) on 7500 Fast Real-Time PCR System (Applied Biosystems™), in five seed development stages of the five accessions. The *NAC* gene expression data from three biological replicates was normalized with the rice actin gene, *ACT1* and a constant *Ct* value was used for calculation of fold changes by the $2^{-\Delta\Delta Ct}$ method.

Promoter Analysis

To analyze the promoters of the three *NAC* genes, a 2 kb genomic region upstream to the translation start site (ATG) of each gene sequence was amplified from all five rice accessions using Phusion® High-Fidelity Polymerase (New England Biolabs®*Inc.*) and was cloned in pJET1.2 (Thermo Scientific). High-quality sequences from a minimum of three independent colonies were analyzed for each. The promoter sequence of each *NAC* gene was analyzed in PLACE database (Higo et al., 1998) for searching *cis*-regulatory elements and these were manually compared with sequence variants discovered among accessions using Clustal X multiple alignment tool (Thompson et al., 1997).

Analysis of SNPs amongst NAC Genes and Their Association with Seed Size/Weight

For large-scale validation and high-throughput genotyping of *NAC* gene-derived sequence variations mined among five rice accessions and to evaluate their trait association potential, the exons, introns and 2 kb URRs (upstream regulatory regions) and 1 kb DRRs (downstream regulatory regions) of three *NAC* genes were targeted for sequencing. Genomic DNA of 192 low and high grain weight rice accessions (belonging to an association panel) was resequenced employing the multiplexed amplicons resequencing method using Illumina MiSeq next-generation sequencing platform. The high-quality *NAC* gene amplicon sequence reads of each accession were mapped to pseudomolecule (version 6.0) of rice genome[2] and the non-erroneous high-quality sequence variants (SNPs and InDels) were detected as described previously (Saxena et al., 2014; Kujur et al., 2015b). The accuracy and reliability of these identified SNPs and InDels were ascertained by comparing that with the gene promoter regions cloned and sequenced from five different accessions as mentioned above.

For genetic association analysis, phenotyping of the above mentioned 192 rice accessions was carried out. These accessions were grown for two consecutive years (as per randomized complete block design with two replications) at two diverse geographical locations (New Delhi and Tamil Nadu) of India and phenotyped for grain weight (g) trait. The weight of 1000-mature dried grains (at 10% moisture content) harvested from 10 to 15 plants of each accession with replications was estimated and diverse statistical parameters pertaining to grain weight were measured using SPSSv17.0 (Saxena et al., 2014). The population genetic structure, principal component analysis (PCA) and LD decay among the accessions using *NAC* gene-derived SNPs were determined and the association mapping was performed using CMLM (compressed mixed linear model) approach of GAPIT (Kujur et al., 2015a; Kumar et al., 2015). The relative distribution of observed and expected $-\log_{10}$(P)-value of each SNP marker-trait association was compared individually according to their derived quantile-quantile plots. The *NAC* gene-derived potential SNP loci exhibiting significant association with rice grain weight trait at highest R^2 (degree of SNP marker-trait association) and lowest FDR adjusted *P*-values (threshold $P < 1 \times 10^{-7}$) were selected. The *NAC* gene-derived SNPs revealing high association with grain weight were validated in a traditional bi-parental F_4 mapping population developed from inter-crossing of a low (SN with 1000 grain weight: 9.9 g) and medium (IR64: 21.4 g) grain weight parental accessions, by establishing their correlation with the phenotypes of low and high grain weight homozygous mapping individuals and were genotyped in four of each low and high grain weight homozygous mapping individuals using MALDI-TOF mass array SNP genotyping assay (Saxena et al., 2014). The genotyping data was used to constitute the haplotypes within a gene. This information was correlated with the grain weight phenotypic data of the association panel.

[1] http://generunner.net/

[2] http://rice.plantbiology.msu.edu

Transactivation and Transrepression Assay of NAC Genes

The coding sequences of all genes and their different isoforms were fused in frame with the GAL4 DNA-binding domain of pGBKT7 vector (Clontech) and a reconstituted GAL4 TF (rGAL4, described below, manuscript submitted) for the activation and the repression assays, respectively. For the transrepression assay, the GAL4 TF of yeast was partly reconstituted by cloning the activation domain (AD) of GAL4 TF in pGBKT7 vector, which already has the binding domain (BD). AD was cloned between *NcoI* and *EcoRI* sites in the pGBKT7 vector. This construct, called as rGAL4, showed a strong transactivation property since it had both BD and AD. In line with effector-reporter assays to test the repressive ability of TFs (Ohta et al., 2001), when a repressor was fused downstream to the rGAL4 TF, it repressed the activation property of GAL4. The repression of the reporter genes depended on the strength of the repressor domain. Here, the effector is the reconstituted GAL4 fused with the protein of interest. The reporter genes are those present in the yeast strain AH109, i.e., *lacZ*, *HIS3,* and *ADE2*. The activity of all three reporter genes was tested for the transactivation and transrepression assays. For the assays, the ORFs of the genes were cloned between *Eco*RI and *Sal*I sites for *ONAC020* and *ONAC026*; and between *Eco*RI and *Pst*I sites for *ONAC023* in pGBKT7 and rGAL4 vectors. The constructs thus made were further transformed into yeast strain AH109 using EZ-Yeast^{TM} Transformation Kit (MP). The transformed yeast cells were selected by plating onto synthetic drop-out (SD) medium (0.667% yeast nitrogen base, 2% glucose and appropriate auxotrophic amino acid supplements) lacking tryptophan. The transactivation and transrepression properties of the constructs were determined by the differential growth of the yeast colonies on SD/-Trp/-His/-Ade media with 10 mM 3-AT in comparison with relevant controls. These were also patched onto SD/-Trp plates containing 80 mg/L of 5-bromo-4-chloro-3-indolyl-α-D-galactopyranoside (X-α-gal). These results were further confirmed by quantitative β-galactosidase enzyme (ONPG) assay as described in Clonetech® Yeast protocol hand book. The β-galactosidase units were calculated for a minimum of three independent colonies for each construct, according to the formula $2000/t * (OD_{420}/OD_{600})$, where t = time elapsed for incubation in minutes. β-galactosidase units thus obtained were further checked for their level of significance by performing an *F*-test (two-sample for variances) followed by a *t*-test with equal or unequal variance as resulted from *F*-test in Microsoft Excel®.

Sub-cellular Localization and Dimerization

The sub-cellular localization of the selected *NAC* genes as well as their isoforms were predicted by analyzing their complete protein sequences using TargetP 1.1[3], CELLO v.2.5[4], WoLF PSORT[5] and Plant-mPLoc[6]. To determine the sub-cellular localization of the *NAC* genes, the full-length coding region of each gene was amplified using gene-specific primers and cloned into the Gateway® entry vector pENTR^{TM}/D-TOPO® (Invitrogen^{TM}). The resulting constructs were transferred by an LR reaction into the destination vector pSITE-3CA (Chakrabarty et al., 2007) for generation of *NAC-YFP* fusion constructs, under the control of a duplicated cauliflower mosaic virus (CaMV) 35S promoter. Similarly, the CDS of each gene was transferred into destination vectors for bimolecular fluorescence complementation (BiFC) assay namely, pSAT5-DEST-C(175-END)EYFP-N1 and pSAT4-DEST-N(1-174)EYFP-C1 (Tzfira et al., 2005) for N and C-terminus tagging respectively, for analyzing the protein-protein interactions. These were further transiently expressed in the onion epidermal cells by biolistics using Biolistic®-PDS-1000/He particle delivery system according to the manufacture's protocol. Following overnight incubation at 28°C, the onion peels were observed in Leica TCS-SP2 Confocal Laser Scanning Microscope for YFP and mCherry signals at 514 and 594 nm, respectively. The nuclear-specific fluorochore, DAPI, was observed at 351/364 nm wavelength.

RESULTS AND DISCUSSION

Three *NAC* Genes Exhibit Extreme Levels of Transcript Abundance in the Developing Rice Seed

Total transcriptome analysis has emerged as a wonderful tool for an overview of the genes controlling a particular process. Online tools and publicly available databases are a convenient source to analyze our genes or conditions of interest (Agarwal et al., 2014); and much work has been done in rice with the aim to understand and improve its yield and stability (Agarwal et al., 2016). Out of 151 genes encoding NAC TFs in rice (Nuruzzaman et al., 2010), we found nine genes [*ONAC020* (*LOC_Os01g01470*), *ONAC026* (*LOC_Os01g29840*), *ONAC023* (*LOC_Os02g12310*), *ONAC055* (*LOC_Os03g01870*), *ONAC096* (*LOC_Os07g04560*), *ONAC025* (*LOC_Os11g31330*), *ONAC127* (*LOC_Os11g31340*), *ONAC128* (*LOC_Os11g31360*), and *ONAC129* (*LOC_Os11g31380*)] to be expressing in a seed-specific manner (Supplementary Figure S1). Three out of these, *ONAC020*, *ONAC026,* and *ONAC023* show extremely high expression levels in our microarray data (Sharma et al., 2012), an indication of their importance to the process. *ONAC020* and *ONAC026* are closely related in the same phylogenetic branch as *CUC3*, an important gene for seed development (Supplementary Figure S2).

Hence, we decided to study the properties of these three *NAC* genes, to elucidate their similarities or differences. All these three genes have a NAC domain and a TRR. They also have the characteristic subdomains A–E and the NAC repression domain (NARD) (Hao et al., 2010). ONAC020 and ONAC026 have a DLN stretch in the B domain, which is a type of the ERF-associated

[3]http://www.cbs.dtu.dk/services/TargetP/

[4]http://cello.life.nctu.edu.tw/

[5]http://www.genscript.com/psort/wolf_psort.html

[6]http://www.csbio.sjtu.edu.cn/bioinf/plant-multi/

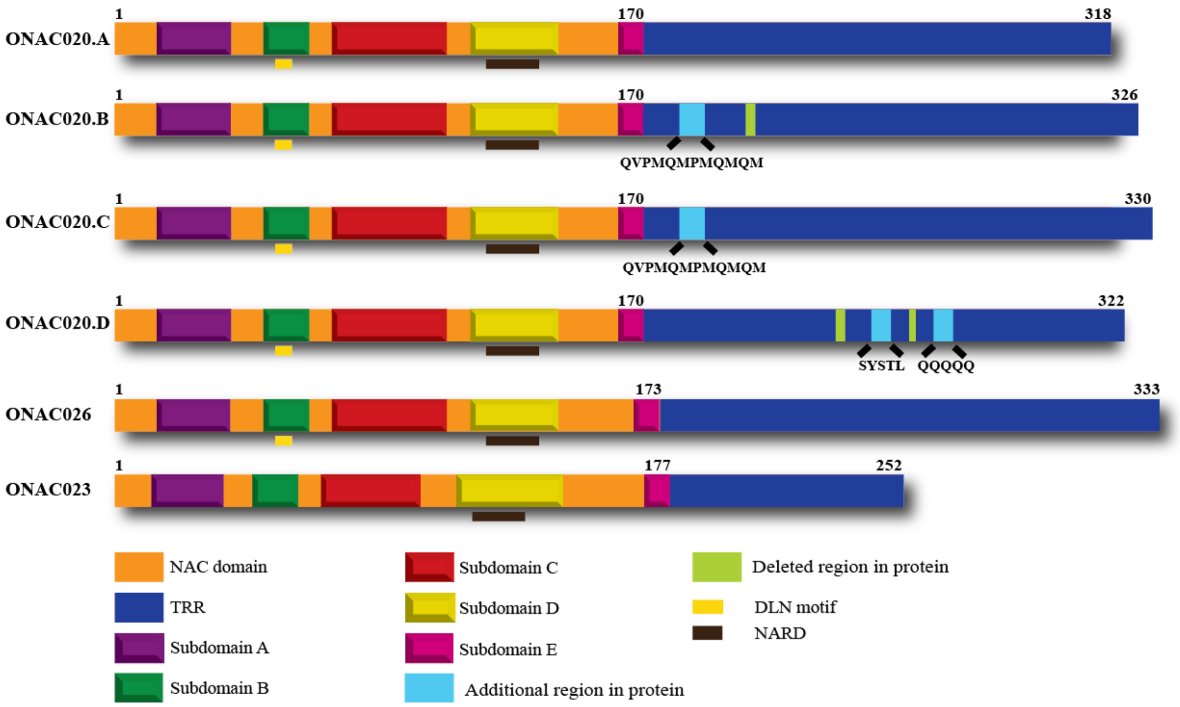

FIGURE 1 | Domains of the three NAC TFs and the different transcript fusion forms of ONAC020. All the proteins have a NAC domain (with subdomains A–E) and a transcriptional regulation region (TRR). ONAC020 and ONAC026 have a putative DLN repressor motif in subdomain B; and all these proteins have a putative NAC repression domain (NARD) in subdomain D. The transcript fusion or *trans*-splicing between ONAC020 and 026 results in three additional proteins (ONAC020.B, 020.C, and 020.D), with insertions/deletions as marked, with respect to ONAC020.A. The color legend indicating the domains is shown at the bottom of the figure.

amphiphilic repression/ EAR repression motif in plants (Ohta et al., 2001; Kagale et al., 2010). Preliminary analysis shows that this motif contributes to the repressive activity of these two proteins. Additionally, we identified three *trans*-spliced forms between ONAC020 and ONAC026, described further on, with changes in TRR only (**Figure 1**).

In order to assess their seed-specificity and relative importance, we examined the expression of the three *NAC* genes in four *indica* and one *japonica* accessions, in five stages of seed development, namely S1–S5 (Agarwal et al., 2007, 2011), by Q-PCR (**Figure 2**). SN has small sized seeds while LGR has heavier and bigger ones. PB1, IR64 and NB have medium weight grains in that order (Supplementary Figure S3). The expression levels for the genes exhibit differences in accessions. NB has the highest expression levels for *ONAC020* and *ONAC026* and lowest for *ONAC023*. The three genes also express in the S1 stage of NB, unlike other accessions (**Figure 2D**). Same is the case for small grained SN, showing a higher expression of *ONAC020* and *ONAC026*, in comparison to *ONAC023*, which is also reflected in the promoter sequence similarity analyzed further in this paper. They also show considerable expression in S2 stage (**Figure 2A**) unlike large grained LGR (**Figure 2E**). S2 represents organ initiation stage (Agarwal et al., 2011) and expression of these genes are indicative of their early role in a small-seeded variety. In PB1, the expression levels for the genes are mostly comparable in S4 and S5 stages while the same is true for S3 stage of IR64 (**Figures 2B,C**). Additionally, all the genes show a drastic drop in expression levels in the S5 stage of IR64 as compared to S3 and S4 (**Figure 2C**). In LGR, the genes increase in expression in S3, at the start of grain filling (Agarwal et al., 2011) and remain high till S5 (**Figure 2E**). Since most of the genes show low/no expression in S1, with higher levels in later stages of seed development, this possibly implies their role in one of the processes occurring during grain filling and hence, a role for grain weight. Also, these subtle variations indicate toward a degree of distinctness along with redundancy in the roles of the three seed specific *NAC* genes during rice seed development. In a similar manner, *CUC* genes involved in the initiation of ovule formation and cotyledon separation in *Arabidopsis*, show partially overlapping expression pattern and also interact amongst themselves (Goncalves et al., 2015). Since the *NAC* genes analyzed here also interact, as is shown further in the paper, the differences in their levels in the five accessions with variable seed weights, may modulate downstream seed development processes to variable extents.

The expression of a gene is controlled directly by the presence of *cis*-elements on its promoter and their arrangement, apart from other factors. So, the 2 kb upstream region of all five genes, in the five accessions was amplified, sequenced and compared and was checked for the presence of *cis*-regulatory elements (**Figure 3**; Supplementary Table S1). For *ONAC020*, the promoter sequence is same for IR64 and PB1, while the deletions and SNPs are similar for the other three accessions.

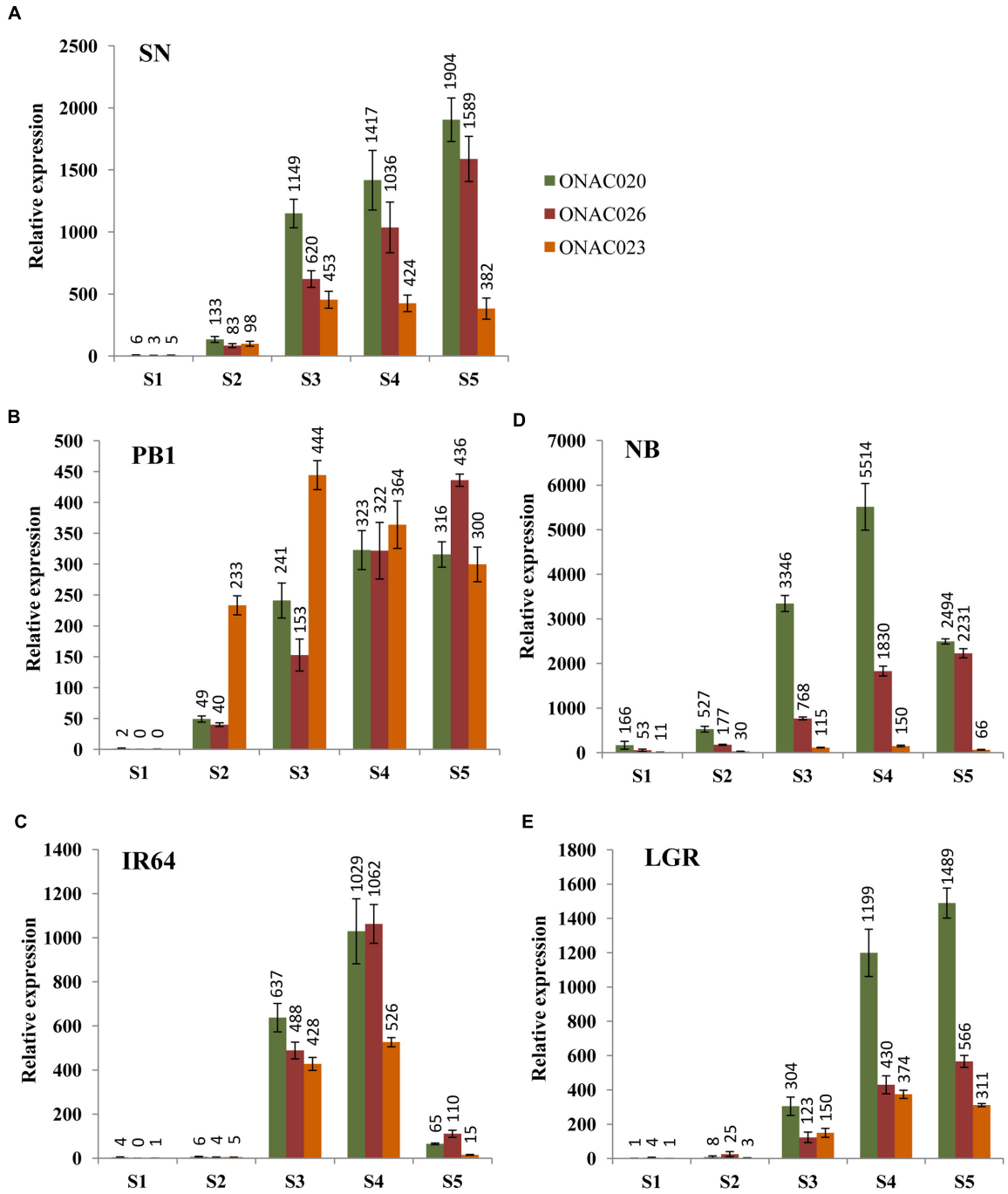

FIGURE 2 | Relative expression levels of three seed-specific *NAC* genes, namely, *ONAC020*, *ONAC026* and *ONAC023*, in five rice accessions with varying seed weights. The accessions are **(A)** Sonasal (SN), **(B)** Pusa Basmati 1 (PB1), **(C)** *indica* Rice 64 (IR64), **(D)** Nipponbare (NB) and **(E)** Large Grain Rice (LGR), arranged according to increasing weights. The relative expression for each gene, in each accession has been studied in all the five stages of seed development (S1–S5), as mentioned in the color legend in **(A)**.

The deletion of a 30 bp stretch at 1118 bp upstream to the translation start site creates an additional RY element (CATGCA) and a Napin motif in IR64 and PB1 (**Figure 3A**). RY elements recognized by the B3 domain transcriptional activators like ABI3 and FUS3, act as important regulators of seed storage protein expression in dicots (Kawakatsu and Takaiwa, 2010). The sequence TACACAT, designated as Napin motif, is an evolutionarily conserved motif activating the expression of the seed storage protein genes in *Brassica* and soybean seeds (Jopcik et al., 2014). Correlation with the expression pattern (**Figure 2**)

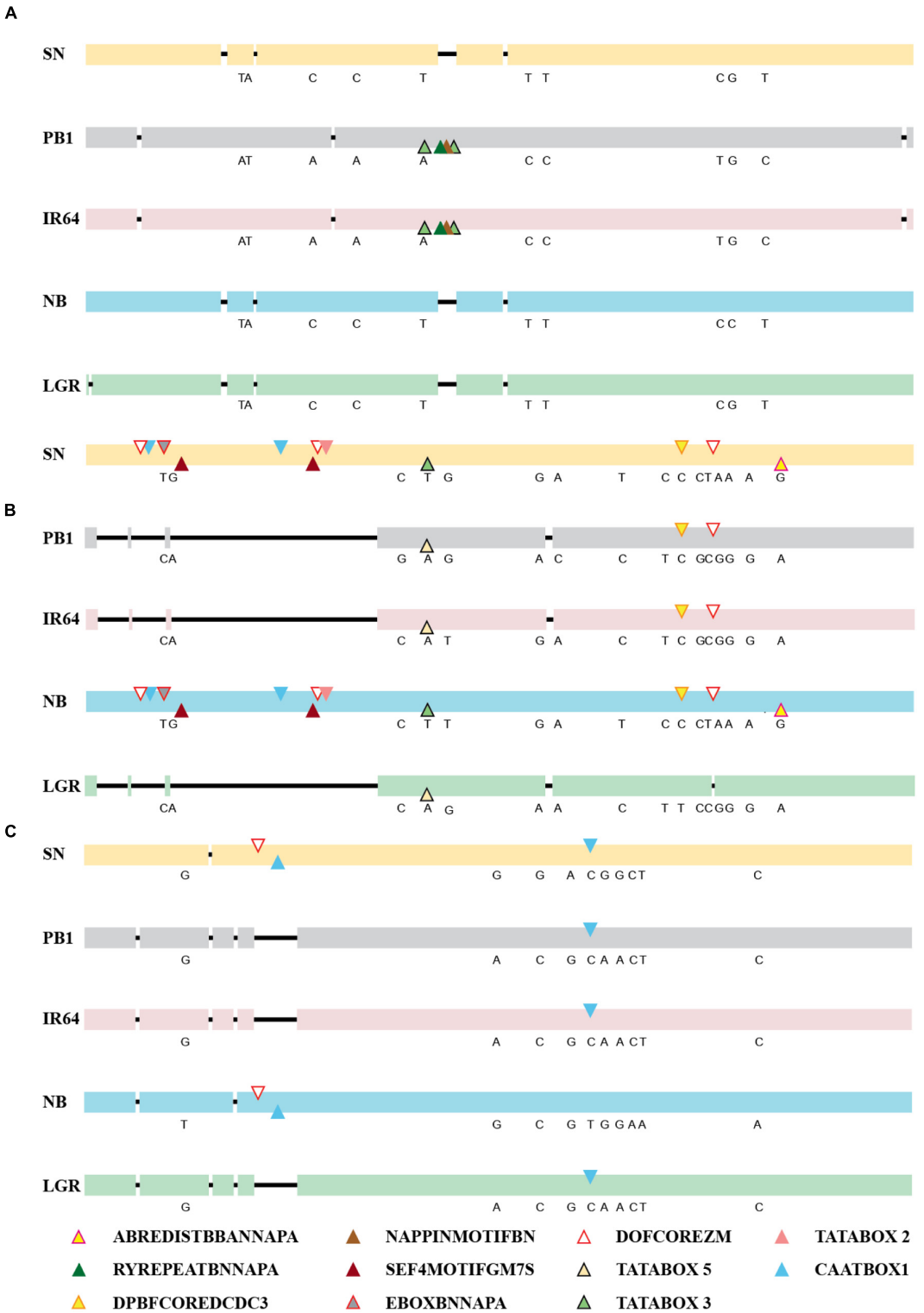

FIGURE 3 | Alignments of the promoter sequences of (A) *ONAC020*, (B) *ONAC026,* and (C) *ONAC023* in five accessions of rice, namely SN, PB1, IR64, NB, and LGR. The deletions have been shown by a black line. The SNPs among the accessions have been written in their corresponding positions (Supplementary Tables S3 and S4). Seed-specific elements with changes amongst different accessions (Supplementary Tables S5 and S6) have been marked with colored triangles, as per the color legend. The promoter elements identified on the positive strand have been marked on the upper side and those on the negative strand, on the lower side. In longer stretches of deletion, in *ONAC026* and *ONAC023*, repeatedly occurring elements in 300 and 75 bp stretches, respectively, have been represented only once by their respective triangles.

indicates that genotypic differences in accessions may be a controlling factor for the expression pattern of this gene. For the promoter of *ONAC026*, SN and NB have an additional stretch of 677 bp, holding multiple copies of DOFCOREZM (AAAG) and EBOXBNNAPA (CANNTG, **Figure 3B**), a probable explanation for the higher expression levels in these two accessions (**Figures 2A,D**). EBOXBNNAPA, conserved in many seed storage protein promoters is critical for directing seed-specific expression (Ellerstrom et al., 1996; Ravel et al., 2014), while DOFCOREZM is the recognition core of DOF proteins, reported to activate several storage protein genes (Yamamoto et al., 2006; Abraham et al., 2016). For *ONAC023*, the promoter sequences for SN and NB are similar, though there are SNPs amongst the two as well, and hence a similar expression pattern though levels are different. The RY element is present in multiple copies in all the promoter sequences analyzed except *ONAC023* (Supplementary Table S1), and could be a reason for its lower expression, comparatively with the other genes, in all accessions, except PB1 (**Figure 2**). On the other hand, *ONAC020*, with relatively higher expression levels, shows an abundance of the RY elements in the promoter (Supplementary Table S1). Promoter sequence analyses have revealed that the promoters of *japonica* and SN are highly similar, even though the latter is an *indica* accession and hence similar expression patterns. These two accessions are related evolutionarily (Parida et al., 2009). The overall nucleotide composition in all the sequences has been maintained due to multiple transversions. The promoter sequences have also been found to be relatively conserved closer to the start codon. Hence, the differences in expression patterns amongst the three genes in the five accessions, can be only partly accounted for by the differences in the promoter sequences. At this juncture, we can hypothesize that these may be due to varietal differences in upstream regulatory factors, amongst these accessions, which are yet to be elucidated. Additionally, chromatin modifiers have been reported to regulate the expression of different TFs involved in embryo and seedling development (Wagner, 2003; Reyes, 2006). Hence, differences in chromatin modifications amongst the five accessions may result in variation in expression in this case, a hypothesis which needs to be verified.

NAC Genes Associate with Rice Grain Size/Weight Phenotype

Since SNPs were observed in the promoter sequences of the three genes amongst five accessions with variable seed weights, we aimed to elucidate any sequence variations in the entire genes, which were associated with seed size/weight character. An integrated genomic strategy by combining SNP-based association analysis, selective genotyping in bi-parental mapping population and molecular haplotyping was employed. For association mapping of grain size/weight traits in rice, targeted NGS-based resequencing of coding and non-coding (intronic and regulatory) sequence components of the three genes amongst 192 diverse low and high grain weight accessions (association panel) was performed. This identified 330 high-quality sequence variants, including 254 SNPs and 76 InDels in the three genes (Supplementary Table S2; Supplementary Figure S4). We have been able to reaffirm most of these SNPs and InDels by the individual sequencing of the promoter regions from five different accessions (Supplementary Tables S3 and S4). Almost 40% SNPs and 43.4% InDels change the *cis*-elements in the URRs (Supplementary Tables S5 and S6), some of which are essential for seed-specific expression (**Figure 3**). The 49 sequence variants mined from the exons of *NAC* genes include both 40 synonymous and non-synonymous SNPs as well as nine InDels showing frameshift mutations (Supplementary Table S2). SNPs discovered from the three *NAC* genes emphasize their utility in targeted genetic mapping and association analysis of important agronomic traits, including seed size/weight, in rice.

The use of these 254 sequence variants in population genetic structure analysis differentiates the association panel into two population clusters, POPI and POPII. The LD patterns exhibit broader LD estimates (r^2: 0.23–0.89) and faster LD decay (r^2 decreased half of its maximum value) nearly at 50–100 kb physical distance of rice chromosomes. This is agreed well with earlier reports on prerequisite of LD decay for effective gene-based and genome-wide association mapping studies in rice (Zhao et al., 2011; Huang et al., 2012, 2013). Hence, this LD decay is adequate enough for association mapping of SNPs with the grain size/weight trait in rice. This trait is known to follow a complex quantitative genetic inheritance pattern as observed by field phenotyping of 192 accessions in the association panel, across two diverse geographical locations, over 2 years. The trait exhibits a normal frequency distribution pattern with a broader phenotypic variation (13.5–42.7 g, mean ± SD: 26.5 ± 4.8, mean CV: 18.1%) as well as a higher heritability for grain weight (mean H^2: 80%) (Supplementary Figures S5A,B). This necessitates essentiality of an integrated genomics-assisted breeding strategy for quantitative dissection of this complex trait. So a combinatorial strategy has been deployed involving association mapping, selective genotyping in bi-parental mapping population and molecular halotyping. Interestingly, the genetic association analysis of the *NAC* genes with the grain size/weight trait in rice identifies two regulatory SNPs located in the URRs of *ONAC026* and *ONAC023* which display remarkable association with the grain size/weight traits at a *P*-value $\leq 10^{-6}$. Also, a 7 bp regulatory InDel present in the URR of *ONAC020* is significantly associated with grain size/weight trait in rice (**Table 1**; **Figure 4A**). This is located within the *cis*-element "CACTFTPPCA1" (Supplementary Table S6), which is responsible for mesophyll-specific expression in C_4 plants (Sharma et al., 2011). This functional regulatory InDel strongly associated with grain size/weight trait can serve as a potential candidate for marker-assisted genetic enhancement of rice. All the three natural sequence variants have diverse association potential for grain size/weight traits in rice on the basis of phenotypic variation (PVE) among the 192 accessions (**Table 1**). The sequence variants have been successfully validated in four of each low (9–12 g) and high (22–25 g) grain weight homozygous individuals of an F_4 mapping population between IR64 and SN (Supplementary Figures S6A–C; **Table 1**). The homologs of *ONAC020* and *ONAC026* have been reported to be associated with the expression of genes encoding grain storage

TABLE 1 | *NAC* gene-derived SNPs associated with grain size/weight traits (length, width, and weight) in rice.

OsNAC genes	Physical positions (bp)	Associated sequence variants (SNPs/InDels)	Functional annotation	*P*-value	PVE[a] (R^2 %)
ONAC020-INDEL01	239034–239040	ATAC/ATACTAC	URR	3.5×10^{-7}	34
ONAC026-SNP50	16722376	A/G	URR	1.2×10^{-6}	31
ONAC023-SNP15	6403097	C/T	URR	2.7×10^{-6}	27

[a]*PVE, phenotypic variation explained.*

protein during endosperm development in wheat (Plessis et al., 2013). The above three genes were selected as target candidates for grain size/weight trait regulation by their further validation through molecular haplotyping in rice. Molecular haplotyping of *ONAC026* reveals 53 SNPs from diverse coding and non-coding (including two coding non-synonymous, three intronic, 25 DRR and 21 URR SNPs) sequence components of the gene, which form six haplotypes in the gene. These exhibit a higher degree of LD ($r^2 > 0.85$ with $P < 1.0 \times 10^{-6}$) resolution (**Figure 4B**; Supplementary Figure S6D). Remarkably, the grain size/weight trait associated SNP in *ONAC026* (**Table 1**) shows strong association potential for high/medium (haplotype I) and low grain weight (haplotype II) differentiation in rice. In addition, four novel haplotypes have been identified (with diverse allelic recombination) revealing differential trait association potential for rice grain size/weight (Supplementary Figure S6D). A number of known genes underlying QTLs regulating grain size/weight have been cloned and characterized till date in rice (Zuo and Li, 2014). Also, many TFs controlling seed development have been documented (Agarwal et al., 2011). Hence, the three seed-specific genes are strongly associated with seed size/weight and are potential markers for genetic enhancement of the rice crop.

ONAC020 Shows Existence of Transcript Fusion with *ONAC026*

For molecular analysis of the three *NAC* genes, they were amplified from rice cDNA, from developing seed stages. Surprisingly, *ONAC020* showed the existence of multiple forms. Analysis showed that they have arisen due to transcript fusion or *trans*-splicing with *ONAC026* (Supplementary Figures S7A,B). Rice transcriptome has been known to exhibit *trans*-splicing (Zhang G. et al., 2010), which is regulatory in nature, and increases the proteome diversity. *ONAC020* and *ONAC026* share 89% homology at the sequence level. *ONAC020.A* is the main transcript of the gene and eventually four different proteins are formed from six transcripts (A to D), with insertions/deletions in TRR only (**Figure 1**). We have named the *trans*-spliced transcripts as forms of *ONAC020* because of their higher homology with this gene. The existence of these was confirmed by semi-quantitative RT-PCR in five stages of seed development (Supplementary Figure S7C). *ONAC020.D* represents the typical form of a chimeric transcript (Dubrovina et al., 2013) between *ONAC020* and *ONAC026*. The other forms, however, seem to be the ones wherein a part of *ONAC026* has been spliced within the *ONAC020* transcript. Such transcripts may be generated due to high homology, amongst the two genes, in spliced regions. We seem to have observed a unique *trans*-splicing event, which needs to be validated further. *ONAC020* and *ONAC026* show maximum expression in the S3 and S4 stage of seed development, so do their *trans*-spliced forms. In other stages, as the levels of *ONAC020* and *ONAC026* decrease, so does the expression of the *trans*-spliced forms (Supplementary Figure S7C; Supplementary Table S7). Gene fusions, brought about by chromosomal rearrangement are reported in many neoplastic cells. Such chimeric transcripts can even result from the fusion of two transcripts without a remarkable DNA rearrangement, and can occur even in normal cells (Jividen and Li, 2014). Even though, the existence of such chimeras are not that well-established in plants, RNA sequencing experiments points to the occurrence of large number of these fusions in plants including rice (Zhang G. et al., 2010). NACs are known to have alternatively spliced forms. *OsSWN2* has an alternatively spliced form which does not cause transactivation (Yoshida et al., 2013). A splice variant of SND1, PtrSND1-A2(IR), does not transactivate or bind to DNA. It rather binds with SND1 and represses its activity (Li et al., 2012). The transcript fusion/*trans*-splicing events may vary under different accessions and needs to be looked into. Nonetheless, the variation only in the TRRs of ONAC020 leads to the speculation that the multiple forms have the same downstream targets with changes in regulatory property only, which has been examined further.

Transcriptional Activation/Repression Ability of the Three NAC TFs Varies

Transcription regulation is a major property of TFs and most TFs harbor an activation and/or repression motif. Transactivation by the three NAC TFs was analyzed by their ability to activate the reporter genes in yeast strain AH109, in both qualitative and quantitative manners. The assay revealed that all the three proteins ONAC020.A, ONAC026, and ONAC023 can activate the reporter genes to a certain extent (**Figure 5A**; Supplementary Figures S8A,C). Since, they also have DLN and NARD motifs (**Figure 1**), their transrepression ability was also checked for, by fusion with a yeast rGAL4 TF. rGAL4 is a partly reconstituted GAL4 TF with both BD and AD domains in the same vector, in frame, which acts as a strong activator (**Figure 5B**). Since GAL4 is a strong activator in yeast, a repressor fused with the same will decrease/nullify its activation property depending on the strength of the repressor (manuscript communicated). Here, the effector is the reconstitued GAL4 fused with the protein of interest. The reporter genes are those present in the yeast strain, i.e., *lacZ*, *HIS3,* and *ADE2*. ONAC020.A and ONAC026 with the DLN repressor motif are able to completely abolish the activation by rGAL4 TF while ONAC023 is not a repressor (**Figure 5B**; Supplementary Figures S8B,D). Thus, ONAC020 and ONAC026 are bifunctional TFs, with low ability of activation and strong repressive activity. On the other hand, ONAC023 is a weak activator. Adding to above properties, is the fact that the presence of ONAC020 and ONAC026 cause the yeast cells

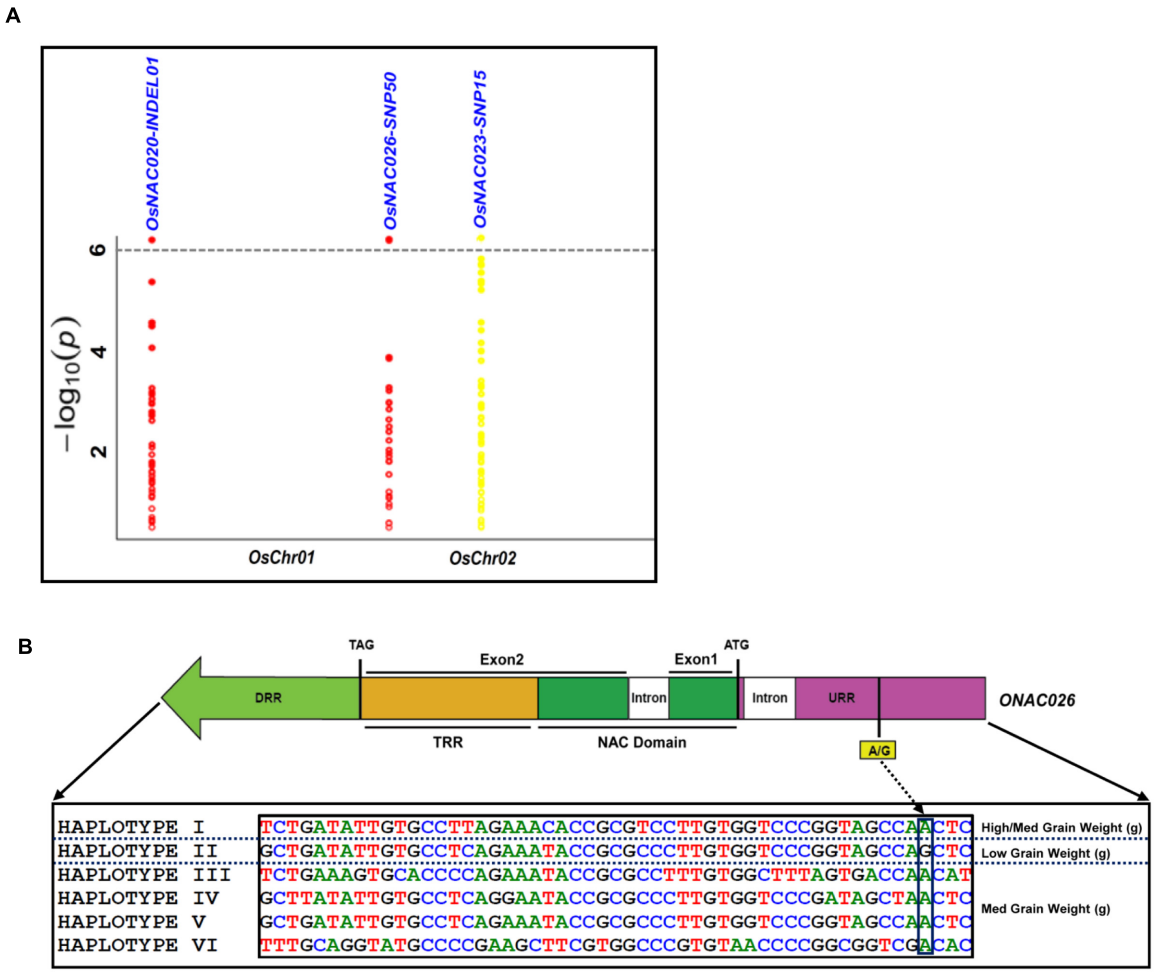

FIGURE 4 | Association of *NAC* genes with grain weight. (A) represents significance of the sequence variants identified for the *NAC* genes, for their association with grain weight trait by the Manhattan plot. The relative density of the sequence variants identified from the *NAC* genes are taken on the X-axis with their $-\log_{10}(P)$ used for scanning the significant trait associated sequence variant on the *Y*-axis. **(B)** Validation of the strong association of the InDel in the URR of *ONAC026* gene with the grain weight trait, where DRR and URR represent the downstream and upstream regulatory region respectively, ATG and TAG correspond to the translation start and stop codons respectively and TRR stands for the transcriptional regulatory region. The InDel clearly differentiates the six haplotypes into low grain weight and medium/high grain weight groups.

to grow at a much slower rate (Supplementary Figure S8E). Again, this may be due to their strong repressive activity in yeast cells. In *Arabidopsis*, the EAR or DLN repressor proteins interact with the co-repressor TOPLESS/TPL through their "DLN" motif (Causier et al., 2012; Oh et al., 2014). TPL related proteins cause repression of target genes by interacting with histone deacetylases/HDA (Zhu et al., 2010). In yeast, repression is caused upon interaction of repressors with HDA (de Bruin et al., 2008; Lorenz et al., 2014), which bind to the promoters of G1 cyclin genes, causing repression of cell cycle (Takahata et al., 2009). It is possible that the NAC TFs causing repression are participating in a similar pathway in yeast, causing repression and hence affecting cell cycle and their growth. Since the various forms of ONAC020 differ in their TRR, they were also tested for their activation/repression properties. The three forms show slight variation in their activation ability. Interestingly, ONAC020.B, with a small deletion in the TRR shows complete abolishment of the repressive ability (**Figures 5A,B**; Supplementary Figures S8A–D). These differences in the activation/repression abilities support our above stated hypothesis of the *trans*-spliced forms having variations in regulatory properties, owing to differences in TRR only.

NAC TFs which function as repressors, possess two types of repression motifs, the NARD motif in the NAC domain and/or the DLN repressive motif. *Arabidopsis* AIF (ANTHER INDEHISCENCE FACTOR) has NARD and suppresses the jasmonic acid biosynthesis pathway to control anther dehiscence, during early flower development (Shih et al., 2014). *Arabidopsis* CBNAC is a calmodulin regulated transcriptional repressor of basal plant defense during normal growth (Kim et al., 2012). VND-INTERACTING1 (VNI1) interacts with VASCULAR-RELATED NAC-DOMAIN7 (VND7) to control the formation of

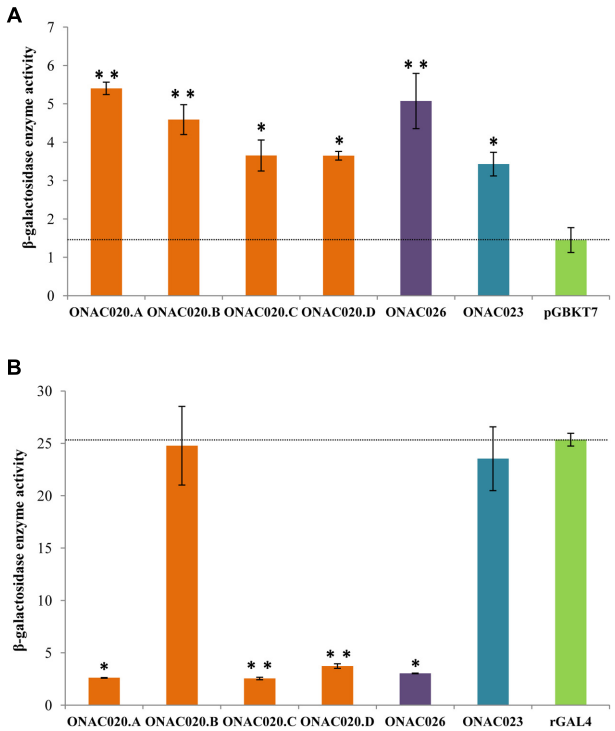

FIGURE 5 | Transcriptional ability of the NAC TFs by (A) transactivation and (B) transrepression assays of the three NAC proteins, and the proteins generated by *trans*-splicing of *ONAC020/ONAC026* (i.e., ONAC020.B–D). The graphs show the relative β-galactosidase activity of the NAC TFs by the ONPG assay on the *Y*-axis. The vector pGBKT7 is the negative control for transactivation assays in **(A)** and the reconstituted GAL4 TF (rGAL4) is the positive control for transrepression assays in **(B)**. The standard error bars have been shown. The values which are significant with a *P*-value cut off ≤0.05 and ≤0.005 are indicated by a single and double asterisks, respectively.

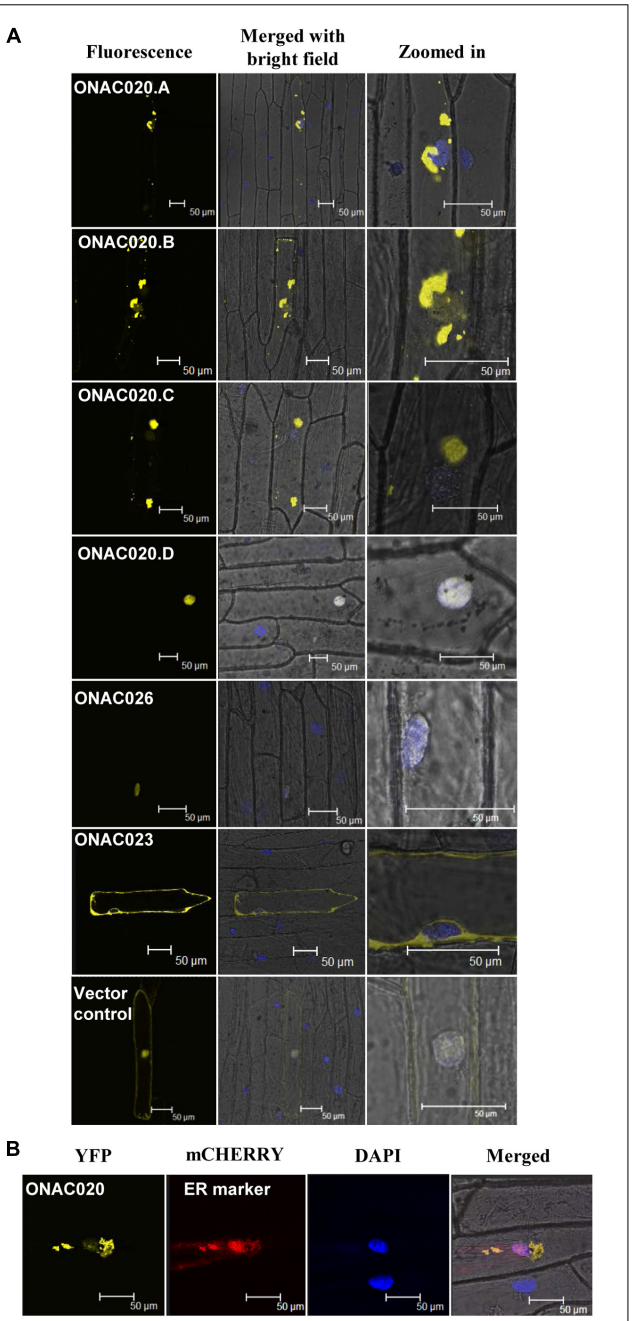

FIGURE 6 | Subcellular localization of NAC-YFP fusion proteins. (A) represents the localization images of the three main NAC TFs and the transcript fusion forms of ONAC020. The left column indicates the localization image of the YFP fusion protein at 514 nm. The middle column is the merger of bright field, fluorescence image and DAPI stained nucleus while the right column shows an enlarged image. **(B)** represents the co-localization of ONAC020 with endoplasmic reticulum (ER) marker. The protein names have been mentioned and the scale bar indicates 50 µm.

Arabidopsis xylem vessels by acting as a repressor. VNI1 has a putative EAR domain in the C-terminal PEST motif (Yamaguchi et al., 2010). ATAF2 negatively regulates pathogenesis related genes (Delessert et al., 2005). Just like the genes examined here, few NAC TFs have been shown to possess both activation and repression domains and hence act as bifunctional TFs, such as GmNAC20 (Hao et al., 2011) and VNI2 (Yang et al., 2011). Apart from NACs, *Arabidopsis* WUSCHEL and Histone-Like NF-Y are bifunctional TFs (Ceribelli et al., 2008; Ikeda et al., 2009). The variation in activation/repression properties and expression patterns of the three NAC TFs provides fuel to our theory of the genes having overlapping as well as independent functions during rice seed development. Moreover, the variations in the properties of the *trans*-spliced forms of the genes further emphasizes their role as competitors (Reddy et al., 2013).

The NAC TFs Localize Differently and Heterodimerize

For TFs to carry out their function, they need to be localized to the nucleus, either independently or in conjunction with other TFs. NAC TFs have been predicted to have transmembrane domains, nuclear localization signals/NLS and nuclear export signals/NES (Olsen et al., 2005). The three NAC TFs analyzed here have been predicted to possess various localization signals (Supplementary Table S8). Their fusion products with YFP

show that only ONAC026 is completely nuclear localized in onion peel cells. ONAC023 is localized in the cytoplasm (**Figure 6A**) and ONAC020.A in the endoplasmic reticulum (**Figures 6A,B**). Amongst the various forms arising due to *trans*-splicing, ONAC020.B and ONAC020.C are localized like the main form. Interestingly, ONAC020.D goes to the nucleus (**Figure 6A**). Splicing of membrane bound TFs, such as ER-membrane bound bZIP60 leads to a functional nuclear-targeted form (Seo, 2014) as is the case with ONAC020.D. CELLO and Plant-mPLoc predict nuclear localization of all six proteins (ONAC020, ONAC026, three *trans*-spliced transcripts and ONAC023; Supplementary Table S8). However, only ONAC020.D and ONAC026 were localized in the nucleus (**Figure 6**). Only sometimes localization predictions are not validated (Chaturvedi et al., 2014). The cause for this variation can only be hypothesized here. The protein conformation may not be exposing the nuclear localization signal efficiently for such results. Additionally, ER is required for protein trafficking (Chen et al., 2012) and the proteins might have accumulated there. Amongst the other NACs, ATAF1, a founding member of the NAC family, is localized to the nucleus (Lu et al., 2007). NTL4 is processed and localized to the nucleus only upon heat stress (Lee et al., 2014). A few NAC TFs localize to organelles other than nucleus. MaNAC6 in banana gets localized to the cell membrane, cytoplasm, and nucleus unlike the nuclear localized MaNAC1–5 (Shan et al., 2012). ANAC of *Arabidopsis* gets localized in both the nucleus and cytoplasm. It interacts with two RING-H2 domain proteins, both of which also show similar localization patterns. The authors hypothesize that the proteins may interact in the nucleus and then be exported outside (Greve et al., 2003). Apart from this, a significant feature of many NAC TFs is the transmembrane domain (TM). ANAC017 is localized to the ER. Subsequent to the cleavage of the TM domain, it gets localized in the nucleus and mediates stress resistance (Ng et al., 2013). Similarly, membrane bound *Arabidopsis* NTM1 and bZIP are activated by proteolytic cleavage leading to a functional nuclear targeted form (Kim et al., 2007; Iwata et al., 2008).

Since only ONAC020.D and ONAC026 are nuclear localized, the others may do so either by proteolytic cleavage during seed development or by heterodimerization. The same was proven when it was found that ONAC026 interacts with three ER localized forms of ONAC020 (ie. ONAC020.A, ONAC020.B, and ONAC020.C), and all the dimers are directed to the nucleus in BiFC experiments. Additionally, ONAC026 and ONAC023 also interact, and the dimers are found in either the nucleus or ER, in an equal number of experiments (**Figure 7**). Surprisingly, ONAC020.D does not interact with ONAC026. Neither does it interact with the other three *trans*-spliced forms of ONAC020. For TFs to activate/supress target genes, they have to be targeted to the nucleus. Often, TFs dimerize with others having a NLS and are thus nuclear localized (Withers et al., 2012; Nayar et al., 2014). Such is the case with the three *trans*-spliced forms of ONAC020 and ONAC023, which interact with ONAC026 (**Figure 7**). However, ONAC023–026 dimers were found in the ER as well. Sometimes, TFs are localized to ER, and get nuclear localized upon proteolytic cleavage (De Clercq et al., 2013; Hofmann, 2013; Ng et al., 2013). This may be because ER serves as the port of entry of many proteins which are destined to other organelles along with the ER resident proteins. It acts as the site of folding and maturation of proteins (Galili et al., 1998). Additionally, it is possible that other interacting partners are required for the complex to completely localize to the nucleus, which can be proven only by further experimentation.

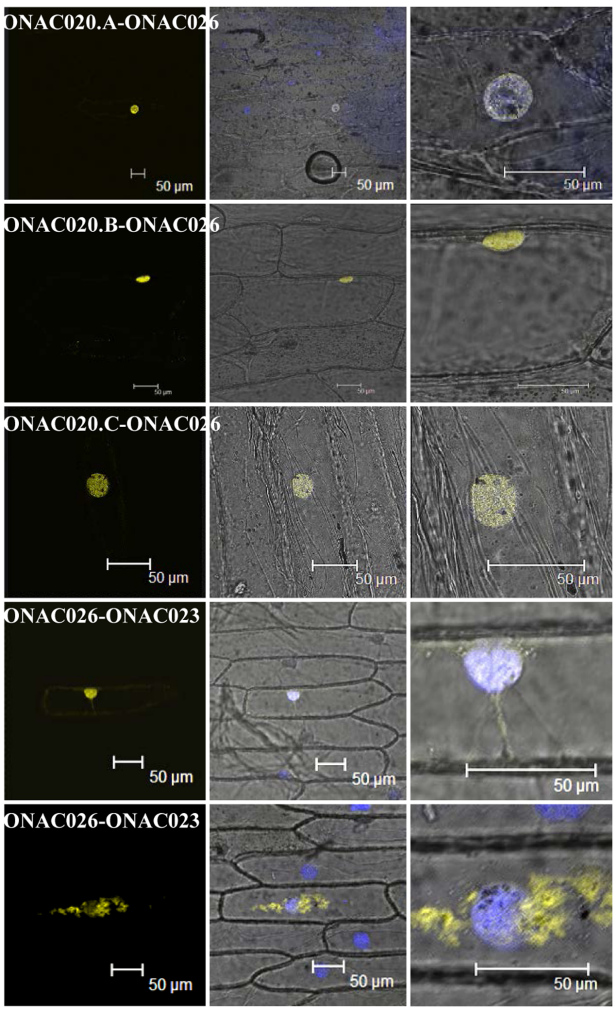

FIGURE 7 | Interactions amongst the NAC proteins by BiFC. The left column indicates the localization image of the reconstituted YFP protein at 514 nm. The middle column is the merger of bright field and fluorescence images while the right column shows an enlarged image of the nucleus. The nucleus has been stained by DAPI dye. The interacting protein names have been mentioned and the scale bar indicates 50 µm.

A protein lacking a few functional domains is called as an interfering protein/small interfering peptide (Seo et al., 2011). Isoforms of a TF may lack one of the functional domains, keeping other functions intact and hence behave as dominant-negative regulators or competitors (Reddy et al., 2013) and hence ONAC020.D may function as a dominant-negative regulator. Also, ONAC020.A and its forms do not interact with ONAC023

in our BiFC experiments. Amongst known NAC interactors, GmNAC30 and GmNAC81, which are bifunctional TFs interact in the nucleus and bind to the promoter of *VPE* to integrate various stress responses (Mendes et al., 2013). ANAC096 interacts with ABF2 and ABF4 to impart abiotic stress resistance (Xu et al., 2013). Regulatory networks are an important aspect of seed development, and often involve formation of homo/heterodimers between TFs. In barley, HvVP1 interacts with HvGAMYB and BPBF and represses their DNA-binding activity (Abraham et al., 2016). Maize OPAQUE2 dimerizes with MADS47 and enhances its activation of zein genes (Qiao et al., 2016). Similarly, the well-known LAFL network of seed maturation involves formation of many hetero and homodimers (Agarwal et al., 2011; Jia et al., 2014). Hence, the observation that ONAC020 and ONAC023 interact with ONAC026 is important. Moreover, the localization of these heteromers to the nucleus, further signifies the transcriptional role of these dimers and makes for an interesting study. Hence, the seed-specific NAC TFs interact amongst each other and subsequently enter the nucleus to bring about their function.

Role-Play amongst NAC Members during the Progression of Rice Seed Development

In spite of huge efforts, very limited number of potential robust genes/QTLs have been deployed in marker-assisted genetic improvement of rice and limited information is available about their functional aspects. Here, we have identified three seed-specific NAC TFs, with variation in their expression patterns in five different rice accessions over a range of seed size/weight trait, which is controlled by sequence variations in the promoter regions to a certain extent. All the three genes *ONAC020*, *026,* and *023* are significantly associated with seed size/weight trait in rice, with the associated sequence variants in the URRs. *ONAC020* and *026* exhibit *trans*-splicing. ONAC026, a strong repressor, dimerizes with ONAC020.A, ONC020.B, ONAC020.C, and ONAC023 and these heterodimers are nuclear localized. Hence, a complex is formed amongst ONAC020, 026, and 023. It is highly probable that the repression/ activation levels of the complex is an overall combination of all the components, which may vary amongst the various stages of seed development amongst accessions, as indicated by expression pattern and sequence variations. For example, the expression pattern for *ONAC020* and *ONAC026* is similar for an accession, though the levels may be different and would suggest the significance of these interactions in controlling seed development. In conclusion, it is proposed here that in the five rice accessions, there is interaction and subsequent nuclear localization amongst three seed-specific NAC TFs, with variable levels of activation and repression. The fusion forms act as competitors or interfering peptides. The genes may be a part of a bigger network controlling seed size/weight in these accessions, based on their differential expression patterns and association analysis. Functional characterization of genotypes/transgenic plants with altered expression of these genes or with alleles controlling high seed size/weight will provide further insight about the same. These genes can be useful for rapid quantitative dissection of complex grain size/weight trait in rice to eventually accelerate the development of rice cultivars with high grain weight and yield.

ACCESSION NUMBERS

The accession IDs at NCBI BankIt database for cDNA sequences from IR64 are KX953272 for ONAC020.A, KX953273 for ONAC020.B, KX953274 for ONAC020.C, KX953275 for ONAC020.D, KX953276 for ONAC020.E, KX953277 for ONAC020.F, KX953278 for ONAC026, and KX953279 for ONAC023. The NCBI BankIt accession IDs for 2kb upstream regions of ONAC020 are KX953280 from SN, KX953281 from LGR, KX953282 from IR64 and KX953283 from PB1; for ONAC026 are KX953284 from SN, KX953285 from LGR, KX953286 from IR64, KX953287 from PB1; for ONAC023 are KX953288 from SN, KX953289 from LGR, KX953290 from IR64 and KX953291 from PB1.

AUTHOR CONTRIBUTIONS

IM performed the experiments. SD performed the transrepression assay of the multiple forms of the genes. AM collected seeds from different accessions and isolated RNA. PA and IM wrote the manuscript and prepared the figures. PA conceptualized and supervised the experiments. All authors have reviewed the manuscript.

FUNDING

IM, SD, and AM acknowledge the Junior and Senior Research Fellowship from University Grants Commission (UGC). PA thanks the Department of Biotechnology (DBT), India for grants supporting research and NIPGR core grant.

ACKNOWLEDGMENTS

The authors thank Prof. Akhilesh K. Tyagi, Delhi University, South Campus, for regular scientific inputs and critical reading of the manuscript. The authors express deep gratitude to Dr. Swarup K. Parida, NIPGR, for helping with the association analysis. The authors are also thankful to Dr. A. K. Singh, IARI, for providing the seeds of LGR accession of rice.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2016.01638/full#supplementary-material

REFERENCES

Abraham, Z., Iglesias-Fernandez, R., Martinez, M., Rubio-Somoza, I., Diaz, I., Carbonero, P., et al. (2016). A developmental switch of gene expression in the barley seed mediated by HvVP1 (Viviparous-1) and HvGAMYB interactions. *Plant Physiol.* 170, 2146–2158. doi: 10.1104/pp.16.00092

Agarwal, P., Arora, R., Ray, S., Singh, A. K., Singh, V. P., Takatsuji, H., et al. (2007). Genome-wide identification of C2H2 zinc-finger gene family in rice and their phylogeny and expression analysis. *Plant Mol. Biol.* 65, 467–485. doi: 10.1007/s11103-007-9199-y

Agarwal, P., Kapoor, S., and Tyagi, A. K. (2011). Transcription factors regulating the progression of monocot and dicot seed development. *Bioessays* 33, 189–202. doi: 10.1002/bies.201000107

Agarwal, P., Parida, S. K., Mahto, A., Das, S., Mathew, I. E., Malik, N., et al. (2014). Expanding frontiers in plant transcriptomics in aid of functional genomics and molecular breeding. *Biotechnol. J.* 9, 1480–1492. doi: 10.1002/biot.201400063

Agarwal, P., Parida, S. K., Raghuvanshi, S., Kapoor, S., Khurana, P., Khurana, J. P., et al. (2016). Rice improvement through genome-based functional analysis and molecular breeding in India. *Rice (N Y).* 9, 1. doi: 10.1186/s12284-015-0073-2

Aida, M., Ishida, T., Fukaki, H., Fujisawa, H., and Tasaka, M. (1997). Genes involved in organ separation in *Arabidopsis*: an analysis of the *cup-shaped cotyledon* mutant. *Plant Cell* 9, 841–857. doi: 10.1105/tpc.9.6.841

Causier, B., Ashworth, M., Guo, W., and Davies, B. (2012). The TOPLESS interactome: a framework for gene repression in *Arabidopsis*. *Plant Physiol.* 158, 423–438. doi: 10.1104/pp.111.186999

Ceribelli, M., Dolfini, D., Merico, D., Gatta, R., Vigano, A. M., Pavesi, G., et al. (2008). The histone-like NF-Y is a bifunctional transcription factor. *Mol. Cell. Biol.* 28, 2047–2058. doi: 10.1128/MCB.01861-07

Chakrabarty, R., Banerjee, R., Chung, S.-M., Farman, M., Citovsky, V., Hogenhout, S. A., et al. (2007). pSITE vectors for stable integration or transient expression of autofluorescent protein fusions in plants: probing *Nicotiana benthamiana*-virus interactions. *Mol. Plant Microbe Interact.* 20, 740–750. doi: 10.1094/MPMI-20-7-0740

Chaturvedi, N. K., Mir, R. A., Band, V., Joshi, S. S., and Guda, C. (2014). Experimental validation of predicted subcellular localizations of human proteins. *BMC Res. Notes* 7:912. doi: 10.1186/1756-0500-7-912

Chen, J., Doyle, C., Qi, X., and Zheng, H. (2012). The endoplasmic reticulum: a social network in plant cells. *J. Integr. Plant Biol.* 54, 840–850. doi: 10.1111/j.1744-7909.2012.01176.x

Christianson, J. A., Dennis, E. S., Llewellyn, D. J., and Wilson, I. W. (2010). ATAF NAC transcription factors: regulators of plant stress signaling. *Plant Signal. Behav.* 5, 428–432. doi: 10.4161/psb.5.4.10847

de Bruin, R. A. M., Kalashnikova, T. I., and Wittenberg, C. (2008). Stb1 collaborates with other regulators to modulate the G(1)-specific transcriptional circuit. *Mol. Cell. Biol.* 28, 6919–6928. doi: 10.1128/MCB.00211-08

De Clercq, I., Vermeirssen, V., Van Aken, O., Vandepoele, K., Murcha, M. W., Law, S. R., et al. (2013). The membrane-bound NAC transcription factor ANAC013 functions in mitochondrial retrograde regulation of the oxidative stress response in *Arabidopsis*. *Plant Cell* 25, 3472–3490. doi: 10.1105/tpc.113.117168

Delessert, C., Kazan, K., Wilson, I. W., Van Der Straeten, D., Manners, J., Dennis, E. S., et al. (2005). The transcription factor ATAF2 represses the expression of pathogenesis-related genes in *Arabidopsis*. *Plant J.* 43, 745–757. doi: 10.1111/j.1365-313X.2005.02488.x

Du, M., Zhai, Q., Deng, L., Li, S., Li, H., Yan, L., et al. (2014). Closely related NAC transcription factors of tomato differentially regulate stomatal closure and reopening during pathogen attack. *Plant Cell* 26, 3167–3184. doi: 10.1105/tpc.114.128272

Dubrovina, A. S., Kiselev, K. V., and Zhuravlev, Y. N. (2013). The role of canonical and noncanonical pre-mRNA splicing in plant stress responses. *Biomed. Res. Int.* 2013, 264314. doi: 10.1155/2013/264314

Ellerstrom, M., Stalberg, K., Ezcurra, I., and Rask, L. (1996). Functional dissection of a napin gene promoter: identification of promoter elements required for embryo and endosperm-specific transcription. *Plant Mol. Biol.* 32, 1019–1027. doi: 10.1007/BF00041385

Fang, Y., You, J., Xie, K., Xie, W., and Xiong, L. (2008). Systematic sequence analysis and identification of tissue-specific or stress-responsive genes of NAC transcription factor family in rice. *Mol. Genet. Genomics* 280, 547–563. doi: 10.1007/s00438-008-0386-6

Fu, F. F., and Xue, H. W. (2010). Coexpression analysis identifies Rice Starch Regulator1, a rice AP2/EREBP family transcription factor, as a novel rice starch biosynthesis regulator. *Plant Physiol.* 154, 927–938. doi: 10.1104/pp.110.159517

Fu, X., Feng, C., Wang, C., Yin, X., Lu, P., Grierson, D., et al. (2014). Involvement of multiple phytoene synthase genes in tissue- and cultivar-specific accumulation of carotenoids in loquat. *J. Exp. Bot.* 65, 4679–4689. doi: 10.1093/jxb/eru257

Galili, G., Sengupta-Gopalan, C., and Ceriotti, A. (1998). The endoplasmic reticulum of plant cells and its role in protein maturation and biogenesis of oil bodies. *Plant Mol. Biol.* 38, 1–29. doi: 10.1023/A:1006011919671

Goncalves, B., Hasson, A., Belcram, K., Cortizo, M., Morin, H., Nikovics, K., et al. (2015). A conserved role for *CUP-SHAPED COTYLEDON* genes during ovule development. *Plant J.* 83, 732–742. doi: 10.1111/tpj.12923

Greve, K., La Cour, T., Jensen, M. K., Poulsen, F. M., and Skriver, K. (2003). Interactions between plant RING-H2 and plant-specific NAC (*NAM/ATAF1/2/CUC2*) proteins: RING-H2 molecular specificity and cellular localization. *Biochem. J.* 371, 97–108. doi: 10.1042/bj20021123

Hao, Y. J., Song, Q. X., Chen, H. W., Zou, H. F., Wei, W., Kang, X. S., et al. (2010). Plant NAC-type transcription factor proteins contain a NARD domain for repression of transcriptional activation. *Planta* 232, 1033–1043. doi: 10.1007/s00425-010-1238-2

Hao, Y. J., Wei, W., Song, Q. X., Chen, H. W., Zhang, Y. Q., Wang, F., et al. (2011). Soybean NAC transcription factors promote abiotic stress tolerance and lateral root formation in transgenic plants. *Plant J.* 68, 302–313. doi: 10.1111/j.1365-313X.2011.04687.x

Higo, K., Ugawa, Y., Iwamoto, M., and Higo, H. (1998). PLACE: a database of plant *cis*-acting regulatory DNA elements. *Nucl. Acids Res.* 26, 358–359. doi: 10.1093/nar/26.1.358

Hofmann, N. R. (2013). Endoplasmic reticulum-localized transcription factors and mitochondrial retrograde regulation. *Plant Cell* 25, 3151. doi: 10.1105/tpc.113.250912

Hu, X. G., Wu, B. H., Liu, D. C., Wei, Y. M., Gao, S. B., and Zheng, Y. L. (2013). Variation and their relationship of *NAM-G1* gene and grain protein content in *Triticum timopheevii* Zhuk. *J. Plant Physiol.* 170, 330–337. doi: 10.1016/j.jplph.2012.10.009

Huang, R., Jiang, L., Zheng, J., Wang, T., Wang, H., Huang, Y., et al. (2013). Genetic bases of rice grain shape: so many genes, so little known. *Trends Plant Sci.* 18, 218–226. doi: 10.1016/j.tplants.2012.11.001

Huang, X., Zhao, Y., Wei, X., Li, C., Wang, A., Zhao, Q., et al. (2012). Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. *Nat. Genet.* 44, 32–39. doi: 10.1038/ng.1018

Ikeda, M., Mitsuda, N., and Ohme-Takagi, M. (2009). *Arabidopsis* WUSCHEL is a bifunctional transcription factor that acts as a repressor in stem cell regulation and as an activator in floral patterning. *Plant Cell* 21, 3493–3505. doi: 10.1105/tpc.109.069997

Iwata, Y., Fedoroff, N. V., and Koizumi, N. (2008). *Arabidopsis* bZIP60 is a proteolysis-activated transcription factor involved in the endoplasmic reticulum stress response. *Plant Cell* 20, 3107–3121. doi: 10.1105/tpc.108.061002

Jia, H., Suzuki, M., and McCarty, D. R. (2014). Regulation of the seed to seedling developmental phase transition by the LAFL and VAL transcription factor networks. *Wiley Interdiscip. Rev. Dev. Biol.* 3, 135–145. doi: 10.1002/wdev.126

Jividen, K., and Li, H. (2014). Chimeric RNAs generated by intergenic splicing in normal and cancer cells. *Genes Chromosomes Cancer* 53, 963–971. doi: 10.1002/gcc.22207

Johansson, O. N., Fantozzi, E., Fahlberg, P., Nilsson, A. K., Buhot, N., Tor, M., et al. (2014). Role of the penetration-resistance genes *PEN1*, *PEN2* and *PEN3* in the hypersensitive response and race-specific resistance in *Arabidopsis thaliana*. *Plant J.* 79, 466–476. doi: 10.1111/tpj.12571

Jopcik, M., Matusikova, I., Moravcikova, J., and Libantova, J. (2014). Expression pattern of *Arabidopsis thaliana* pollen- and embryo-specific promoter in transgenic tobacco plants. *Acta Biol. Cracov. Bot.* 56, 73–79. doi: 10.2478/abcsb-2014-0009

Kagale, S., Links, M. G., and Rozwadowski, K. (2010). Genome-wide analysis of ethylene-responsive element binding factor-associated amphiphilic repression motif-containing transcriptional regulators in *Arabidopsis*. *Plant Physiol.* 152, 1109–1134. doi: 10.1104/pp.109.151704

Kawakatsu, T., and Takaiwa, F. (2010). Cereal seed storage protein synthesis: fundamental processes for recombinant protein production in cereal grains. *Plant Biotechnol. J.* 8, 939–953. doi: 10.1111/j.1467-7652.2010.00559.x

Kawakatsu, T., Yamamoto, M. P., Touno, S. M., Yasuda, H., and Takaiwa, F. (2009). Compensation and interaction between RISBZ1 and RPBF during grain filling in rice. *Plant J.* 59, 908–920. doi: 10.1111/j.1365-313X.2009.03925.x

Kikuchi, K., Ueguchi-Tanaka, M., Yoshida, K. T., Nagato, Y., Matsusoka, M., and Hirano, H. Y. (2000). Molecular analysis of the *NAC* gene family in rice. *Mol. Gen. Genet.* 262, 1047–1051. doi: 10.1007/PL00008647

Kim, H. S., Park, H. C., Kim, K. E., Jung, M. S., Han, H. J., Kim, S. H., et al. (2012). A NAC transcription factor and SNI1 cooperatively suppress basal pathogen resistance in *Arabidopsis thaliana*. *Nucleic Acids Res.* 40, 9182–9192. doi: 10.1093/nar/gks683

Kim, S. G., Kim, S. Y., and Park, C. M. (2007). A membrane-associated NAC transcription factor regulates salt-responsive flowering via *FLOWERING LOCUS T* in *Arabidopsis*. *Planta* 226, 647–654. doi: 10.1007/s00425-007-0513-3

Kou, X., Liu, C., Han, L., Wang, S., and Xue, Z. (2016). NAC transcription factors play an important role in ethylene biosynthesis, reception and signaling of tomato fruit ripening. *Mol. Genet. Genomics* 291, 1205–1217. doi: 10.1007/s00438-016-1177-0

Kujur, A., Bajaj, D., Upadhyaya, H. D., Das, S., Ranjan, R., Shree, T., et al. (2015a). A genome-wide SNP scan accelerates trait-regulatory genomic loci identification in chickpea. *Sci. Rep.* 5, 11166. doi: 10.1038/srep11166

Kujur, A., Bajaj, D., Upadhyaya, H. D., Das, S., Ranjan, R., Shree, T., et al. (2015b). Employing genome-wide SNP discovery and genotyping strategy to extrapolate the natural allelic diversity and domestication patterns in chickpea. *Front. Plant Sci.* 6:162. doi: 10.3389/fpls.2015.00162

Kumar, V., Singh, A., Mithra, S. V., Krishnamurthy, S. L., Parida, S. K., Jain, S., et al. (2015). Genome-wide association mapping of salinity tolerance in rice (*Oryza sativa*). *DNA Res.* 22, 133–145. doi: 10.1093/dnares/dsu046

Kunieda, T., Mitsuda, N., Ohme-Takagi, M., Takeda, S., Aida, M., Tasaka, M., et al. (2008). NAC family proteins NARS1/NAC2 and NARS2/NAM in the outer integument regulate embryogenesis in *Arabidopsis*. *Plant Cell* 20, 2631–2642. doi: 10.1105/tpc.108.060160

Lee, S., Lee, H. J., Huh, S. U., Paek, K. H., Ha, J. H., and Park, C. M. (2014). The *Arabidopsis* NAC transcription factor NTL4 participates in a positive feedback loop that induces programmed cell death under heat stress conditions. *Plant Sci.* 227, 76–83. doi: 10.1016/j.plantsci.2014.07.003

Li, Q., Lin, Y. C., Sun, Y. H., Song, J., Chen, H., Zhang, X. H., et al. (2012). Splice variant of the SND1 transcription factor is a dominant negative of SND1 members and their regulation in *Populus trichocarpa*. *Proc. Natl. Acad. Sci. U.S.A.* 109, 14699–14704. doi: 10.1073/pnas.1212977109

Li, S., Gao, F., Xie, K., Zeng, X., Cao, Y., Zeng, J., et al. (2016). The OsmiR396c-OsGRF4-OsGIF1 regulatory module determines grain size and yield in rice. *Plant Biotechnol. J.* 14, 2134–2146. doi: 10.1111/pbi.12569

Liang, C., Wang, Y., Zhu, Y., Tang, J., Hu, B., Liu, L., et al. (2014). OsNAP connects abscisic acid and leaf senescence by fine-tuning abscisic acid biosynthesis and directly targeting senescence-associated genes in rice. *Proc. Natl. Acad. Sci. U.S.A.* 111, 10013–10018. doi: 10.1073/pnas.1321568111

Lorenz, D. R., Meyer, L. F., Grady, P. J. R., Meyer, M. M., and Cam, H. P. (2014). Heterochromatin assembly and transcriptome repression by Set1 in coordination with a class II histone deacetylase. *Elife* 3:e04506. doi: 10.7554/eLife.04506

Lu, P. L., Chen, N. Z., An, R., Su, Z., Qi, B. S., Ren, F., et al. (2007). A novel drought-inducible gene, *ATAF1*, encodes a NAC family protein that negatively regulates the expression of stress-responsive genes in *Arabidopsis*. *Plant Mol. Biol.* 63, 289–305. doi: 10.1007/s11103-006-9089-8

Mendes, G. C., Reis, P. A., Calil, I. P., Carvalho, H. H., Aragao, F. J., and Fontes, E. P. (2013). GmNAC30 and GmNAC81 integrate the endoplasmic reticulum stress- and osmotic stress-induced cell death responses through a vacuolar processing enzyme. *Proc. Natl. Acad. Sci. U.S.A.* 110, 19627–19632. doi: 10.1073/pnas.1311729110

Nayar, S., Kapoor, M., and Kapoor, S. (2014). Post-translational regulation of rice MADS29 function: homodimerization or binary interactions with other seed-expressed MADS proteins modulate its translocation into the nucleus. *J. Exp. Bot.* 65, 5339–5350. doi: 10.1093/jxb/eru296

Nayar, S., Sharma, R., Tyagi, A. K., and Kapoor, S. (2013). Functional delineation of rice *MADS29* reveals its role in embryo and endosperm development by affecting hormone homeostasis. *J. Exp. Bot.* 64, 4239–4253. doi: 10.1093/jxb/ert231

Ng, S., Ivanova, A., Duncan, O., Law, S. R., Van Aken, O., De Clercq, I., et al. (2013). A membrane-bound NAC transcription factor, ANAC017, mediates mitochondrial retrograde signaling in *Arabidopsis*. *Plant Cell* 25, 3450–3471. doi: 10.1105/tpc.113.113985

Nuruzzaman, M., Manimekalai, R., Sharoni, A. M., Satoh, K., Kondoh, H., Ooka, H., et al. (2010). Genome-wide analysis of NAC transcription factor family in rice. *Gene* 465, 30–44. doi: 10.1016/j.gene.2010.06.008

Oh, E., Zhu, J. Y., Ryu, H., Hwang, I., and Wang, Z. Y. (2014). TOPLESS mediates brassinosteroid-induced transcriptional repression through interaction with BZR1. *Nat. Commun.* 5:4140. doi: 10.1038/ncomms5140

Ohashi-Ito, K., Saegusa, M., Iwamoto, K., Oda, Y., Katayama, H., Kojima, M., et al. (2014). A bHLH complex activates vascular cell division via cytokinin action in root apical meristem. *Curr. Biol.* 24, 2053–2058. doi: 10.1016/j.cub.2014.07.050

Ohta, M., Matsui, K., Hiratsu, K., Shinshi, H., and Ohme-Takagi, M. (2001). Repression domains of class II ERF transcriptional repressors share an essential motif for active repression. *Plant Cell* 13, 1959–1968. doi: 10.1105/tpc.13.8.1959

Olsen, A. N., Ernst, H. A., Leggio, L. L., and Skriver, K. (2005). NAC transcription factors: structurally distinct, functionally diverse. *Trends Plant Sci.* 10, 79–87. doi: 10.1016/j.tplants.2004.12.010

Ooka, H., Satoh, K., Doi, K., Nagata, T., Otomo, Y., Murakami, K., et al. (2003). Comprehensive analysis of NAC family genes in *Oryza sativa* and *Arabidopsis thaliana*. *DNA Res.* 10, 239–247. doi: 10.1093/dnares/10.6.239

Parida, S. K., Dalal, V., Singh, A. K., Singh, N. K., and Mohapatra, T. (2009). Genic non-coding microsatellites in the rice genome: characterization, marker design and use in assessing genetic and evolutionary relationships among domesticated groups. *BMC Genomics* 10:140. doi: 10.1186/1471-2164-10-140

Pereira-Santana, A., Alcaraz, L. D., Castano, E., Sanchez-Calderon, L., Sanchez-Teyer, F., and Rodriguez-Zapata, L. (2015). Comparative genomics of NAC transcriptional factors in angiosperms: implications for the adaptation and diversification of flowering plants. *PLoS ONE* 10:e0141866. doi: 10.1371/journal.pone.0141866

Plessis, A., Ravel, C., Bordes, J., Balfourier, F., and Martre, P. (2013). Association study of wheat grain protein composition reveals that gliadin and glutenin composition are trans-regulated by different chromosome regions. *J. Exp. Bot.* 64, 3627–3644. doi: 10.1093/jxb/ert188

Qiao, Z., Qi, W., Wang, Q., Feng, Y. N., Yang, Q., Zhang, N., et al. (2016). ZmMADS47 regulates zein gene transcription through interaction with Opaque2. *PLoS Genet.* 12:e1005991. doi: 10.1371/journal.pgen.1005991

Ravel, C., Fiquet, S., Boudet, J., Dardevet, M., Vincent, J., Merlino, M., et al. (2014). Conserved *cis*-regulatory modules in promoters of genes encoding wheat high-molecular-weight glutenin subunits. *Front. Plant Sci.* 5:621. doi: 10.3389/fpls.2014.00621

Reddy, A. S., Marquez, Y., Kalyna, M., and Barta, A. (2013). Complexity of the alternative splicing landscape in plants. *Plant Cell* 25, 3657–3683. doi: 10.1105/tpc.113.117523

Reyes, J. C. (2006). Chromatin modifiers that control plant development. *Curr. Opin. Plant Biol.* 9, 21–27. doi: 10.1016/j.pbi.2005.11.010

Ricachenevsky, F. K., Menguer, P. K., and Sperotto, R. A. (2013). kNACking on heaven's door: how important are NAC transcription factors for leaf senescence and Fe/Zn remobilization to seeds? *Front. Plant Sci.* 4:226. doi: 10.3389/fpls.2013.00226

Saxena, M. S., Bajaj, D., Das, S., Kujur, A., Kumar, V., Singh, M., et al. (2014). An integrated genomic approach for rapid delineation of candidate genes regulating agro-morphological traits in chickpea. *DNA Res.* 21, 695–710. doi: 10.1093/dnares/dsu031

Seo, P. J. (2014). Recent advances in plant membrane-bound transcription factor research: emphasis on intracellular movement. *J. Integr. Plant Biol.* 56, 334–342. doi: 10.1111/jipb.12139

Seo, P. J., Hong, S. Y., Kim, S. G., and Park, C. M. (2011). Competitive inhibition of transcription factors by small interfering peptides. *Trends Plant Sci.* 16, 541–549. doi: 10.1016/j.tplants.2011.06.001

Shan, W., Kuang, J. F., Chen, L., Xie, H., Peng, H. H., Xiao, Y. Y., et al. (2012). Molecular characterization of banana NAC transcription factors and their

interactions with ethylene signalling component EIL during fruit ripening. *J. Exp. Bot.* 63, 5171–5187. doi: 10.1093/jxb/ers178

Sharma, N., Russell, S. D., Bhalla, P. L., and Singh, M. B. (2011). Putative *cis*-regulatory elements in genes highly expressed in rice sperm cells. *BMC Res. Notes* 4:319. doi: 10.1186/1756-0500-4-319

Sharma, R., Agarwal, P., Ray, S., Deveshwar, P., Sharma, P., Sharma, N., et al. (2012). Expression dynamics of metabolic and regulatory components across stages of panicle and seed development in *indica* rice. *Funct. Integr. Genomics* 12, 229–248. doi: 10.1007/s10142-012-0274-3

Shih, C. F., Hsu, W. H., Peng, Y. J., and Yang, C. H. (2014). The *NAC*-like gene *ANTHER INDEHISCENCE FACTOR* acts as a repressor that controls anther dehiscence by regulating genes in the jasmonate biosynthesis pathway in *Arabidopsis*. *J. Exp. Bot.* 65, 621–639. doi: 10.1093/jxb/ert412

Souer, E., van Houwelingen, A., Kloos, D., Mol, J., and Koes, R. (1996). The *no apical meristem* gene of Petunia is required for pattern formation in embryos and flowers and is expressed at meristem and primordia boundaries. *Cell* 85, 159–170. doi: 10.1016/S0092-8674(00)81093-4

Sun, X., Ling, S., Lu, Z., Ouyang, Y. D., Liu, S., and Yao, J. (2014). *OsNF-YB1*, a rice endosperm-specific gene, is essential for cell proliferation in endosperm development. *Gene* 551, 214–221. doi: 10.1016/j.gene.2014.08.059

Takahata, S., Yu, Y., and Stillman, D. J. (2009). The E2F functional analogue SBF recruits the Rpd3(L) HDAC, via Whi5 and Stb1, and the FACT chromatin reorganizer, to yeast G1 cyclin promoters. *EMBO J.* 28, 3378–3389. doi: 10.1038/emboj.2009.270

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., and Higgins, D. G. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 25, 4876–4882. doi: 10.1093/nar/25.24.4876

Tzfira, T., Tian, G. W., Lacroix, B., Vyas, S., Li, J., Leitner-Dagan, Y., et al. (2005). pSAT vectors: a modular series of plasmids for autofluorescent protein tagging and expression of multiple genes in plants. *Plant Mol. Biol.* 57, 503–516. doi: 10.1007/s11103-005-0340-5

Wagner, D. (2003). Chromatin regulation of plant development. *Curr. Opin. Plant Biol.* 6, 20–28. doi: 10.1016/S1369526602000079

Wang, S., Wu, K., Yuan, Q., Liu, X., Liu, Z., Lin, X., et al. (2012). Control of grain size, shape and quality by *OsSPL16* in rice. *Nat. Genet.* 44, 950–954. doi: 10.1038/ng.2327

Withers, J., Yao, J., Mecey, C., Howe, G. A., Melotto, M., and He, S. Y. (2012). Transcription factor-dependent nuclear localization of a transcriptional repressor in jasmonate hormone signaling. *Proc. Natl. Acad. Sci. U.S.A.* 109, 20148–20153. doi: 10.1073/pnas.1210054109

Xu, Z. Y., Kim, S. Y., Hyeon do, Y., Kim, D. H., Dong, T., Park, Y., et al. (2013). The *Arabidopsis* NAC transcription factor ANAC096 cooperates with bZIP-type transcription factors in dehydration and osmotic stress responses. *Plant Cell* 25, 4708–4724. doi: 10.1105/tpc.113.119099

Yamaguchi, M., Ohtani, M., Mitsuda, N., Kubo, M., Ohme-Takagi, M., Fukuda, H., et al. (2010). VND-INTERACTING2, a NAC domain transcription factor, negatively regulates xylem vessel formation in Arabidopsis. *Plant Cell* 22, 1249–1263. doi: 10.1105/tpc.108.064048

Yamamoto, M. P., Onodera, Y., Touno, S. M., and Takaiwa, F. (2006). Synergism between RPBF Dof and RISBZ1 bZIP activators in the regulation of rice seed expression genes. *Plant Physiol.* 141, 1694–1707. doi: 10.1104/pp.106.082826

Yang, H., Krebs, M., Stierhof, Y. D., and Ludewig, U. (2014). Characterization of the putative amino acid transporter genes *AtCAT2*, *3* &*4*: the tonoplast localized *AtCAT2* regulates soluble leaf amino acids. *J. Plant Physiol.* 171, 594–601. doi: 10.1016/j.jplph.2013.11.012

Yang, S. D., Seo, P. J., Yoon, H. K., and Park, C. M. (2011). The *Arabidopsis* NAC transcription factor VNI2 integrates abscisic acid signals into leaf senescence via the *COR/RD* genes. *Plant Cell* 23, 2155–2168. doi: 10.1105/tpc.111.084913

Yin, L. L., and Xue, H. W. (2012). The *MADS29* transcription factor regulates the degradation of the nucellus and the nucellar projection during rice seed development. *Plant Cell* 24, 1049–1065. doi: 10.1105/tpc.111.094854

Yoshida, K., Sakamoto, S., Kawai, T., Kobayashi, Y., Sato, K., Ichinose, Y., et al. (2013). Engineering the *Oryza sativa* cell wall with rice NAC transcription factors regulating secondary wall formation. *Front. Plant Sci.* 4:383. doi: 10.3389/fpls.2013.00383

Zhang, C. Q., Xu, Y., Lu, Y., Yu, H. X., Gu, M. H., and Liu, Q. Q. (2011). The WRKY transcription factor OsWRKY78 regulates stem elongation and seed development in rice. *Planta* 234, 541–554. doi: 10.1007/s00425-011-1423-y

Zhang, G., Guo, G., Hu, X., Zhang, Y., Li, Q., Li, R., et al. (2010). Deep RNA sequencing at single base-pair resolution reveals high complexity of the rice transcriptome. *Genome Res.* 20, 646–654. doi: 10.1101/gr.100677.109

Zhang, J., Nallamilli, B. R., Mujahid, H., and Peng, Z. (2010). *OsMADS6* plays an essential role in endosperm nutrient accumulation and is subject to epigenetic regulation in rice (*Oryza sativa*). *Plant J.* 64, 604–617. doi: 10.1111/j.1365-313X.2010.04354.x

Zhang, L., Gu, L., Ringler, P., Smith, S., Rushton, P. J., and Shen, Q. J. (2015). Three WRKY transcription factors additively repress abscisic acid and gibberellin signaling in aleurone cells. *Plant Sci.* 236, 214–222. doi: 10.1016/j.plantsci.2015.04.014

Zhang, X., Dou, L., Pang, C., Song, M., Wei, H., Fan, S., et al. (2015). Genomic organization, differential expression, and functional analysis of the *SPL* gene family in *Gossypium hirsutum*. *Mol. Genet. Genomics* 290, 115–126. doi: 10.1007/s00438-014-0901-x

Zhao, K., Tung, C. W., Eizenga, G. C., Wright, M. H., Ali, M. L., Price, A. H., et al. (2011). Genome-wide association mapping reveals a rich genetic architecture of complex traits in *Oryza sativa*. *Nat. Commun.* 2, 467. doi: 10.1038/ncomms1467

Zhu, Y., Cai, X. L., Wang, Z. Y., and Hong, M. M. (2003). An interaction between a MYC protein and an EREBP protein is involved in transcriptional regulation of the rice *Wx* gene. *J. Biol. Chem.* 278, 47803–47811. doi: 10.1074/jbc.M302806200

Zhu, Z., Xu, F., Zhang, Y., Cheng, Y. T., Wiermer, M., and Li, X. (2010). *Arabidopsis* resistance protein SNC1 activates immune responses through association with a transcriptional corepressor. *Proc. Natl. Acad. Sci. U.S.A.* 107, 13960–13965. doi: 10.1073/pnas.1002828107

Zuo, J., and Li, J. (2014). Molecular genetic dissection of quantitative trait loci regulating rice grain size. *Annu. Rev. Genet.* 48, 99–118. doi: 10.1146/annurev-genet-120213-092138

Conflict of Interest Statement: 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.