



# Annotation, phylogeny and expression analysis of the nuclear factor Y gene families in common bean (*Phaseolus vulgaris*)

Carolina Rípodas, Mélisse Castaingts, Joaquín Clúa, Flavio Blanco and María Eugenia Zanetti\*

Instituto de Biotecnología y Biología Molecular, Universidad Nacional de La Plata, CCT-CONICET, La Plata, Argentina

## Edited by:

Jun Yu, Beijing Institute of Genomics, China

## Reviewed by:

Stewart Gillmor, CINVESTAV-IPN, Mexico  
Swarup Kumar Parida, National Institute of Plant Genome Research, India

## \*Correspondence:

María Eugenia Zanetti, Instituto de Biotecnología y Biología Molecular, Universidad Nacional de La Plata, CCT-CONICET, Calle 115 y 47, C. P. 1900, La Plata, Argentina  
e-mail: ezanetti@biol.unlp.edu.ar

In the past decade, plant nuclear factor Y (NF-Y) genes have gained major interest due to their roles in many biological processes in plant development or adaptation to environmental conditions, particularly in the root nodule symbiosis established between legume plants and nitrogen fixing bacteria. NF-Ys are heterotrimeric transcriptional complexes composed of three subunits, NF-YA, NF-YB, and NF-YC, which bind with high affinity and specificity to the CCAAT box, a *cis* element present in many eukaryotic promoters. In plants, NF-Y subunits consist of gene families with about 10 members each. In this study, we have identified and characterized the NF-Y gene families of common bean (*Phaseolus vulgaris*), a grain legume of worldwide economical importance and the main source of dietary protein of developing countries. Expression analysis showed that some members of each family are up-regulated at early or late stages of the nitrogen fixing symbiotic interaction with its partner *Rhizobium etli*. We also showed that some genes are differentially accumulated in response to inoculation with high or less efficient *R. etli* strains, constituting excellent candidates to participate in the strain-specific response during symbiosis. Genes of the NF-YA family exhibit a highly structured intron-exon organization. Moreover, this family is characterized by the presence of upstream ORFs when introns in the 5' UTR are retained and miRNA target sites in their 3' UTR, suggesting that these genes might be subjected to a complex post-transcriptional regulation. Multiple protein alignments indicated the presence of highly conserved domains in each of the NF-Y families, presumably involved in subunit interactions and DNA binding. The analysis presented here constitutes a starting point to understand the regulation and biological function of individual members of the NF-Y families in different developmental processes in this grain legume.

**Keywords:** CCAAT box, gene regulation, legumes, nodulation, transcription factors, symbiosis

## INTRODUCTION

Nuclear factor Ys (NF-Ys) are heterotrimeric transcription factors evolutionary conserved in yeast, mammals and plants that have emerged as important regulators of gene expression. The complex is composed of three subunits: NF-YA, NF-YB, and NF-YC (also known as HAP2, HAP3, and HAP5, respectively). The current model for assembling of the NF-Y complex in mammals involves the dimerization of the NF-YB and NF-YC subunits through the interaction of their histone-fold domains (HFDs) in the cytoplasm, the translocation of the heterodimer to the nucleus and the association with the NF-YA subunit to form the mature and functional complex (Kahle et al., 2005). This trimer recognizes and binds CCAAT boxes, *cis* elements present in many eukaryotic promoters, activating or repressing transcription of the downstream genes (Ceribelli et al., 2008).

In yeast and metazoan, each NF-Y subunit is encoded by one or two genes, whereas plant families have largely expanded (Petroni et al., 2012; Laloum et al., 2013). The *Arabidopsis thaliana* genome encodes 10 NF-YA, 10 NF-YB, and 10 NF-YC

subunits, whereas in rice (*Oryza sativa*) these families have 10, 11, and 7 genes, respectively (Siefers et al., 2009; Petroni et al., 2012; Laloum et al., 2013). The analysis of these families in leguminous plants identified 8, 19, and 11 members for the NF-YA, -YB, and -YC families, respectively, in the model *Medicago truncatula*, whereas in soybean (*Glycine max*) these families have expanded to 20, 39, and 27 members for each subunit, respectively (Laloum et al., 2013). However, the annotation of *M. truncatula* and soybean included two groups of more divergent HFD containing proteins (referred to as NC2 $\beta$  and NC2 $\alpha$ ) within the NF-YB and -YC families, respectively.

The expansion of the NF-Y gene family in the plant lineage could result theoretically in more than a thousand alternative unique heterotrimeric combinations (Siefers et al., 2009). This has led to the speculation that this combinatorial association between NF-Y subunits may provide a versatile and flexible regulatory system that integrates developmental and environmental inputs to fine-tune transcriptional outputs. Two systematic studies analyzed the interactions between Arabidopsis NF-Y subunits

using the yeast two hybrid system (Calvenzani et al., 2012; Hackenberg et al., 2012), leading to the conclusion that most NF-YB and NF-YC subunits can interact promiscuously with each other, and that heterodimer formation through the HFD seems to be a prerequisite for translocation into the nucleus, association with the NF-YA subunit and the subsequent binding to DNA.

The function of the NF-Y complex has been extensively characterized in mammals, where it plays a central role in the control of cell proliferation and early stages of development. In plants, forward and reverse genetic approaches have led to the identification of several genes encoding NF-Y subunits that are required for the correct development of plant programs such as embryogenesis (Lotan et al., 1998; Kwong et al., 2003), seed germination (Warpeha et al., 2007; Kumimoto et al., 2013), chloroplast biogenesis (Miyoshi et al., 2003), flowering (Ben-Naim et al., 2006; Kumimoto et al., 2010; Tiwari et al., 2010) and root elongation (Ballif et al., 2011). Genes encoding NF-Y subunits have been implicated also in the organogenesis and development of symbiotic root nodules in leguminous plants, where they act as components of a hierarchical transcriptional activation cascade in the nodulation signaling pathway (Combiere et al., 2006; Zanetti et al., 2010; Soyano et al., 2013; Laloum et al., 2014; Laporte et al., 2014).

Expression analyses of individual members of each NF-Y family conducted in *Arabidopsis*, rice, wheat (*Triticum aestivum*), *Brachypodium distachion* and canola (*Brassica napus* L.) have revealed organ-specific expression patterns for several members of each subunit family (Gusmaroli et al., 2002; Stephenson et al., 2007; Thirumurugan et al., 2008; Siefers et al., 2009; Cao et al., 2011; Liang et al., 2014). In legumes, expression analyses have been conducted for individual NF-Y members at different stages of nodule formation or during arbuscular mycorrhization (Combiere et al., 2006, 2008; Zanetti et al., 2010; Soyano et al., 2013; Laloum et al., 2014; Laporte et al., 2014; Hogeekamp et al., 2011; Schaarschmidt et al., 2013). However, a systematic analysis of gene expression for each member of individual subunits in a single legume has not been reported.

Common bean is the most important grain legume for human consumption, representing the major source of proteins and essential nutrients in the diets of developing countries. This species is original of America, where two main centers of genetic diversification have been proposed: the Mesoamerican and the South Andean centers (Gepts, 1998). As other legumes, common bean establishes a nitrogen-fixing symbiotic interaction with soil bacteria collectively known as rhizobia. This interaction results in the formation of a new organ in the root, the nodule, which allocates bacteria and provides the appropriate environment for nitrogen fixation. Among the several *Rhizobium* species recognized as microsymbionts of common bean, *R. etli* is the predominant species associated to nodules of wild and domesticated common bean plants in America (Aguilar et al., 1998). Moreover, it has been shown that Mesoamerican accessions of common bean show a preferential and more efficient nodulation with *R. etli* strains that are predominant in Mesoamerican soils respect to those that are more abundant in Andean soils (Aguilar et al., 2004; Peltzer Meschini et al., 2008; Mazziotta et al., 2013). A member of the NF-YC family, named PvNF-YC1, was shown to play a role in nodule organogenesis and rhizobial infection, as well as in

the preferential and more efficient association observed between Mesoamerican accession and sympatric rhizobial strains (Zanetti et al., 2010). Moreover, PvNF-YC1 physically interacts (in yeast and *in planta*) with another transcription factor belonging to the GRAS family, which also plays a role in nodule development and in the progression of the infection process in the common bean-*R. etli* interaction (Battaglia et al., 2014; Ripodas et al., 2014).

Considering the important role played by NF-Y genes in root nodule symbiosis, we took advantage of the recently released genome of common bean v1.0 (Schmutz et al., 2014) to annotate and analyze the gene families of each NF-Y subunit in this grain legume. The analysis identified 9, 14, and 7 members of the NF-YA, NF-YB, and NF-YC gene families, respectively. Genes of the NF-YA family are highly structured, showing a complex array of exons and introns in the coding and untranslated regions, a feature that was not observed in the NF-YB and NF-YC families. The quantitative analysis of the expression allowed us to identify transcripts that accumulate at high levels in nodules or in roots at early time points after inoculation with rhizobia, whereas some others exhibit specific expression in aerial tissues. In addition, the use of *R. etli* strains with different degrees of nodulation efficiency led us to identify some NF-Y members whose mRNA accumulates at higher levels in either the more or the less efficient symbiotic interactions.

## MATERIALS AND METHODS

### SEQUENCE ANALYSIS, CHROMOSOME MAPPING, ALIGNMENTS AND PHYLOGENETIC TREES

Individual subunits were identified by TBLASTN searches against the common bean genome V1.0 database (<http://phytozome.jgi.doe.gov>) using the full length amino acid sequences of the 30 members of Arabidopsis NF-Y gene family. The physical position of PvNF-Y members was mapped on common bean chromosomes using Map-view function available at Plant Genome Duplication Database (PGDD) as implemented by Lee et al. (2013). Alignments and phylogenies were created from full-length protein sequences. *Arabidopsis thaliana*, *Medicago truncatula* and *Glycine max* sequences were obtained from the GenBank, TAIR or Phytozome v1.0 websites. Multiple sequence alignments were performed with ClustalW (Thompson et al., 1994) and shaded with BOXSHADE 3.21 software ([http://www.ch.embnet.org/software/BOX\\_form.html](http://www.ch.embnet.org/software/BOX_form.html)). Protein alignments were imported into MEGA5, where rooted phylogenetic trees were created (Tamura et al., 2011). The locus search function of PGDD was used to identify syntenic regions (colinear blocks) containing PvNF-Y loci using default parameters (Lee et al., 2013). miR169 target sites in NF-YA family members were identified with Target-align (Xie and Zhang, 2010) using default parameters. The presence of uORF in *P. vulgaris* NF-YA genes was analyzed using Vector NTI (Invitrogen, [www.lifetechnologies.com](http://www.lifetechnologies.com)), allowing a minimum length of 60 nt.

### PLANT GROWTH AND RHIZOBIAL INOCULATION

Seeds of common bean (NAG12 accession) were surface sterilized and germinated on wet paper for 2–3 days at 26°C in the dark as previously described (Peltzer Meschini et al., 2008). Germinated seedlings were transferred to acrylic boxes containing slanted

agar-Fahraeus media free of nitrogen (Fahraeus, 1957). Seedlings were grown in a MLR-350HT growth chamber (Sanyo Electric Co., [www.panasonic.net/sanyo](http://www.panasonic.net/sanyo)) maintained at 26°C with a 16/8 h day/night cycle and 80% humidity. Seven days after transplantation to acrylic boxes, leaf, stem and root tissue was collected. For nodulation, roots of 7 day-old seedlings were inoculated with 5 ml of a culture of *R. etli* strain SC15 or 55N1 grown in liquid YEM media until the OD<sub>600</sub> reached 0.8. Nodules were collected at 7 and 14 days post inoculation (dpi). In all cases, tissue from 3 plants was pooled for each biological replicate, frozen in liquid N<sub>2</sub> and stored at -80°C.

### RNA EXTRACTION AND RT-qPCR

Total RNA extraction was performed with Trizol following manufacturer's instructions (Invitrogen). RNA concentration was determined by measuring absorbance at 260 nm in a Nanodrop ND-1000 (Nanodrop Technologies Inc., [www.nanodrop.com](http://www.nanodrop.com)) and quality was evaluated using the Agilent RNA 6000 nano kit and the Agilent Technologies 2100 Bioanalyzer ([www.agilent.com](http://www.agilent.com)). One µg of total RNA was treated with RNase free-DNAse I and used in a reverse transcription reaction with 0.5 µg of oligo dT<sub>15</sub> and 200 units of M-MLV-RT following instructions provided by the manufacturer (Promega, [www.promega.com](http://www.promega.com)). Individual cDNA samples were used in quantitative PCR (qPCR) reactions as previously reported (Zanetti et al., 2010) with gene specific primers (see Table S1). Amplification of common bean elongation factor 1α (*EF1α*) was used to normalize the amount of template cDNA as reported (Peltzer Meschini et al., 2008). Statistical significance of differences between samples was determined by unpaired two tailed *t*-tests using at least two biological replicates.

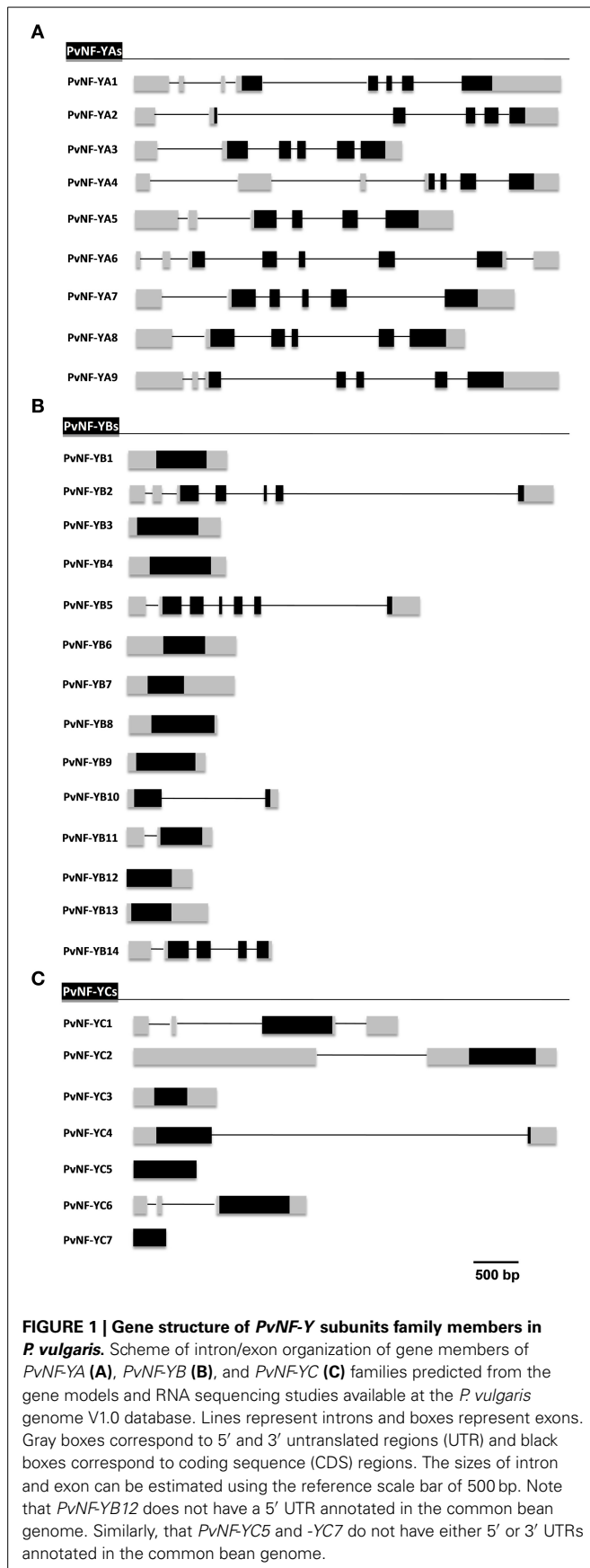
## RESULTS

### IDENTIFICATION, ANNOTATION, AND STRUCTURE OF *P. VULGARIS* NF-Y GENES

We used the full length amino acid sequences of the 30 members of Arabidopsis NF-Y gene family to sequentially BLAST search the *P. vulgaris* genome V1.0 database (publicly available at <http://phytozome.jgi.doe.gov>) using an *E*-value cutoff of 10<sup>-5</sup>. We identified 30 members of these gene families distributed as 9 NF-YAs, 14 NF-YBs, and 7 NF-YCs (Table 1). Based on the recent recommendations to assign names for plant NF-Y families (Laloum et al., 2013), we have designated each member with a two letter code corresponding to the species initials (Pv for *P. vulgaris*) followed by NF-Y, the letter for each subunit (A, B, or C) and finally by a number. Because there is no previous characterization of any member of the NF-YB family in common bean, there is not a previous scheme to number members of this family. In the cases of NF-YA and NF-YC families, PvNF-YA1 and PvNF-YC1, -YC2, and -YC3 were previously characterized (Zanetti et al., 2010; Battaglia et al., 2014); therefore, we continued numbering after these genes. Table 1 presents assigned names together with the genomic gene ID and its best homologs/putative orthologs in *Arabidopsis*. Table S2 presents, in addition, the best homologs/putative orthologs in two other legume species (*M. truncatula* and soybean). The gene structure of each NF-Y gene predicted from the available gene models are presented in Figure 1. The physical position of the 30 PvNF-Y members was mapped on the chromosomes available

Table 1 | Annotation of common bean NF-Y families.

PvNF-YA family				PvNF-YB family				PvNF-YC family			
Name	Locus name	Best At homolog	Name	Locus name	Best At homolog	Name	Locus name	Best At homolog	Name	Locus name	Best At homolog
PvNF-YA1	Phvul.001G196800	NF-YA10 (AT5G06510)	PvNF-YB1	Phvul.008G278900	NF-YB3 (AT4G14540)	PvNF-YC1	Phvul.006G093200	NF-YC9 (AT1G08970)			
PvNF-YA2	Phvul.002G246600	NF-YA7 (AT1G30500)	PvNF-YB2	Phvul.007G134000	NF-YB8 (AT2G37060)	PvNF-YC2	Phvul.005G050000	NF-YC2 (AT1G56170)			
PvNF-YA3	Phvul.005G156100	NF-YA8 (AT1G17590)	PvNF-YB3	Phvul.006G139400	NF-YB3 (AT4G14540)	PvNF-YC3	Phvul.010G102300	NF-YC4 (AT5G63470)			
PvNF-YA4	Phvul.003G133100	NF-YA3 (AT1G72830)	PvNF-YB4	Phvul.002G264300	NF-YB3 (AT4G14540)	PvNF-YC4	Phvul.009G166000	NF-YC1 (AT3G48590)			
PvNF-YA5	Phvul.008G283100	NF-YA9 (AT3G20910)	PvNF-YB5	Phvul.007G163100	NF-YB10 (AT3G53340)	PvNF-YC5	Phvul.006G152400	NF-YC2 (AT1G56170)			
PvNF-YA6	Phvul.011G211300	NF-YA9 (AT3G20910)	PvNF-YB6	Phvul.007G165100	NF-YB5 (AT2G47810)	PvNF-YC6	Phvul.007G181400	NF-YC9 (AT1G08970)			
PvNF-YA7	Phvul.006G062200	NF-YA1 (AT5G12840)	PvNF-YB7	Phvul.003G086800	NF-YB3 (AT4G14540)	PvNF-YC7	Phvul.010G082500	NF-YC3 (AT1G54830)			
PvNF-YA8	Phvul.010G133300	NF-YA3 (AT1G72830)	PvNF-YB8	Phvul.002G273800	NF-YB7 (AT2G13570)						
PvNF-YA9	Phvul.007G267100	NF-YA10 (AT5G06510)	PvNF-YB9	Phvul.009G155500	NF-YB7 (AT2G13570)						
			PvNF-YB10	Phvul.009G246400	NF-YB5 (AT2G47810)						
			PvNF-YB11	Phvul.003G103500	NF-YB5 (AT2G47810)						
			PvNF-YB12	Phvul.007G234800	NF-YB6 (AT5G47670)						
			PvNF-YB13	Phvul.007G195900	NF-YB4 (AT1G09030)						
			PvNF-YB14	Phvul.001G174600	NF-YB10 (AT3G53340)						



from the common bean genome. *PvNF-Y* genes were distributed on 10 out of the 11 common bean chromosomes (Figure S1). In particular, the 9 *PvNF-YA* genes were uniformly distributed on several chromosomes, whereas members of the *PvNF-YB* family showed a higher degree of association, with five members located on chromosome Pv Chr07. Members of the *PvNF-YC* family mapped only in five chromosomes, with one or two genes per chromosome.

### Structure of *PvNF-YA* genes

The gene structure analysis revealed that members of the *NF-YA* family have a complex intron/exon organization with 5–7 introns per gene, distributed mainly along the 5' UTRs and the coding regions, being *PvNF-YA6* the only member with an intron in the 3' UTR region (Figure 1A). Notably, all *PvNF-YA* members contain at least 1 intron in the 5' UTR region. A previous report has shown that the *NF-YA1* gene of *M. truncatula* (formerly *MtHAP2-1*) is subjected to a post-transcriptional regulatory mechanism that involves a short peptide that destabilizes the mRNA. This peptide is encoded by an upstream open reading frame (uORF) produced by an alternative spliced variant of *NF-YA1* that retains a long intron in the 5' UTR region (Combiér et al., 2008). The presence of introns in the 5' UTR regions of all common bean *NF-YA* members prompted us to investigate whether uORFs are encoded within the sequence upstream of the main ORF when introns are retained. We identified uORFs ( $\geq 60$  nt) in 5 out of 9 *NF-YA* genes. *PvNF-YA1* and *-YA2* contain 2 uORFs, *PvNF-YA4* contains 4 uORFs, and *PvNF-YA8* and *-YA9* contain only 1 uORF when introns present at the 5' UTR are retained (Figure S2). *NF-YA* transcripts are also regulated at postranscriptional level by the action of microRNA169 (miR169), which binds to the 3' UTR region of the *NF-YA* transcripts and promotes their cleavage by the slicing protein Argonaute 1 (Jones-Rhoades and Bartel, 2004; Combiér et al., 2006; Li et al., 2008). This regulatory mechanism is conserved among angiosperms, including the *Leguminosae* family (Cuperus et al., 2011). A recent report identified 7 isoforms of the miR169 family in common bean, being miR169b, miR169c and miR169d the most abundant 21 nt isoforms (Pelaez et al., 2012). Therefore, we investigated whether the 9 annotated members of the *PvNF-YA* family (Table 1) contained putative target sites for miR169b/c/d. For this analysis we used the Target-align algorithm (Xie and Zhang, 2010), allowing up to four mismatches in the mRNA:miRNA duplex. With the exception of *PvNF-YA2*, *-YA6*, and *-YA8*, all members of the *PvNF-YA* family contain at least one target site for miR169b, miR169c, and/or miR169d in their 3' UTRs (Figure S3). As previously observed in *NF-YA1* of *M. truncatula* (Combiér et al., 2006), the 3' UTR of *PvNF-YA1* exhibits the so called “two hit” model of dual miRNA target sites for either miR169b or miR169d.

### Structure of *PvNF-YB* and *PvNF-YC* genes

In contrast to the *PvNF-YA* family, genes of the *PvNF-YB* family were less structured, showing a variable intron/exon organization among the different members of this family (Figure 1B). Introns were detected only in 5 out of the 14 genes of this family, being *PvNF-YB2* and *-YB5* the most structured ones with 6 introns each, followed by *PvNF-YB14* with 4 and *PvNF-YB10* and *-YB11* with only 1 intron each. These introns were distributed along the 5'

UTR and the coding region of these genes. On the other hand, the analysis of the NF-YC family revealed that 4 (*PvNF-YC1*, *-YC2*, *-YC4*, or *-YC6*) out of the 7 *PvNF-YC* members contain at least one intron. *PvNF-YC1* contains 3 introns, 2 of them located in the 5' UTR region and the third one located in the 3' UTR. In the case of *PvNF-YC2* and *-YC6*, introns are detected exclusively in the 5' UTR region, whereas the only intron detected in *PvNF-YC4* is located in the coding region (Figure 1C). For *PvNF-YC7*, 5' or 3' UTR regions were not defined in the gene models proposed in the current version of the common bean genome or in the EST database.

## MULTIPLE ALIGNMENT AND PHYLOGENETIC ANALYSIS OF PvNF-Y FAMILIES

Multiple alignments for each PvNF-Y subunit families were constructed using ClustalW (Thompson et al., 1994). As previously described in other plant species, PvNF-Y proteins of each family show a central core region of extensive homology flanked by non-conserved sequences (Figures S4–S6). Based on studies performed in yeast and mammals, this central region contains the conserved domains required for subunit interactions and DNA binding (Xing et al., 1993, 1994; Kim et al., 1996; Sinha et al., 1996; McNabb et al., 1997). Alignments of the central core regions are depicted in Figure 2. The multiple alignments of entire protein sequences generated for each PvNF-Y family were used to develop neighbor-joining phylogenetic trees (Figure 3). In addition, we developed phylogenetic trees for each NF-Y family comparing proteins of *Arabidopsis*, *M. truncatula*, soybean and common bean (Figure S7).

### PvNF-YA family

PvNF-YA proteins are variable in length, ranging between 234 and 349 amino acids. The conserved central region is composed of 53 amino acids and contains two characteristic domains with separated functions (Figure 2A). Studies carried out in metazoans and yeast indicated that the first domain is required for subunit interactions, whereas the second one is important for DNA binding and sequence recognition specificity (Xing et al., 1993, 1994; Mantovani et al., 1994). Outside this conserved central region, limited amino acid conservation was observed among different members of this family. Previous studies have shown that regions flanking the conserved central domain of NF-YA proteins are characterized by the presence of Gln (Q) and Ser/Thr (S/T) rich- regions, which presumably act in transcriptional activation (Coustry et al., 1996; De Silvio et al., 1999). As in other eukaryotes, the alignment presented in Figure S4 indicates that regions flanking the conserved domain of PvNF-YA proteins contain a high composition of Q and S/T residues; however, the number and position of these residues are variable between the different members of this family. The phylogenetic tree of the PvNF-YA family using full length proteins identified two main clades, one containing only 2 members and another one containing 7 members (Figure 3A). The identity matrix of full length PvNF-YA proteins generated by ClustalW (Table S3) indicated that some pairs of proteins exhibit higher amino acid identity with each other (e.g., PvNF-YA5 and -YA7, PvNF-YA1 and -YA9, or PvNF-YA3 and -YA8 exhibit 71, 63, or 61% of identity, respectively) than

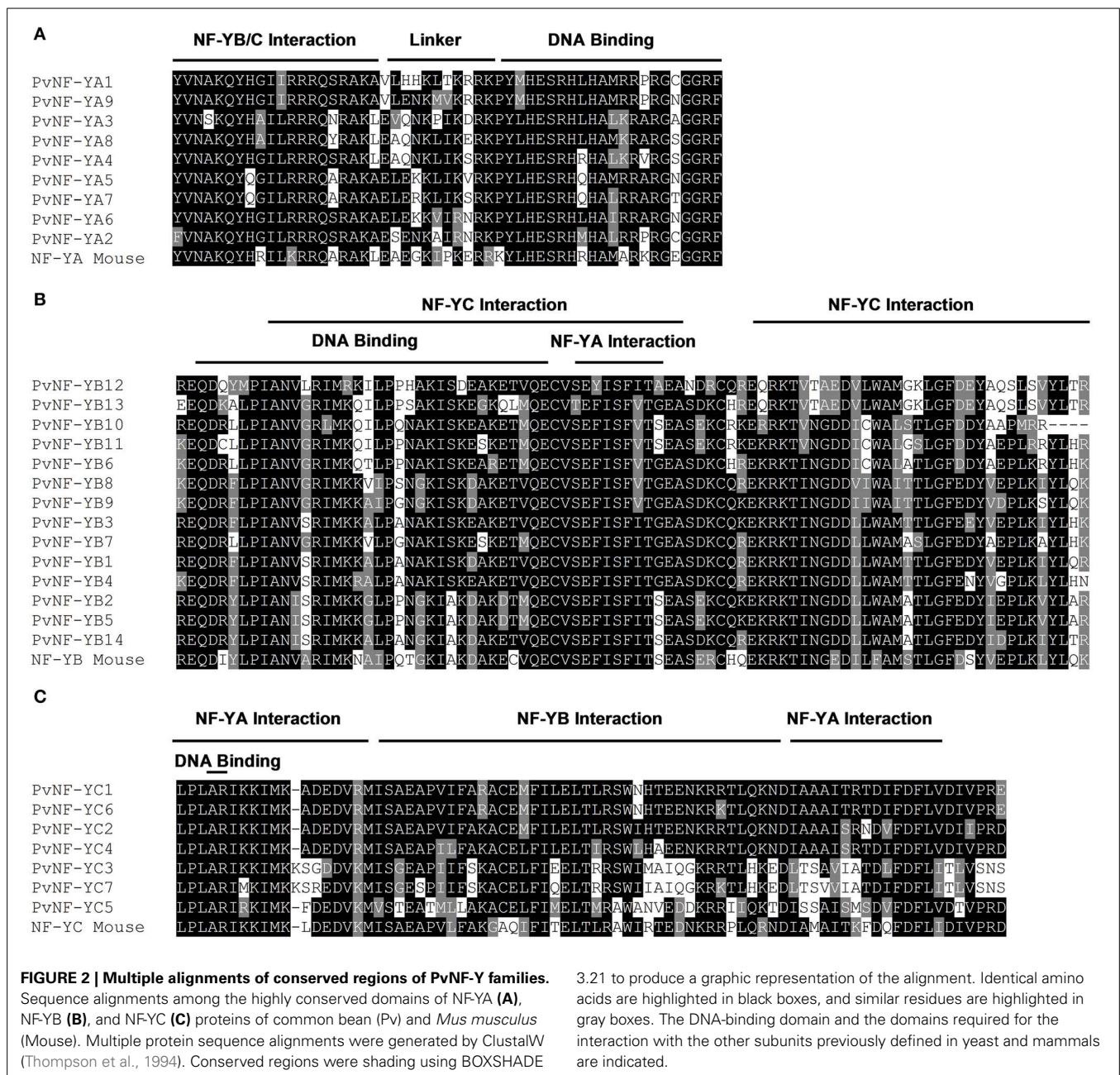
with the rest of the members in this family. The microsynteny analysis supported these phylogenetic relationships among PvNF-YA duplicates (Figure S8A). The ratio of non-synonymous ( $K_a$ ) and synonymous ( $K_s$ ) substitution rates estimated was less than 1 for the majority of the PvNF-YA duplicates found in syntenic regions, suggesting that these genes were subjected to purifying or negative selection (e.g., most amino acid changes are deleterious and, therefore, are selected against).

### PvNF-YB family

The PvNF-YB family is composed of 14 members of variable length (132–229 amino acids), which are highly similar in the 96 amino acids that define the conserved central domain involved in the interaction with NF-YC and NF-YA subunits and DNA binding (Figure 2B). This conserved domain exhibits structural and sequence similarity to the HFD of histone H2B. Outside the central region, the sequences of the different PvNF-YB members are variable in length and amino acid composition (Figure S5). The phylogenetic tree of the common bean NF-YB family is divided into 2 main clades (Figure 3B). One of these clades contains only 3 members, 2 of which are identical in amino acid composition (PvNF-YB2 and -YB5). The second clade contains 11 members, with only few of them showing high sequence similarity at amino acid level among each other, e.g., PvNF-YB8 and -YB9 or PvNF-YB10 and -YB11 with 77 and 80% of identity, respectively (Table S4). The microsynteny results illustrated in Figure S8B indicated that PvNF-YB8 and -YB9, as well as PvNF-YB10 and -YB11 loci, are contained in large syntenic regions ( $E$ -values  $3e^{-150}$  and 0.0, respectively), supporting the phylogenetic relationships among these members. Previous studies in other plant species have showed that some NF-YB members are closely related to the metazoan counterpart, whereas other are more divergent, suggesting that this family might have experienced relaxed selective constraints that resulted in an asymmetric evolution of NF-YB duplicates in plants (Yang et al., 2005; Cao et al., 2011). The results presented in Figure 3B and Figure S7B indicated that this might be also the case in common bean.

### PvNF-YC family

The PvNF-YC family contains a central domain of 80 amino acids that is highly conserved across the different members, but also with the NF-YC mouse counterpart (Figure 2C). This central domain contains an HFD closed related to histone H2A. The use of mutants in yeast and mammals confirmed the importance of this domain in both DNA binding and NF-Y subunit interactions (Romier et al., 2003). As previously observed in other plant species (Siefers et al., 2009; Cao et al., 2011; Petroni et al., 2012), both the N- and C-terminal domains of most PvNF-YC proteins are characterized by an overall enrichment in Q residues, being PvNF-YC3 and -YC7 the exceptions (Figure S6). Functions in transcriptional activation have been assigned to the Q rich regions of mammal and yeast NF-YC proteins (Coustry et al., 1996; De Silvio et al., 1999). The phylogenetic tree presented in Figure 3C revealed 2 different clades. One of these clades includes 3 members, being PvNF-YC3 and -YC7 highly similar to each other (87% of identity); whereas the second clade contains 4 members, where PvNF-YC1 and -YC6 are more closely related,



with 86% of identity (Table S5). The microsynteny analysis of this family indicated that PvNF-YC1 and -YC6 loci are contained in a large syntenic region, supporting the close phylogenetic relationship between these PvNF-YC duplicates (Figure S8C). The estimated Ka/Ks ratios for these PvNF-Y duplicates was 0.13, suggesting that these genes were also subjected to purifying selection.

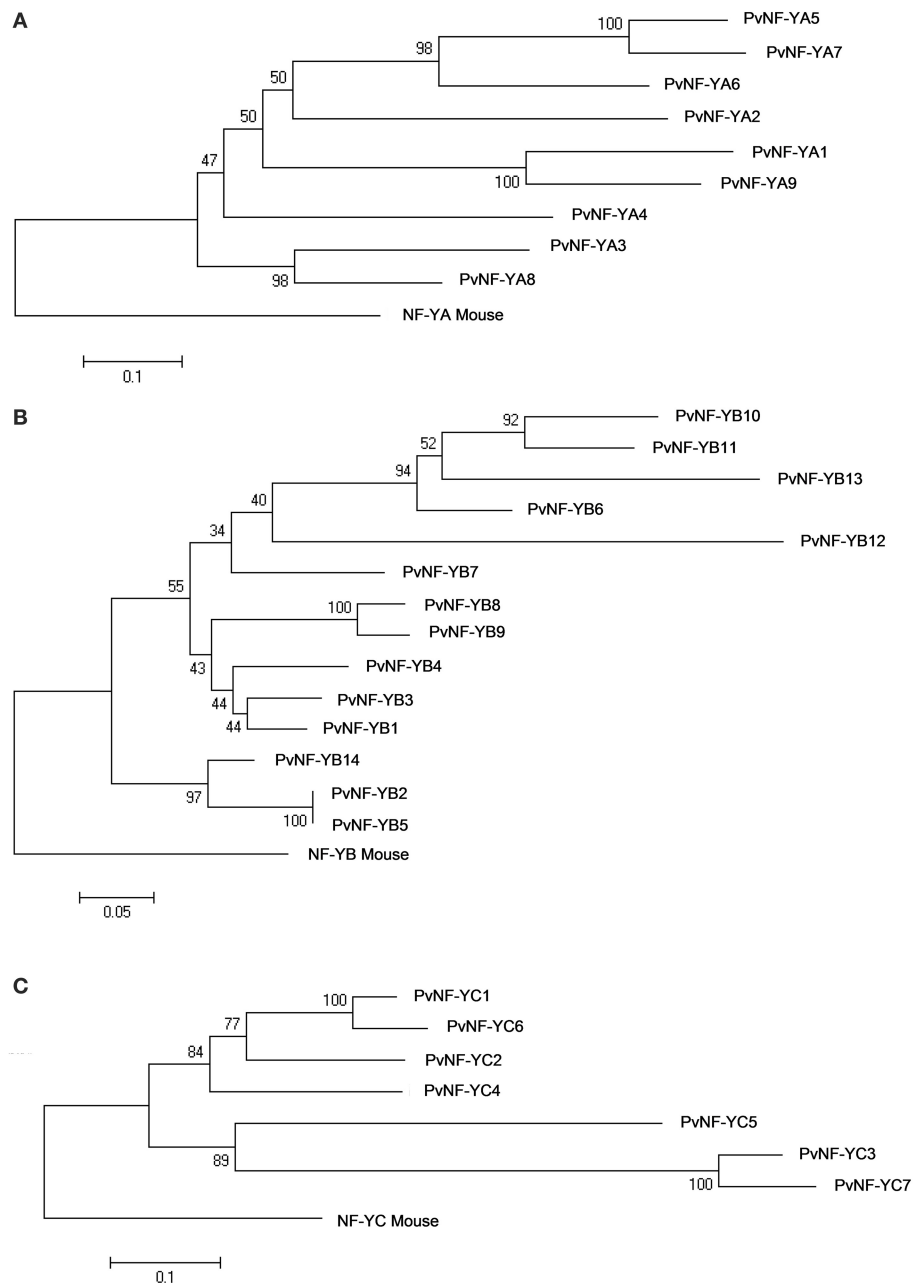
#### EXPRESSION ANALYSIS OF PvNF-Ys IN DIFFERENT ORGANS AND DURING THE SYMBIOTIC ASSOCIATION WITH *R. ETLI*

Individual members of the NF-Y subunits have shown a tissue specific expression pattern in other plant species (Stephenson et al., 2007; Thirumurugan et al., 2008; Siefers et al., 2009;

Cao et al., 2011; Liang et al., 2014). Therefore, we aimed to characterize the expression of different members of each NF-Y subunit in both photosynthetic and non-photosynthetic organs of common bean plants by reverse transcription followed by quantitative PCR (RT-qPCR) using gene specific primers (Table S1). In addition, we aimed to identify PvNF-Y genes that are differentially expressed in a more efficient interaction with rhizobia as compared with a less efficient one.

#### PvNF-YA expression

Analysis of the PvNF-YA family revealed that all members have detectable transcript levels in at least one of the organs examined (Figure 4A). PvNF-YA2, -YA4, -YA5, and -YA7 mRNAs

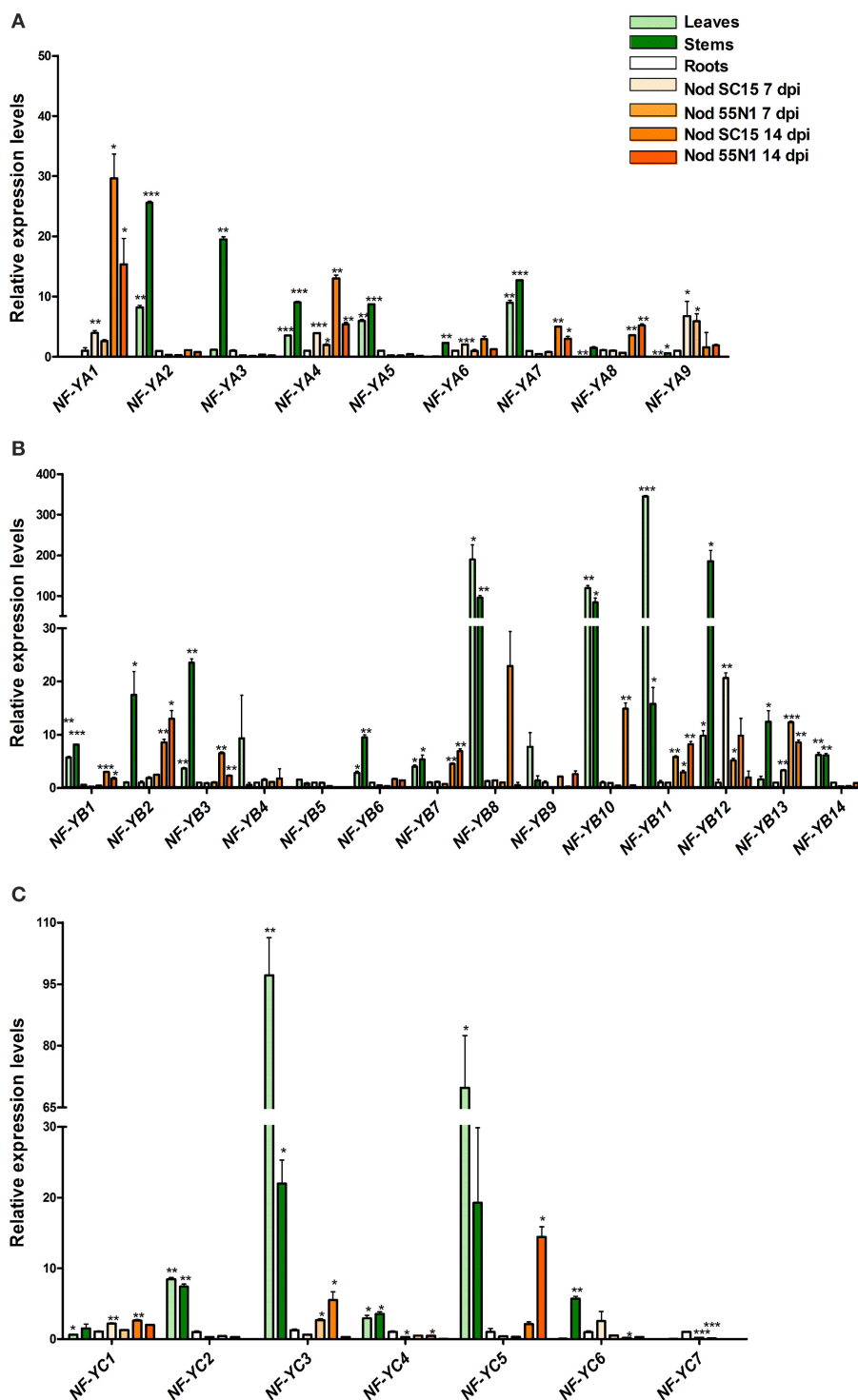


**FIGURE 3 | Phylogenetic trees for the PvNF-Y families.** Phylogenetic trees for the NF-YA (A), NF-YB (B), and NF-YC (C) families were constructed using the neighbor-joining method based on the multiple sequence alignment analysis of the full-length proteins obtained from the common bean genome

V1.0 database, including the mouse NF-Y amino acid sequences as root (rooted tree). All trees were determined and constructed using MEGA 5 from a ClustalW analysis. Reliability values at each branch represent bootstrap samples (10,000 replicates).

accumulated at higher levels in leaves (>6-fold) and stems (>9-fold) as compared to roots. On the other hand, transcript levels of *PvNF-YA3* were 20 fold higher in stems than in leaves or roots. *PvNF-YA6*, *-YA8*, and *-YA9* transcripts were detected in stems and roots, but barely detected in leaf tissues. Notably, transcript accumulation of *PvNF-YA1* was observed in roots, but not in photosynthetic tissues. On the other hand, nodules of 14 dpi with *R. etli* SC15 accumulated significantly higher mRNAs levels of

*PvNF-YA1*, *-YA4*, *-YA7*, and *-YA8* as compared to roots or nodules of 7 dpi, whereas *PvNF-YA9* mRNA levels were higher in nodules of 7 dpi than in roots or older nodules. Interestingly, expression of both *PvNF-YA1* and *-YA4* was lower in nodules formed with the less efficient strain 55N1, whereas no significant differences in the expression of *PvNF-Y9* were observed between nodules occupied by SC15 or 55N1. In order to identify genes that are regulated during early times of the symbiotic interaction, we



**FIGURE 4 | Expression analysis of NF-Y family members in different organs.** Relative transcript levels obtained by reverse-transcriptase quantitative polymerase chain reaction (RT-qPCR) of *PvNF-YA* (A), *PvNF-YB* (B), and *PvNF-YC* (C) gene family members in different organs: leaves, stems, roots, and nodules of 7 or 14 days post-inoculation (dpi) with *Rhizobium etli* SC15 (Nod SC15 7dpi) or 55N1 (Nod 55N1 7dpi). Expression levels were normalized to elongation factor 1 $\alpha$  (eEF1 $\alpha$ ) values and are presented as relative

to the root sample. Error bars represent standard deviation (SD) of at least three technical replicates. Results are representative of two independent biological experiments. Single, double, and triple asterisks indicate that values are significantly different from the control value in an unpaired two-tailed *t*-test with  $p < 0.05$ ,  $0.01$ , and  $0.001$ , respectively. Note that in (B,C) the Y axis was segmented to accommodate the expression values of genes that are expressed at much higher levels in shoot than in root tissue.



analyzed gene expression in roots at 24 hpi with *R. etli*. *PvNF-YA1* and *-YA9* significantly increased in response to rhizobia, whereas the rest of the members did not significantly change in the tested conditions (**Figure 5A**). *PvNF-YA1* mRNA levels increased about five-fold upon inoculation with the high efficient strain SC15, but did not significantly change with 55N1. On the other hand, *PvNF-YA9* increased with both strains, although the magnitude of the increase was higher with SC15 than with 55N1 (>90- vs. 40-fold).

### ***PvNF-YB* expression**

Analysis of the *PvNF-YB* family revealed that seven members, *PvNF-YB1*, *-YB6*, *-YB8*, *-YB10*, *-YB11*, *-YB12*, and *-YB14*, were expressed at higher levels in photosynthetic tissues as compared with below-ground plant organs; i.e., roots and nodules (**Figure 4B**). On the other hand, *PvNF-YB2* and *-YB3* exhibited higher levels of transcripts in stems than in leaves, roots and nodules, whereas expression of *PvNF-YB1*, *-YB2*, *-YB7* was higher in nodules of 14 dpi than in roots or nodules of 7 dpi. Expression of *PvNF-YB10*, *-YB12*, and *-YB13* was very low in roots, but higher in nodules of 14 dpi formed by the strain SC15 as compared to those formed by the strain 55N1. In particular, *PvNF-YB12* transcript levels were also higher in young nodules (7 dpi) formed by SC15 than by 55N1. The expression analysis at an early time point of the symbiotic interaction identified several NF-YB members that significantly increased their transcript levels in response to rhizobia. Notably, most of these transcripts increased to higher levels (e.g., *PvNF-YB4*, *-YB5*, and *-YB13*) or exclusively (e.g., *PvNF-YB8*, *-YB11*, and *-YB14*) with the strain 55N1 as compared to strain SC15, at least at the time point analyzed here (**Figure 5B**). On the other hand, two members of this family, *PvNF-YB1* and *-YB3*, significantly decreased their levels upon inoculation with strain SC15.

### ***PvNF-YC* expression**

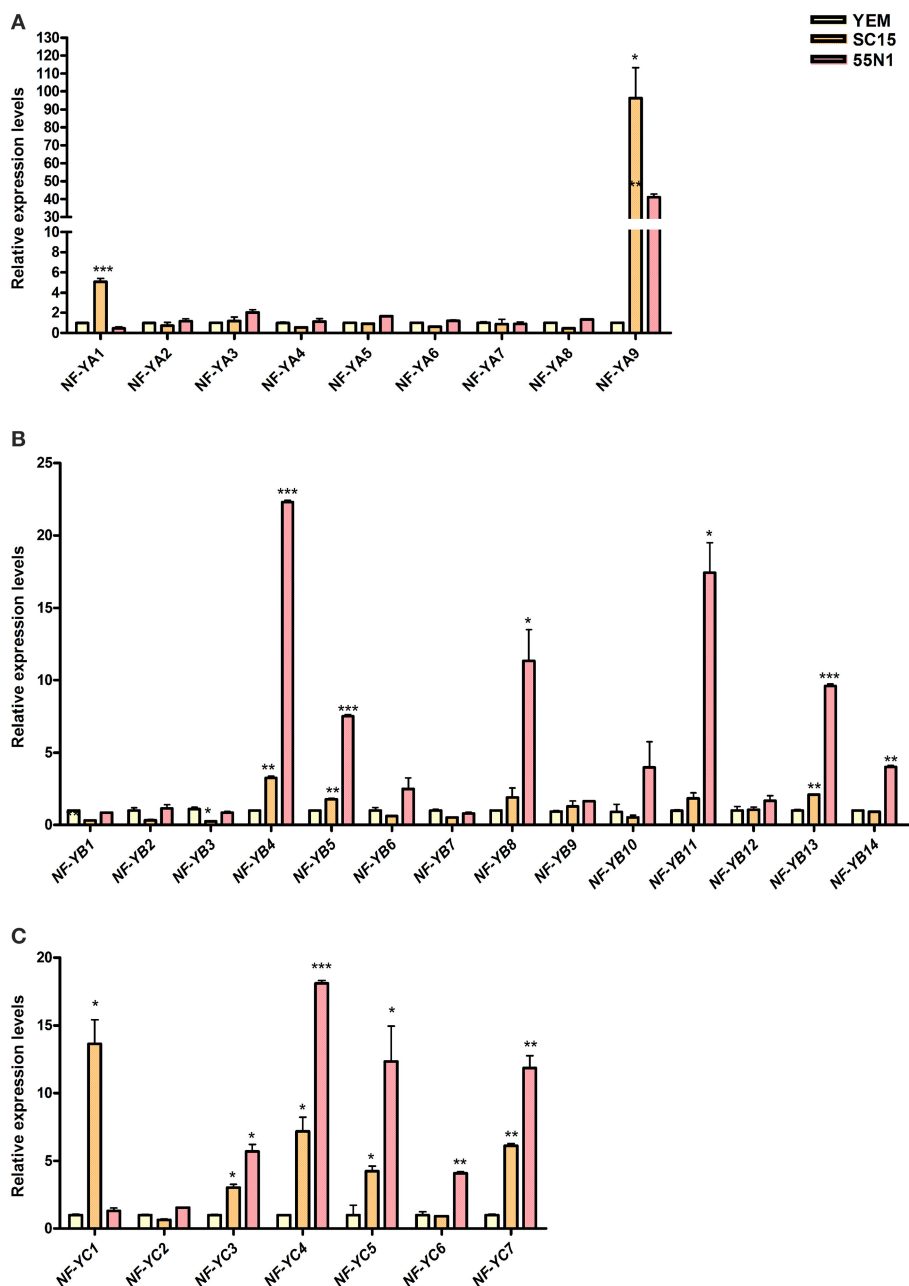
The expression analysis of *PvNF-YC* members in different organs (**Figure 4C**) revealed that *PvNF-YC1* is expressed in all selected tissues, in agreement with that previously reported (Peltzer Meschini et al., 2008). Transcripts of three members (*PvNF-YC2*, *-YC3*, and *-YC5*) of this family accumulated at much higher levels (from 10- to 100-fold) in leaf and stem tissues than in roots or nodules. On the other hand, *PvNF-YC4* transcript levels were moderately higher in leaves and stems as compared to roots, and decreased in nodules of 7 and 14 dpi. *PvNF-YC6* transcripts were almost undetectable in leaves, but showed high levels in stems, whereas *PvNF-YC7* mRNAs were detected at higher levels in roots than in nodules or photosynthetic tissues. The analysis of the *PvNF-YC* family at early stages of the symbiotic interaction revealed that *PvNF-YC1* increased exclusively in response to strain SC15 as previously described (Peltzer Meschini et al., 2008; Zanetti et al., 2010), whereas *PvNF-YC6* was induced only in roots inoculated with strain 55N1 (**Figure 5C**). On the other hand, four *PvNF-YC* members (*PvNF-YC3*, *-YC4*, *-YC5*, and *-YC7*) presented augmented transcript levels in response to both strains; although, as observed for some *PvNF-YB* members, the fold change in gene expression was higher in roots inoculated with strain 55N1 than in those inoculated with SC15.

## **DISCUSSION**

In the past decade, plant NF-Y genes have gained major interest due to their roles in plant development or in the response to changing environmental conditions, particularly in the root nodule symbiosis. In this work, we have annotated and characterized 9, 14 and 7 members of *PvNF-YA*, *PvNF-YB*, and *PvNF-YC* families, respectively (**Table 1**). The number of genes identified for each family in common bean is similar to that of rice, but lower than that previously reported in *Arabidopsis* and in the model legume *M. truncatula* (Laloum et al., 2014). This might be explained by a contraction of these gene families in common bean, as well as by the incompleteness of the common bean genome sequence (Schmutz et al., 2014).

The gene-structure analysis of the *PvNF-YA* gene family revealed that all members of this family contain five or more introns distributed in their 5', 3' UTR or within the coding region (**Figure 1**). Consistently, it has been recently reported that *NF-YA* genes of *Arabidopsis* and canola are also highly structured (Liang et al., 2014). Five members of the *PvNF-YA* family contain uORFs in their 5' UTR when introns are retained (**Figure S2**). These uORFs might either act limiting the translation of the main ORFs (Juntawong et al., 2014) or lead to the synthesis of putative small peptides that destabilize *NF-YA* transcripts (Comber et al., 2008). In *M. truncatula*, *Arabidopsis* and rice, most members of the *NF-YA* family have at least 1 intron in the 5' UTR region (Laloum et al., 2013), suggesting that this might be a conserved regulatory mechanisms among plant *NF-YA* genes. It has been described that some uORFs encode peptides that are conserved among flowering plants (CPuORFs), which might control translation of downstream ORF in response to small molecules (Hayden and Jorgensen, 2007; Jorgensen and Dorantes-Acosta, 2012). Transcription factors represent approximately 30% of the genes that possess CPuORFs. However, we found that the peptide encoded by the uORF1 of *M. truncatula NF-YA1* gene (Comber et al., 2008) is non-conserved in other legumes (e.g., common bean and soybean) and non-legume species (e.g., *Arabidopsis* and rice), since a TBLASTN search using the *M. truncatula* uORF1 predicted amino acid sequence against the genome databases of these other species did not retrieve any hit. The analysis of the intron/exon organization of *PvNF-YB* and *PvNF-YC* family members indicated that most of these genes are less structured than *NF-YA* members (**Figure 1**). This result is in accordance to that observed in *Arabidopsis* and canola. On the contrary, rice *NF-YAs*, *NF-YBs* and *NF-YCs* have different gene structures from *Arabidopsis*, canola and common bean (Liang et al., 2014).

*NF-Y* genes showed high variability in their expression patterns (**Figures 4, 5**). Some of them showed an ubiquitous expression (e.g., *NF-YA4* and *NF-YC1*), whereas others exhibited an organ specific expression pattern (e.g., *NF-YA1* and *NF-YA9* in roots and nodules, *NF-YA2*, *NF-YB14*, and *NF-YC2* in photosynthetic tissues). Based on the phylogenetic trees presented in **Figure S7**, some of these *NF-Y* genes expressed at higher levels in leaves or stems as compared to root tissues (e.g., *PvNF-YB1*, *-YB3*, *-YB4*, *-YB8*, *-YC2*, and *-YC5*) have high sequence similarity with *Arabidopsis* members implicated in the response to the endoplasmic reticulum stress and/or in the promotion of flowering,



**FIGURE 5 | Expression analysis of PvNF-Y family members in common bean roots upon inoculation with *R. etli* strains SC15 or 55N1.** Relative transcript levels of *PvNF-YA* (A), *PvNF-YB* (B), and *PvNF-YC* (C) in roots 24 h post-inoculation with either strain SC15 or 55N1 of *R. etli* or with the medium used to grow the bacteria (YEM) as control, as determined by RT-qPCR. Expression data were normalized

to eEF1 $\alpha$  and presented as relative to roots inoculated with YEM. Error bars represent SD of at least three technical replicates. Results are representative of two independent biological experiments. Single, double, and triple asterisks indicate that values are significantly different from the control value in an unpaired two-tailed *t*-test with  $p < 0.05$ , 0.01, and 0.001, respectively.

such as *AtNF-YB3*, *AtNF-YB2* and *AtNF-YC2* (Petroni et al., 2012; Laloum et al., 2013).

In the context of the root nodule symbiosis, we identified two members of the *PvNF-YA* family, *PvNF-YA1* and *-YA9*, which are highly expressed in nodules as compared to other tissues (Figure 4A). Both *PvNF-YA1* and *-YA9* are differentially induced by the high efficient strain at an early stage of the interaction

(Figure 5A), suggesting that these genes might play a role in partner selection, as previously observed for *PvNF-YC1* (Zanetti et al., 2010). *PvNF-YA1* and *PvNF-YA9* proteins are closely related to the products of *MtNF-YA1* and *MtNF-YA2* genes, respectively, and to the product of the *LjNF-YA1* gene (see Figure S7 and Soyano et al., 2013), which are required for the development of indeterminate and determinate types of nodules in *M. truncatula*

and *L. japonicus*, respectively (Combiér et al., 2006, 2008; Soyano et al., 2013). In addition, *MtNF-YA1* and *MtNF-YA2* are required for bacterial infection and for the induction of early nodulation genes, such as *ERN1* and *ENOD11* (Laloum et al., 2014; Laporte et al., 2014). However, down regulation of *NF-YA1* did not alter the number of infection events in *L. japonicus* roots. These evidences indicated that *NF-YA1* genes might have evolved to play slightly different functions in these two legumes. Thus, it will be of interest to elucidate the function of *PvNF-YA1* and *PvNF-YA9* at different stages of the symbiotic association between common bean and *R. etli*. Similarly, we identified three members of the *PvNF-YB* family, *PvNF-YB8*, *-YB10*, and *-YB12*, whose transcripts exhibited high levels in nodules formed by strain SC15 as compared to roots, although they are not up-regulated by SC15 at an early stage of the interaction (Figures 4B, 5B). These *PvNF-YB* genes are likely candidates to play a role in nodule development at late stages of the highly efficient association between *P. vulgaris* and the *R. etli* strain SC15. Interestingly, a phylogenetic analysis reported by Soyano et al. (2013) indicated that *PvNF-YB10* protein is in the same clade that *LjNF-YB1*, which physically interacts and functions with *LjNF-YA1* in the promotion of root cell divisions and the transcriptional activation of the cell cycle gene *cyclin B*. Previously, we have shown that *PvNF-YC1* is required for the up-regulation of G2/M transition cell cycle genes- including *cyclin B*- in response to rhizobia, and the over-expression of *PvNF-YC1* increased nodule number in *P. vulgaris* roots (Zanetti et al., 2010). Our analysis failed to identify *PvNF-YB* family members that are exclusively or extensively induced by strain SC15 at 24 hpi; however, we can not exclude that some members might be differentially regulated by this strain at other time points, since *PvNF-YB8*, *-YB10*, and *-YB12* mRNAs accumulated at higher levels in nodules formed by SC15 than in those formed by 55N1 (Figure 4B). Future studies focused on subunit interactions will help to discover whether *PvNF-YC1*, *PvNF-YA1*, or *-YA9* and *PvNF-YB8*, *-YB10*, or *-YB12* function in the same heterotrimeric complex to control cell cycle genes and promote the cortical cell divisions that lead to nodule primordia formation.

## CONCLUSIONS

In the past 10 years, an increasing amount of genetic and biochemical evidences have supported the importance of the NF-Y family of proteins in different developmental processes along the lifespan of plants, as well as in their adaptation to adverse environmental conditions. Here, we have generated an initial dataset of sequences, intron/exon arrangements and expression patterns of genes encoding NF-Y subunits in *P. vulgaris*, the most important grain legume used for direct human consumption. The phylogenetic relationships and expression patterns generated here constitute a starting point to elucidate the function of the different members of the NF-Y family in morphogenetic programs in *P. vulgaris*, such as the epidermal infection events and nodule organogenesis, as well as to understand how Mesoamerican plants discriminate and select the most efficient rhizobial strains. In addition, the identification of NF-Y subunits that are co-expressed in the same organ or at the same stage of the symbiotic interaction will facilitate the analysis of putative

subunit combinations that form functional NF-Y complexes in this legume.

## ACKNOWLEDGMENTS

We thank Ulises Villagra, Claudio Mazo, Paula Giménez and Silvana Tongiani for technical assistance. We also thank Andreas Niebel and Maël Baudin for fruitful discussions. This work was financially supported by grants from ANPCyT, Argentina (PICT 2008-0443, PICT 2010-2431, and PICT 2010-2722) and from a CONICET-CNRS cooperative program.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/journal/10.3389/fpls.2014.00761/abstract>

**Figure S1 | Chromosomal locations of PvNF-Ys.** The chromosomal position of each NF-YA (red), NF-YB (green) and NF-YC (blue) family members was mapped to the common bean genome using the Map-View function of PGDD. The chromosome number is indicated at the top of each chromosome. The scale is 2 Mb.

**Figure S2 | Presence of uORFs in the PvNF-YA family members.** Representation of upstream and main open reading frames (uORFs and mORFs, respectively) of individual members of *PvNF-YA* subunit. The analysis was performed with the VectorNTI Advance 10 program using default parameters and a minimum ORF size  $\geq 60$  nucleotides (nt). Lines represent introns, gray boxes represent 5' UTRs, black arrows represent the mORFs and blue arrows represent uORFs present when intron are retained. Sizes of intron, exons, and ORFs can be estimated using the scale bar (200 bp).

**Figure S3 | miR169 target sites in the 3' UTR of PvNF-YA transcripts.** Putative target sites of miR169b/c/d within the 3' UTR of *PvNF-YA* transcripts were predicted with Target-align (Xie and Zhang, 2010) using default parameters and a maximum of 4 mismatches. The target sites of *PvNF-YA* transcripts are shown at the top and the miRNA at the bottom. Sequence location of the transcript is indicated in gray over the sequence. Perfectly complementary bases are shown in black and connected with solid lines, G:U wobble are indicated in blue and connected with dotted lines. Mismatches are shown in pink.

**Figure S4 | Multiple alignments of entire protein sequences of PvNF-YA family.** Multiple alignment of the full length NF-YA proteins of common bean and mouse. The analysis was made with the ClustalW program (Thompson et al., 1994) and shaded using BOXSHADE 3.21 to produce a graphic representation of the alignment. Identical amino acids are highlighted in black boxes, and similar residues are highlighted in gray boxes. Numbers on the left indicate the amino acid position on the protein.

**Figure S5 | Multiple alignments of entire protein sequences of PvNF-YB family.** Multiple alignment of the full length NF-YB proteins of common bean and mouse. The analysis was performed as described in Figure S3.

**Figure S6 | Multiple alignments of entire protein sequences of PvNF-YC family.** Multiple alignment of the full length NF-YC proteins of common bean and mouse. The analysis was performed as described in Figure S3.

**Figure S7 | Phylogenetic trees comparing NF-Y families in P. vulgaris, M. truncatula, G. max, and Arabidopsis.** Phylogenetic relationship of the

NF-YA (A), NF-YB (B), and NF-YC (C) families of common bean, *M. truncatula*, soybean and *Arabidopsis*. Multiple sequence alignments were generated on full-length proteins with ClustalW (Thompson et al., 1994). Phylogenetic trees were constructed using the neighbor-joining method based on the multiple sequence alignment analysis using MEGA5. Reliability values at each branch represent bootstrap values (10,000 replicates). The mouse NF-Y subunits were used to root the trees. Gene ID codes from the genome database available at TAIR or Phytozome were included. For MtNF-YB18 and MtNF-YC4, which do not have a Gene ID at the *M. truncatula* genome database, GenBank accession numbers were included. The higher number of NF-Y members in soybean as compared to the other plant species can be explained by a recent whole-genome duplication event.

#### Figure S8 | Colinear blocks of Pv-NF-Y loci in the common bean genome.

Colinear blocks containing PvNF-YA (A), PvNF-YB (B), and PvNF-YC (C) loci were obtained using the Locus Search function of PGDD. Structures of colinear blocks are shown. In each case, PvNF-Y loci are indicated by red arrows. Blue arrows are other anchor genes in a display region of  $\pm 100$  kb. Non-synonymous (Ka) and synonymous (Ks) substitution rates are indicated.

## REFERENCES

- Aguilar, O. M., Lopez, M. V., Riccillo, P. M., Gonzalez, R. A., Pagano, M., Grasso, D. H., et al. (1998). Prevalence of the *Rhizobium etli*-like allele in genes coding for 16S rRNA among the indigenous rhizobial populations found associated with wild beans from the Southern Andes in Argentina. *Appl. Environ. Microbiol.* 64, 3520–3524.
- Aguilar, O. M., Riva, O., and Peltzer, E. (2004). Analysis of *Rhizobium etli* and of its symbiosis with wild *Phaseolus vulgaris* supports coevolution in centers of host diversification. *Proc. Natl. Acad. Sci. U.S.A.* 101, 13548–13553. doi: 10.1073/pnas.0405321101
- Ballif, J., Endo, S., Kotani, M., Macadam, J., and Wu, Y. (2011). Over-expression of HAP3b enhances primary root elongation in *Arabidopsis*. *Plant Physiol. Biochem.* 49, 579–583. doi: 10.1016/j.plaphy.2011.01.013
- Battaglia, M., Ripodas, C., Clúa, J., Baudin, M., Aguilar, O. M., Niebel, A., et al. (2014). A nuclear factor Y interacting protein of the GRAS family is required for nodule organogenesis, infection thread progression, and lateral root growth. *Plant Physiol.* 164, 1430–1442. doi: 10.1104/pp.113.230896
- Ben-Naim, O., Eshed, R., Parnis, A., Teper-Bamnolker, P., Shalit, A., Coupland, G., et al. (2006). The CCAAT binding factor can mediate interactions between CONSTANS-like proteins and DNA. *Plant J.* 46, 462–476. doi: 10.1111/j.1365-313X.2006.02706.x
- Calvenzani, V., Testoni, B., Gusmaroli, G., Lorenzo, M., Gnesutta, N., Petroni, K., et al. (2012). Interactions and CCAAT-binding of *Arabidopsis thaliana* NF-Y subunits. *PLoS ONE* 7:e42902. doi: 10.1371/journal.pone.0042902
- Cao, S., Kumimoto, R. W., Siriwardana, C. L., Risinger, J. R., and Holt, B. F. 3rd (2011). Identification and characterization of NF-Y transcription factor families in the monocot model plant *Brachypodium distachyon*. *PLoS ONE* 6:e21805. doi: 10.1371/journal.pone.0021805
- 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
- Comber, J. P., De Billy, F., Gamas, P., Niebel, A., and Rivas, S. (2008). Trans-regulation of the expression of the transcription factor MtHAP2-1 by a uORF controls root nodule development. *Genes Dev.* 22, 1549–1559. doi: 10.1101/gad.461808
- Comber, J. P., Frugier, F., De Billy, F., Boualem, A., El-Yahyaoui, F., Moreau, S., et al. (2006). MtHAP2-1 is a key transcriptional regulator of symbiotic nodule development regulated by microRNA169 in *Medicago truncatula*. *Genes Dev.* 20, 3084–3088. doi: 10.1101/gad.402806
- Coustry, F., Maity, S. N., Sinha, S., and De Crombrughe, B. (1996). The transcriptional activity of the CCAAT-binding factor CBF is mediated by two distinct activation domains, one in the CBF-B subunit and the other in the CBF-C subunit. *J. Biol. Chem.* 271, 14485–14491. doi: 10.1074/jbc.271.24.14485
- Cuperus, J. T., Fahlgren, N., and Carrington, J. C. (2011). Evolution and functional diversification of MIRNA genes. *Plant Cell* 23, 431–442. doi: 10.1105/tpc.110.082784
- De Silvo, A., Imbriano, C., and Mantovani, R. (1999). Dissection of the NF-Y transcriptional activation potential. *Nucleic Acids Res.* 27, 2578–2584. doi: 10.1093/nar/27.13.2578
- Fahraeus, G. (1957). The infection of clover root hairs by nodule bacteria studied by a simple glass slide technique. *J. Gen. Microbiol.* 16, 374–381. doi: 10.1099/00221287-16-2-374
- Gepts, P. (1998). Origin and evolution of common bean: past events and recent trends. *Hort. Sci.* 33, 1124–1130.
- Gusmaroli, G., Tonelli, C., and Mantovani, R. (2002). Regulation of novel members of the *Arabidopsis thaliana* CCAAT-binding nuclear factor Y subunits. *Gene* 283, 41–48. doi: 10.1016/S0378-1119(01)00833-2
- Hackenberg, D., Wu, Y., Voigt, A., Adams, R., Schramm, P., and Grimm, B. (2012). Studies on differential nuclear translocation mechanism and assembly of the three subunits of the *Arabidopsis thaliana* transcription factor NF-Y. *Mol. Plant* 5, 876–888. doi: 10.1093/mp/ssr107
- Hayden, C. A., and Jorgensen, R. A. (2007). Identification of novel conserved peptide uORF homology groups in *Arabidopsis* and rice reveals ancient eukaryotic origin of select groups and preferential association with transcription factor-encoding genes. *BMC Biol.* 5:32. doi: 10.1186/1741-7007-5-32
- Hogekamp, C., Arndt, D., Pereira, P. A., Becker, J. D., Hohnjec, N., and Kuster, H. (2011). Laser microdissection unravels cell-type-specific transcription in arbuscular mycorrhizal roots, including CAAT-box transcription factor gene expression correlating with fungal contact and spread. *Plant Physiol.* 157, 2023–2043. doi: 10.1104/pp.111.186635
- Jones-Rhoades, M. W., and Bartel, D. P. (2004). Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. *Mol. Cell* 14, 787–799. doi: 10.1016/j.molcel.2004.05.027
- Jorgensen, R. A., and Dorantes-Acosta, A. E. (2012). Conserved peptide upstream open reading frames are associated with regulatory genes in angiosperms. *Front. Plant Sci.* 3:191. doi: 10.3389/fpls.2012.00191
- Juntawong, P., Girke, T., Bazin, J., and Bailey-Serres, J. (2014). Translational dynamics revealed by genome-wide profiling of ribosome footprints in *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A.* 111, E203–E212. doi: 10.1073/pnas.1317811111
- Kahle, J., Baake, M., Doenecke, D., and Albig, W. (2005). Subunits of the heterotrimeric transcription factor NF-Y are imported into the nucleus by distinct pathways involving importin beta and importin 13. *Mol. Cell. Biol.* 25, 5339–5354. doi: 10.1128/MCB.25.13.5339-5354.2005
- Kim, I. S., Sinha, S., De Crombrughe, B., and Maity, S. N. (1996). Determination of functional domains in the C subunit of the CCAAT-binding factor (CBF) necessary for formation of a CBF-DNA complex: CBF-B interacts simultaneously with both the CBF-A and CBF-C subunits to form a heterotrimeric CBF molecule. *Mol. Cell. Biol.* 16, 4003–4013.
- Kumimoto, R. W., Siriwardana, C. L., Gayler, K. K., Risinger, J. R., Siefers, N., and Holt, B. F. 3rd. (2013). NUCLEAR FACTOR Y transcription factors have both opposing and additive roles in ABA-mediated seed germination. *PLoS ONE* 8:e59481. doi: 10.1371/journal.pone.0059481
- Kumimoto, R. W., Zhang, Y., Siefers, N., and Holt, B. F. 3rd. (2010). NF-YC3, NF-YC4 and NF-YC9 are required for CONSTANS-mediated, photoperiod-dependent flowering in *Arabidopsis thaliana*. *Plant J.* 63, 379–391. doi: 10.1111/j.1365-313X.2010.04247.x
- Kwong, R. W., Bui, A. Q., Lee, H., Kwong, L. W., Fischer, R. L., Goldberg, R. B., et al. (2003). LEAFY COTYLEDON1-LIKE defines a class of regulators essential for embryo development. *Plant Cell* 15, 5–18. doi: 10.1105/tpc.006973
- Laloum, T., Baudin, M., Frances, L., Lepage, A., Billault-Penneteau, B., Cerri, M. R., et al. (2014). Two CCAAT box-binding transcription factors redundantly regulate early steps of the legume-rhizobia endosymbiosis. *Plant J.* 79, 757–768. doi: 10.1111/tpj.12587
- Laloum, T., De Mita, S., Gamas, P., Baudin, M., and Niebel, A. (2013). CCAAT-box binding transcription factors in plants: Y so many? *Trends Plant Sci.* 18, 157–166. doi: 10.1016/j.tplants.2012.07.004
- Laporte, P., Lepage, A., Fournier, J., Catrice, O., Moreau, S., Jardinaud, M. F., et al. (2014). The CCAAT box-binding transcription factor NF-YA1 controls rhizobial infection. *J. Exp. Bot.* 65, 481–494. doi: 10.1093/jxb/ert392
- Lee, T. H., Tang, H., Wang, X., and Paterson, A. H. (2013). PGDD: a database of gene and genome duplication in plants. *Nucleic Acids Res.* 41, D1152–D1158. doi: 10.1093/nar/gks1104

- Li, W. X., Oono, Y., Zhu, J., He, X. J., Wu, J. M., Iida, K., et al. (2008). The Arabidopsis NFYA5 transcription factor is regulated transcriptionally and post-transcriptionally to promote drought resistance. *Plant Cell* 20, 2238–2251. doi: 10.1105/tpc.108.059444
- Liang, M., Yin, X., Lin, Z., Zheng, Q., Liu, G., and Zhao, G. (2014). Identification and characterization of NF-Y transcription factor families in Canola (*Brassica napus* L.). *Planta* 239, 107–126. doi: 10.1007/s00425-013-1964-3
- Lotan, T., Ohto, M., Yee, K. M., West, M. A., Lo, R., Kwong, R. W., et al. (1998). Arabidopsis LEAFY COTYLEDON1 is sufficient to induce embryo development in vegetative cells. *Cell* 93, 1195–1205. doi: 10.1016/S0092-8674(00)81463-4
- Mantovani, R., Li, X. Y., Pessara, U., Hooft Van Huisdijnen, R., Benoist, C., and Mathis, D. (1994). Dominant negative analogs of NF-YA. *J. Biol. Chem.* 269, 20340–20346.
- Mazziotta, L., Reynoso, M. A., Aguilar, O. M., Blanco, F. A., and Zanetti, M. E. (2013). Transcriptional and functional variation of NF-YC1 in genetically diverse accessions of *Phaseolus vulgaris* during the symbiotic association with *Rhizobium etli*. *Plant Biol. (Stuttg.)* 15, 808–818. doi: 10.1111/j.1438-8677.2012.00683.x
- Mcnabb, D. S., Tseng, K. A., and Guarente, L. (1997). The *Saccharomyces cerevisiae* Hap5p homolog from fission yeast reveals two conserved domains that are essential for assembly of heterotetrameric CCAAT-binding factor. *Mol. Cell. Biol.* 17, 7008–7018.
- Miyoshi, K., Ito, Y., Serizawa, A., and Kurata, N. (2003). OsHAP3 genes regulate chloroplast biogenesis in rice. *Plant J.* 36, 532–540. doi: 10.1046/j.1365-313X.2003.01897.x
- Pelaez, P., Trejo, M. S., Iniguez, L. P., Estrada-Navarrete, G., Covarrubias, A. A., Reyes, J. L., et al. (2012). Identification and characterization of microRNAs in *Phaseolus vulgaris* by high-throughput sequencing. *BMC Genomics* 13:83. doi: 10.1186/1471-2164-13-83
- Peltzer Meschini, E. P., Blanco, F. A., Zanetti, M. E., Beker, M. P., Kuster, H., Puhler, A., et al. (2008). Host genes involved in nodulation preference in common bean (*Phaseolus vulgaris*)-rhizobium *etli* symbiosis revealed by suppressive subtractive hybridization. *Mol. Plant Microbe Interact.* 21, 459–468. doi: 10.1094/MPMI-21-4-0459
- Petroni, K., Kumimoto, R. W., Gnesutta, N., Calvenzani, V., Fornari, M., Tonelli, C., et al. (2012). The promiscuous life of plant NUCLEAR FACTOR Y transcription factors. *Plant Cell* 24, 4777–4792. doi: 10.1105/tpc.112.105734
- Ripodas, C., Clúa, J., Battaglia, M., Baudin, M., Niebel, A., Zanetti, M. E., et al. (2014). Transcriptional regulators of legume-rhizobia symbiosis: nuclear factors Ys and GRAS are two for tango. *Plant Signal. Behav.* 9:e28847. doi: 10.4161/psb.28847
- Romier, C., Cocchiarella, F., Mantovani, R., and Moras, D. (2003). The NF-YB/NF-YC structure gives insight into DNA binding and transcription regulation by CCAAT factor NF-Y. *J. Biol. Chem.* 278, 1336–1345. doi: 10.1074/jbc.M209635200
- Schaarschmidt, S., Gresshoff, P. M., and Hause, B. (2013). Analyzing the soybean transcriptome during autoregulation of mycorrhization identifies the transcription factors GmNF-YA1a/b as positive regulators of arbuscular mycorrhization. *Genome Biol.* 14:R62. doi: 10.1186/gb-2013-14-6-r62
- Schmutz, J., Mclean, P. E., Mamidi, S., Wu, G. A., Cannon, S. B., Grimwood, J., et al. (2014). A reference genome for common bean and genome-wide analysis of dual domestications. *Nat. Genet.* 46, 707–713. doi: 10.1038/ng.3008
- Siefers, N., Dang, K. K., Kumimoto, R. W., Bynum, W. E. T., Tayrose, G., and Holt, B. F. 3rd (2009). Tissue-specific expression patterns of Arabidopsis NF-Y transcription factors suggest potential for extensive combinatorial complexity. *Plant Physiol.* 149, 625–641. doi: 10.1104/pp.108.130591
- Sinha, S., Kim, I. S., Sohn, K. Y., De Crombrughe, B., and Maity, S. N. (1996). Three classes of mutations in the A subunit of the CCAAT-binding factor CBF delineate functional domains involved in the three-step assembly of the CBF-DNA complex. *Mol. Cell. Biol.* 16, 328–337.
- Soyano, T., Kouchi, H., Hirota, A., and Hayashi, M. (2013). Nodule inception directly targets NF-Y subunit genes to regulate essential processes of root nodule development in *Lotus japonicus*. *PLoS Genet.* 9:e1003352. doi: 10.1371/journal.pgen.1003352
- Stephenson, T. J., McIntyre, C. L., Collet, C., and Xue, G. P. (2007). Genome-wide identification and expression analysis of the NF-Y family of transcription factors in *Triticum aestivum*. *Plant Mol. Biol.* 65, 77–92. doi: 10.1007/s11103-007-9200-9
- Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28, 2731–2739. doi: 10.1093/molbev/msr121
- Thirumurugan, T., Ito, Y., Kubo, T., Serizawa, A., and Kurata, N. (2008). Identification, characterization and interaction of HAP family genes in rice. *Mol. Genet. Genom.* 279, 279–289. doi: 10.1007/s00438-007-0312-3
- Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22, 4673–4680. doi: 10.1093/nar/22.22.4673
- Tiwari, S. B., Shen, Y., Chang, H. C., Hou, Y., Harris, A., Ma, S. F., et al. (2010). The flowering time regulator CONSTANS is recruited to the FLOWERING LOCUS T promoter via a unique cis-element. *New Phytol.* 187, 57–66. doi: 10.1111/j.1469-8137.2010.03251.x
- Warpeha, K. M., Upadhyay, S., Yeh, J., Adamiak, J., Hawkins, S. I., Lapik, Y. R., et al. (2007). The GCRI, GPA1, PRN1, NF-Y signal chain mediates both blue light and abscisic acid responses in Arabidopsis. *Plant Physiol.* 143, 1590–1600. doi: 10.1104/pp.106.089904
- Xie, F., and Zhang, B. (2010). Target-align: a tool for plant microRNA target identification. *Bioinformatics* 26, 3002–3003. doi: 10.1093/bioinformatics/btq568
- Xing, Y., Fikes, J. D., and Guarente, L. (1993). Mutations in yeast HAP2/HAP3 define a hybrid CCAAT box binding domain. *EMBO J.* 12, 4647–4655.
- Xing, Y., Zhang, S., Olesen, J. T., Rich, A., and Guarente, L. (1994). Subunit interaction in the CCAAT-binding heteromeric complex is mediated by a very short alpha-helix in HAP2. *Proc. Natl. Acad. Sci. U.S.A.* 91, 3009–3013. doi: 10.1073/pnas.91.8.3009
- Yang, J., Xie, Z., and Glover, B. J. (2005). Asymmetric evolution of duplicate genes encoding the CCAAT-binding factor NF-Y in plant genomes. *New Phytol.* 165, 623–631. doi: 10.1111/j.1469-8137.2004.01260.x
- Zanetti, M. E., Blanco, F. A., Beker, M. P., Battaglia, M., and Aguilar, O. M. (2010). A C subunit of the plant nuclear factor NF-Y required for rhizobial infection and nodule development affects partner selection in the common bean *Rhizobium etli* symbiosis. *Plant Cell* 22, 4142–4157. doi: 10.1105/tpc.110.079137

**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.

Received: 29 September 2014; accepted: 10 December 2014; published online: 14 January 2015.

Citation: Ripodas C, Castaingts M, Clúa J, Blanco F and Zanetti ME (2015) Annotation, phylogeny and expression analysis of the nuclear factor Y gene families in common bean (*Phaseolus vulgaris*). *Front. Plant Sci.* 5:761. doi: 10.3389/fpls.2014.00761

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

Copyright © 2015 Ripodas, Castaingts, Clúa, Blanco and Zanetti. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.