Heliyon 10 (2024) e36370

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Genome-wide identification of nuclear factor -Y (NF–Y) transcription factor family in finger millet reveals structural and functional diversity

Varsha Rani [a,b], Vinay Kumar Singh [c], D.C. Joshi [d,**], Rajesh Singh [e], Dinesh Yadav [a,*]

[a] Department of Biotechnology, Deen Dayal Upadhyaya Gorakhpur University, Gorakhpur, 273009, Uttar Pradesh, India
[b] Department of Biotechnology, School of Engineering and Technology, Sandip University, Nashik, 422213, Maharashtra, India
[c] Centre for Bioinformatics, School of Biotechnology, Banaras Hindu University, Varanasi, 221005, Uttar Pradesh, India
[d] ICAR-Vivekananda Institute of Hill Agriculture, Almora, 263601, Uttarakhand, India
[e] Department of Genetics and Plant Breeding, Institute of Agricultural Sciences, Banaras Hindu University, Varanasi, Uttar Pradesh, 221 005, India

ARTICLE INFO

Keywords:
Abiotic stress
Finger millet
miRNA
Magnaporthe oryzae
Nuclear factor-Y
Transcription factor

ABSTRACT

The Nuclear Factor Y (NF–Y) is one of the widely explored transcription factors (TFs) family for its potential role in regulating molecular mechanisms related to stress response and developmental processes.

Finger millet (*Eleusine coracana* (L.) Gaertn) is a hardy and stress-tolerant crop where partial efforts have been made to characterize a few transcription factors. However, the NF–Y TF is still poorly explored and not well documented. The present study aims to identify and characterize *NF–Y* genes of finger millet using a bioinformatics approach. Genome mining revealed 57 EcNF-Y (*Eleusine coracana* Nuclear Factor-Y) genes in finger millet, comprising 18 NF-YA, 23 NF-YB, and 16 NF-YC genes. The gene organization, conserved motif, cis-regulatory elements, miRNA target sites, and three-dimensional structures of these NF-Ys were analyzed. The nucleotide substitution rate and gene duplication analysis showed the presence of 7 EcNF-YA, 10 EcNF-YB, and 8 EcNF-YC paralogous genes and revealed the possibilities of synonymous substitution and stabilizing selection during evolution. The role of NF-Ys of finger millet in abiotic stress tolerance was evident by the presence of relevant *cis*-elements such as ABRE (abscisic acid-responsive elements), DRE (dehydration-responsive element), MYB (myeloblastosis) or MYC (myelocytomatosis). Twenty-three isoforms of miR169, mainly targeting a single NF–Y gene, i.e., the EcNF-YA13 gene, were observed. This interaction could be targeted for finger millet improvement against *Magnaporthe oryzae* (blast fungus). Therefore, by this study, the putative functions related to biotic and abiotic stress tolerance for many of the EcNF-Y genes could be explored in finger millet.

* Corresponding author.
** Corresponding author.
E-mail addresses: Dinesh.Joshi@icar.gov.in, dinesh.pbl@gmail.com (D.C. Joshi), dinesh_yad@rediffmail.com, dinesh.biotech@ddugu.ac.in (D. Yadav).

https://doi.org/10.1016/j.heliyon.2024.e36370
Received 19 December 2023; Received in revised form 11 August 2024; Accepted 14 August 2024
Available online 15 August 2024

1. Introduction

The NF–Y TF is a heterotrimeric complex, comprised of three subunits NF-YA, NF-YB, and NF-YC [1,2] that controls the gene expression by binding to the CCAAT promoter regions of target genes [3–6]. Unlike yeast and humans, where single genes are known to code for each subunit of NF–Y, plants have multiple genes that are known to code for each subunit of NF-Ys [7,8]. Genomic investigations related to NF-Ys have been conducted in many monocot and dicot species with the availability of genome sequences [9]. Several roles unique to plants associated with growth, development, and survival under stress conditions have been assigned to the NF-Ys [10,11]. The role of NF–Y has recently been thoroughly investigated in a variety of crops, particularly with moisture and salinity stress [12–17]. For instance, AtNF-YA9 and OsNF-YA7 play a significant role associated with moisture stress tolerance in Arabidopsis and rice respectively [12,13]. Similarly, NF-YC13 overexpression enhanced salt tolerance in transgenic rice plants [21], whereas expression of TaNF-YA10 made Arabidopsis susceptible to salt stress [18]. Among the millets, NF-Ys in sorghum [19] and foxtail millet [23] have been extensively characterized, and their association with enhanced tolerance to salt, drought, and cold has been elucidated.

Millets are a group of cereals mainly cultivated in harsh climatic conditions in Asia and Africa [19,20]. Their high nutritional value and hardy nature make them potential crops for a "New Green Revolution" [21]. Finger millet is one important millet crop, widely grown in drought hit areas, and serves as an easily available and cheap source of food for humans and fodder for animals [27]. It has an abundance of essential amino acids, including lysine and methionine, with minerals Fe, Zn, P, and K [22,23]. Furthermore, finger millet is an inexpensive food and contains high levels of calcium as compared to other cereals [24,25]. Even though finger millet is generally tolerant of abiotic stresses [26], newer sources of stress tolerance need to be identified so that crop improvement programs can utilize these resources. Research has shown that many transcription factors regulate plant development and stress tolerance. It has already been discussed that plant-specific NF–Y TFs possess master regulatory properties and are connected to various biological processes with a unique plant function, which offers new opportunities to engineer crops for better performance.

Despite its orphan status, in recent years genomic information has emerged in finger millet to perform studies at genomic and transcriptome levels [23,27–29]. Recently, the identification and characterization of nutrient transporters and TFs like DoF [30], NAC [31], and MYB [32] have been reported in finger millet. However, the NF–Y gene family remains to be studied in finger millet. To fill this research gap, extensive studies of the NF–Y TF gene family in finger millet have been performed using bioinformatics tools to reveal its potential for crop improvement.

2. Materials and methods

2.1. Database mining

Two strategies were used for the mining of the NF–Y gene family in the finger millet genome, one through NCBI and another by using Phytozome (https://phytozome-next.jgi.doe.gov/) [33]. In the first strategy, the rice (*Oryza sativa* subsp. *Indica)* NF-YA, NF-YB and NF-YC protein sequences were obtained from PlantTFDB v5.0 (http://plantfdb.cbi.pku.edu.cn/) [34] and were utilized to fish out NF–Y genes in finger millet genome by using TBLASTN [35] tool of NCBI (https://www.ncbi.nlm.nih.gov/) against TSA database (WGS: LXGH01000001-LXGH01525627, WGS: BQKI01000001-BQKI01001196) [23,29], limiting to organism *Eleusine coracana* (L.) Gaertn. (taxid:4511). While in the second strategy, rice NF–Y protein sequences were used for BlastP search in Phytozome v13 database to fish out NF–Y genes in finger millet genome (*Eleusine coracana* v1.1, Phytozome genome Id: 560, NCBI Id: 4511), In addition, the Phytozome database was also searched by using keywords NF–Y, limiting to finger millet genome (*Eleusine coracana* v1.1). The retrieved hits were further processed by Venny 2.1 (https://bioinfogp.cnb.csic.es/tools/venny/) to remove redundancies [36].

2.2. Identification of conserved domain and motif analysis

The identified non-redundant protein sequences were checked for the presence of NF–Y conserved domain associated with HMM profile through Pfam (http://pfam.xfam.org/) database, InterProscan (https://www.ebi.ac.uk/interpro/) [37], and prosite (https://prosite.expasy.org/) [38]. Further, to check the identity and reliability of the conserved domain, the putative NF-Ys proteins were subjected to motif analysis *via* MEME Suite Version 5.5.0 (https://meme-suite.org/meme/tools/meme) with a maximum number of motifs ten and width range from 6 to 40 [39].

2.3. Protein sequence analysis, chromosome mapping, and gene structure

Since the genome of finger millet is not yet annotated for the NF-Ys gene family, so we assign their nomenclature as (EcNF-Yx), in which Ec stands for *Eleusine coracana*, NF–Y for Nuclear factor Y and x for the consecutive number of NF–Y genes appeared in each subfamily. Furthermore, the gff3 annotation file, downloaded from phytozome database was used for physical mapping of NF-Ys gene of finger millet by MapGene2Chromosomev2.1 (http://mg2c.iask.in/mg2c_v2.1/) [33,40], and Gene structure display server (GSDS 2.0) http://gsds.gao-lab.org/was used to display the exon-intron structure of NF–Y genes by comparing CDS sequences and genomic sequences [41]. The predicted EcNF-Y protein were further subjected to characterized for several attributes like the number of amino acids, richness of amino acids, molecular weight, Isoelectric points (pI), aliphatic and instability index together with Grand average of hydropathicity index (GRAVY) as part of physicochemical analysis using the ExPASyProtParam server [42]. Transmembrane domain prediction was carried out by DeepTMHMM server [43]. However, the subcellular and nuclear localization was predicted by the

WOLFPSORT (http://wolfpsort.org/) [44]and NLSDB (https://rostlab.org/services/nlsdb/) [45], respectively.

2.4. Multiple sequence alignment and phylogenetic analysis

EcNF-Y protein sequences were aligned through ClustalW [46], and an evolutionary analysis was conducted by ML method using MEGA 11 [47] by setting parameters to 1000 bootstrap, substitution type of amino acids with Poisson model, rates among the site of gamma distribution with partial deletion. The constructed phylogenetic tree was analyzed for the genetic divergence in the sequence based on branch length, motif distribution, and conserved domain. The closely related members of the same groups and subgroups were analyzed through a distance matrix, which estimates the evolutionary divergence between members of EcNF-Ys proteins. A MEGA11 distance matrix computation tool was used to build an all-to-all matrix from the sequence query set that describes the separation between each pair of sequences. Members of closely related groups appeared at the same distance, indicating that they were structurally, functionally, and evolutionary similar. For such analysis, the distance matrix was constructed through MEGA 11, keeping constant parameter evolutionary analysis. A comparative evolutionary study was also attempted between rice, finger millet, foxtail millet, and sorghum NF-Ys protein sequences.

A gene duplication event was observed during evolutionary analysis, so the Ka/Ks calculation tool (http://services.cbu.uib.no/tools/kaks) [48] was used to predict the nucleotide substitution rate among EcNF-Y and the duplication time was calculated by using ($T = Ks/2\lambda$) where the mean value of clock-like rates (λ)= ($6.5 \times 10-9$) [49]. Gene selection pressure was determined using the ratio of non-synonymous (Ka) to synonymous (Ks) genes. If $Ka/Ks = 1$ Neutral Selection and $Ka/Ks < 1$ = Purifying or Stabilizing Selection (Acting against change), then ($Ka/Ks > 1$) = Positive or Darwinian Selection (Driving Change).

2.5. Cis-regulatory elements and miRNA target site analysis

The 2000 bp upstream sequences from the translation start site (TSS) of all of the EcNF-Ys genes were obtained from the phytozome database [33], and the putative cis-regulatory elements were predicted by using the PlantCARE server (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) [50]. The finger millet miRNA dataset was obtained from the PmiREN database (https://www.pmiren.com/) in order to determine the miRNA target sites [57] and the CDS sequence of EcNF-Ys genes was examined against the miRNA dataset using the pSRNATarget server [51]. The cytoscape platform was used to construct the interacting network between the NF–Y and miRNA genes [59].

2.6. In-silico structure prediction of EcNF-Ys TF

The *in-silico* structure of EcNF-Y proteins were predicted using the SWISS MODEL [52]. To assess the robustness of the predicted structures, various features were scanned, such Ramachandran map, hydrogen bond, hydrophilic and hydrophobic residues, and the number of bad atom-atom contacts. Several quality assessment programs, including ERRAT [53], PROCHECK [54], and QMEAND [55], were used to assess the accuracy of model structures predicted from experimental data. Further, PDBsum [56] and SAVES v6.0 server were used to evaluate the reliability of build model. The ResProx (http://www.resprox.ca/) was used to estimate the resolution of the model, and Super Pose v. 1 and Biovia Discovery Studio 2019 were utilized to superposition and visualize the model.

3. Results

3.1. Identification, characterization, and mapping of finger millet NF–Y protein family

The keyword BLAST search and HMM profile predicted that the finger millet genome encodes about 57 NF–Y proteins comprised of NF-YA (18), NF-YB(23), and NF-YC (16) members. These NF–Y members possess a conserved domain with accession numbers PF02045 and PF00808. Without proper annotation, the existing gene identifiers were highly disordered. As a result, we have internally assigned them consecutive numbering as (EcNF-YA1-18), (EcNF-YB1-23) and (EcNF-YC1-16) to make it more convenient for the research community. All the EcNF-Y proteins varied greatly in length and were distributed among nine pairs of chromosomes (Table 1). Among the nine chromosomes of finger millet, the smallest is chromosome 4 (41 Mb), and the largest is chromosome 5. According to the *in-silico* mapping of EcNF-Ys on finger millet chromosomes (Fig. 1A), the genes are dispersed irregularly throughout the nine chromosomes. Among all, chromosome 5 comprised of the highest number of EcNF-Y genes (22 %), whereas chromosomes 4 and 9 have the lowest total genes (3 %) (Fig. 1B).

3.2. Conserved motif and gene structure analysis

Different NF–Y subunits revealed different motif compositions, as shown in Fig. 2A–C. Each EcNF-YA member showed conserved motif 1 (PIYVNAKQYHGILRRRQARAKAEAENKLVKVRKPYLHESRHLHAMKRARG) and motif 2 (SGGRFLNTKKLQEGKQAGASG) as shown in Fig. 2A. Motif 1 have two functional domains within it, first is subunit association domain (YVNAKQYHGILRRRQARAKAEAEN) and another is DNA binding domain/NF-YA domain (KPYLHESRHLHAMKRARG). In the case of the EcNF-YB subfamily, all members showed conserved motif 1 (AKETVQECVSEFISFVTGEASDKCQREKRKTINGDDLLWAMTTLGFEDYV). At the same time, motif 3 (EPLKLYLQRYRESEGEAAAASSGSGS) was also present in all except EcNF-YB22 (Fig. 2B). Motif 1 has a signature sequence (CVSEFISFVTGEASDKC) representing NF-YB domain. In contrast, the EcNF-YC subfamily has only one conserved motif 1

Table 1
Gene structure, Physico-chemical and domain analysis of NF-Ys family members of finger millet.

Gene Name	Gene Bank ACC NO.	NCBI TSA ACC NO.	Gene Name (Phytozomedatabse)	Chromosome Number	Strand	Gene Start (Bp)	Gene End (Bp)	PFAM ID	MW (Kda)	PI	RICHNESS Of AA
EcNF-YA1	KAK3149857.1	GGMA01037673.1	ELECO.r07.3AG0223950	3A	1	14180381	14183849	PF02045	36234.5	8.66	Gly (G)
EcNF-YA2	KAK3146978.1	GGMA01025560.1	ELECO.r07.3BG0276110	3B	−1	43127507	43132031	PF02045	34171.7	8.86	Ser (S)
EcNF-YA3	KAK3146391.1	GGMA01049907.1	ELECO.r07.3BG0265500	3B	−1	8820320	8822905	PF02045	28158.4	9.42	Ala (A)
EcNF-YA4	KAK3150630.1	GGMA01049297.1	ELECO.r07.3AG0235660	3A	−1	38326070	38327529	PF02045	26420.4	8.96	Ser (S)
EcNF-YA5	KAK3140267.1	GGMA01048791.1	ELECO.r07.5AG0398390	5A	−1	55606636	55610033	PF02045	36429.1	8.85	Ser (S)
EcNF-YA6	KAK3125652.1	GGMA01049858.1	ELECO.r07.7BG0607990	7B	−1	55017274	55019694	PF02045	32681	9.51	Ser (S)
EcNF-YA7	KAK3140229.1	GGMB01046041.1	ELECO.r07.5AG0397870	5A	−1	54999465	55002147	PF02045	23594.8	7.23	Ala (A)
EcNF-YA8	KAK3137052.1	GGMB01031753.1	ELECO.r07.5BG0447020	5B	−1	72033554	72036539	PF02045	31095	9.03	Ser (S)
EcNF-YA9	KAK3127657.1	GGMB01026261.1	ELECO.r07.7AG0576540	7A	−1	43907545	43909939	PF02045	32484.8	9.41	Ser (S)
EcNF-YA10	KAK3148230.1	GGMB01045732.1	ELECO.r07.3BG0292400	3B	1	60772429	60774038	PF02045	26463.5	8.98	Ser (S)
EcNF-YA11	KAK3149648.1	GGLY01048916.1	ELECO.r07.3AG0220400	3A	−1	9585769	9588371	PF02045	24592.4	9.34	Ser (S)
EcNF-YA12	KAK3151879.1	GGLZ01006028.1	ELECO.r07.3AG0251930	3A	1	53388265	53391537	PF02045	31897.2	9.3	Ser (S)
EcNF-YA13	KAK3124474.1	GGLZ01009043.1	ELECO.r07.7BG0587020	7B	1	3539201	3541221	PF02045	38910	6.94	Ser (S)
EcNF-YA14	KAK3155429.1	GGPD01062720.1	ELECO.r07.2BG0203220	2B	−1	68770830	68773747	PF02045	17992.4	9.52	Ala (A)
EcNF-YA15	KAK3137008.1	GGPD01021784.1	ELECO.r07.5BG0446430	5B	−1	71527415	71530070	PF02045	23608.9	7.23	Ala (A)
EcNF-YA16	KAK3146615.1	GGPE01005353.1	ELECO.r07.3BG0268940	3B	1	13453798	13457178	PF02045	7207.54	10.72	Arg (R
EcNF-YA17	KAK3159250.1	GGPE01143578.1	ELECO.r07.2AG0147810	2A	−1	58904161	58907067	PF02045	15360.9	9.51	Leu (L)
EcNF-YA18	KAK3137052.1	GGPE01099362.1	ELECO.r07.5BG0447020	5B	−1	72033554	72036539	PF02045	28864.5	9.24	Ser (S)
EcNF-YB1	KAK3138544.1	GGMA01037069.1	ELECO.r07.5AG0370310	5A	1	6642248	6644875	PF00808	18733.9	6.31	Gly (G)
EcNF-YB2	KAK3165885.1	GGMA01039224.1	ELECO.r07.1AG0038940	1A	−1	51542980	51544688	PF00808	17768.8	5.81	Gly (G)
EcNF-YB3	KAK3121126.1	GGMA01039226.1	ELECO.r07.8BG0646560	8B	−1	4227595	4228350	PF00808	19135.1	5.27	Gly (G)
EcNF-YB4	KAK3162389.1	GGMB01032686.1	ELECO.r07.1BG0088970	1B	−1	67110908	67112642	PF00808	23812.8	7.77	Ser (S)
EcNF-YB5	KAK3146967.1	GGMB01033765.1	ELECO.r07.3BG0275920	3B	−1	42811707	42812384	PF00808	23852.2	6.37	Gly (G)
EcNF-YB6	KAK3134779.1	GGMB01032685.1	ELECO.r07.5BG0410620	5B	−1	840174	841728	PF00808	16712.7	4.97	Glu E
EcNF-YB7	KAK3122838.1	GGMB01033761.1	ELECO.r07.8AG0619080	8A	−1	4795857	4796571	PF00808	18693.7	5.42	Gly (G)
EcNF-YB8	KAK3125643.1	GGLY01050951.1	ELECO.r07.7BG0607860	7B	−1	54964851	54965366	PF00808	14546.3	6.42	Lys (K)
EcNF-YB9	KAK3127648.1	GGLY01050952.1	ELECO.r07.7AG0576420	7A	−1	43851059	43851688	PF00808	22514.8	5.31	Gly (G)
EcNF-YB10	KAK3135344.1	GGLY01010339.1	ELECO.r07.5BG0417830	5B	1	5584646	5587310	PF00808	18190.6	7.74	Gly (G)
EcNF-YB11	KAK3137944.1	GGLY01050950.1	ELECO.r07.5AG0362430	5A	−1	1337901	1339499	PF00808	15214.1	4.83	Glu E
EcNF-YB12	KAK3138817.1	GGLZ01025604.1	ELECO.r07.5AG0373770	5A	−1	9325230	9325721	PF00808	18633.8	6.52	Gly (G)
EcNF-YB13	KAK3166437.1	GGLZ01025188.1	ELECO.r07.1AG0045790	1A	−1	55848068	55848544	PF00808	14632.4	6.11	Gly (G), Lys (L)
EcNF-YB14	KAK3162952.1	GGLZ01025605.1	ELECO.r07.1BG0095760	1B	−1	71098001	71098477	PF00808	18239.2	5.27	Gly
EcNF-YB15	KAK3158966.1	GGLZ01025187.1	ELECO.r07.2AG0143990	2A	−1	56210919	56211620	PF00808	83934.7	8.93	Gln (Q)
EcNF-YB16	KAK3155143.1	GGPD01061059.1	ELECO.r07.2BG0199400	2B	−1	66053931	66054632	PF00808	17768.8	5.81	Gly (G)
EcNF-YB17	KAK3132204.1	GGPE01234691.1	ELECO.r07.6AG0517430	6A	1	11266441	11267214	PF00808	13527.4	5.87	Gly (G)
EcNF-YB18	KAK3162951.1	GGPE01216158.1	ELECO.r07.1BG0095750	1B	−1	71092636	71093100	PF00808	15200.1	4.82	Glu E, Gly (G)
EcNF-YB19	KAK3135711.1	GGPE01216157.1	ELECO.r07.5BG0422490	5B	1	9179659	9180186	PF00808	16725.7	4.94	Glu E
EcNF-YB20	KAK3166436.1	GGPE01234698.1	ELECO.r07.1AG0045780	1A	−1	55842810	55843151	PF00808	15173	4.83	Glu E
EcNF-YB21	KAK3158967.1	GGPE01234697.1	ELECO.r07.2AG0144010	2A	−1	56220236	56220966	PF00808	12052.6	5.06	Lys (K)
EcNF-YB22	KAK3129404.1	GGPE01234702.1	ELECO.r07.6BG0478970	6B	1	51921038	51922577	PF00808	20099.2	7.83	Gly (G)
EcNF-YB23	KAK3155144.1	GGPE01234699.1	ELECO.r07.2BG0199430	2B	−1	66063786	66064486	PF00808	13592.3	5.26	Glu E

(*continued on next page*)

Table 1 (*continued*)

Gene Name	Gene Bank ACC NO.	NCBI TSA ACC NO.	Gene Name (Phytozomedatabse)	Chromosome Number	Strand	Gene Start (Bp)	Gene End (Bp)	PFAM ID	MW (Kda)	PI	RICHNESS Of AA
EcNF-YC1	GJM98477.1	GGMA01047497.1	ELECO.r07.6AG0539090	6A	1	51308086	51308688	PF00808	21330.3	5.37	Ala (A)
EcNF-YC2	KAK3151064.1	GGMA01035097.1	ELECO.r07.3AG0241250	3A	−1	42482325	42483089	PF00808	26780	5.41	Ala (A)
EcNF-YC3	KAK3143326.1	GGMA01015586.1	ELECO.r07.4AG0299000	4A	1	568415	568774	PF00808	12774.6	8.74	Ala (A)
EcNF-YC4	KAK3120350.1	GGMA01045992.1	ELECO.r07.9AG0686210	9A	−1	34730228	34733990	PF00808	27705.8	5.28	Gly (G)
EcNF-YC5	KAK3141108.1	GGMB01007669.1	ELECO.r07.4BG0329660	4B	1	537816	538154	PF00808	11757.4	6.82	Ala (A)
EcNF-YC6	KAK3119016.1	GGMB01031775.1	ELECO.r07.9BG0712390	9B	−1	53067349	53070998	PF00808	27785.9	5.39	Gly (G)
EcNF-YC7	KAK3138401.1	GGMB01045339.1	ELECO.r07.5AG0368440	5A	−1	5319577	5324995	PF00808	32199.6	5.18	Asp (D)
EcNF-YC9	KAK3122059.1	GGLY01018380.1	ELECO.r07.8BG0664580	8B	1	59216059	59216682	PF00808	15202.4	5.35	Ala (A)
EcNF-YC10	KAK3130352.1	GGLY01048101.1	ELECO.r07.6BG0492410	6B	1	65101749	65102351	PF00808	15382.6	5.36	Ala (A)
EcNF-YC11	KAK3135207.1	GGLY01029892.1	ELECO.r07.5BG0416090	5B	−1	4424037	4428682	PF00808	7645.11	9.66	Lys (K), Lys (K)
EcNF-YC13	KAK3122918.1	GGLZ01008654.1	ELECO.r07.8AG0620240	8A	−1	6293443	6296254	PF00808	16572.7	8.5	Lys (K)
EcNF-YC14	KAK3134317.1	GGPD01021903.1	ELECO.r07.6AG0547610	6A	1	57039366	57040139	PF00808	11778.4	5.06	Ala (A)
EcNF-YC15	KAK3152243.1	GGPD01004732.1	ELECO.r07.2BG0156290	2B	−1	2784732	2785517	PF00808	11836.5	4.89	Ala (A)
EcNF-YC16	KAK3156110.1	GGPE01098361.1	ELECO.r07.2AG0102940	2A	−1	2860670	2861455	PF00808	11656.3	4.92	Ala (A)
EcNF-YC17	GJN04677.1	GGPE01221602.1	ELECO.r07.8AG0620240	8A	−1	6293443	6296254	PF00808	5544.41	4.65	Leu (L)
EcNF-YC18	KAK3122990.1	GGPE01070831.1	ELECO.r07.8AG0621600	8A	1	9281504	9282178	PF00808	9284.79	8.06	Leu (L)

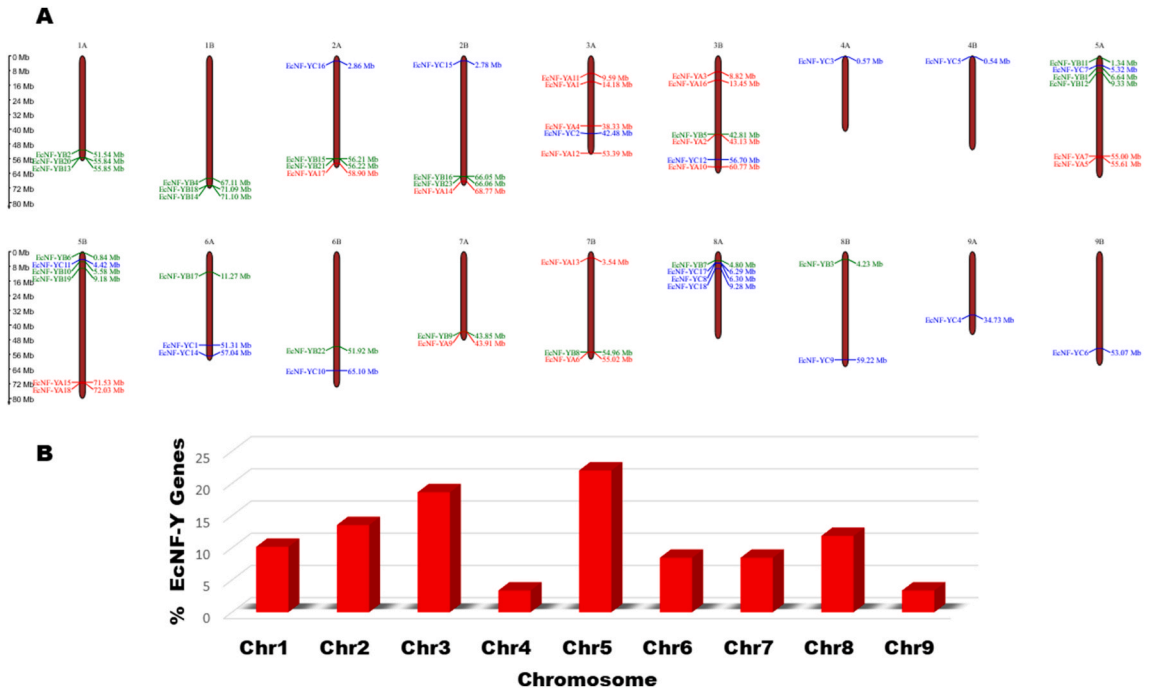

Fig. 1. Distribution of NF–Y genes onto nine finger millet chromosomes. (A))*in-silico* chromosome mapping of NF–Y genes on chromosome (numbered 1–9). (B) Percentage of NF–Y genes on each the finger millet chromosome to show their distribution abundance. Chromosomal distances are given in Mb.

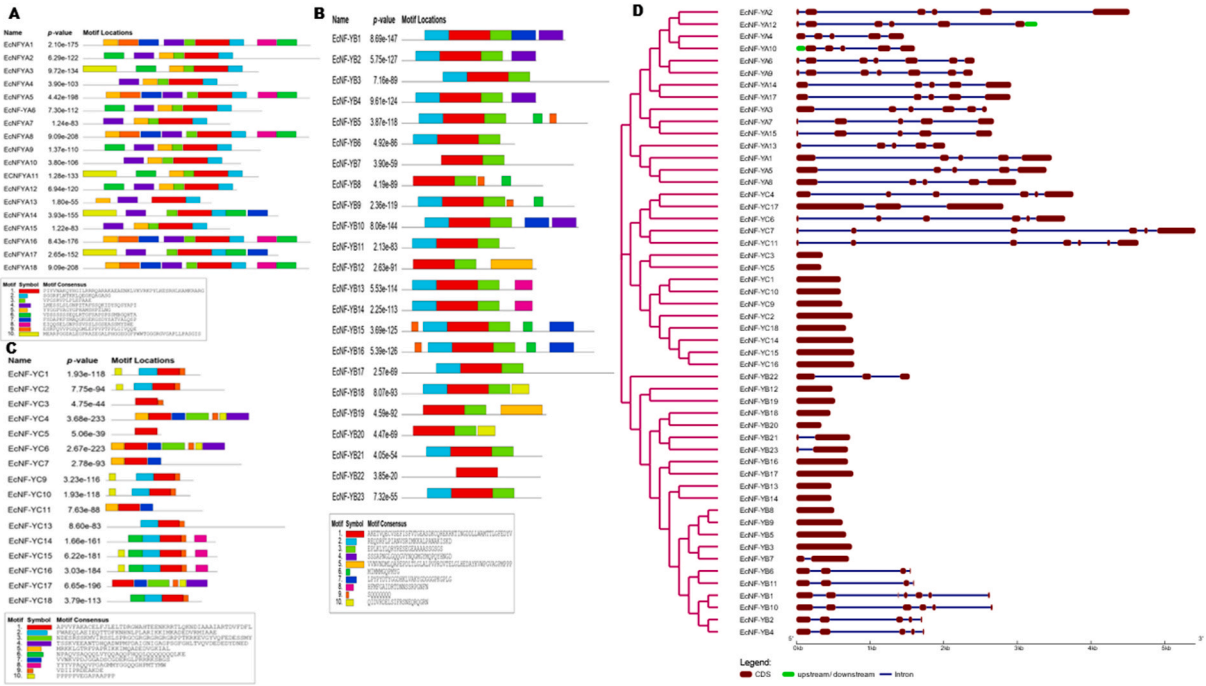

Fig. 2. Schematic representation of finger millet NF–Y subfamily with motif and gene structure (A) Motif analysis of EcNF-YA genes (B) Motif analysis of EcNF-YB genes (C) Motif analysis of EcNF-YC genes, (D) Evolutionary analysis and gene structure organizations of EcNF-Y genes, Exons and introns are represented by maroon boxes and blue lines, respectively. Scale represents the sizes of exons and introns. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

(APVVFAKACELFJLELTDRGWAHTEENKRRTLQKNDIAAAIARTDVFDFL) while others were relatively diverse (Fig. 2C). Each of the ubiquitous conserved motifs of EcNF-YA and EcNF-YB protein sequence revealed the presence of unique NF–Y domain through InterPro accession number IPR001289 and IPR003958.

The gene structure organization of the EcNF-Ys gene revealed that the finger millet genome undergoes significant evolutionary changes based on the highly diverse distribution of intronic regions (ranging from 0 1 2 in number), as shown in Fig. S1. The shortest EcNF-Y gene belongs to the NF-YC Subfamily and is 338 bp long (EcNF-YC5). The longest EcNF-Y gene (EcNF-YC7) possesses 5.5 kb of genomic sequence. Furthermore, 23 genes of the EcNF-Y family (40 %) lacked introns, while the number of CDS regions varies from gene to gene, as shown in Fig. 2D.

3.3. Physico-chemical analysis of EcNF-Y proteins

The number of physicochemical attributes of NF–Y genes of finger millet has been determined using the ExPASy ProtParam tool, as shown in Table 1, and a detailed analysis is presented in Table S1. The length of EcNF-YA protein ranges from 62 (EcNF-YA16) to 348

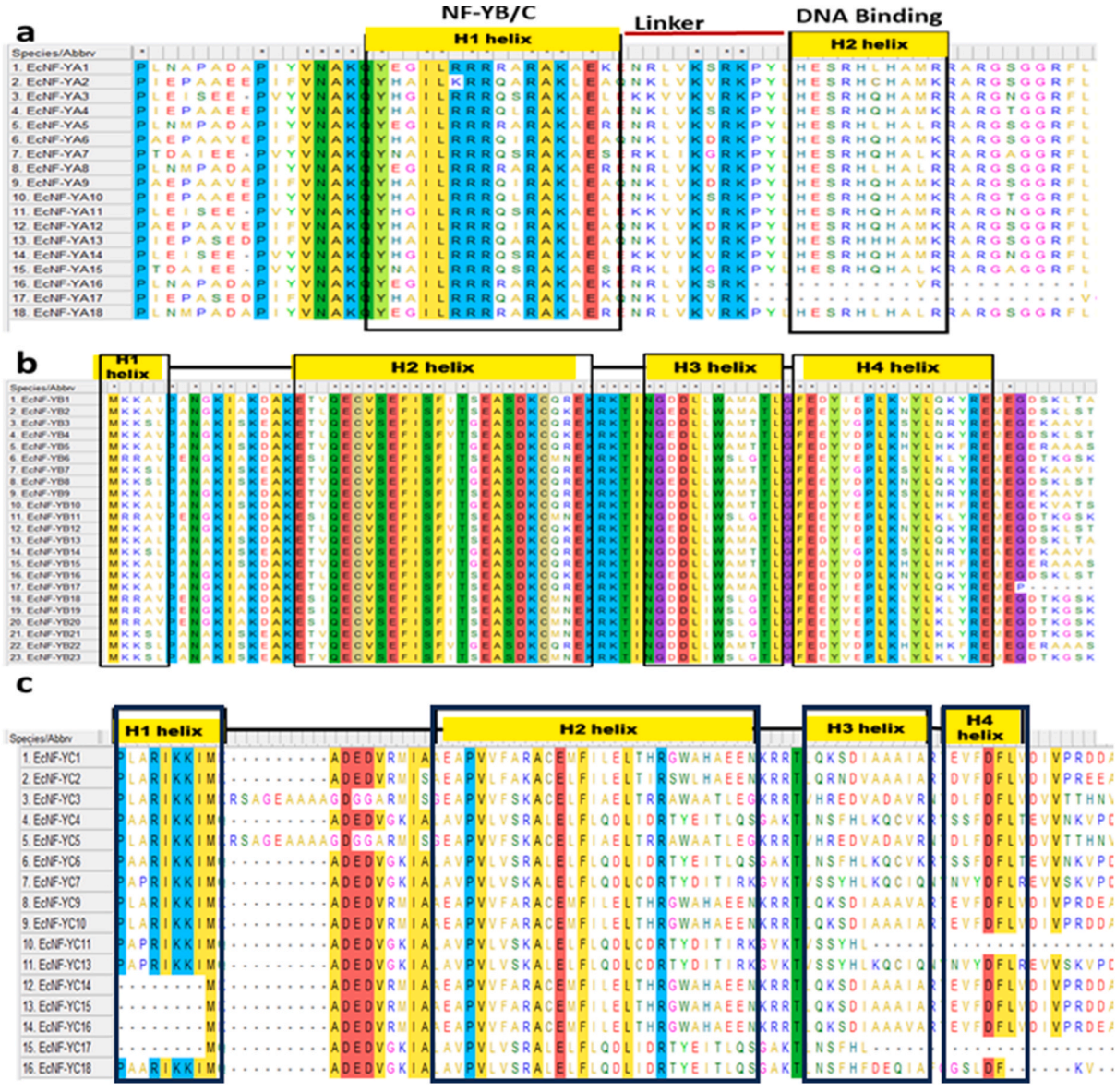

Fig. 3. Multiple Sequence alignment (MSA) of finger millet NF–Y genes. The top of each alignment shows conserved residues in secondary structures of each NF-Ys genes. Helixes are represented by yellow rectangles and loops/helix-helix interaction by black lines. Whereas, red line indicate linker which connect H1 helix with H2 helix in NF-YA members.**(A)** MSA ofEcNF-YA subfamily **(B)** MSA of EcNF-YB subfamily **(C)** MSA of EcNF-YC subfamily. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

(EcNF-YA13) amino acids, with a molecular weight of 7 kDa–38 kDa. The range of isoelectric point (PI values) was found to vary from 6.94 (slightly acidic) to 9.52 (base). Additionally, all the EcNF-YA proteins showed an instability index of more than 40, reflecting unstable proteins. The GRAVY values for all the ECNF-YA proteins were found in the negative range except EcNF-YA17, revealing a predominance of hydrophilic nature. In the case of EcNF-YB, protein ranges from 108 (EcNF-YB21) to 764 (EcNF-YB15) amino acid residue, with molecular weight from 12 kDa to 83 kDa. The pI values ranged from 4.82 (acidic) to 7.83 (slightly basic). Here, most of the EcNF-YB proteins showed an instability index of less than 40, revealing a stable nature, while nine proteins, namely EcNFYB1/4/5/10/11/12/15/17/22 were found to be unstable. All the EcNF-YB proteins were hydrophilic based on the negative values of GRAVY. In the case of EcNF-YC protein, the amino acid ranges from 50 (EcNF-YC17) to 292 (EcNF-YC7) with molecular weight ranging from 5 kDa to 32 kDa. The pI values determined were from 4.65 (acidic) to 10.52 (basic). The value of the instability index was found to be more than 40 for most of the EcNF-YC proteins, revealing an unstable nature, though few of them, namely EcNFYC3/5/11/13/18, were found to be stable proteins. The hydrophilic nature of these proteins was predominant except for EcNF-YC17, which showed a hydrophobic nature. Additionally, the finger millets' NF–Y proteins showed a predominance of Gly, Ser, Ala, Arg, Leu, Glu, Gln, Lys, and Asp amino acids. Based on DeepTMHMM prediction, it can be stated that these EcNF-Y proteins are predominately present in the inner region of the cell and are not observed in the transmembrane region.

3.4. Subcellular and nuclear localization prediction

The WOLFPSORT analysis revealed that most of the EcNF-YA proteins are present in the nucleus except EcNF-YA16 and EcNF-YA17, which are found in the chloroplast. Similarly, all EcNF-YB proteins are also found in the nucleus, though only a few EcNF-YC proteins are observed in the nucleus, and predominately, they are found in the cytoplasm and chloroplast. The detailed analysis of the subcellular localization of all the EcNF-Y proteins is shown in Table S2. Further, the NLSDB database revealed that all the members of the EcNF-YA subfamily have nuclear localization signals (NLS). Detailed output of the NLSDB prediction is shown in Table S3. It consists of the MKRAR sequence, a conserved sequence of motif 1 of the EcNF-YA subfamily, as shown in Fig. 2A.

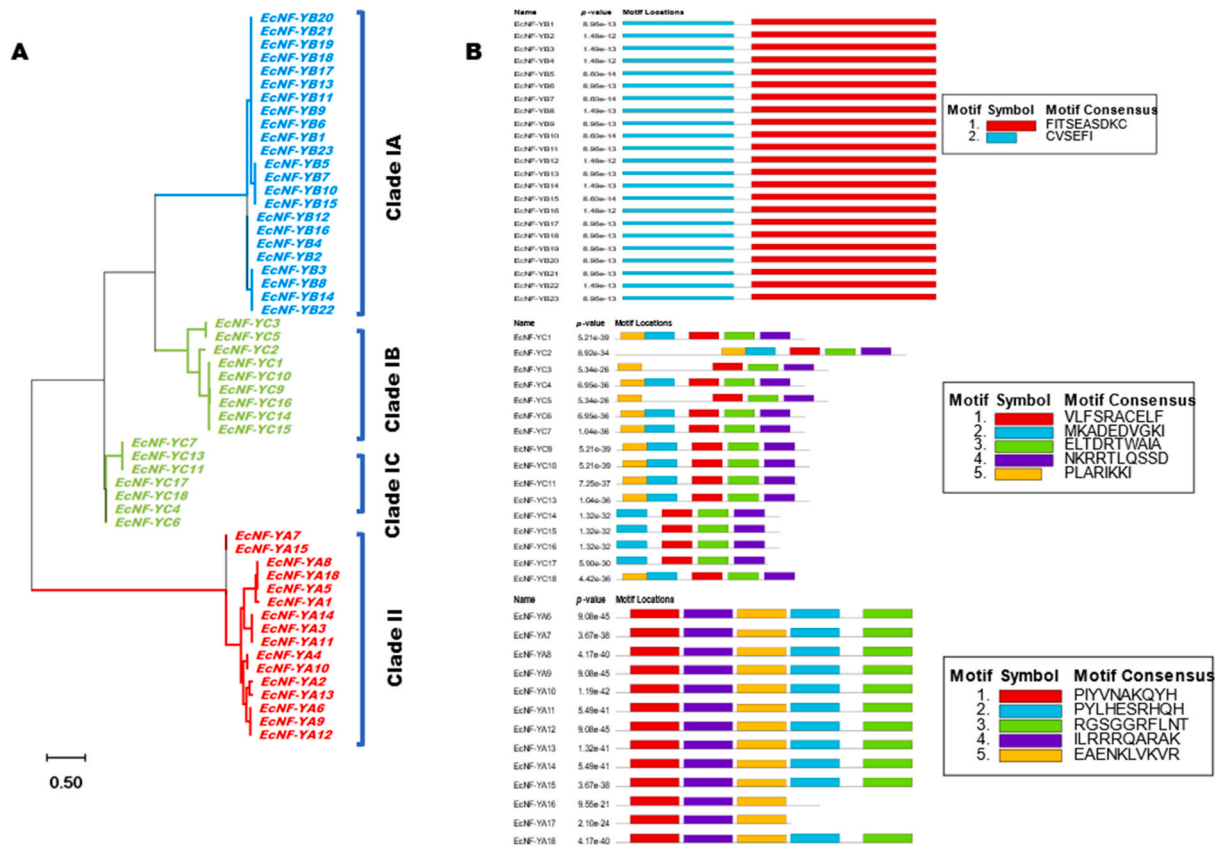

Fig. 4. (A) Evolutionary analysis between NF–Y genes based on conserved domain, different color represents three subfamily of NF–Y genes, EcNF-YB represented by cyan, EcNF-YC by green and EcNF-YA represented with red colour. (B) conserved motif present in domain sequence of EcNF-YB, EcNF-YC and EcNF-YA subfamily. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

3.5. Multiple sequence alignment (MSA) and phylogenetic analysis

The MSA revealed that EcNF-YA proteins have 62 % entirely conserved amino acid residues, whereas 38 % residues were variable (with 36 % Parsimony informative sites and 0.0001 % singletons site). This subfamily possesses two conserved domains, one for NF-YB/C interaction and the other for DNA-binding domains (Fig. 3A). The EcNF-YB proteins showed 88 % conserved amino acid residues and 12 % variable amino acid residues (with 11 % parsimony sites and zero percent singleton sites) (Fig. 3B). Compared to EcNF-YA and EcNF-YB subfamilies, EcNF-YC proteins showed less conserved amino acid residues. The MSA of different subfamilies of EcNF-Y is shown in Fig. 3A-C.

Based on the conserved domain sequences of each EcNF-Y subfamily, an evolutionary tree was constructed (Fig. 4A). In an evolutionary tree, the genetic divergence can be estimated by branch lengths, i.e. the longer the branch, the greater the genetic changes. The evolutionary tree revealed two major clades designated as Clades I and II. Further, Clade I was subdivided into subclades A, B, and C. Subclade IA consists of members of the EcNF-YB subfamily with a conserved motif namely, CVSEFI and FITSEASDKC (Fig. 4A and B). Subclade IB and IC, comprising members of the NF-YC family, showed several conserved motifs like VLFSRACELF, MKADEDVGKI, ELTDRTWAIA, NKRRTLQSSD, and PLARIKKI (Fig. 4A and B). The Clade-II comprises members of EcNF-YA with conserved motifs EPIYVNAKQY, PYLHESRHQH, RGSGGRFLNT, ILRRRQARAK, and EAENKLVKVR (Fig. 4A and B). Generally, the predicted EcNF-Y family members have been arranged in mainly two clades based on conserved domain, IPR001289 and IPR003958, and subclades based on motif arrangement and evolutionary relations (Fig. 4A–C). Furthermore, all 57 EcNF-Y genes shared one conserved motif, "FVTSKASDKC" (Fig. S2). The estimates of evolutionary divergence between EcNF-Y genes are shown in Table S4.

During the evolutionary analysis, it was noted that many of the finger millet NF–Y genes had been duplicated; therefore, as part of the evolutionary analysis, the paralogous genes and nucleotide substitution rate were analyzed. Based on the Ka/Ks calculation tool, 7, 10, and 8 paralogous genes were identified for EcNF-YA/B/C, respectively. The selective pressure on genes due to which encoded protein shows functional alteration can be predicted by nucleotide substitution rate, i.e., the ratio of Ka (nonsynonymous) vs. Ks (synonymous), as shown in Table 2. The twenty-one gene pairs with Ka/Ks < 1 indicate the possibility of synonymous substitution and stabilizing selection during evolution. On the other hand, three gene pairs having Ka/Ks > 1 indicate Darwinian selection during evolution. Moreover, the average evolution time for these paralogous genes was estimated to be 8.23 million years ago.

To further investigate the comparative evolutionary relations among finger millet, rice, foxtail millet, and sorghum NF-Ys, a total of 228 NF-Ys (57 EcNF-Ys, 39 OsNF-Ys, 65 SiNF-Ys, and 65 SbNF-Ys) protein sequences were retrieved and an unrooted phylogenetic tree were constructed. The tree representing NF–YAs comprising 18 EcNF-Ys, 10 OsNF-Ys, 24 SiNF-Ys, and 27SbNF-Ys showed distinct nine clades designated as clade A to I (Fig. 5A). Similarly, NF-YBs with 23 EcNF-Ys, 14OsNF-Ys, 20SiNF-Ys, and 18SbNF-Ys grouped into 11 clades, marked as A to K (Fig. 5B). A total of 9 clades were observed for NF–YCs comprising of 16EcNF-YCs, 15OsNF–YCs, 21SiNF–YCs and 20SbNF–YCs (Fig. 5C). The EcNF-YA members represented a series of paralogs, equidistant from each other, in all the clades except clade-C and clade-G, revealing that finger millet EcNF-YA members have a diverse set of amino acid residues. The members of finger millet EcNF-YAs showed a distinct subclade in the phylogenetic tree except EcNF-YA16, which is closely placed with rice OsNF-YA (BGIOSGA010142/OsNF-YA10) and distantly related with EcNF-YA1 (Fig. 5A). The OsNF-YA10 is a perfect homolog of

Table 2
The gene duplication time identified in different paralogous pairs of NF–Y gene families of finger millet. The Ka represents the number of non-synonymous substitution per non-synonymous site while Ks is the number of synonymous substitutions per synonymous site and Ka/Ks represents the ratios of non-synonymous (Ka) versus synonymous (Ks) mutations.

Gene Pair	Ka	Ks	Ka/Ks ratio	Time (MYA)	Type of selection
EcNF-YC5, EcNF-YC3	0.0023202	0.01281305	0.181081007	0.985619231	Stabilizing selection
EcNF-YC7, EcNF-YC2	0.432	0.6304	0.685279188	48.49230769	Stabilizing selection
EcNF-YC16, EcNF-YC9	0	0.005736295	0	0.441253462	Stabilizing selection
EcNF-YC15, EcNF-YC10	0	0.023066705	0	1.774361923	Stabilizing selection
EcNF-YC1, EcNF-YC14	0.00228312	0.00874082	0.261202037	0.672370773	Stabilizing selection
EcNF-YC18, EcNF-YC6	0.07364029	0.06127879	1.201725589	4.713753077	Darwinian selection
EcNF-YC4, EcNF-YC11	0.05418132	0.37755	0.143507668	29.04230769	Stabilizing selection
EcNF-YB14, EcNF-YB6	0.00712088	0.014976235	0.47547865	1.152018077	Stabilizing selection
EcNF-YB9, EcNF-YB3	0.012578495	0.05302187	0.237232202	4.078605385	Stabilizing selection
EcNF-YB19, EcNF-YB6	0.005703025	0.00326019	1.749292192	0.25078385	Darwinian selection
EcNF-YB8, EcNF-YB11	0.001442385	0.003595325	0.401183476	0.276563465	Stabilizing selection
EcNF-YB23, EcNF-YB21	0.006434445	0.03078815	0.208990959	2.368319231	Stabilizing selection
EcNF-YB4, EcNF-YB2	0	1E-10	0	7.69231E-09	Stabilizing selection
EcNF-YB16, EcNF-YB12	0.01748499	0.05126283	0.341085149	3.943294615	Stabilizing selection
EcNF-YB13, EcNF-YB1	0.00625884	0.06715501	0.093199897	5.16577	Stabilizing selection
EcNF-YA10, EcNF-YA4	0.002947275	0.01667083	0.176792337	1.282371538	Stabilizing selection
EcNF-YA15, EcNF-YA7	0.00212085	0.01125918	0.188366293	0.866090769	Stabilizing selection
EcNF-YA9, EcNF-YA6	0.012535495	0.01384678	0.905300366	1.065136923	Stabilizing selection
EcNF-YA13, EcNF-YA2	0.038051575	0.0657691	0.578563109	5.059161538	Stabilizing selection
EcNF-YB22, EcNF-YB5	0.006643645	0.02755316	0.241120982	2.119473846	Stabilizing selection
EcNF-YA14, EcNF-YA3	0.002945615	1E-10	29456150	7.69231E-09	Darwinian selection
EcNF-YA16, EcNF-YA6	0.076321975	0.100918305	0.75627484	7.762946538	Stabilizing selection
EcNF-YB10, EcNF-YA5	0.65799474	0.7318344	0.899103322	56.29495385	Stabilizing selection
EcNF-YA8, EcNF-YA18	0.00068597	0.00120659	0.568519547	0.092814615	Stabilizing selection

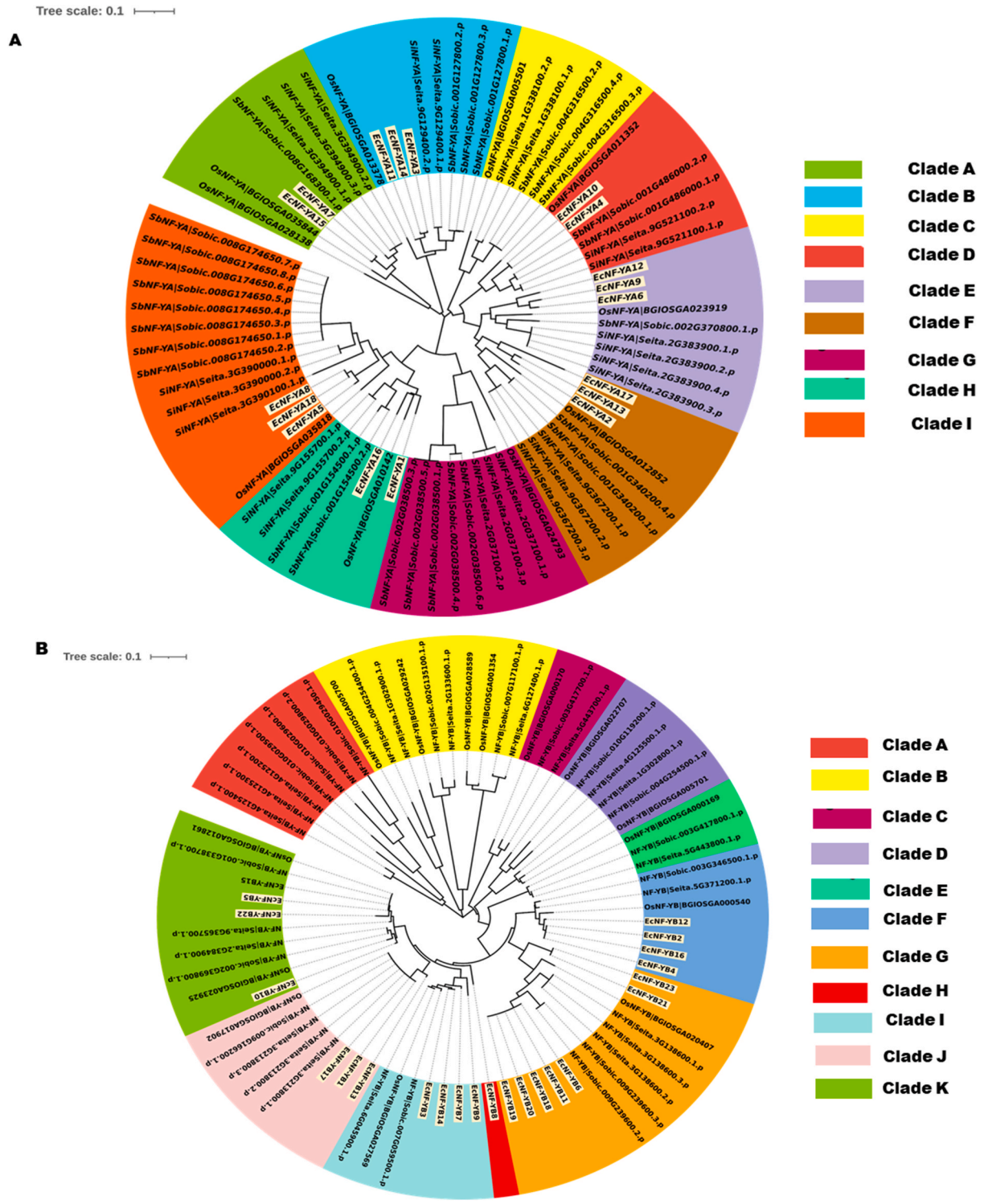

Fig. 5. Comparative phylogenetic analysis of finger millet NF–Y genes with NF–Y genes of rice, sorghum and foxtail millet. Full-length NF–Y proteins from finger millet, rice, sorghum and foxtail millet were aligned using CLUSTALW, and the phylogenetic tree was created with MEGA11 with 1000 bootstrap replicates. Interactive tree of life (iTOL; https://itol.embl.de/upload.cgi) was used to display the tree. (A) phylogenetic tree of NF-YA subfamily (B) phylogenetic tree of NF-YB subfamily (C) phylogenetic tree of NF-YC subfamily.

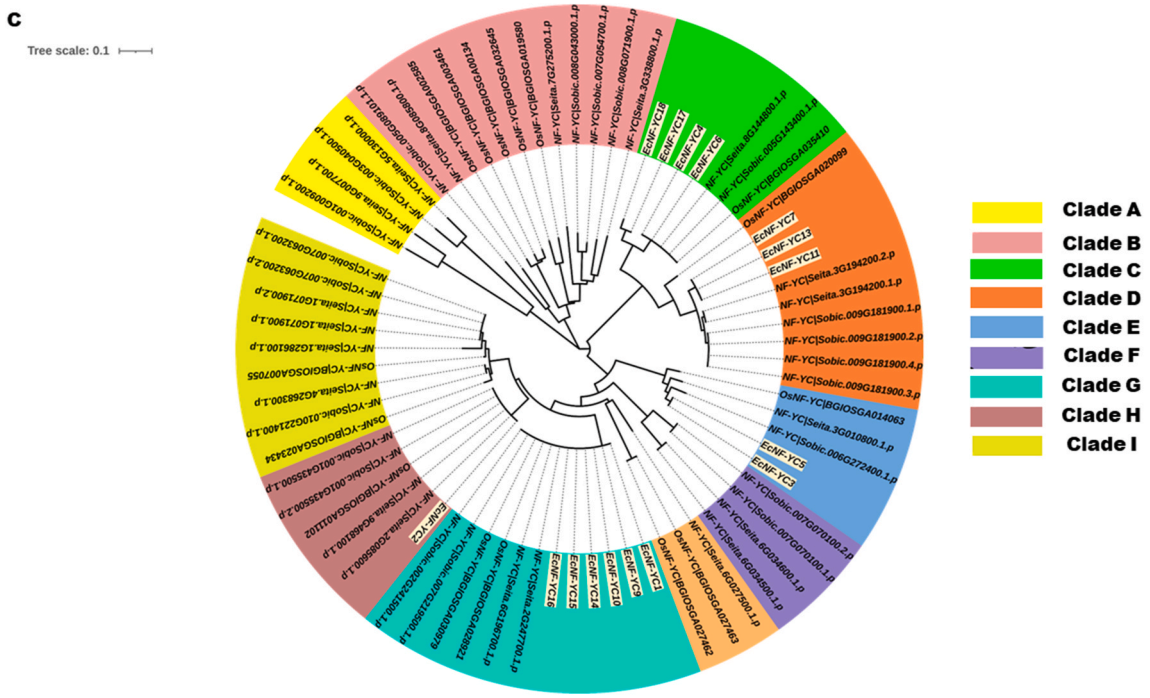

Fig. 5. (*continued*).

Arabidopsis (AtNF-YA10) [57], known to be associated with leaf development, could be utilized for genetic engineering of vegetable biomass and crop productivity.

The NF-YB members of finger millet were distributed among clades F, G, and K while NF-YB members of rice, sorghum, and foxtail millet were closely associated with clades A to E. The EcNF-YB8 is an exception and is placed in clade H (Fig. 5B), making it unique. Similarly, the members of EcNF-YC were found to be distributed among five clades namely C, D, E, G, and, H out of 9 clades observed in the phylogenetic tree (Fig. 5C). One of the members of EcNF-YC2 was closely related with foxtail millet NF-YC (seita.2G085600.1p) (Fig. 5C), assists in the growth and development of flowers by negatively regulating the expression of EHD1 and HD3A genes [58]. Hence, EcNF-YC2 might be playing the same role in finger millet.

3.6. Cis-regulatory elements (CRE) of EcNF-Y protein

The CREs in the promoter region of the protein sequences reflect the putative functions of the gene of interest and can be determined using the PlantCare program. To get an insight into the putative roles of identified EcNF-Y proteins, especially for stress tolerance, the presence of relevant CREs was searched by considering sequences of 2000 bp upstream of the promoter sequence. The identified cis-elements were well known for their functional role in various abiotic stresses (DRE, MYB, STRE, GT-1, SP1, CAAT Motif, G-Box, TCT motif, LTR, MYC, and I-Box), hormonal signaling (ABRE, TCA-elements, AuxRR-core, TGA-box, GARE-motif, TGACG-motif, and CGTCA-motif), plants developmental processes (A-box, Circadian, GATA) and biotic stress (W-Box, TC-rich repeats) besides these O2 sites and GC motif were also identified in finger millet NF–Y genes. The detailed genomic organization of CREs with their motif sequences, present in the promoter region of NF–Y protein responsible for biotic stress, abiotic stress, hormonal signal, and development, have been provided in Table S5. The ubiquitous presence of the cis-acting element CCAAT in promoter and enhancer regions of EcNF-Y was a distinguishing feature. In contrast, the bulk of EcNF-Y proteins consists of ABA-responsive element (ABRE), MYB, MYC, and dehydration-responsive element (DRE) in their promoter region. The element STRE, activated in osmotic stress, low pH, and nutrient starvation conditions, is also found in most of the EcNF-Ys proteins. In contrast, motif CAAAGATATC engaged in circadian regulation was observed in EcNF-YB4, EcNF-YB8, EcNF-YB14, and EcNF-YB21, while A-box encoding plant development-related activities was deciphered in many EcNF-Ys genes. CGTCA motifs (involved in signal transduction in plant defense mechanism) were found in most of the EcNF-Ys genes. However, W-Box, a defensive CRE against the fungal pathogen, was primarily seen in EcNF-YA and EcNF-YB genes. Similarly, MYB, STRE, and G-Box, associated with abiotic stresses, were found in most of the EcNF-Ys genes. The DRE (dehydration response element) is primarily found in promoters of cold-responsive genes in EcNF-YC9 and EcNF-YC16. The presence of different CREs in the promoter region of finger millet reveals its significant role in plant development and stress responses.

3.7. miRNAtarget sites analysis

To produce proteins and regulate gene expression, miRNAs cleave mRNA or inhibit translation. Thirty-six miRNA families of finger millet were retrieved from the PmiREN database. Among these, a total 23 isoforms of miRNA belong to Eco-miR169, and remains from Eco-miR394, Eco-miR396, Eco-miRN5579, Eco-miRN5583, Eco-miRN5584, Eco-miRN5586 and Eco-miRN5588 family, which were approximately perfect target sequences for EcNF-YA1/2/13, EcNF-YB3/5/7/9/14/15/22 and EcNF-YC1/4/6 proteins. All the miRNA sequences predicted and aligned against CDS of EcNF-Y proteins are shown in Table S6 and represented by Fig. 6. The 23 isoforms of Eco-miR169 contain target sites in EcNFYA13 with cleavage activity (Fig. 6). Whereas, EcNF-YB and EcNF-YC genes are involved with other remaining miRNA families (Eco-miR394, Eco-miR396, Eco-miRN5579, Eco-miRN5583, Eco-miRN5584, Eco-miRN5586 and Eco-miRN5588). Stress-responsive miRNAs have been found to influence plant growth and development mainly by targeting transcription factors. By identifying miRNA target sites, the regulation of EcNF-Ys under various biotic and abiotic stresses can be elucidated.

3.8. In silico structure prediction of EcNF-Ys TF

The widely preferred method for modeling protein structures is to compare a sequence of interest, i.e., the target sequence (unknown structure), with a template (known structure) and look for structural similarities. SWISS-MODEL was utilized to ascertain the structure of EcNF-Y proteins, namely EcNF-YA2, EcNF-YB3, and EcNFYC1 (Fig. 7A–C). The PDB ID 6r2v, 7c9p, 6r0m was used to determine the structure of EcNF-YA2, EcNF-YB3, and EcNF-YC1. The findings of the PROCHECK assessment for the EcNF-YA2, EcNF-YB3, and EcNF-YC1 proteins revealed that over 90 % of the residues of each NF–Y protein were found in the most preferred region, and none of the residues fell in the disallowed region of the Ramachandran plot, suggesting the reliability of the model (Supplementary Fig. S3). By measuring ERRAT score (>50 %) and degree of similarity to template using SuperPose V.1, the model was considered acceptable. The RMSD (Root Mean Square Deviation) of the alpha carbon and backbone atoms and NRES (Number of Overlapped Residue) were analyzed for the superposed structure. According to the superposition analysis, it can be stated that the model quality is good with the lowest amount of residue divergence from the template (Fig. S4). Further validation of the predicted model was performed with ResProx resolution criteria (Good: 0–1.5, Middle: 1.5–2.5, Bad: >2.5), and accordingly, the present models are of good resolution. Table 3 summarizes the detailed analysis of homology modeling using SWISS-MODEL, structural assessment, validation using the SAVES server, and quality and resolution estimation using SuperPOSE v.1 and ResProx. The 2-D structural analysis revealed that finger millet NF-YA members contain two helices and six beta-turns (Fig. 7A), while NF-YB members have four helices, three helix-helix interactions, and one beta-turn (Fig. 7B). In contrast, NF-YC members contain four helices, four helix-helix interactions, and five beta turns (Fig. 7C), and the conserved sequences of helices of each member of the EcNF-Ys subfamily are shown in Fig. 3A–C.

4. Discussion

The genome-wide studies for NF–Y transcription factors that contribute to essential biological processes and adaptability to stresses reveal variability in the number of different subfamilies of NF-Ys in several crops. For example, in Arabidopsis, 10 NF-YA, 13 NF-YB,

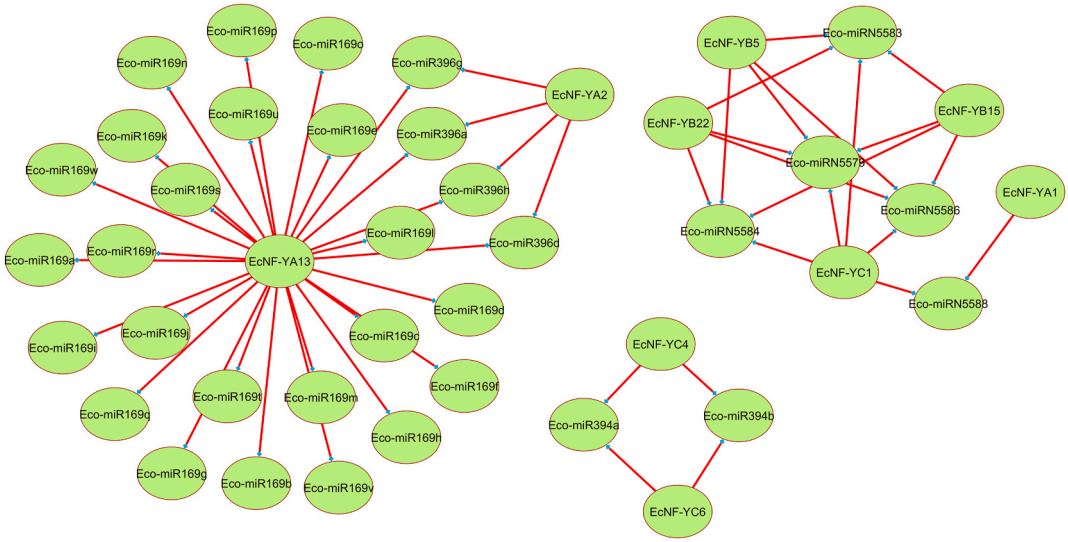

Fig. 6. Finger millet, NF–Y genes, and miRNA interaction, carried out by using cystoscope, reveals most of the miR169 and miR396 isoforms shows interaction with EcNF-YA13 genes, whereas, EcNF-YA2 genes of finger millet also shows interaction with four isoforms of miR396, additionally NF-YC6, NF-YC1 and NF-YA1 shows interaction with other form of miRNA (mir5583/5579/5586/5584/5588 and two isoforms of miR394.

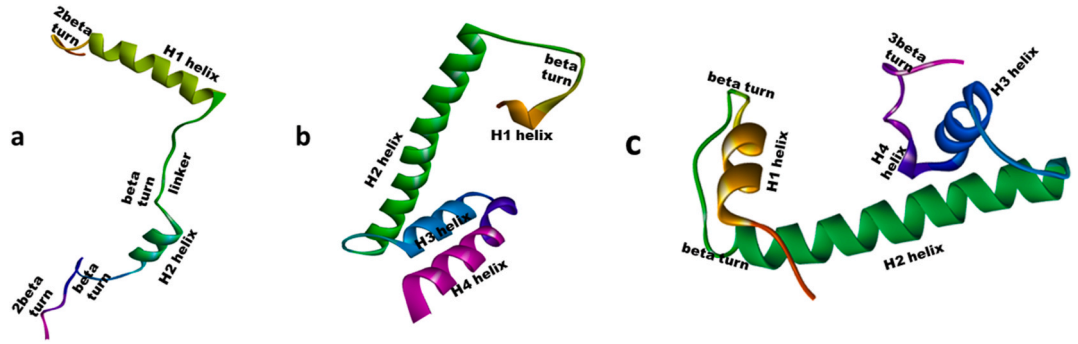

Fig. 7. *In-silico* structure prediction of NF–Y genes of finger millet. (A) *in-silico*model ofEcNF-YA2 consist H1 helix (lime green) and H2 helix (green) with six beta turn (orange, green and purple). The H1 and H2 helix are connected with linker (green) (B) *in-silico*model ofEcNF-YB3 consist H1 helix (orange), H2 helix (green),H3 helix (cyan) and H4 helix (purple) with one beta turn (green) (C) *in-silico*model ofEcNF-YC1 consist H1 helix (orange), H2 helix (green), H3 helix (blue) and H4 helix (purple) with five beta turn (green and purple). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

and 13NF-YC have been reported [59], while the *Oryza sativa* genome revealed 10, 11, and 7 genes for NF-YA, NF-YB, and NF-YC, respectively [13]. In the case of *Setaria italica* 10, 15, and 14, NF-YA, NF-YB, and NF-YC have been observed, respectively [60]. At the same time, the *Sorghum bicolor* genome predicted 8, 19, and 15 genes for NF-YA, NF-YB, and NF-YC subunits respectively [16,17]. The emerging genome sequence information paved the way for exploring NF–Y and other versatile transcriptome families in finger millet [23,29,40]. In this study, genome mining of finger millet was performed to identify and characterize the NF–Y TFs. Using the bioinformatics approach, a total of 57 EcNF-Ys comprising 18 NF–YAs, 23 NF-YBs, and 16 NF–YCs were deciphered from the finger millet genome. Further, several attributes like conserved domain, motif, subcellular localization, phylogenetic, nucleotide substitutions rate, miRNA target sites, and *cis*-acting element were characterized. These computational characterizations could be utilized to identify the potential NF-Ys, which could be targeted for crop improvement using biotechnological tools.

The multiple sequence alignments of finger millet NF–Y revealed that each of the individual families of EcNF-YA/B/C possesses several conserved regions associated with DNA and protein interactions. These have been attempted in several other plants [16,17,61,62]. Several conserved sequences observed in multiple sequence alignment are closely associated with the CCAAT sequence and have a crucial role in DNA interactions [13,16,17,59,63,64]. The ubiquitous presence of highly conserved three arginine and histidine residues (Fig. 3) among the members of the NF-YA subfamily of finger millet was similar as observed for several crops like Arabidopsis, rice, sorghum, wheat, walnut, potato, and barley [13,16,17,59,62,65,66]. Similarly, a few conserved motifs, namely QECVSEFISF, EASDKC, and EKKVTING among EcNF-YB members, have an important role to play in DNA, NF-YA, and NF-YC interactions. The genome-wide study for NF–Y in foxtail millet revealed several conserved motifs in NF-YB which interact with DNA and NF-YA/C. A similar conserved motif has also been reported in the sorghum and walnut genomes [17,65].

The phylogenetic tree of 57 EcNF-Ys revealed three distinct subfamilies with a common ancestor (Fig. 4) and members of the EcNF-YC subfamily showed more proximity to the EcNF-YB subfamily as compared to the EcNF-YA. A total of 25 EcNF-Y members were paralogous to each other, and 21 showed stabilizing selection during evolution. In sorghum and foxtail millet, most of the paralogous genes showed stabilizing selection, which might be due to a similar pattern of evolutionary selection revealed among members of the Poaceae family [16,17,60].

The subcellular localization of most of the NF-YA/B members of finger millet is the nucleus. In contrast, NF-YC members are predominately observed in cytoplasm and chloroplast, though EcNF-YB proteins lack NLS signal. The *in-silico* study revealed that EcNF-YB transcription factors are responsible for the Nuclear Export Signal (NES), a signal used to transport proteins from the nucleus to the

Table 3
Structural assessment, validation, and quality estimation of finger millet NF–Y genes model.

Homology Modeling				Structural assessment and validation				Quality and Resolution estimation through				
EcNF-Y Protein/Acc. No.	Template PDB ID	Identity	Q MEAN D	Residues in most favoured regions	Residues in additional allowed regions	Residues in generously allowed and disallowed regions	ERRAT	Template resolution	Resolution of Predicted model	Resolution of Superposed structure	RMSD	NRES
EcNF-YA (GGMA01025560.1)	6r2v	80.33 %	0.62 ± 0.11	52	2	0	88.89	2.5 Å	1.6 Å	2.5	0.45	60
EcNFYB (GGMA01039226.1)	7c9p	98.68 %	0.66 ± 0.11	66	2	0	100	2.0 Å	1.5 Å	2	0.10	76
EcNFYC (GGMA01035097.1)	6r0m	84.04 %	0.73 ± 0.10	68	6	0	92.53	2.3 Å	1.5 Å	2.3	0.29	80

cytoplasm [67]. The members of the EcNF-YA subfamily showed nuclear localization signals (NLS), which mediate the transportation of proteins from the cytoplasm to the nucleus [45]. The localization of NF-YB3/10 in Arabidopsis studied under normal and stress conditions revealed that both NF-YB members are present in the cytoplasm under normal conditions, while under stress, AtNF-YB3 moved to the nucleus. In the case of AtNF-YB10, it could be transported to the nucleus after interaction with AtNF-YC2. Therefore, it can be hypothesized that in finger millet, NF-YB subcellular localization may be influenced by NF-YC availability or post-modification. The translocation mechanism of finger millet NF–Y protein could be further confirmed after expression analysis under stress conditions. This could have a direct impact on NF–Y transcriptional activity. Transcription factors regulate plant growth and developmental processes, which may be targeted by stress-responsive miRNAs. Using *in-silico* tools, attempts have been made to identify miRNA target sites and associated NF–Y gene regulation in finger millet. It has been demonstrated that miR169 and NF-YA gene regulatory networks influence the plant response to pathogens [11,68]. Different miR169 isoforms accumulated differentially in rice plants inoculated with a blast (*M. oryzae*) pathogen [75]. Interestingly, all the miR169 isoforms have similar sequences and target the common NF–Y gene. Their distinct responses to elicitor treatments suggest a function for miR169/NF-YA modules in rice defense against blast pathogens [75]. Further, interactions of miRNA and transcription factors such as NF-Ys could also reveal their relevance in root growth, branching, nutrient signaling [69], and abiotic stress tolerance [70]. In the present study, all the twenty-three miR169 isoforms target the same transcription factor EcNF-YA13 with cleavage activity, indicating EcNF-YA13 to be associated with several biotic and abiotic stress regulations. Furthermore, plants, in which NF-YA5 and miR169 were knocked down are hypersensitive to moisture stress, suggesting their crucial role in contributing to moisture stress tolerance. By reducing the stomata opening of tomato plants, miR169c overexpression enhanced tolerance toward drought stress [71]. Additionally, miR169 imparts resistance to blast fungus by suppressing the expression of the NF-YA gene in rice and could also be targeted for finger millet improvement against blast disease [11].

These findings clearly state that miRNAs play an array of regulatory functions, and their *in-silico* screening and validation for stresses are extremely important. The CREs identified in the promoter regions of NF–Y genes showed their functional role in various abiotic stresses (DRE, MYB, STRE, GT-1, SP1, CAAT Motif, G-Box, TCT motif, LTR, MYC and I-Box), hormonal signaling (ABRE, TCA-elements, AuxRR-core, TGA-box, GARE-motif, TGACG-motif and CGTCA-motif), plants developmental processes (A-box, Circadian, GATA) and biotic stress (W-Box, Tc-rich repeats) [8,16,72–76].

5. Conclusion

For the first time, the finger millet genome has been mined to decipher the diversity of NF–Y genes, which varies from crop to crop, provided there is availability of the whole genome sequences. Elucidating the diversity of NF–Y TFs in crops is important to reveal the possibility of functional redundancy. A total of 57EcNF-Y genes were analyzed, which vary from other crops. The diversity analysis of NF-Ys and prevalence of ABRE and MYB, CREs in the promoter region associated with drought tolerance could be one of the reasons for the finger millet to be a drought-tolerant crop. Further, several unique attributes of EcNF-Ys have been reported in the present study. The *in-silico* investigation provides a roadmap for wet-lab-based experimentation, and several attributes characterized by bioinformatics tools could be utilized for the genetic enhancement of finger millet. Genome-wide analysis revealed the presence of a total of 57 EcNF-Y genes comprising 18 EcNF-YA, 23 EcNF-YB, and 16 EcNF-YC genes in the finger millet genome. Further, it was observed that variability exists in terms of conserved amino acid residues in different subunits. EcNF-YA/B/C subfamily members showed 62 %, 88 %, and 14 % conserved amino acid residues, respectively. The phylogenetic studies showed the average evolution time for the paralogous genes of finger millet NF-Ys was 8.23 million years ago, and they possess highly conserved functional motif FVTSKASDKC within their domain. The assessment of CREs of finger millet NF–Y genes clearly shows the possibility of its significant role in biotic, abiotic stress and development. The predominance of the elements includes W-Box, STRE, MYC, MYB, DRE, ABRE, A-Box, Circadian, and GATA cis-elements in most of the EcNF-Ys genes. In addition, the interaction between twenty-three isoforms of mir169 and NF-YA13 contributes to developing strategies to protect finger millet from *M.oryzae* (a fungal blast pathogen). The EcNF-Ys need to be studied for real-time gene expression profiling after imposing stress to reveal the potential NF–Y genes, specific for different stresses, which might be further utilized for developing stress-tolerant crops. Therefore, the bioinformatics assessment of these NF-Ys of finger millets provides an opportunity for developing stress-tolerant crops by adopting an amalgamation of conventional breeding and multi-omics approaches.

Ethical approval

This article does not contain any studies with human participants or animals performed by any of the authors.

Data availability

The datasets are available in the NCBI database (Accession Number TPA: BK063099-BK063157) and provided as supplementary files.

CRediT authorship contribution statement

Varsha Rani: Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation. **Vinay Kumar Singh:** Visualization, Validation, Software. **D.C. Joshi:** Writing – review & editing, Supervision. **Rajesh Singh:**

Supervision. **Dinesh Yadav:** Writing – review & editing, Supervision, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e36370.

References

[1] H. Zhao, D. Wu, F. Kong, K. Lin, H. Zhang, G. Li, The Arabidopsis thaliana nuclear factor Y transcription factors, Front. Plant Sci. 7 (2017) 1–11, https://doi.org/10.3389/fpls.2016.02045.

[2] R. Mantovani, The Molecular Biology of the CCAAT-Binding Factor NF-Y, vol. 239, 1999, pp. 15–27.

[3] D. Hackenberg, Y. Wu, A. Voigt, R. Adams, P. Schramm, B. Grimm, Studies on differential nuclear translocation mechanism and assembly of the three subunits of the arabidopsis thaliana transcription factor NF-Y, Mol. Plant 5 (2012) 876–888, https://doi.org/10.1093/mp/ssr107.

[4] D.-H. Yan, X. Xia, W. Yin, NF-YB family genes identified in a poplar genome-wide analysis and expressed in populus euphratica are responsive to drought stress, Plant Mol. Biol. Rep. 31 (2013) 363–370, https://doi.org/10.1007/s11105-012-0508-5.

[5] J.-J. Xu, X.-F. Zhang, H.-W. Xue, Rice aleurone layer specific OsNF-YB1 regulates grain filling and endosperm development by interacting with an ERF transcription factor, J. Exp. Bot. 67 (2016) 6399–6411, https://doi.org/10.1093/jxb/erw409.

[6] S. Cao, R.W. Kumimoto, N. Gnesutta, A.M. Calogero, R. Mantovani, B.F. Holt III, A distal CCAAT/NUCLEAR FACTOR Y complex promotes Chromatin looping at the FLOWERING LOCUS T promoter and regulates the timing of flowering in arabidopsis, Plant Cell 26 (2014) 1009–1017, https://doi.org/10.1105/tpc.113.120352.

[7] X. yan Li, R. Mantovani, R. Hooft Van Huijsduijnen, I. Andre, C. Benoist, D. Mathis, Evolutionary variation of the CCAAT-binding transcription factor NF-Y, Nucleic Acids Res. 20 (1992) 1087–1091, https://doi.org/10.1093/nar/20.5.1087.

[8] K. Petroni, R.W. Kumimoto, N. Gnesutta, V. Calvenzani, M. Fornari, C. Tonelli, B.F. Holt, R. Mantovani, The promiscuous life of plant NUCLEAR FACTOR Y transcription factors, Plant Cell 24 (2013) 4777–4792, https://doi.org/10.1105/tpc.112.105734.

[9] L. Chen, Y. Zhou, W. Lai, L. Hu, L. Jiang, S. Liu, In silico identification and expression analysis of nuclear factor Y (NF-Y) transcription factors in cucumber, Agronomy 10 (2020), https://doi.org/10.3390/agronomy10020236.

[10] P.B. Kavi, K. Showkat, A. Ganie, S.H. Wani, R. Guddimalli, Nuclear factor - Y (NF - Y): developmental and stress - responsive roles in the plant lineage, J. Plant Growth Regul. (2022), https://doi.org/10.1007/s00344-022-10739-6.

[11] V. Rani, D.C. Joshi, P. Joshi, R. Singh, D. Yadav, "Millet Models" for harnessing nuclear factor-Y transcription factors to engineer stress tolerance in plants: current knowledge and emerging paradigms, Planta (2023) 258, https://doi.org/10.1007/s00425-023-04186-0.

[12] D.E. Nelson, P.P. Repetti, T.R. Adams, R.A. Creelman, J. Wu, D.C. Warner, D.C. Anstrom, R.J. Bensen, P.P. Castiglioni, M.G. Donnarummo, B.S. Hinchey, R. W. Kumimoto, D.R. Maszle, R.D. Canales, K.A. Krolikowski, S.B. Dotson, N. Gutterson, O.J. Ratcliffe, J.E. Heard, Plant nuclear factor Y (NF-Y) B subunits confer drought tolerance and lead to improved corn yields on water-limited acres, Proc. Natl. Acad. Sci. U.S.A. 104 (2007) 16450–16455, https://doi.org/10.1073/pnas.0707193104.

[13] D. Lee, H. Il, G. Jang, P. Joong, J. Seo, Y. Shic, S. Woon, H. Jung, Y. Do, J. Kim, Plant Science the NF-YA transcription factor OsNF-YA7 confers drought stress tolerance of rice in an abscisic acid independent manner, Plant Sci. 241 (2015) 199–210, https://doi.org/10.1016/j.plantsci.2015.10.006.

[14] H. Su, Y. Cao, L. Ku, W. Yao, Y. Cao, Z. Ren, D. Dou, H. Wang, Z. Ren, H. Liu, L. Tian, Y. Zheng, C. Chen, Y. Chen, Dual functions of ZmNF-YA3 in photoperiod-dependent flowering and abiotic stress responses in maize, J. Exp. Bot. 69 (2018) 5177–5189, https://doi.org/10.1093/jxb/ery299.

[15] M. Yang, Y. Zhao, S. Shi, X. Du, J. Gu, K. Xiao, Wheat nuclear factor Y (NF-Y) B subfamily gene TaNF-YB3;l confers critical drought tolerance through modulation of the ABA-associated signaling pathway, Plant Cell Tissue Organ Cult. 128 (2017) 97–111, https://doi.org/10.1007/s11240-016-1088-0.

[16] P. Maheshwari, D. Kummari, S.R. Palakolanu, U. Nagasai Tejaswi, M. Nagaraju, G. Rajasheker, G. Jawahar, N. Jalaja, P. Rathnagiri, P.B. Kavi Kishor, Genome-wide identification and expression profile analysis of nuclear factor Y family genes in Sorghum bicolor L. (Moench), PLoS One 14 (2019) 1–27, https://doi.org/10.1371/journal.pone.0222203.

[17] N. Malviya, P. Jaiswal, D. Yadav, Genome- wide characterization of Nuclear Factor Y (NF-Y) gene family of sorghum [Sorghum bicolor (L.) Moench]: a bioinformatics approach, Physiol. Mol. Biol. Plants 22 (2016) 33–49, https://doi.org/10.1007/s12298-016-0349-z.

[18] X. Ma, C. Li, M. Wang, Wheat NF-YA10 functions independently in salinity and drought stress, Bioengineered 6 (2015) 245–247, https://doi.org/10.1080/21655979.2015.1054085.

[19] M. Ramakrishnan, S.A. Ceasar, K.K. Vinod, V. Duraipandiyan, T.P. Ajeesh Krishna, H.D. Upadhyaya, N.A. Al-Dhabi, S. Ignacimuthu, Identification of putative QTLs for seedling stage phosphorus starvation response in finger millet (Eleusine coracana L. Gaertn.) by association mapping and cross species synteny analysis, PLoS One 12 (2017) e0183261, https://doi.org/10.1371/journal.pone.0183261.

[20] T. Maharajan, S.A. Ceasar, T.P. Ajeesh Krishna, Finger millet (Eleusine coracana (L.) Gaertn): nutritional importance and nutrient transporters, CRC Crit. Rev. Plant Sci. 41 (2022) 1–31, https://doi.org/10.1080/07352689.2022.2037834.

[21] T.L. Goron, M.N. Raizada, in: Genetic Diversity and Genomic Resources Available for the Small Millet Crops to Accelerate a New Green Revolution, vol. 6, 2015, https://doi.org/10.3389/fpls.2015.00157.

[22] B. Shashi, S. Sharan, S. Hittalamani, A.G. Shankar, T. Nagarathna, Micronutrient composition, antinutritional factors and bioaccessibility of iron in different finger millet (Eleusine coracana) genotypes, Karnataka J. Agric. Sci. 20 (2007) 583–585.

[23] S. Hittalmani, H.B. Mahesh, M.D. Shirke, H. Biradar, G. Uday, Y.R. Aruna, H.C. Lohithaswa, A. Mohanrao, Genome and Transcriptome sequence of Finger millet (Eleusine coracana (L.) Gaertn.) provides insights into drought tolerance and nutraceutical properties, BMC Genom. 18 (2017) 1–16, https://doi.org/10.1186/s12864-017-3850-z.

[24] A. Kumar, M. Metwal, S. Kaur, A.K. Gupta, S. Puranik, S. Singh, M. Singh, S. Gupta, B.K. Babu, S. Sood, R. Yadav, Nutraceutical value of finger millet [Eleusine coracana (L.) Gaertn.], and their improvement using omics approaches, Front. Plant Sci. 7 (2016) 1–14, https://doi.org/10.3389/fpls.2016.00934.

[25] T. Maharajan, S.A. Ceasar, T. Palayullaparambil, A. Krishna, S. Ignacimuthu, in: Finger Millet [Eleusine Coracana (L .) Gaertn]: an Orphan Crop with a Potential to Alleviate the Calcium Deficiency in the Semi-arid Tropics of Asia and Africa, vol. 5, 2021, pp. 1–8, https://doi.org/10.3389/fsufs.2021.684447.

[26] A. Ceasar, Genome-editing in millets: current knowledge and future perspectives, Mol. Biol. Rep. 49 (2022) 773–781, https://doi.org/10.1007/s11033-021-06975-w.

[27] S. Antony Ceasar, T. Maharajan, T.P. Ajeesh Krishna, M. Ramakrishnan, G. Victor Roch, L. Satish, S. Ignacimuthu, Finger millet [Eleusine coracana (L.) Gaertn.] improvement: current status and future interventions of whole genome sequence, Front. Plant Sci. 9 (2018), https://doi.org/10.3389/fpls.2018.01054.

[28] S. Sood, D.C. Joshi, A.K. Chandra, A. Kumar, Phenomics and genomics of finger millet: current status and future prospects, Planta 250 (2019) 731–751, https://doi.org/10.1007/s00425-019-03159-6.

[29] M. Hatakeyama, S. Aluri, M.T. Balachadran, S.R. Sivarajan, A. Patrignani, S. Grüter, L. Poveda, R. Shimizu-Inatsugi, J. Baeten, K.-J. Francoijs, K.N. Nataraja, Y. A.N. Reddy, S. Phadnis, R.L. Ravikumar, R. Schlapbach, S.M. Sreeman, K.K. Shimizu, Multiple hybrid de novo genome assembly of finger millet, an orphan allotetraploid crop, DNA Res. an Int. J. Rapid Publ. Reports Genes Genomes. 25 (2018) 39–47, https://doi.org/10.1093/dnares/dsx036.

[30] S. Gupta, R.K. Pathak, S.M. Gupta, V.S. Gaur, N.K. Singh, A. Kumar, Identification and molecular characterization of Dof transcription factor gene family preferentially expressed in developing spikes of Eleusine coracana L, 3 Biotech. 8 (2018) 1–13, https://doi.org/10.1007/s13205-017-1068-z.

[31] H. Rahman, V. Ramanathan, J. Nallathambi, S. Duraialagaraja, M. Raveendran, Over-expression of a NAC 67 transcription factor from finger millet (Eleusine coracana L.) confers tolerance against salinity and drought stress in rice, BMC Biotechnol. 16 (2016), https://doi.org/10.1186/s12896-016-0261-1.

[32] P. Jadhav, Expression of ECMYB transcription factor gene under different abiotic stress conditions in Eleusine coracana, Int. J. Agric. Environ. Biotechnol. 11 (2018) 30954, https://doi.org/10.30954/0974-1712.10.2018.12.

[33] D.M. Goodstein, S. Shu, R. Howson, R. Neupane, R.D. Hayes, J. Fazo, T. Mitros, W. Dirks, U. Hellsten, N. Putnam, D.S. Rokhsar, Phytozome: a comparative platform for green plant genomics, Nucleic Acids Res. 40 (2012) D1178–D1186, https://doi.org/10.1093/nar/gkr944.

[34] F. Tian, D.C. Yang, Y.Q. Meng, J. Jin, G. Gao, PlantRegMap: Charting functional regulatory maps in plants, Nucleic Acids Res. 48 (2020) D1104–D1113, https://doi.org/10.1093/nar/gkz1020.

[35] S.F. Altschul, T.L. Madden, A.A. Schäffer, J. Zhang, Z. Zhang, W. Miller, D.J. Lipman, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, Nucleic Acids Res. 25 (1997) 3389–3402, https://doi.org/10.1093/nar/25.17.3389.

[36] J.C. Oliveros, Venny 2.1. 0, Venny. An interact, Tool Comp. List. with Venn's Diagrams 2015 (2007).

[37] M. Blum, H.Y. Chang, S. Chuguransky, T. Grego, S. Kandasaamy, A. Mitchell, G. Nuka, T. Paysan-Lafosse, M. Qureshi, S. Raj, L. Richardson, G.A. Salazar, L. Williams, P. Bork, A. Bridge, J. Gough, D.H. Haft, I. Letunic, A. Marchler-Bauer, H. Mi, D.A. Natale, M. Necci, C.A. Orengo, A.P. Pandurangan, C. Rivoire, C.J. A. Sigrist, I. Sillitoe, N. Thanki, P.D. Thomas, S.C.E. Tosatto, C.H. Wu, A. Bateman, R.D. Finn, The InterPro protein families and domains database: 20 years on, Nucleic Acids Res. 49 (2020) D344–D354, https://doi.org/10.1093/nar/gkaa977.

[38] C.J.A. Sigrist, E. De Castro, L. Cerutti, B.A. Cuche, N. Hulo, A. Bridge, L. Bougueleret, I. Xenarios, New and continuing developments at PROSITE, Nucleic Acids Res. 41 (2013) 344–347, https://doi.org/10.1093/nar/gks1067.

[39] T.L. Bailey, J. Johnson, C.E. Grant, W.S. Noble, The MEME suite, Nucleic Acids Res. 43 (2015) W39–W49, https://doi.org/10.1093/nar/gkv416.

[40] K.M. Devos, P. Qi, B.A. Bahri, D.M. Gimode, K. Jenike, S.J. Manthi, D. Lule, T. Lux, L. Martinez-bello, T.H.P. Iv, C. Plott, D. Saha, G.S. Sidhu, A. Sreedasyam, E. E. Mneney, H.F. Ojulong, M.C. Schatz, Genome analyses reveal population structure and a purple stigma color gene candidate in fi nger millet. https://doi.org/10.1038/s41467-023-38915-6, 2023.

[41] B. Hu, J. Jin, A.Y. Guo, H. Zhang, J. Luo, G. Gao, Gsds 2.0: an upgraded gene feature visualization server, Bioinformatics 31 (2015) 1296–1297, https://doi.org/10.1093/bioinformatics/btu817.

[42] E. Gasteiger, C. Hoogland, A. Gattiker, S. Duvaud, M.R. Wilkins, R.D. Appel, A. Bairoch, in: J.M. Walker (Ed.), Protein Identification and Analysis Tools on the ExPASy Server BT - the Proteomics Protocols Handbook, Humana Press, Totowa, NJ, 2005, pp. 571–607, https://doi.org/10.1385/1-59259-890-0:571.

[43] J. Hallgren, K.D. Tsirigos, M. Damgaard Pedersen, J. Juan, A. Armenteros, P. Marcatili, H. Nielsen, A. Krogh, O. Winther, in: DeepTMHMM Predicts Alpha and Beta Transmembrane Proteins Using Deep Neural Networks, BioRxiv., 2022, 2022.04.08.487609, https://www.biorxiv.org/content/10.1101/2022.04.08.487609v1%0A.

[44] P. Horton, K.-J. Park, T. Obayashi, N. Fujita, H. Harada, C.J. Adams-Collier, K. Nakai, WoLF PSORT: protein localization predictor, Nucleic Acids Res. 35 (2007) W585–W587, https://doi.org/10.1093/nar/gkm259.

[45] R. Nair, P. Carter, B. Rost, NLSdb: database of nuclear localization signals, Nucleic Acids Res. 31 (2003) 397–399, https://doi.org/10.1093/nar/gkg001.

[46] M.A. Larkin, G. Blackshields, N.P. Brown, R. Chenna, P.A. Mcgettigan, H. McWilliam, F. Valentin, I.M. Wallace, A. Wilm, R. Lopez, J.D. Thompson, T.J. Gibson, D.G. Higgins, Clustal W and Clustal X version 2.0, Bioinformatics 23 (2007) 2947–2948, https://doi.org/10.1093/bioinformatics/btm404.

[47] K. Tamura, G. Stecher, S. Kumar, MEGA11: molecular evolutionary genetics analysis version 11, Mol. Biol. Evol. 38 (2021) 3022–3027, https://doi.org/10.1093/molbev/msab120.

[48] Z. Zhang, J. Li, X.-Q. Zhao, J. Wang, G.K.-S. Wong, J. Yu, KaKs_Calculator: calculating Ka and Ks through model selection and model averaging, Dev. Reprod. Biol. 4 (2006) 259–263, https://doi.org/10.1016/S1672-0229(07)60007-2.

[49] B.S. Gaut, B.R. Morton, B.C. McCaig, M.T. Clegg, Substitution rate comparisons between grasses and palms: synonymous rate differences at the nuclear gene Adh parallel rate differences at the plastid gene rbcL, Proc. Natl. Acad. Sci. U.S.A. 93 (1996) 10274–10279, https://doi.org/10.1073/pnas.93.19.10274.

[50] M. Lescot, P. Déhais, G. Thijs, K. Marchal, Y. Moreau, Y. Van de Peer, P. Rouzé, S. Rombauts, PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences, Nucleic Acids Res. 30 (2002) 325–327, https://doi.org/10.1093/nar/30.1.325.

[51] X. Dai, P.X. Zhao, psRNATarget: a plant small RNA target analysis server, Nucleic Acids Res. 39 (2011) W155–W159, https://doi.org/10.1093/nar/gkr319.

[52] A. Waterhouse, M. Bertoni, S. Bienert, G. Studer, G. Tauriello, R. Gumienny, F.T. Heer, T.A.P. de Beer, C. Rempfer, L. Bordoli, R. Lepore, T. Schwede, SWISS-MODEL: homology modelling of protein structures and complexes, Nucleic Acids Res. 46 (2018) W296–W303, https://doi.org/10.1093/nar/gky427.

[53] C. Colovos, T. Yeates, Verification of Protein Structures : Patterns of Nonbonded Atomic Interactions, 1993, pp. 1511–1519.

[54] R.A. Laskowski, M.W. MacArthur, D.S. Moss, J.M. Thornton, PROCHECK: a program to check the stereochemical quality of protein structures, J. Appl. Crystallogr. 26 (1993) 283–291.

[55] G. Studer, C. Rempfer, A.M. Waterhouse, R. Gumienny, J. Haas, T. Schwede, QMEANDisCo—distance constraints applied on model quality estimation, Bioinformatics 36 (2020) 1765–1771, https://doi.org/10.1093/bioinformatics/btz828.

[56] R.A. Laskowski, J. Jabłońska, L. Pravda, R.S. Vařeková, J.M. Thornton, PDBsum: structural summaries of PDB entries, Protein Sci. 27 (2018) 129–134, https://doi.org/10.1002/pro.3289.

[57] M. Zhang, X. Hu, M. Zhu, M. Xu, L. Wang, Transcription factors NF-YA2 and NF-YA10 regulate leaf growth via auxin signaling in Arabidopsis, Sci. Rep. 7 (2017) 1395, https://doi.org/10.1038/s41598-017-01475-z.

[58] S.-K. Kim, H.-Y. Park, Y.H. Jang, K.C. Lee, Y.S. Chung, J.H. Lee, J.-K. Kim, OsNF-YC2 and OsNF-YC4 proteins inhibit flowering under long-day conditions in rice, Planta 243 (2016) 563–576, https://doi.org/10.1007/s00425-015-2426-x.

[59] C.C.W. Oa, N. Siefers, K.K. Dang, R.W. Kumimoto, W. Edwards, B. Iv, G. Tayrose, B.F.H. Iii, N. Carolina, C. Hill, N.C.N. S, Tissue-Specific Expression Patterns of Arabidopsis NF-Y Transcription Factors Suggest Potential for Extensive, vol. 149, 2009, pp. 625–641, https://doi.org/10.1104/pp.108.130591.

[60] M. Chen, Z. Feng, G. He, W. Zheng, P. Lu, M. Chen, Foxtail millet NF-Y families : genome-wide survey and evolution analyses foxtail millet NF-Y families : genome-wide survey and evolution analyses identified two functional genes important in abiotic stresses. https://doi.org/10.3389/fpls.2015.01142, 2015.

[61] Z.J. Feng, G.H. He, W.J. Zheng, P.P. Lu, M. Chen, Y.M. Gong, Y.Z. Ma, Z.S. Xu, Foxtail millet NF-Y families: genome-wide survey and evolution analyses identified two functional genes important in abiotic stresses, Front. Plant Sci. 6 (2015) 1–19, https://doi.org/10.3389/fpls.2015.01142.

[62] D. Yadav, Y. Shavrukov, N. Bazanova, L. Chirkova, N. Borisjuk, N. Kovalchuk, A. Ismagul, B. Parent, P. Langridge, M. Hrmova, S. Lopato, Constitutive Overexpression of the TaNF-YB4 Gene in Transgenic Wheat Significantly Improves Grain Yield, vol. 66, 2015, pp. 6635–6650, https://doi.org/10.1093/jxb/erv370.

[63] T. Thirumurugan, Y. Ito, Identi W cation, characterization and interaction of HAP family genes in rice (2008) 279–289, https://doi.org/10.1007/s00438-007-0312-3.

[64] C. Romier, F. Cocchiarella, R. Mantovani, D. Moras, The NF-YB/NF-YC structure gives insight into DNA binding and transcription regulation by CCAAT factor NF-Y, J. Biol. Chem. 278 (2003) 1336–1345, https://doi.org/10.1074/jbc.M209635200.

[65] S. Quan, J. Niu, L. Zhou, H. Xu, L. Ma, Y. Qin, Identification and Characterization of NF-Y Gene Family in Walnut (Juglans Regia L .), 2018, pp. 1–17.

[66] Z. Liu, Y. Li, J. Zhu, W. Ma, Z. Li, Z. Bi, C. Sun, J. Bai, J. Zhang, Y. Liu, Genome-wide identification and analysis of the NF-Y gene family in potato (Solanum tuberosum L.), Front. Genet. 12 (2021) 1–18, https://doi.org/10.3389/fgene.2021.739989.

[67] D. Xu, A. Farmer, G. Collett, N. V Grishin, Y.M. Chook, Sequence and structural analyses of nuclear export signals in the NESdb database, Mol. Biol. Cell 23 (2012) 3677–3693, https://doi.org/10.1091/mbc.E12-01-0046.
[68] Y. Li, S.L. Zhao, J.L. Li, X.H. Hu, H. Wang, X.L. Cao, Y.J. Xu, Z.X. Zhao, Z.Y. Xiao, N. Yang, J. Fan, F. Huang, W.M. Wang, Osa-miR169 negatively regulates rice immunity against the blast fungus magnaporthe oryzae, Front. Plant Sci. 8 (2017) 1–13, https://doi.org/10.3389/fpls.2017.00002.
[69] C. Sorin, M. Declerck, A. Christ, T. Blein, L. Ma, C. Lelandais-Brière, M.F. Njo, T. Beeckman, M. Crespi, C. Hartmann, A miR169 isoform regulates specific NF-YA targets and root architecture in Arabidopsis, New Phytol. 202 (2014) 1197–1211, https://doi.org/10.1111/nph.12735.
[70] L. Xing, M. Zhu, M. Luan, M. Zhang, L. Jin, Y. Liu, J. Zou, L. Wang, M. Xu, miR169q and NUCLEAR FACTOR YA8 enhance salt tolerance by activating PEROXIDASE1 expression in response to ROS, Plant Physiol. 188 (2022) 608–623, https://doi.org/10.1093/plphys/kiab498.
[71] X. Zhang, Z. Zou, P. Gong, J. Zhang, K. Ziaf, H. Li, F. Xiao, Z. Ye, Over-expression of microRNA169 confers enhanced drought tolerance to tomato, Biotechnol. Lett. 33 (2011) 403–409, https://doi.org/10.1007/s10529-010-0436-0.
[72] S. Basu, A. Roychoudhury, D.N. Sengupta, Deciphering the role of various cis-acting regulatory elements in controlling samDC gene expression in rice, Plant Signal. Behav. 9 (2014) 1–5, https://doi.org/10.4161/psb.28391.
[73] J.M. Franco-Zorrilla, I. López-Vidriero, J.L. Carrasco, M. Godoy, P. Vera, R. Solano, DNA-binding specificities of plant transcription factors and their potential to define target genes, Proc. Natl. Acad. Sci. U.S.A. 111 (2014) 2367–2372, https://doi.org/10.1073/pnas.1316278111.
[74] S. Wenkel, F. Turck, K. Singer, L. Gissot, J. Le Gourrierec, A. Samach, G. Coupland, CONSTANS and the CCAAT box binding complex share a functionally important domain and interact to regulate flowering of Arabidopsis, Plant Cell 18 (2006) 2971–2984, https://doi.org/10.1105/tpc.106.043299.
[75] T. Yao, J. Zhang, M. Xie, G. Yuan, T.J. Tschaplinski, W. Muchero, J.G. Chen, Transcriptional regulation of drought response in arabidopsis and Woody plants, Front. Plant Sci. 11 (2021) 1–12, https://doi.org/10.3389/fpls.2020.572137.
[76] A. Kaur, P.K. Pati, A.M. Pati, A.K. Nagpal, In-silico analysis of cis-acting regulatory elements of pathogenesis-related proteins of Arabidopsis thaliana and Oryza sativa, PLoS One 12 (2017), https://doi.org/10.1371/journal.pone.0184523.