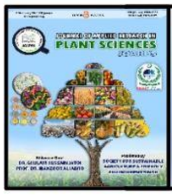
J. Appl. Res Plant Sci. Vol. 5(1), 63-78, 2024, https://doi.org/10.38211/joarps.2024.05.179

Research Article

OPEN ACCESS



Available on https://www.joarps.org
Journal of Applied Research in Plant Sciences
(JOARPS)
ISSN: 2708-3004 (Online), 2708-2997 (Print)

Genome-wide Analysis of Plant Specific *YABBY* Transcription Factor Gene Family in Watermelon (*Citrullus lanatus*) and Arabidopsis

Mehr-ul-Nisa[1], Muhammad Shafiq[1], Mujahid Manzoor*[2], Muhammad Bilal[3], Tariq Manzoor[6], Muhammad Ali[2], Malik Muazzam Anees[1], Muhammad Rizwan[4], Muhammad Zeeshan Haider[5], Adnan Sami[5], Muhammad Saleem Haider[6]

[1]Department of Horticulture, Faculty of Agriculture sciences, University of the Punjab, Lahore, Pakistan
[2]Department of Entomology, Faculty of Agriculture sciences, University of the Punjab, Lahore, Pakistan
[3]Department of Agronomy, Faculty of Agriculture sciences, University of the Punjab, Lahore, Pakistan
[4]Department of Food Science and Technology, Faculty of Agriculture sciences, University of the Punjab, Lahore, Pakistan
[5]Department of Plant Breeding and Genetics, Faculty of Agricultural Sciences, University of the Punjab, P.O BOX. 54590, Lahore, Pakistan
[6]Department of Plant Pathology, Faculty of Agricultural Sciences, University of the Punjab, P.O BOX. 54590, Lahore, Pakistan
*Corresponding Authors Email; mujahid.iags@pu.edu.pk

Article Received 10-03-2023, Article Revised 10-10-2023, Article Accepted 21-10-2023.

Abstract

The *YABBY* gene family is a specific transcription factor for plants and a DNA binding domain that carries out several different functions, such as regulating the length of blooming plant styles and the polarity of lateral organ development. The *YABBY* gene family members were identified in the watermelon (*Citrullus lanatu*s *subsp. vulgaris var*. 97103 V1) genome using a set of bioinformatics techniques. Protein motifs, protein architectures, protein sequences, miRNA targets, and tissue-specific expression patterns were all examined. All chromosomes had an uneven distribution of about eight putative *YABBY* genes. Inner No Outer *INO*, *CRC* (Crabs Claw), *YAB2*, *YAB3/AFO*, and *YAB5* were the five subgroups that the *YABBY* proteins in watermelon fall within, in accordance with the accepted Arabidopsis categorization which is based on International Standards of Nomenclature. Segmental duplication was more frequent than tandem duplication, and it was predominantly responsible for the growth of the *YABBY* gene family in watermelon. The results of tissue-specific expression profiling of *ClYAABY* genes showed that the vast majority of these genes were substantially expressed in roots and seedlings. In this study, cis-regulatory element (CRE) analyses were employed to identify elements in seedlings and roots that are highly responsive to light,wound, drought, auxin, stress, salicylic acid, and abscisic acid (ABA). The findings reveal specific CREs within the promoter regions of genes associated with these responses. Five groups or sub-families have also been identified by comparing the *YABBY* genes in watermelon and Arabidopsis, however the *CRC* and *YAB2* groups do not share gene pairing among the other groups. This research contributes to a deeper understanding of plant adaptability and stress response mechanisms, with implications for agriculture and plant science.

Keywords: Watermelon, Gene family, Genomic analysis, Specific plant transcription factors, *YABBY*.

Introduction

Being a sessile organism (which cannot move) by nature, plants are constantly subjected to a variety of external stresses, which impair the biochemical and functional processes occurring within them (Sharif *et al.*, 2020). Several plant transcription factors (Tfs) cause the corresponding gene expression in response to such biotic and abiotic stresses (Li *et al.*, 2019; Zhao *et al.*, 2017). A tiny gene family known as the *YABBY* TFs is unique to seed plants (Yin *et al.*, 2022), typically classified as a subfamily of the zinc-finger super family (Y.-Y. Chen *et al.*, 2020). A collection of TFs with two conserved domains were encoded by the *YABBY* protein. The N-terminal (Cys2 Cys2) zinc-finger domain and the C-terminal (helix-loop-helix) domain are the two conserved domains (Song, Joshi, DiPiazza, & Joshi, 2020) . Both domains retain highly sealed amino acid (AA) residues and participate in the particular binding of DNA (Sawa *et al.*, 1999). The growth and development of lateral organs (Bowman, Smyth, & Meyerowitz, 1989), establishment of adaxial-abaxial polarity (Kumaran, Bowman, & Sundaresan, 2002), expansion of leaves (Eckardt, 2010), development of leaf edges (Finet *et al*., 2016), and the response to stress are significantly influenced by *YABBY* TFs (Zhao *et al*., 2017). The *YABBY* family was divided into five subfamilies by phylogenetic analysis: the CRABS CLAW (CRC), FILAMENTOUS FLOWER (FIL)/*YABBY*3 (YAB3), INNER NO OUTER (INO), *YABBY*2 (YAB2), and *YABBY*5 (YAB5) (Bowman, 2000; Yamada, Ito, & Kato, 2004). Moreover, CRC helps to create the polarity of developing carpels and nectarines, while FIL, YAB3, YAB2, and YAB5 excessively encourage the development of lateral organs (Siegfried *et al*., 1999). The remaining INO has a variety of functions that help the ovule's outer integument, a layer of cells that surrounds the nucleus, grow into the seed coat (Villanueva *et al*., 1999). Based on their crucial developmental roles, *YABBY* genes have undergone extensive genomic analysis in plants since it's crucial to understand how they relate to one another among species or even within a single species. 6 *YABBY* genes have been found in *Arabidopsis thaliana*, 9 in Pineapple (Li *et al*., 2019), 7 in Grapevine (Zhang *et al*., 2019), 11 *CmoYABBYs* in *Cucurbita moschata*, 12 *CmaYABBYs* in *Cucurbita maxima*, 11 *CpeYABBYs* in *Cucurbita pepo* (Yuan *et al*., 2020), 12 in *Brassica rapa ssp. Chinensis* (Hou, Wu, Gao, Zhang, & Hou, 2019), 21 in *Triticum aestivum* (wheat) (Buttar *et al*., 2020), 9 in *Tuberosum lycopersicum* (Tomato) (Huang, Van Houten, Gonzalez, Xiao, & van der Knaap, 2013), 8 in *Oriza sativa* (Rice) (Toriba *et al*., 2007), 12 in *Gossypium arboreum*, 12 *G. raimondii* and 23 in *G. hirsutum* (Upland Cotton) (Zhaoen Yang *et al*., 2018). The Cucurbitaceae family includes the watermelon, *Citrullus lanatus subsp. vulgaris var. 97103 V1*. It originates from tropical regions of Africa close to the Kalahari Desert (Kyriacou, Leskovar, Colla, & Rouphael, 2018; Naz, Butt, Sultan, Qayyum, & Niaz, 2014). Botanists typically refer to it as a "pepo," a fruit with a thick skin and fleshy core (Mehra, Pasricha, & Gupta, 2015). Because of its refreshing texture, appealing colour, delicate and sweet flavor, and high water content to alleviate summer thirst, it is typically consumed as a pleasant summer fruit that is well cherished by customers (Romdhane *et al*., 2017). Watermelon fruits produce 55.3% juice, 31.5% rind, and 10.4% pomace (Maoto, Beswa, & Jideani, 2019; Oberoi & Sogi, 2017). Using various bioinformatics tools, the primary goal of this study was to locate and characterize the genes belonging to the *YABBY* TFs family in the Watermelon genome. To put it briefly, *YABBY* genes from the watermelon genome were identified using methodical way i.e. Phylogenetic Analysis, Synteny Analysis and Conserved Motif Analysis etc.Investigations were also conducted into their chromosomal distribution, intron/exon distribution pattern, presence of conserved domains, and cis-regulatory elements. The wide-ranging genome assessment of *YABBY* genes presents an orientation for cloning and functioning properties in watermelon. Investigations revealed genome-wide distribution and the role of *YABBY* genes in watermelon with the best-fit comparison to Arabidopsis thaliana.

Materials and methods

Database search and sequence retrieval: By using ID PF04690 (http://pfam.xfam.org/) the AA sequence for the *YABBY* PSTrFs (Plant specific Transcription factors) was obtained from the peptide genome of *Arabidopsis thaliana* (Accession No. A0A1P8APE2) and contains 164 amino acids (Mistry *et al*., 2021). The Cucurbit Genomics Database (CuGenDB) (http://cucurbitgenomics.org/) was used to Blast-p (Protein against Protein search tool) this sequence against the watermelon genome. The obtained amino acid sequences were investigated using the default parameters of the NCBI CDD (Conserved Domain Database) (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) (Lu *et al*., 2020) and the simple modular architecture research tool (SMART) available at (http:// smart.embl-heidelberg.de)(Letunic & Bork, 2018). Any identified proteins that did not include the *YABBY* conserved domain were ruled out.

Investigation of physio-chemical characteristics of *ClYABBY* proteins: Through using ProtParam tool (http://web.expasy.org/protparam/), the length (AA residues), molecular weight (MW), and theoretical pI of *ClYABBY* proteins were estimated (Gasteiger *et al*., 2005). The (CuGenDB) was utilized to obtain information on gene IDs, chromosomal locations, and protein and gene sequences(http://cucurbitgenomics.org/). According to the order of their physical placements, these *ClYABBY* genes were given new names such as Cl initials of Scientific Name of Watermelon and Yabby As an transcription factor WoLF PSORT (https://wolfpsort.hgc.jp/), an online tool, was used to estimate *ClYABBY's* subcellular localization (Horton, Park, Obayashi, & Nakai, 2006).

Gene structure analysis: In order to investigate the intron/exon organization of Cl*YABBY*, the genomic and coding sequences of identified genes were collected from the Cucurbit Genomics Database (CuGenDB)(http://cucurbitgenomics.org). Additionally, the watermelon genome's gf3 file was acquired from the Cucurbit Genomics Database (CuGenDB). With the use of the Gene Structure Display Server (GSDS v2.0) (Hu *et al*., 2015) (located at http://gsds.cbi.pku.edu.cn/), these sequences were then used to depict the gene structure.

Multiple sequence alignment and phylogenetic analysis: The *YABBY* protein's amino acid sequences were aligned with Clustal W version 2.1 (Thompson, Higgins, & Gibson, 1994), and the phylogeny was built using MEGA X v2.0 (Kumar, Stecher, Li, Knyaz, & Tamura, 2018) with neighbour-joining (NJ) and bootstrapping set at 1000 replications with partial deletion. Eight watermelons, thirteen musk melons, seven potatoes, and six *Arabidopsis YABBY* protein sequences were employed for the phylogenetic study.

CRE and conserved motif analysis: For each putative *ClYABBY* gene, a sequence 1000 bp upstream of the initiation codon was obtained for the investigation of the promoter regions. Following this, CREs in these sequences were predicted using the PlantCare database (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) (Rombauts, Déhais, Van Montagu, & Rouzé, 1999), and their validity was checked against the PLACE databases (http://www.dna.afrc.go.jp/PLACE/) (Higo, Ugawa, Iwamoto, & Higo, 1998) (Higo, Ugawa, Iwamoto, & Korenaga, 1999). The predicted protein sequences of the *ClYABBY* were used to assess motifs using Multiple EM for Motif Elicitation (MEME) (http://meme.nbcr.net/meme/) with a maximum number of motifs set at 20 (Bailey, Johnson, Grant, & Noble, 2015). Along with other variables, the default values for motif widths were set to 6 and 50, respectively.

Gene duplication and synteny analysis of Watermelon (Cucurbita Family): Using Ks and Ka values, the watermelon *YABBY* gene family's timing of divergence was calculated. TB Tbtools was used to determine the Ka and Ks substitution rates and computed Ka/Ks ratios according to the instructions in the software package manuals. For each pair of paralogous genes, the rates of molecular evolution were determined using the Ka/Ks ratios. T=Ks/2 was used to calculate the time of divergence (T), where λ is equal to $1.5 \times 10{-}8$ (P. Wang *et al.*, 2019). With the default settings, the Multiple Collinearity Scan toolkit (MCScanX) was employed to investigate the gene duplication occurrences (Y. Wang, Li, & Paterson, 2013). The syntenic analysis map was created using Tbtools' Micro Synteny view software to show the synteny association of the paralogous *YABBY* genes isolated from the watermelon (C. Chen *et al.*, 2020). By Using Tb Tools dual synetnic comparison between watermelon and melon, watermelon and musa, watermelon and Arabidopsis, watermelon and bottle gourd was found out.

Transcriptome analysis: We gathered previously generated RNA sequence data using high through put sequencing for watermelon plant tissues, including leaves and stems at seedling stage under semi-controlled environment and open field, to examine the organ-specific expression profile of *ClYABBY* at various development stages (Song, Joshi, DiPiazza, & Joshi, 2020). The Reads Per Kilobases per Million mapped reads (RPKM) values from RNA-seq data were log2 converted for expression profiling. Heatmap Illustrator in TBtools displays expression patterns with hierarchical clustering (Chen *et al.*, 2020).

Putative microRNA and target site analysis: The CDS sequences of all watermelon *ClYABBY* genes were searched for sequences complementary to miRNAs using psRNATarget (https://plantgrn.noble.org/psRNATarget/analysis?function=3/) (Samad, 2017) using default parameters in order to find miRNAs that potentially target the watermelon Cl*YABBY* genes. Later, the NCBI website was searched for these putative microRNAs' functions, and references were noted.

Results

Identification of the *YABBY* genes in watermelon: The sequence of the *YABBY* domain was BLAST searched against the watermelon's entire genome sequence that was received from the (CuGenDB) in order to discover the *YABBY* genes. There were 27 Proteins found during an initial study. Proteins with a truncated *YABBY* DNA-binding domain and those produced by the same gene isoform were excluded from the study. Eight distinct *ClYABBY* genes were discovered and further investigated. The highly conserved *YABBY* domain was present in these non-redundant *YABBY* protein sequences from watermelons. Eight of the 27 amino acids reported to be present in the highly conserved watermelon *YABBY* domain sequences were found to be 100% conserved in all *YABBY* domain sequences (Fig. S1), while the remaining 21 amino acids were found to be varied in all *ClYABBY* proteins. The *ClYABBY* genes code proteins with a MW range of 18.74 to 26.91 kDa and a length of 169 to 242 AA, with *ClYABBY8* being the shortest and *ClYABBY1* is the longest protein (Table 1). The discovered proteins have Ip ranging from 7.70 to 9.14.

Gene structure and recognition of conserved motifs and domains: The structure of exons and introns serves as the framework for genes and facilitates the study of evolutionary links among genes or species (Koralewski & Krutovsky, 2011). A gene family can be identified by their numbers and dispersion patterns due to evolution. Phylogenetic analysis and a thorough demonstration of the exon-intron architectures of the watermelon *YABBY* genes showed that the gene structural pattern coincided with the phylogenetic analysis. In watermelon, there were four to six different introns (Fig.S2; Table 1). Four genes, namely *ClYABBY3*, *ClYABBY5*, *ClYABBY6*, and *ClYABBY7*, each include five introns, whereas only two genes, Cl*YABBY*8, and Cl*YABBY*4, each contain six introns. Only one gene, *ClYABBY2*, contains four introns (Table 1; Fig.S2). While the *ClYABBY* genes in subfamily YAB5 have six introns, all of the *ClYABBY* genes in subfamily INO have five introns. Some of the watermelon's *YABBY* genes have four introns, similar to the *YABBY* genes examined in many species, but others have up to six introns (Table 1; Fig. S2). The MEME tool was used to determine the locations of 20 motifs across all of the watermelon *YABBY* proteins (Fig. 2). All of the *ClYABBY* proteins have the *YABBY* domain, which was present in all of them. The observation that the *YABBY* genes in the same group encode similar motifs leads to the conclusion that these conserved motifs play a crucial role in the activities that are unique to a group or subgroup. With the exception of subgroup YAB5, *ClYABBY8* has certain motifs that are distinct from . *ClYABBY4*. The identical *YABBY* domain is encoded by all 8 watermelon *YABBY* genes. Similar pattern distribution among different *YABBY* genes shows that these genes might have evolved as a result of gene expansion

Table 1: Information about discovered 8 YABBY genes in watermelon genome.Accession Number, Chromosome number and location, gene direction, Amino acid sequence length, Molecular weight, Isoelectric point (Pi-value), No. of Introns and Exons.

***YABBY* gene**	**Accession Number**		**Chromosome Number**	**Chromosome Location (bp)**		**Direction**	**NO. of Amino Acids**	**Molecular Weight (kD)**	**Pi-Value**	**No.of Introns**	**No.of Exons**
Name	**V1**	**WCG**		**Start**	**End**						
ClYABBY1	Cla011557	ClCG01G003810	1	3768830	3771373	R	242	26.91	8.73	6	7
ClYABBY2	Cla005676	ClCG10G003000	10	3521920	3525337	R	188	20.77	8.75	4	5
ClYABBY3	Cla008610	ClCG02G022510	2	32629974	32634962	R	170	18.74	9.08	5	6
ClYABBY4	Cla016700	ClCG11G003140	11	23955965	23959639	F	192	21.62	8.82	6	7
ClYABBY5	Cla000807	ClCG08G017890	0	18956326	18957519	R	186	20.70	7.70	5	6
ClYABBY6	Cla010073	ClCG05G025460	3	32287212	32288282	R	169	19.07	8.41	5	6
ClYABBY7	Cla010231	ClCG05G023780	5	31061561	31062707	F	188	20.90	9.14	5	6
ClYABBY8	Cla012226	ClCG06G011300	6	20045189	20048379	R	169	18.82	8.24	6	7

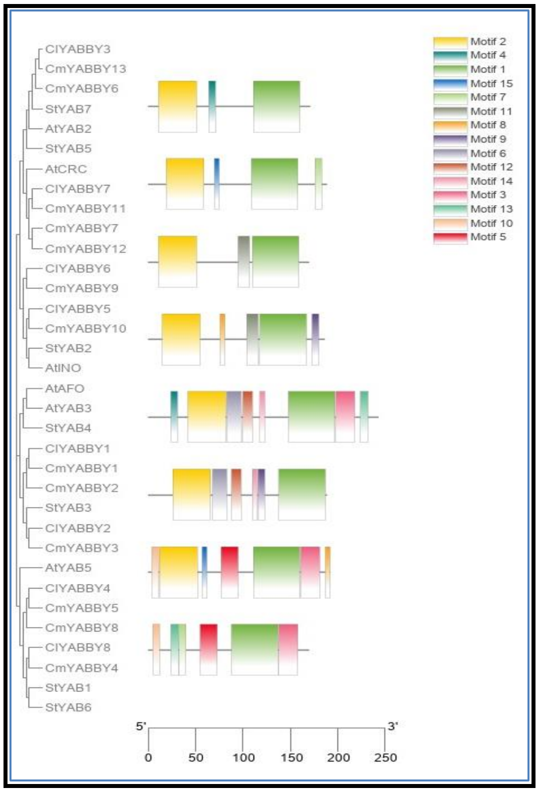


Figure. 2. The distribution of 15 motifs on 8 *YABBY* proteins of Watermelon by using MEME version 4.9.0. The bars represent motifs with different color codes for different types of motifs.

Comparative phylogenetic relatedness of watermelon gene family with Arabidopsis: (Table 2; Fig. 1). The Eight *ClYABBY* proteins were found to be spread across five groupings, according to the results of phylogenetic study as AtINO, AtCRC, AtYAB5, AtAFO/AtYAB3, and AtYAB2 (Table 2; Fig. 1). Six *YABBY* proteins, including one from Arabidopsis (*AtINO*), were found in the group *AtINO*. The other members of the group were *ClYABBY5*, *ClYABBY6*, *CmYABBY9*, *CmYABBY10*, and *StYABBY2*. *AtFIL*, *ClYABBY7*, *CmYABBY12*, *CmYABBY11*, and *CmYABBY7* were all members of the *AtCRC* group, which consists of five *YABBY*-like proteins. Seven *YABBY* proteins were found in the AtYAB5 group, including one from one from the Arabidopsis *AtYAB5* plant, two from the watermelon *ClYABBY8*, *ClYABBY4*, three from the musk melon *CmYABBY4*, *CmYABBY5*, *CmYABBY8*, and one from the potato *StYABBY1*. There were nine *YABBY*-like proteins in the AtAFO, including two from the Arabidopsis AtAFO, AtYAB3, two from watermelon, *ClYABBY1*, *ClYABBY2*, three from musk melon, *CmYABBY1*, *CmYABBY2*, *CmYABBY3* and two from the potato, *StYABBY3*, *StYABBY4*. The last group, AtYAB2, contained six *YABBY*-like proteins, of which one was from the Arabidopsis *AtYAB2* plant, one is from the watermelon *ClYABBY3*, two were from the musk melon *CmYABBY6*, *CmYABBY13* and two were from the potato (*StYABBY5*, *StYABBY7*). Individuals belonging to the same clade exhibit the same structure and behavior (Fig 1). Therefore, it was determined that proteins from related clades have comparable functions **Table 2.**

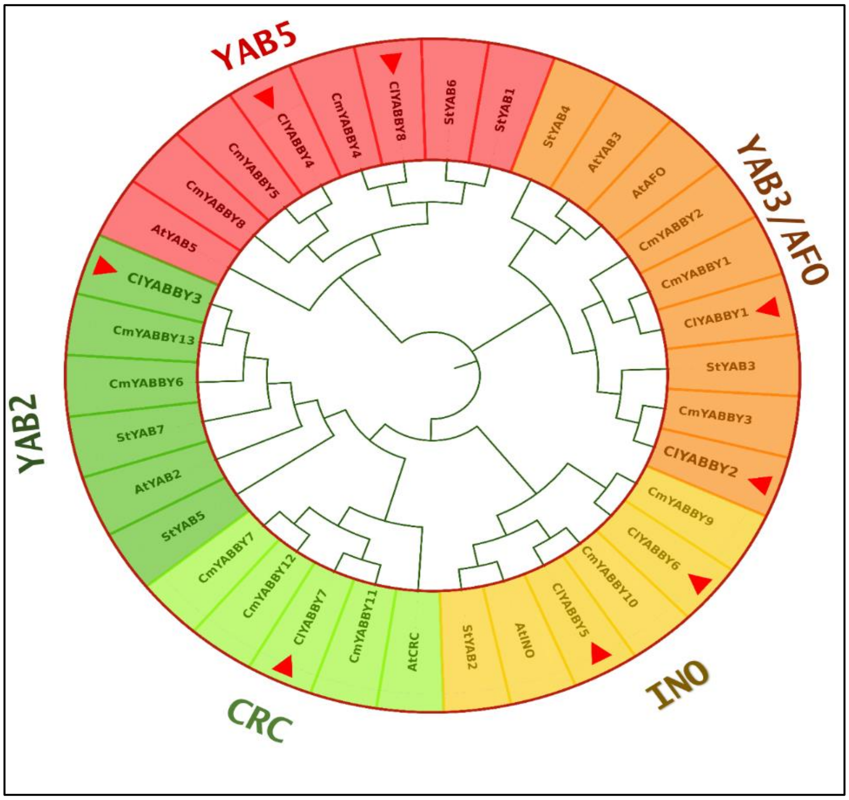


Fig 1: Phylogenetic Relationship among *YABBY* genes of *C. lanatus, S. tuberosum, A. thaliana, C. sativus,* and *C. maxima* was studied. *C. lanatus* genes are marked with red triangle. The evolutionary history was inferred using the NJ method with 1000 Bootstrap. This analysis involved 34 *YABBY* genes. Evolutionary analyses were conducted in MEGA 11.

Table 2: Gene ontology enrichment analysis of *ClYABBY* genes their GO functions, Sub-Cellular localization Signal Genes expression, Orthologs in Arabidopsis and their functions.

Gene ID	**GO Function**			**Sub-Cellular Localization**	**Gene expression**	**Ortholog in Arabidopsis**	
	Molecular Function	**Biological process**	**Cellular component**			**Gene Name**	**Function**
ClYABBY1	DNA-binding transcription factor activity	chromatin assembly or disassembly, regulation of transcription, DNA-templated, stomatal complex	transcription factor complex	Peroxisomes	Leaves	*YABBY3*	Axial regulator

		morphogenesis, iron-sulfur cluster assembly, plant ovule development, abaxial cell fate specification					
ClYABBY2		specification of floral organ identity, abaxial cell fate specification, specification of animal organ position, inflorescence meristem growth		Nucleus	Leaves and seedlings	*YABBY3*	Axial regulator
ClYABBY3	DNA-binding transcription factor activity	regulation of transcription, DNA-templated, multicellular organism development, abaxial cell fate specification	transcription factor complex	Nucleus	Leaves and seedlings	*YABBY2*	Putative axial-regulator
ClYABBY4	DNA-binding transcription factor activity	chromatin assembly or disassembly, regulation of transcription, DNA-templated, regulation of translation,	transcription factor complex	Nucleus	Leaves and seedlings	*YABBY5*	Axial regulator
ClYABBY5		multicellular organism development		Cytoplasm	leaves	INO	Outer integument Growth
ClYABBY6		multicellular organism development, cellular process		chloroplast	Leaves and seedlings	INO	Outer integument Growth
ClYABBY7	DNA-binding transcription factor activity	nectary development, regulation of anthocyanin biosynthetic process, specification of floral organ identity, Stamen development, style development, polarity specification of adaxial/abaxial axis	transcription factor complex	Nucleus		CRC	Polarity establishment in carpel and nectary development.
ClYABBY8		regulation of gene expression, cellular nitrogen compound biosynthetic process, regulation of nitrogen compound metabolic process, regulation of cellular macromolecule biosynthetic process		Nucleus	Leaves and seedlings	YABBY 5	Axial regulator

Table 3*: ClYABBY*genes targeting Putative miRNA functions along with their targeted genes.

miRNA	Targetting Gene	Function	Reference
Cla-miR156c	*ClYABBY6*	Suppress adventitious root development.	(M. Xu *et al*., 2016a)
Cla-miR157b	*ClYABBY6*	Responsible for the temporal expression pattern of most SPL genes.	(M. Xu *et al*., 2016b)
Cla-miR159a	*ClYABBY6*	Mediate strong silencing of GAMYB to enable normal growth.	(M. Xu *et al*., 2016a)
Cla-miR162	*ClYABBY7*	Not Available	
Cla-miRN820	*ClYABBY6*	Associated with epigenetic modifications	(Nosaka *et al*., 2013)
Cla-miRN825	*ClYABBY8*	Regulate plant defence responses to pathogens	(M. Xu *et al*., 2016a)
Cla-miR157a	*ClYABBY6*	Responsible for the temporal expression pattern of most SPL genes.	(M. Xu *et al*., 2016b)
Cla-miR159b	*ClYABBY6*	Mediate strong silencing of GAMYB to enable normal growth.	(Nosaka *et al*., 2013)
Cla-miRN812	*ClYABBY1*	Not available	

Location of chromosomes and assessment of gene duplication of watermelon *YABBY* genes: The distribution of the examined *Citrullus lanatus subsp. vulgaris var.* 97103 *YABBY* genes on chromosomes showed that different chromosomes included *ClYABBY* genes. The most *YABBY* genes were found on chromosome 5, and there were 2 on it.. There was just one *YABBY* gene on chromosomes 1, 2, 6, 8, 10, and 11. Contrarily, it was discovered that chromosomes 3, 4, 7, and 9 did not have the *YABBY* gene (Supplementary Fig. S4). In order to evaluate segmental and tandem duplication of the *ClYABBY* gene family in chromosomal position, synteny analysis for *ClYABBY* genes was also carried out **(Fig 2a).** While *ClYABBY7* and *ClYABBY6* genes were clustered together on the same chromosome, it was possible that these genes evolved via tandem duplication in the watermelon *YABBY* genes because three homologs gene pairs were spread unevenly throughout the watermelon genome (Fig. S4). The comparison research shows that the genomic family has expanded due to the lower number of genes that are shared between watermelon and other members of the Cucurbit family. Watermelon and Arabidopsis only share eight genes, but watermelon and bottle gourd share 12 genes, according to a dual syntenic analysis. Additionally, watermelon and melon share 7 genes, whereas watermelon and musa share just 1, revealing their evolutionary ties within the cucurbita family (Fig 2b)

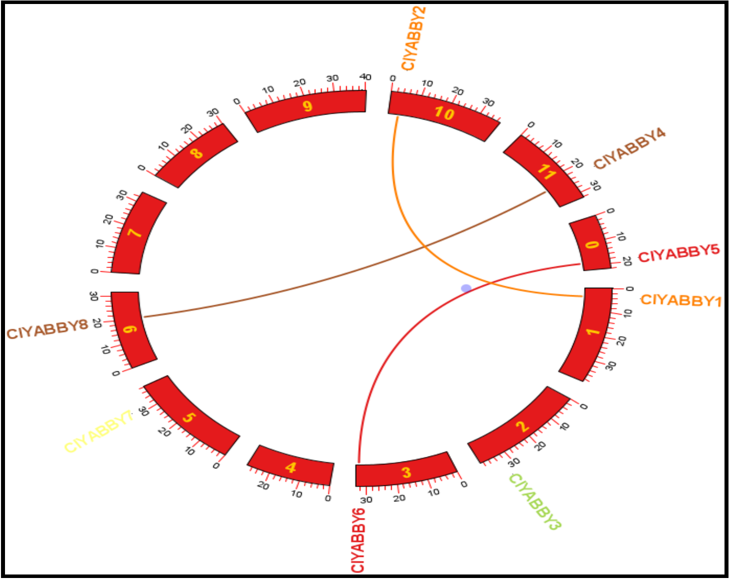
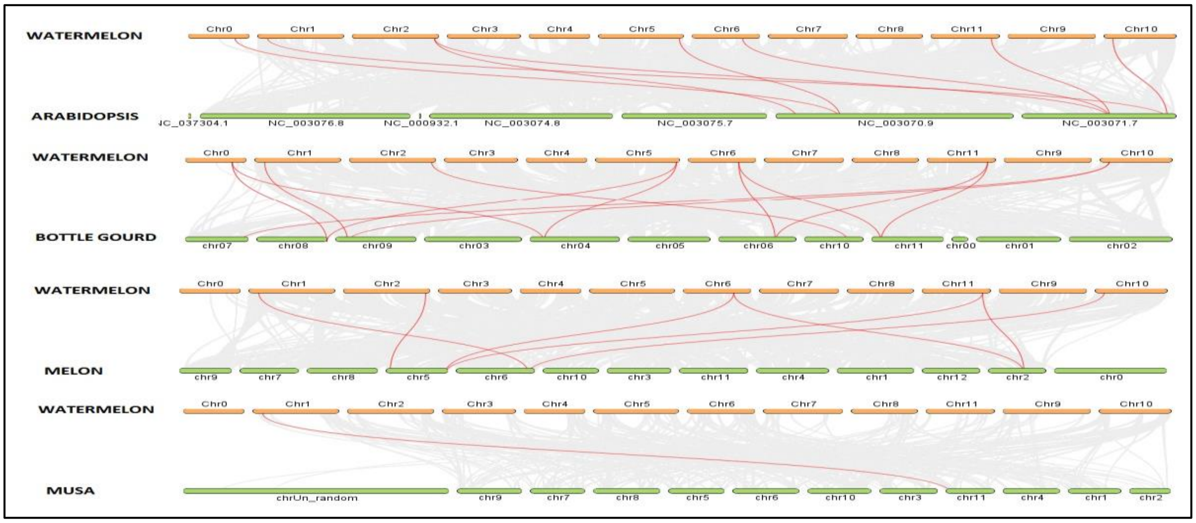


Fig 2(a): Genome-wide synteny analysis of *StYAB*genes showing the dominance of segmental duplication and rare occurrence of tandem duplication. Joining lines showing the duplicated *StYAB*genes in the genome.**Fig 2(b): Dual synteny analysis of Potato-Arabidopsis, Potato-Tomato, and Potato-Chili. Orange bars represents chromosomes of potato while green bars represent chromosomes of Arabidopsis, Tomato, and Chili respectively. Red lines showing the duplicated genes in the respective genomes.**



Additionally, TB Tools software was used to determine the date of gene duplication using pairwise alignment, which produced Ks and Ka values, followed by a manual calculation of Ka/Ks (Fig 3). The ratio of non synonymous (Ka) to synonymous (Ks) mutation was indicated as Ka/Ks. Ks displays the number of synonymous substitutions per synonymous site, whereas Ka shows the number of non synonymous substitutions per non synonymous site. The *ClYABBY5/ClYABBY6* pair had a ratio of 0.36, while the *ClYABBY2/ClYABBY1* pair had a ratio of 0.17. The homologs group *ClYABBY2/ClYABBY1* tandem gene duplication date was anticipated to be 43.2 Mya, which was the greatest, while the segmental duplication date for the paralogous pair *ClYABBY8/ClYABBY4* was predicted to be 30.7 Mya, which was the lowest. Segmental duplication date for the final paralogous pair, *ClYABBY5/ClYABBY6*, was estimated to be 34.2 Mya (Table S3). In the watermelon, all 3 paralogous group pairs had Ka/Ks ratios larger than 0.3 but lower than 1, which raises the possibility of significant functional divergence following duplication brought on by purifying selection

.

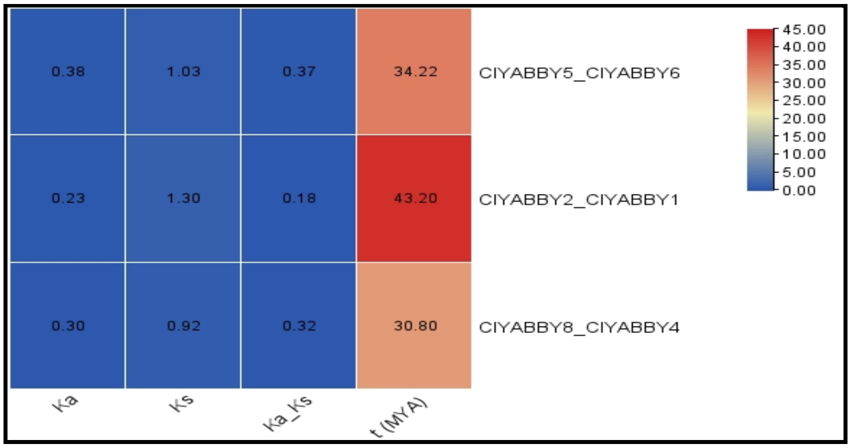


Figure 3: Time of gene duplication estimated for different paralogous pairs of watermelon *ClYAB* genes on the basis of Ks and Ka values. Analyses were conducted using Tbtools (Software). Ka/Ks represents the ratio of nonsynonymous (Ka) versus Synonymous (Ks) mutations.

Analysis of CREs: The presence and arrangement of different CREs at the binding site of TFs on the promoter region influence the spatiotemporal transcriptomic expression of genes. To assess the potential roles of genes, numerous CREs can be analyzed in silico (Bulow & Hehl, 2016) (Jones & Vandepoele, 2020). CR components with functions like anaerobic induction, ethylene production, sensitivity to light, wound healing, hormone-specific stress, and drought were noticed (Fig. 5;Table. S4). Interestingly, the ARE element, which was necessary for anaerobic induction, was present in 6 of the 8 *ClYABBY* genes (*ClYABBY1*, *ClYABBY2*, *ClYABBY4*, *ClYAABY6*, *ClYABBY7*, and *ClYABBY8*). There were also 6 light responsive elements. Only one Cl*YABBY*1 gene displayed the GT1 motif, whereas the G-box has three *ClYABBY* genes (*ClYABBY1*, *ClYABBY2*, and *ClYABBY4*), the AAAC motif has only one *ClYAABY2* gene, the TCT motif includes four *ClYABBY* genes (*ClYABBY2*, *ClYABBY6*, *ClYABBY7*, and *ClYAABY8*), the TCCC motif has two *ClYABBY* genes (*ClYABBY7*, *ClYABBY8*), and I-box possess 1 *ClYABBY7* gene, Box 4 elements, a piece of a conserved DNA module involved in light responsiveness, were only found in 2 *ClYABBY* genes (*ClYABBY4 and ClYABBY7*). The ABRE element was present in three *ClYABBY* genes (*ClYABBY1*, *ClYABBY2*, and *ClYABBY4*) and was involved in the abscisic acid response. 4 *ClYABBY* genes (*ClYABBY1*, *ClYABBY2*, *ClYABBY5*, and *ClYABBY8*) have the salicylic acid responsive TCA element, while 4 other *ClYABBY* genes (*ClYABBY2*, *ClYABBY3*, *ClYABBY6*, and *ClYABBY7*) have the salicylic acid responsive TCA element. Only 2 *ClYABBY* genes (*ClYABBY4*, *ClYABBY6*) showed the wound-responsive WUN motif. *ClYABBY2*, *ClYABBY5*, and *ClYABBY6* showed TC-rich repeats that demonstrate responses in defence and stress, respectively. The MBS element was present in 2 *ClYABBY* genes (*ClYABBY2*, *ClYABBY4*) and was connected to drought-inducibility, The auxin responsive TGA element was present in 2 *ClYABBY* genes (*ClYABBY6*, *ClYABBY8*). Ethylene response element ERE is present in 3 *ClYABBY* genes (*ClYABBY1*, *ClYABBY5*, and *ClYABBY8*). The MYB, which was involved in plant development, is present in 6 *ClYABBY* genes (*ClYABBY1*, *ClYABBY4*, *ClYABBY5*, *ClYABBY6*, *ClYABBY7*, and *ClYABBY8*), but only one *ClYABBY8* gene possesses the gibberellin responsive GARE-motif, and another gibberellin responsive element p-box contains only 1 *ClYABBY2* gene. The STRE, which was engaged in stress-related processes, was demonstrated by 5 *ClYABBY* genes (*ClYABBY1*, *ClYABBY2*, *ClYABBY3*, *ClYABBY7*, and *ClYABBY8*). Only 1 *ClYABBY2* gene includes the ABRE4 which mediates ABA-dependent stress responses, as well as ABRE3a, a positive regulator of abiotic stress and ABA signalling. The MYC, which stimulates the expression of proliferative genes, was found in 2 *ClYABBY* genes (*ClYABBY2*, *ClYABBY8*). No Cl*YABBY* genes, however, have been discovered to be deficient in the previously stated cis regulatory regions.

The eight watermelon *YABBY* genes' identified CREs are displayed in (Fig 4) and Table S4 along with their functional annotations.

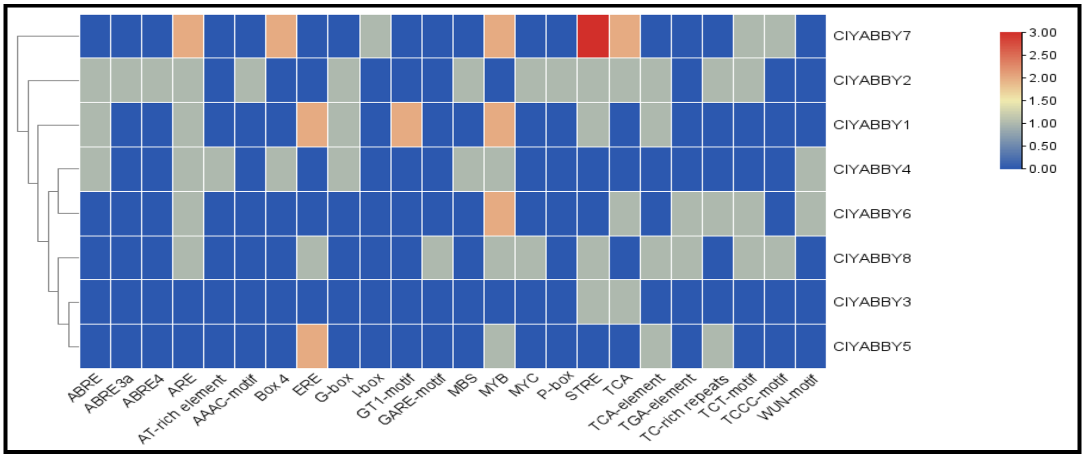


Fig 4: *Cis*-regulatory elements in putative *ClYAB* promoters which are associated with different plant developmental process.

Transcriptomic analysis of watermelon *YABBY* genes: With the aid of high throughput sequencing, the differentiating expression patterns of all watermelon *YABBY* genes at distinct developmental stages were also examined (Song, Joshi, DiPiazza, *et al*., 2020) (Song, Joshi, & Joshi, 2020). There were two different kinds of studies on the effects of salt and drought stress. Watermelon seedlings were subjected to short-term salt stress in order to study and discover genes and pathways associated with response to salt stress treatments, such as differences in photosystem II photosynthetic efficiency and free amino acids. A heat map displaying the expression of the watermelon *ClYABBY* genes (Fig 5). Out of the eight genes that were subjected to salt stress, *ClYABBY2*, *ClYAABY3*, *ClYABBY4*, *ClYABBY6*, and *ClYABBY8* showed expression in five of them. Cl*YABBY*1 nevertheless displayed expression, but in the control group. The ethylene biosynthesis signaling pathways were activated during salt stress treatments, according to the hypothesis, and photosystem efficiency was decreased under salt stress. Therefore, the genes that were expressed play a role in the ethylene biosynthesis signaling pathway, which explains why they were highly expressed. However, ClYABBY1, ClYABBY7, and ClYABBY5 play a role in pathways that increase photosynthesis by enhancing the efficiency of photosystem II. This is why they were not expressed in watermelon seedlings under salt stress (Song, Joshi, & Joshi, 2020)

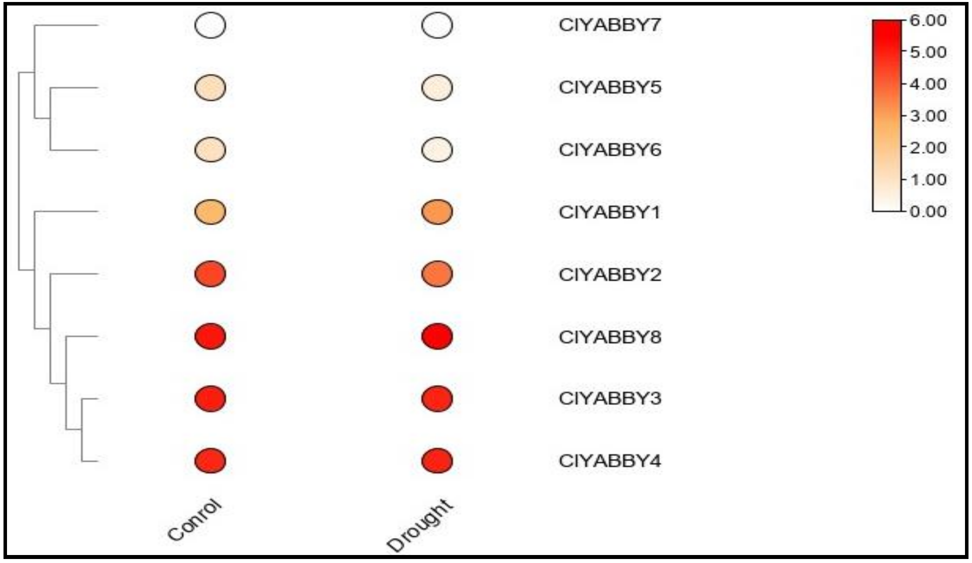


Fig 5: Heat map shows the expression profile of the *ClYAB* genes in leaves of watermelon cultivar under different levels of Drought stress.

The goal of the drought stress experiment was to explore the accumulation of drought-induced citrulline in watermelon leaves by monitoring stress treatments using physiological measurements. Watermelon seedlings were stressed in a semi-controlled environment and open field. A heat map represents the expression profiles of the watermelon *ClYABBY* genes (Fig 6) Because only 5 hits from the available RNA seq data were identified, the expression of 5 of the 8 *ClYABBY* genes, namely *ClYABBY1*, *ClYABBY2*, *ClYABBY3*, *ClYABBY4*, and *ClYABBY8*, was detected. Based on the assumption, these genes play a part in the biosynthesis of citrulline, which accumulated in the leaves of watermelon seedlings under drought stress conditions, while *ClYABBY5*, *ClYABBY6*, and *ClYABBY7* are thought to play a minor role in the biosynthesis of citrulline or may not be directly associated to its creation in young watermelon plants, which may explain why they did not exhibit expression (Song, Joshi, DiPiazza, *et al*., 2020).

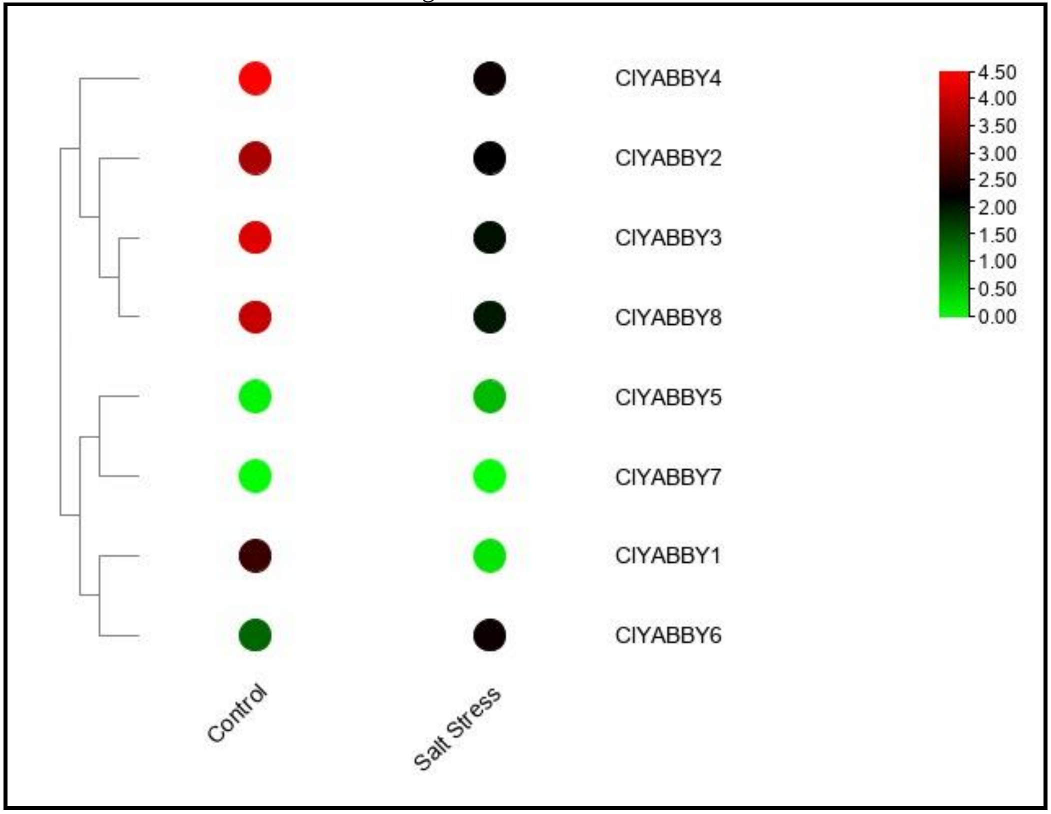


Fig 6: Heat map shows the expression profile of the *ClYAB* genes in seedlings of Watermelon cultivar under different levels of salt stress.

Putative miRNA targets in watermelon: The results showed that nine miRNAs in total were discovered, and four of the eight *ClYABBY* genes were their targets. None of these miRNAs targeted the remaining four *ClYABBY* genes (*ClYABBY2*, *ClYABBY3*, *ClYABBY4*, and *ClYABBY5*) (Table S6). These miRNAs were between 20 and 22 amino acids long. Per the *ClYABBY* gene, there were between 156 and 825 miRNAs that targeted these genes. Cla-miR156c, Cla-miR157b, Cla-miR159, Cla-miR820, Cla-miR157a, and Cla-miR159b were six of the nine mature miRNAs that target *ClYABBY6,* while *ClYABBY7*, *ClYABBY8*, and *ClYABBY1* were each targeted by just one mature miRNA. Cla-miR162, Cla-miR825, and Cla-miR812 all target *ClYABBY7*, *ClYABBY8*, and *ClYABBY1*, respectively (Table S6). Therefore, Cl*YABBY*6 was the sole gene that had the greatest number of miRNAs targeting it. The most targeted group was Group INO, which had six mature miRNAs target it. In contrast, just one miRNA targeted Group CRC, Group YAB5, and Group YAB3, whereas no miRNAs targeted Group YAB2.

Discussion

YABBY PSTrFs were divided into 5 families (Group *AtINO*, *AtCRC*, *AtYAB5*, *AtAFO/AtYAB3*, *AtYAB2*), following the phylogenetic and domain analysis patterns of *A. thaliana* (D. Lijavetzky, P. Carbonero, & J. Vicente-Carbajosa, 2003), citrus (Wu, Fu, & Yi, 2016), and eggplant (Wei *et al*., 2018). The Cucurbit Genomics Database's recently released data was utilized to identify Cl*YABBY* genes at the genome level (CuGenDB) (http://cucurbitgenomics.org/)(Table. 1). Using phylogenetic analysis, the watermelon's 8 *YABBY* genes were divided into five groups (Group *AtINO*, *AtCRC*, *AtYAB5*, *AtAFO/AtYAB3*, *AtYAB2*) (Fig. 1, Table. 1). Watermelon had less *YABBY* genes than rice (30 *OsYABBY*) (X. Yang & Tuskan, 2006), Arabidopsis (36 *AtYABBY*) (X. Yang & Tuskan, 2006), tomato (34 *SiYABBY*) (Cai *et al*., 2013), banana (74 *MaYABBY*) (Dong, Hu, & Xie, 2016), and Chinese cabbage (76 *BrATYABBY*) (Ma, Li, Wang, Tang, & Xiong, 2015).

Exons and introns were present in all the genes, albeit in varying amounts (Table. 1). Genes with the same number of exons and introns were classified into the same clade following analysis. The watermelon *ClYABBY* genes found in the same family (Groups) generally had the similar exon-intron architecture, however variants were found in other families. Similar intron-exon configurations have also been found in Arabidopsis, rice, and soybean (Diego Lijavetzky, Pilar Carbonero, & Jesús Vicente-Carbajosa, 2003) (Gu *et al*., 2013) suggesting that these structures were evolutionary preserved.

The examination of conserved motifs added more support to the classification of *ClYABBY* genes. To find the conserved motifs, the MEME analysis tool was used to upload all of the *ClYABBY* protein sequences. There were consequently a total of fifteen conserved motifs found. The length of the motifs of *ClYABBY* proteins discovered by MEME ranged from 6 to 50 amino acids. All of the watermelon *YABBY* proteins shared Motif-1, or *YABBY* (Fig. 2). Motif-1 and Motif-2 are present in all Group INO proteins as well as all Group YAB2, YAB3, CRC, and YAB5 proteins with the exception of *ClYAABY8* which lack Motif-2. Extra particular motifs found in several *YABBY* proteins may be important for various functions. The most intricate motif pattern, Motif-5, was found in the *YABBY* proteins from Group YAB5, and Motif-10 is unique to this subgroup. In contrast to Group YAB5, the members of the YAB2 group have a simpler motif pattern, although some of the group members also feature group-specific patterns, like Motif-4. Additionally, the YAB3 group included group-specific motifs like Motif-3 and Motif-6. In order to comprehend the probable functions of the Group CRC-specific motifs, the GO annotations of the Group CRC genes in Arabidopsis were first examined. We discovered something intriguing: In contrast to the Arabidopsis *YABBY* genes (Table 2) in other groups, AtCRC gene in Group CRC have responsibilities in "Polarity establishment in carpel and nectary development," which is comparable to GO annotations of watermelon *ClYABBY7* gene in Group CRC which have roles in "nectary development," "definition of floral organ identity," "Stamen development," "style development," and "polarity specification of adaxial/a (Table 2). Rice (Yamaguchi *et al*., 2004), Arabidopsis (Alvarez & Smyth, 1999; Bowman & Smyth, 1999; Lee *et al*., 2005) and cucumber (Yin *et al*., 2022) have all shown similar functions for genes belonging to the CRC clade. This suggests that these genes function similarly in Arabidopsis, rice, cucumber, and watermelon.

The distribution of motifs in the watermelon *YABBY* proteins (Fig. 2) shows how they evolved, as inferred from the phylogenetic tree (Gupta *et al*., 2015) (Malviya *et al*., 2015). The alignment of the watermelon *ClYABBY* protein sequences and the motif data analysis by Motif finder and domain analysis by NCBI CDD all pointed to a highly conserved *YABBY* domain. Evolutionarily, *YABBY* PSTrFs have been preserved in a variety of plants. In addition to the *YABBY* domain, 14 newly discovered unique motifs were identified and distributed differently among the *ClYABBY* genes (Fig. 2). Consequently, one or two motifs were revealed to be preserved in the same group of distinct comparative crop species; however, some differences were observed between motifs of other groups, suggesting that members of the same group functioned similarly. Furthermore, the *ClYABBY* genes' structural arrangement was retained throughout all five separated groups, including other species such as Arabidopsis, Cucumber, and Musk melon (Dong *et al*., 2016; Diego Lijavetzky *et al*., 2003; Nasim, Malviya, Kumar, & Yadav, 2016; X. Yang & Tuskan, 2006). In addition, an analysis of the subcellular localization of Cl*YABBY* proteins using the online tool WoLF PSORT (https://wolfpsort.hgc.jp/) revealed that all Cl*YABBY* proteins had nuclear localization, with only minor variances among the cytoplasm, peroxysomes, and chloroplasts, all of which were typically found in the nucleus. (Table 2, Fig S3). A gene's position on a chromosome can be used to evaluate gene duplication. It is known that segmental duplication occurs when two or more genes from the same species are located on distinct chromosomes of the same species. Tandem duplication occurs when two or more genes are present on the same chromosome (Panchy, Lehti-Shiu, & Shiu, 2016). Despite the fact that the high number of watermelon *YABBY* genes on chromosome 5 (Fig. S4) indicates the occurrence of tandem duplication, some of the *ClYABBY* genes were also displaying segmental duplication (Fig. S4). In the *YABBY* family of genes, segmental duplication predominated in chickpea (Nasim *et al*., 2016) and pigeon pea (Malviya *et al*., 2015). While domain duplication may have contributed to the larger number of *YABBY* genes throughout the evolution of eukaryotic plants, gene duplication is the primary method involved in the proliferation of gene families (Moore & Purugganan, 2005; Taylor & Raes, 2004).

Understanding amino acid substitution requires an understanding of the Ka/Ks ratio (Fig. 3). When Ka/Ks is less than 1, purifying selection takes place, and when Ka/Ks is greater than 1, positive selection takes place (Hurst, 2002; Ziheng Yang & Bielawski, 2000). Variations in selection pressure typically result in changes to a protein's unique amino acid pattern, which were crucial for understanding how various proteins operate (Morgan, Loughran, Walsh, Harrison, & O'Connell, 2010). There was just a small amount of variation in Ka/Ks ratios between *ClYABBY* genes. Despite the changes, all anticipated values of Ka/Ks ranged from 0.17 to 0.36, which is less than 1, indicating that the sequences of *YABBY* present in all groups underwent strong purifying selection pressure and that only a small number of sites may have been impacted by positive selection only during the process of evolution (Fig.3).

Using the available RNA seq data, the expression of all *ClYABBY* genes in several watermelon experiments was examined. Five *ClYABBY* genes were found to express in the leaf tissue of seedlings during a drought stress experiment. As stated in the results, the cis-regulatory element MBS in *ClYABBY2* and *ClYABBY4* has a role in drought inducibility and expressed itself in this experiment. Furthermore, Go annotations of these genes in comparison to Arabidopsis showed how these gene expressions occurs in leaf and stem tissue. The functions of *ClYABBY2*, "abaxial cell fate specification" and "inflorescence meristem growth," which are consistent with those of Arabidopsis (Siegfried *et al*., 1999), confirm their function in the leaves of early watermelon seedlings under drought stress. Cl*YABBY*2 is a member of Group YAB3, which also includes the axial regulators AtYAB3 and AtAFO Arabidopsis genes. When watermelon seedlings are subjected to drought stress, above mentioned genes are engaged in pathways that create citrulline, which was used in biosynthesis (Song, Joshi, DiPiazza, *et al*., 2020)

However, it was discovered in additional studies on salt stress that mechanisms for the manufacture of ethylene are activated under salt stress. Therefore, 5 *ClYABBY* genes were expressed in watermelon seedlings, indicating that they may be involved in ethylene responsiveness and stress-mediated response. *ClYABBY6* has CREs TC-rich repeats, which have functions associated in defence and stress responsiveness, and *ClYABBY8* has CRE ERE, which has functions involved in ethylene responsiveness, as described in the results. Since photosystem II efficiency decreases under salt stress, it follows that the genes that did not exhibit any expression should have a role in light sensitivity and photosynthetic activity. This supports the cause for the expression of these genes. *ClYABBY1* has the light responsive CREs G-box and Gt1 motif, while *ClAYBBY7* has the light responsive CREs I-box, TCCC-motif, and TCT-motif (Song, Joshi, & Joshi, 2020)

From the developmental stage to pathogen defence and maintaining healthy internal conditions, microRNAs play a critical role in controlling plant growth (Carbone *et al*., 2019; Samad, 2017; Spanudakis, 2014; Terzi, 2008). Regardless of the type of species they were found in, miRNAs are present in the majority of plant species in a manner that is preserved, which explains their specific function. The majority of the *ClYABBY* genes have transcriptional activity linked with them, which suggests that they inhibit miRNA activity. Due to this, members of the Cla-miR156 and Cla-miR825 families only targeted one out of the eight *ClYABBY* genes (Table. S6). Only one *ClYABBY6* gene was targeted by each of the two Cla-miR157 (Cla-miR157a, Cla-miR157b) and Cla-miR159 (Cla-miR159a, Cla-miR159b), whereas Cla-miR162 was targeted by *ClYABBY7*. Because these two genes are on the same chromosome 5, it is likely that the majority of their origin and activity is on chromosome 5. Numerous plant species include Cla-miR156c, which inhibits the growth of adventitious roots (Xu, Chen, Ying, & Cai, 2016a).The majority of SPL genes in different plant species exhibit a temporal expression pattern controlled by the miR157 family (Xu, Chen, Ying, & Cai, 2016b), whereas the miR159 family plays a role in the strong suppression of GAMYB necessary for proper growth (P. Xu *et al*., 2016a) Most plant species have Cla-miR820, which *ClYABBY6* targets, associated with epigenetic modifications. (Nosaka *et al*., 2013), While the role of Cla-miR825 that Cl*YABBY*8 targets is to control how plants respond to pathogens in terms of defence (P. Xu *et al*., 2016a). We can infer potential roles for the targeted genes in watermelon by studying the actions of miRNAs. Due to a dearth of research resources, cla-miR812 and cla-miR162, which were the targets of *ClYAABY7* and *ClYABBY1*, have no known functions. (Table. 3) To help us understand the roles played by miRNAs, more investigation is needed. Table 3:

Conclusion

The ClYABBY PSTrFs gene family inside the watermelon genome was explored in this comprehensive study. The research focused on the structural and functional properties of the eight ClYABBY genes, classifying them into five separate groups based on Arabidopsis classification. These genes were essential not only for normal watermelon growth and development, but also for alleviating salt and drought-induced stress. Furthermore, the study examined miRNAs that target ClYABBY genes, revealing their role in defense mechanisms and stress responses, notably under salt and drought stress. This study's computational results offer potential for applications in molecular cloning, gene expression profiling, and exploring relationships with various transcription factors (TFs). The surprising variations in *YABBY* gene counts among different plant species are an intriguing part of our research. While certain plants, such as tomato, pepper, and potato, have almost comparable quantities of *YABBY* genes, others, such as bottle gourd, melon, wax gourd, and watermelon, have far fewer. The hypothesis suggests that the differences in gene numbers were a result of gene duplication and loss events that occurred during plant evolution. Furthermore, it was discovered that segmental duplications were more important in the expansion of the *YABBY* gene family in watermelon than tandem duplications.. This finding highlights the mechanisms responsible for the diversification of this gene family within the watermelon genome. The research also investigated tissue-specific expression patterns, revealing that *ClYABBY* genes exhibited high expression levels in roots and seedlings. This information sheds light on the critical roles that these genes play in early plant growth and root function. Finally, distinct CREs linked to responses to light, wound, drought, auxin, stress, salicylic acid, and abscisic acid (ABA) in seedlings and roots were identified through cis-regulatory element (CRE) studies. This study enhances the understanding of plant adaptation and stress response systems, with direct implications for agriculture and plant science. In conclusion, the research offers valuable insights into the activities and regulatory networks of *ClYABBY* genes, serving as a valuable resource for future plant biology research.

Authors' Contribution

The manuscript was written by MN*, WT and MH. BS, H.S.M, MH and M.W.H played a significant role in data analysis. AI provided overall direction and experimental design. reviewed the manuscript and the final manuscript was proofread by all the authors and was accepted for further publication.

Acknowledgments

This work was supported by the University of the Punjab, Lahore, Pakistan.

Conflict of Interest

The authors have no potential conflict of interest.

References

Alvarez, J., & Smyth, D. R. (1999). CRABS CLAW and SPATULA, two Arabidopsis genes that controcarpel development in parallel with AGAMOUS. *Development,* **126**(11), 2377-2386.

Bailey, T. L., Johnson, J., Grant, C. E., & Noble, W. S. (2015). The MEME suite. *Nucleic acids research,* **43**(W1), W39-W49.

Bondarenko, V. S., & Gelfand, M. S. (2016). Evolution of the Exon-Intron Structure in Ciliate Genomes. *PLoS One,* **11**(9), e0161476. 10.1371/journal.pone.0161476

Bowman, J. L. (2000). The *YABBY* gene family and abaxial cell fate. *Current opinion in plant biology,* **3**(1), 17-22.

Bowman, J. L., & Smyth, D. R. (1999). CRABS CLAW, a gene that regulates carpel and nectary development in Arabidopsis, encodes a novel protein with zinc finger and helix-loop-helix domains. *Development,* **126(**11), 2387-2396.

Bowman, J. L., Smyth, D. R., & Meyerowitz, E. M. (1989). Genes directing flower development in Arabidopsis. *The Plant Cell,* **1**(1), 37-52.

Bulow, L., & Hehl, R. (2016). Bioinformatic Identification of Conserved Cis-Sequences in Coregulated Genes. [Research Support, Non-U.S. Gov't]. *Methods Mol Biol,* **1482**, 233-245.

Buttar, Z. A., Yang, Y., Sharif, R., Nan Wu, S., Xie, Y., & Wang, C. (2020). Genome wide identification, characterization, and expression analysis of *YABBY*-gene family in wheat (***Triticum aestivum*** L.). *Agronomy,* **10**(8), 1189.

Cai, X., Zhang, Y., Zhang, C., Zhang, T., Hu, T., Ye, J., Ye, Z. (2013). Genome-wide analysis of plant-specific Dof transcription factor family in tomato. *J Integr Plant Biol, 55*(6), 552-566.

Cannon, S. B., Mitra, A., Baumgarten, A., Young, N. D., & May, G. (2004). The roles of segmental and tandem gene duplication in the evolution of large gene families in Arabidopsis thaliana. *BMC plant biology,* **4**(1), 1-21.

Carbone, F., Bruno, L., Perrotta, G., Bitonti, M. B., Muzzalupo, I., & Chiappetta, A. (2019). Identification of miRNAs involved in fruit ripening by deep sequencing of Olea europaea L. transcriptome. PLoS One, **14**(8), e0221460.*PLoS One, 14*(8), e0221460.

Chen, C., Chen, H., Zhang, Y., Thomas, H. R., Frank, M. H., He, Y., & Xia, R. (2020). TBtools: an integrative toolkit developed for interactive analyses of big biological data. Molecular plant, **13**(8), 1194-1202.

Chen, Y.-Y., Hsiao, Y.-Y., Chang, S.-B., Zhang, D., Lan, S.-R., Liu, Z.-J., & Tsai, W.-C. (2020). Genome-wide identification of *YABBY* genes in Orchidaceae and their expression patterns in Phalaenopsis orchid. *Genes,* **11**(9), 955.

Dong, C., Hu, H., & Xie, J. (2016). Genome-wide analysis of the DNA-binding with one zinc finger (Dof) transcription factor family in bananas. *Genome,* **59**(12), 1085-1100.

Eckardt, N. A. (2010). *YABBY* genes and the development and origin of seed plant leaves: American Society of Plant Biologists.

Finet, C., Floyd, S. K., Conway, S. J., Zhong, B., Scutt, C. P., & Bowman, J. L. (2016). Evolution of the *YABBY* gene family in seed plants. *Evolution & development,* **18**(2), 116-126.

Floyd, S. K., & Bowman, J. L. (2007). The ancestral developmental tool kit of land plants. *International journal of plant sciences,* **168**(1), 1-35.

Gasteiger, E., Hoogland, C., Gattiker, A., Wilkins, M. R., Appel, R. D., & Bairoch, A. (2005). Protein identification and analysis tools on the ExPASy server *The proteomics protocols handbook* (pp. 571-607): Springer.

Golz, J. F., Roccaro, M., Kuzoff, R., & Hudson, A. (2004). GRAMINIFOLIA promotes growth and polarity of Antirrhinum leaves.

Gu, X., Zou, Y., Su, Z., Huang, W., Zhou, Z., Arendsee, Z., & Zeng, Y. (2013). An update of DIVERGE software for functional divergence analysis of protein family. *Molecular biology and evolution, 30*(7), 1713-1719.

Gupta, S., Malviya, N., Kushwaha, H., Nasim, J., Bisht, N. C., Singh, V., & Yadav, D. (2015). Insights into structural and functional diversity of Dof (DNA binding with one finger) transcription factor. *Planta,* **241**(3), 549-562.

Higo, K., Ugawa, Y., Iwamoto, M., & Higo, H. (1998). PLACE: a database of plant cis-acting regulatory DNA elements. [Research Support, Non-U.S. Gov't]. *Nucleic Acids Res,* **26**(1), 358-359.

Higo, K., Ugawa, Y., Iwamoto, M., & Korenaga, T. (1999). Plant cis-acting regulatory DNA elements (PLACE) database: 1999. [Research Support, Non-U.S. Gov't]. *Nucleic Acids Res,* **27**(1), 297-300

Horton, P., Park, K.-J., Obayashi, T., & Nakai, K. (2006). *Protein subcellular localization prediction with Wolf PSORT.* Paper presented at the Proceedings of the 4th Asia-Pacific bioinformatics conference.

Hou, H., Wu, P., Gao, L., Zhang, C., & Hou, X. (2019). Characterization and expression profile analysis of *YABBY* family genes in Pak-choi (Brassica rapa ssp. chinensis) under abiotic stresses and hormone treatments. *Plant Growth Regulation,* **87**(3), 421-432.

http://cucurbitgenomics.org/. http://pfam.xfam.org/.

Hu, B., Jin, J., Guo, A. Y., Zhang, H., Luo, J., & Gao, G. (2015). GSDS 2.0: an upgraded gene feature visualization server. [Research Support, Non-U.S. Gov't]. *Bioinformatics,* **31**(8), 1296-1297.

Huang, Z., Van Houten, J., Gonzalez, G., Xiao, H., & van der Knaap, E. (2013). Genome-wide identification, phylogeny and expression analysis of SUN, OFP and *YABBY* gene family in tomato. *Molecular genetics and genomics,* **288**(3), 111-129.

Hurst, L. D. (2002). The Ka/Ks ratio: diagnosing the form of sequence evolution. *TRENDS in Genetics, 9*(18), 486-487.

Jones, D. M., & Vandepoele, K. (2020). Identification and evolution of gene regulatory networks: insights from comparative studies in plants. *Curr Opin Plant Biol,* **54**, 42-48.

Kanaya, E., Nakajima, N., & Okada, K. (2002). Non-sequence-specific DNA binding by the Filamentous Flower protein from Arabidopsis thaliana is reduced by EDTA. *Journal of Biological Chemistry,* **277**(14), 11957-11964.

Koralewski, T. E., & Krutovsky, K. V. (2011). Evolution of exon-intron structure and alternative splicing. PloS one, **6**(3), e18055.

Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Molecular Biology and Evolution,* **35**(6), 1547-1549.

Kumaran, M. K., Bowman, J. L., & Sundaresan, V. (2002). *YABBY* polarity genes mediate the repression of KNOX homeobox genes in Arabidopsis. *The Plant Cell***, 14(**11), 2761-2770.

Kyriacou, M. C., Leskovar, D. I., Colla, G., & Rouphael, Y. (2018). Watermelon and melon fruit quality: The genotypic and agro-environmental factors implicated. *Scientia Horticulturae,* **234,** 393-408.

Lee, J.-Y., Baum, S. F., Alvarez, J., Patel, A., Chitwood, D. H., & Bowman, J. L. (2005). Activation of CRABS CLAW in the nectaries and carpels of Arabidopsis. *The Plant Cell, 17*(1), 25-36.

Letunic, I., & Bork, P. (2018). 20 years of the SMART protein domain annotation resource. *Nucleic acids research, 46*(D1), D493-D496.

Li, Z., Li, G., Cai, M., Priyadarshani, S. V., Aslam, M., Zhou, Q., Qin, Y. (2019). Genome-wide analysis of the *YABBY* transcription factor family in pineapple and functional identification of Ac*YABBY*4 involvement in salt stress. *International journal of molecular sciences,* **20(**23), 5863.

Lijavetzky, D., Carbonero, P., & Vicente-Carbajosa, J. (2003). Genome-wide comparative phylogenetic analysis of the rice and Arabidopsis Dof gene families. *BMC evolutionary biology,* **3**(1), 17.

Lijavetzky, D., Carbonero, P., & Vicente-Carbajosa, J. (2003). Genome-wide comparative phylogenetic analysis of the rice and Arabidopsis Dof gene families. [Comparative Study

Lu, S., Wang, J., Chitsaz, F., Derbyshire, M. K., Geer, R. C., Gonzales, N. R., Marchler-Bauer, A. (2020). CDD/SPARCLE: the conserved domain database in 2020. [Research Support, N.I.H., Intramural]. *Nucleic Acids Res,* **48(**D1), D265-D268.

Ma, J., Li, M. Y., Wang, F., Tang, J., & Xiong, A. S. (2015). Genome-wide analysis of Dof family transcription factors and their responses to abiotic stresses in Chinese cabbage. *BMC Genomics,* **16**, 33-33.

Malviya, N., Gupta, S., Singh, V., Yadav, M., Bisht, N., Sarangi, B., & Yadav, D. (2015). Genome wide in silico characterization of Dof gene families of pigeonpea (Cajanus cajan (L) Millsp.). *Molecular biology reports,* **42**(2), 535-552.

Maoto, M. M., Beswa, D., & Jideani, A. I. (2019). Watermelon as a potential fruit snack. *International Journal of food properties,* **22**(1), 355-370.

Mehra, M., Pasricha, V., & Gupta, R. K. (2015). Estimation of nutritional, phytochemical and antioxidant activity of seeds of musk melon (Cucumis melo) and water melon (Citrullus lanatus) and nutritional analysis of their respective oils. *Journal of Pharmacognosy and Phytochemistry,* **3**(6), 98-102.

Moore, R. C., & Purugganan, M. D. (2005). The evolutionary dynamics of plant duplicate genes. *Current opinion in plant biology,* **8**(2), 122-128.

Morgan, C. C., Loughran, N. B., Walsh, T. A., Harrison, A. J., & O'Connell, M. J. (2010). Positive selection neighboring functionally essential sites and disease-implicated regions of mammalian reproductive proteins. *BMC evolutionary biology*, **10**(1), 39.

Nasim, J., Malviya, N., Kumar, R., & Yadav, D. (2016). Genome-wide bioinformatics analysis of Dof transcription factor gene family of chickpea and its comparative phylogenetic assessment with Arabidopsis and rice. *Plant Systematics and Evolution,* **302**(8), 1009-1026.

Naz, A., Butt, M. S., Sultan, M. T., Qayyum, M. M. N., & Niaz, R. S. (2014). Watermelon lycopene and allied health claims. *EXCLI journal*, **13**, 650.

Nosaka, M., Ono, A., Ishiwata, A., Shimizu-Sato, S., Ishimoto, K., Noda, Y., & Sato, Y. (2013). Expression of the rice microRNA miR820 is associated with epigenetic modifications at its own locus. *Genes Genet Syst,* **88(**2), 105-112.

Oberoi, D. P. S., & Sogi, D. S. (2017). Utilization of watermelon pulp for lycopene extraction by response surface methodology. *Food Chemistry,* **232**, 316-321.

Panchy, N., Lehti-Shiu, M., & Shiu, S. H. (2016). Evolution of gene duplication in plants. Plant physiology, **171**(4), 2294-2316.

Research Support, Non-U.S. Gov't]. *BMC Evol Biol, 3*, 17.

Research Support, U.S. Gov't, Non-P.H.S.]. *PLoS One,* **6**(3), e18055.

Review]. *Plant Physiol,* **171**(4), 2294-2316.

Rombauts, S., Déhais, P., Van Montagu, M., & Rouzé, P. (1999). PlantCARE, a plant cis-acting regulatory element database. *Nucleic acids research,* **27**(1), 295-296.

Romdhane, M. B., Haddar, A., Ghazala, I., Jeddou, K. B., Helbert, C. B., & Ellouz-Chaabouni, S. (2017). Optimization of polysaccharides extraction from watermelon rinds: Structure, functional and biological activities. *Food Chemistry,* **216**, 355-364.

Samad, A. F. A. (2017). MicroRNA and Transcription Factor: Key Players in Plant Regulatory Network. *Frontiers in Plant Science*. **8**, 565.

Sawa, S., Watanabe, K., Goto, K., Kanaya, E., Morita, E. H., & Okada, K. (1999). Filamentous Flower, a meristem and organ identity gene of Arabidopsis, encodes a protein with a zinc finger and HMG-related domains. *Genes & development,* **13**(9), 1079-1088.

Sharif, R., Xie, C., Wang, J., Cao, Z., Zhang, H., Chen, P., & Li, Y. (2020). Genome wide identification, characterization and expression analysis of HD-ZIP gene family in Cucumis sativus L. under biotic and various abiotic stresses. *International journal of biological macromolecules,* **158**, 502-520.

Siegfried, K. R., Eshed, Y., Baum, S. F., Otsuga, D., Drews, G. N., & Bowman, J. L. (1999). Members of the *YABBY* gene family specify abaxial cell fate in Arabidopsis. *Development,* **126**(18), 4117-4128.

Song, Q., Joshi, M., & Joshi, V. (2020). Transcriptomic Analysis of Short-Term Salt Stress Response in Watermelon Seedlings. *International journal of molecular sciences, 21*(17), 6036-6036.

Song, Q., Joshi, M., DiPiazza, J., & Joshi, V. (2020). Functional relevance of citrulline in the vegetative tissues of watermelon during abiotic stresses. *Frontiers in Plant Science,* **11**, 512-512.

Song, Q., Joshi, M., DiPiazza, J., & Joshi, V. (2020). Functional relevance of citrulline in the vegetative tissues of watermelon during abiotic stresses. Frontiers in Plant Science, 11, 512.

Spanudakis, E. (2014). The role of microRNAs in the control of flowering time. *J Exp Bot*.

Taylor, J. S., & Raes, J. (2004). Duplication and divergence: the evolution of new genes and old ideas. *Annu. Rev. Genet.,* **38**, 615-643.

Terzi, L. C. (2008). Regulation of Flowering Time by RNA Processing. *Springer Link*.

Thompson, J. D., Higgins, D. G., & 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 research,* **22**(22), 4673-4680.

Toriba, T., Harada, K., Takamura, A., Nakamura, H., Ichikawa, H., Suzaki, T., & Hirano, H.-Y. (2007). Molecular characterization the *YABBY* gene family in Oryza sativa and expression analysis of Os*YABBY*1. *Molecular genetics and genomics,* **277**(5), 457-468.

Villanueva, J. M., Broadhvest, J., Hauser, B. A., Meister, R. J., Schneitz, K., & Gasser, C. S. (1999). INNER NO OUTER regulates abaxial–adaxial patterning in Arabidopsis ovules. *Genes & development,* **13**(23), 3160-3169.

Wang, P., Wang, S., Chen, Y., Xu, X., Guang, X., & Zhang, Y. (2019). Genome-wide analysis of the MADS-Box gene family in watermelon. *Computational biology and chemistry,* **80,** 341-350.

Wang, Y., Li, J., & Paterson, A. H. (2013). MCScanX-transposed: detecting transposed gene duplications based on multiple colinearity scans. [Research Support, U.S. Gov't, Non-P.H.S.]. *Bioinformatics,* **29**(11), 1458-1460

Wei, Q., Wang, W., Hu, T., Hu, H., Mao, W., Zhu, Q., & Bao, C. (2018). Genome-wide identification and characterization of Dof transcription factors in eggplant (Solanum melongena L.). *PeerJ,* **6**, e4481.

Wen, C. L., Cheng, Q., Zhao, L., Mao, A., Yang, J., Yu, S.,Xu, Y. (2016). Identification and characterisation of Dof transcription factors in the cucumber genome. [Research Support, Non-U.S. Gov't]. *Sci Rep,* **6**, 23072.

Wu, J., Fu, L., & Yi, H. (2016). Genome-Wide Identification of the Transcription Factors Involved in Citrus Fruit Ripening from the Transcriptomes of a Late-Ripening Sweet Orange Mutant and Its Wild Type. [Research Support, Non-U.S. Gov't]. *PLoS One,* **11**(4), e0154330.

Xu, M., Hu, T., Zhao, J., Park, M.-Y., Earley, K. W., Wu, G., Poethig, R. S. (2016a). Developmental Functions of miR156-Regulated Squamosa Promoter Binding Protein-Like (SPL) Genes in Arabidopsis thaliana. *PLoS genetics,* **12**(8), e1006263-e1006263.

Xu, M., Hu, T., Zhao, J., Park, M.-Y., Earley, K. W., Wu, G., Poethig, R. S. (2016b). Developmental Functions of miR156-Regulated Squamosa Promoter Binding Protein-Like (SPL) Genes in Arabidopsis thaliana. *PLOS Genetics,* **12(**8), e1006263.

Xu, P., Chen, H., Ying, L., & Cai, W. (2016a). AtDOF5.4/OBP4, a DOF Transcription Factor Gene that Negatively Regulates Cell Cycle Progression and Cell Expansion in Arabidopsis thaliana. [Research Support, Non-U.S. Gov't]. *Sci Rep,* **6,** 27705.

Xu, P., Chen, H., Ying, L., & Cai, W. (2016b). AtDOF5.4/OBP4, a DOF Transcription Factor Gene that Negatively Regulates Cell Cycle Progression and Cell Expansion in Arabidopsis thaliana. *Sci Rep,* **6**, 27705-27705.

Yamada, T., Ito, M., & Kato, M. (2004). *YABBY*2-homologue expression in lateral organs of Amborella trichopoda (Amborellaceae). *International journal of plant sciences, 165*(6), 917-924.

Yamaguchi, T., Nagasawa, N., Kawasaki, S., Matsuoka, M., Nagato, Y., & Hirano, H.-Y. (2004). The *YABBY* gene DROOPING LEAF regulates carpel specification and midrib development in Oryza sativa. *The Plant Cell,* **16**(2), 500-509.

Yanagisawa, S., & Schmidt, R. J. (1999). Diversity and similarity among recognition sequences of Dof transcription factors. *The Plant Journal,* **17**(2), 209-214.

Yang, X., & Tuskan, G. A. (2006). Divergence of the Dof gene families in poplar, Arabidopsis, and rice suggests multiple modes of gene evolution after duplication. *Plant physiology,* **142**(3), 820-830.

Yang, Z., & Bielawski, J. P. (2000). Statistical methods for detecting molecular adaptation. *Trends in ecology & evolution,* **15**(12), 496-503.

Yang, Z., Gong, Q., Wang, L., Jin, Y., Xi, J., Li, Z., . Chen, Q. (2018). Genome-wide study of *YABBY* genes in upland cotton and their expression patterns under different stresses. *Frontiers in genetics,* **9**, 33.

Yin, S., Li, S., Gao, Y., Bartholomew, E. S., Wang, R., Yang, H., Liu, X. (2022). Genome-Wide Identification of *YABBY* Gene Family in Cucurbitaceae and Expression Analysis in Cucumber (Cucumis sativus L.). *Genes,* **13**(3), 467.

Yin, S., Li, S., Gao, Y., Bartholomew, E. S., Wang, R., Yang, H., Liu, X. (2022). Genome-wide identification of YABBY gene family in Cucurbitaceae and expression analysis in Cucumber (Cucumis sativus L.). Genes, **13**(3), 467.

Yuan, J., Shen, C., Xin, J., Li, Z., Li, X., & Zhou, J. (2020). Genome-wide Characterization and Expression Analysis of *YABBY* Gene Family in Three Cultivars of Cucurbita Linn. And Their Response of Salt Stress in Cucurbita Moschata.

Zhang, S., Wang, L., Sun, X., Li, Y., Yao, J., Nocker, S. v., & Wang, X. (2019). Genome-wide analysis of the *YABBY* gene family in grapevine and functional characterization of Vv*YABBY*4. *Frontiers in plant science,* **10,** 1207.

Zhao, S.-P., Lu, D., Yu, T.-F., Ji, Y.-J., Zheng, W.-J., Zhang, S.-X., . Cui, X.-Y. (2017). Genome-wide analysis of the *YABBY* family in soybean and functional identification of Gm*YABBY*10 involvement in high salt and drought stresses. *Plant Physiology and Biochemistry***, 119,** 132-146.