J. Appl. Res Plant Sci. Vol. 3(2), 325-341, 2022 eISSN 2708-3004; pISSN: 2708-2997
www.joarps.org https://doi.org/10.38211/joarps.2022.3.2.40

Genome-wide Identification and Characterization of Plant-specific Transcription Factor YABBY Gene Family in Cucumber (*Cucumis sativus*) and its Comparison with *Arabidopsis* to Reveal its Role in Abiotic Stress Responses

Muhammad Muneeb Hashmi[1], Zuha Kamran[1], Mujahid Manzoor*[2], Muhammad Shafiq[1], Masah Qamar[1], Mehr-Un-Nisa[1], Muhammad Ashfaq[5], Muhammad Adnan Shahid[3], Muhammad Saleem Haider[4]

[1]Department of Horticulture, Faculty of Agricultural Sciences, University of the Punjab, Lahore, Pakistan
[2]Department of Entomology, Faculty of Agricultural Sciences, University of the Punjab, Lahore, Pakistan

[3]North Florida Research and Education Center 155 Research Rd. Quincy, FL 32351
[4]Department of Plant Pathology, Faculty of Agricultural Sciences, University of the Punjab, Lahore, Pakistan
[5]Department of Plant Breeding and Genetics. Faculty of Agricultural Sciences, University of the Punjab Lahore Pakistan.
*Corresponding author's email; mujahid.iags@pu.edu.pk

Article Received 11-10-2022, Article Revised 01-11-2022, Article Accepted 07-11-2022

Abstract

Plant-specific transcription factor (PSTFs) YABBY is one of the vital transcription factors that play a crucial role in abaxial organ development, carpel formation and abiotic stress. Although the Cucumber genome (*Cucumis sativus*) has been published, functional studies are still needed to understand cucumber. The cucumber genome was used in this study to identify YABBY gene family member by using a set of various bioinformatic tools. Eight YABBY gene family members were identified that were unevenly distributed on different chromosomes. Eight members of the YABBY gene family in cucumber were divided into five subgroups (FIL/YAB3), CRC, INO, YAB2, and YAB5 based on the published Arabidopsis YABBY gene classification. The structure of PSTF YABBY was seen to be conserved throughout the process of evolution through Motif analysis, Conserved Domain Analysis and Gene structure Intron Exon Display. PSTF YABBY has roles in wound healing, abiotic stress like cold, heat and drought stress, phytohormone responses and transcription initiation. *CsYABBY*4 was seen to be over-expressed under long day and heat stress conditions, implying its significant role in heat stress.

Keywords: Cucumber, YABBY, Plant specific transcription factor, Genome wide, Bioinformatics.

Introduction

YABBY is the unique plant-specific transcription factor (PSTFs) Zinc Finger domain encoding gene family and its role is indispensable in lateral organ and the development of primordia of leaves (Li *et al.*, 2019); (Hudson *et al.*, 2004). PSTFs YABBY gene member has the YABBY domain in C-terminal and C_2C_2 Zinc finger (ZF) at N-terminal (Yuan *et al.*, 2020). The C-terminal helix-loop-helix domains (YABBY) coincide with a high-mobility group (HMG) box and the YABBY family proteins are named for this specific feature. The highly conserved amino acid residues in the two domains of YABBY proteins have shown alliance with DNA specificity. The YABBY gene family has several subfamilies. Each subfamily encodes YABBY proteins with other conserved structural domains, resulting in the peculiar feature of that family appearing (T Zhang *et al.*, 2020). The genome-wide analysis of A. thaliana scrambles 6 YABBY members, further divided into five subfamilies; YAB5, CRC, YAB3 or FIL, YAB2 and INO. YAB5, (FIL) Filamentous Flower and YAB3 assist lateral organ development in a recurring order. CRABS CLAW (CRC) plays its role in initiating the development of poles in the nectaries and carpels under the different stages of development (S Zhang *et al.*, 2019a). Inner No Outer (INO) is responsible for the synthesis of ovule outer integuments and seed coat (S Zhang *et al.*, 2019a). In the case of *A. thaliana*, YABBY genes are also involved in controlling the abaxial identity of lateral organs (Hudson *et al.*, 2004). AFO/FIL involved in the formation of the shoot apical and flower meristems, filamentous flower, YAB3 and YAB2 are confined to Crabs's Claw; according to homology, these genes express themselves in a polar manner in almost all the primordia of lateral organs YABBY transcription factor plays extensive functions in the expansion of leaves, development of reproductive organs, dorsoventral polarity establishment in stress and phytohormone responses (T Zhang *et al.*, 2020). Many observations have depicted that YABBY genes in monocots might function differently in lateral organ formation in contrast to their homologs in *A. thaliana* (Hudson *et al.*, 2004) The most remarkable about YABBY is that it acts both as an activator and inhibitor in regulating abiotic stress in plants (Yuan *et al.*, 2020). Earlier studies revealed that it was noted that the function of the YABBY gene was only restricted to FIL and YAB3 in leaves but recent studies tell us that during the development of leaf, it's possible to observe the similar function of YAB5 and YAB2 with the other two

vegetative YABBY genes. In a genetic approach, it is illustrated that YAB5, YAB3, FIL and YAB2 which are the vegetative genes of the YABBY gene family, regulates the patterning of the embryo and the growth of the lamina of the leaf. The absence of YABBY activities affects the earlier stages of lamina development and leaf primordia maturation programs (Sarojam *et al.*, 2010). In Pineapple, RT-qPCR revealed the expression of 3 members of the YABBY gene family to have the function during abiotic stress. According to some experiments, under NaCl stress, the AcYABBY4 overexpressed itself in *A. thaliana* and out turned into short roots, which denoted that AcYABBY4 functions as an inhibitor in salt tolerance (Yuan *et al.*, 2020). Globally, there are two most important cultivated vegetable crop species belonging to the genus Cucumis (family Cucurbitaceae), *C. sativus* (cucumber: 2n=2x=14) and *C. melo* (melon: 2n=2x=24) (Cavagnaro *et al.*, 2010). The origin of cucumber was in India about 3000 years ago and after that it was cultivated in the East and South of the Himalayas. Later then, Cucumber was brought to Greece, Italy and China (Bisognin, 2002). *C. sativus* is a frost-susceptible plant grown best at a temperature above 20℃ (Tatlioglu, 1993). The cucumber is an annually growing vine that has setaceous hairy climbing type. Hairy margin denticulate in both surfaces of simple alternate deeply cordate 3-5 lobed leaves (Mukherjee *et al.*, 2013). The cucumber fruit is green in color and the size depends upon the variety. It consists of 96% water and a lot of other vitamins, minerals and an organic acid, increasing its nutritional value (Mousavizadeh *et al.*, 2010). The main aim of this research was to classify and identify YABBY genes belonging to gene family YABBY in C. sativus with the help of various bioinformatics tools. For the identification of YABBY gene in *C. sativus*, a systematic approach was adopted. The distribution of YABBY genes on chromosomes, the distribution pattern of exons or introns, the *cis-regulatory* elements and conserved domains were discovered. To determine the orthologous relationship of the Cucumber with *C. maxima, M. acuminata and A. thaliana*, a phylogenetic tree with neighbor joining format was also constructed with the YABBY genes of all these species and also to predict their possible functions. YABBY genes' genome-wide analysis in *C. sativus* might lead to the development of abiotic stress-resistant varieties.

Materials and methods

Retrieval of sequence and Database search: Amino acid sequence of YABBY domain was retrieved from Pfam database (http://pfam.xfam.org/) (Finn *et al.*, 2014). Phytozome (https://phytozome.jgi.doe.gov/pz/portal.html) was used to retrieve peptide sequences of Cucumber using BLAST-P (Protein-basic local alignment search tool) program (Goodstein *et al.*, 2014). The peptide sequence which was retrieved was submitted to NCBI Conserved Domain's Database's simple modular architecture. (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) with the default parameters. The protein sequences which lacked real conserved domain of YABBY in their sequence were excluded. (PF04690.13) (http://pfam.xfam.org/family/PF04690.13) (Lu *et al.*, 2020).

Determination of physiochemical properties of *CsYABBY*: Protoparam tool was used to predict the physiochemical properties; length of protein, value of theoretical pi and relative molecular weight of *CsYABBY* proteins. (http://web.expasy.org/protparam/) (Gasteiger *et al.*, 2005).

Gene structure analysis: To identify the arrangement of exons and introns of identified YABBY genes member in cucumber, Phytozom Cucumber genome database (https://phytozome.jgi.doe.gov/pz/portal.html). was used to retrieve the coding and genomic sequence. These sequences were further used to draw the gene structure using Gene Structure Display Server (GSDS v2.0) (available at http://gsds.cbi.pku.edu.cn/) (Hu *et al.*, 2015).

Domain and motif recognition: Multiple EM for Motif Elicitation (MEME) program was used to analyse the motif structure and organization in YABBY peptide sequences of *C. sativus, C. maxima, A. thaliana* and *M. acuminate* (http://meme.nbcr.net/meme/) with the maximum number of motifs set as 20. The minimum and maximum width of motif were set to 6 and 50 (Bailey *et al.*, 2015).

Multiple sequence alignment and phylogenetic analysis: The amino acid sequences of *YABBY* proteins were aligned using Clustal W version 2.1 (Thompson *et al.*, 2003; Thompson *et al.*, 1994) and the phylogeny analysis was done through MEGA X v2.0 (Kumar *et al.*, 2018). A neighbor-joining tree with bootstrap value 1000 was constructed using phylogeny through Mega X.v2.0. Eight sequences of cucumber, 13 sequences of *C. maxima*, 25 sequences of *M. acuminata* and six sequences of *A. thaliana* were used for phylogenetic analysis.

Cis-regulatory elements Analysis: For the analysis of promoter regions, a sequence of 1000 bp upstream was retrieved from the initiation codon for each putative CsYABBY gene. Plant Care database (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) (Rombauts *et al.*, 1999) was then used to predict cis-regulatory elements in these sequences and validated in the PLACE databases (http://www.dna.afrc.go.jp/PLACE/) (Higo *et al.*1998, 1999). For the analysis of promoter region, 1000-bp upstream was retrieved from the Plant Care database (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) (Rombauts *et al.*, 1999) was used to retrieve 1000 bp upstream the promoter region. Sequences were retrieved from the initiation codon of YABBY genes. The sequences were further validated in PLACE databases

(http://www.dna.affrc.go.jp/PLACE/) (Higo *et al.*, 1998; Higo *et al.*, 1999). Evaluation of the rate of synonymous (Ks) and non-synonymous (Ks) substitution: The values of Ka and Ks were also figured out for *CsYABBYs.* For this purpose, *CsYABBY* gene pair duplicates, for the calculation of Ka and Ks values arising from different modes of duplication were used for the calculation of the Ka and Ks values substitution rates. TBtools was used for the estimation of Ka and Ks values of the duplicated gene pairs along with their ratios of Ka/Ks, using the Simple Ka/Ks calculator option. For this, the Ka/Ks ratio was evaluated to ferret the molecular evolutionary rates of each gene pair. Moreover, the estimation of the time of divergence of these pairs of genes was done using the formula "T=Ks/2λ (λ=6.5×10 e-9)'' represents neutral substitution.

Synteny analysis: Advanced Circos Toolkit along with default parameters was used to interpret the gene duplication events (Wang *et al.*, 2013). To exhibit the synteny relationship of the paralogous *YABBY* genes obtained from Cucumber, the maps of syntenic analysis were made using the Micro Synteny view software in TBtools (Chen *et al.*, 2020).

Transcriptome analysis: The data which was previously generated on RNA-seq cucumber shoot apex under various conditions of heat and photoperiodism was used for analyzing the organ-specific expression at various stages of development (Ramirez-Tejero *et al.*, 2020). RPKM (Reads per Kilobases per million mapped reads) values were converted to $\log_2$. The patterns of expressions under various conditions are illustrated in the heatmap through TBtools (Chen *et al.*, 2020).

Analysis of target sites of putative microRNA: NCBI Gene Expression Omnibus was used to retrieve the data sets of micro RNAs of cucumber in an experiment involving micro-RNAs and their targets in cucumber shoot apices in response to temperature and photoperiod (X Zhang *et al.*, 2018). There, to find out the miRNAs which target the *CsYABBY* genes, CDS sequences of all *CsYABBY* genes were retrieved for the complementary sequences of miRNAs with the help of RNA target (https://plantgrn.noble.org/psRNATarget/analysis?function=3) with default parameters (Samad, 2017).

Subcellular localization: Wolf Port was used to retrieve the subcellular localization of *CsYABBYs.* (https://wolfpsort.hgc.jp/). By entering the multicast format protein sequence of the desired protein, the required database of subcellular localization was redirected.

Dual Syntenny Analysis: The peptide, genomic and gff files of *C. sativus*, *C. maxima*, and *C. lanatus* were retrieved from Cucurbita Genomic Database (http://cucurbitgenomics.org/blast), and that of *A. thaliana*, *M. acuminata*, and *Glycine max* were retrieved from Phytozome (https://phytozome-next.jgi.doe.gov/). One step Multiple Collinear Scanning Toolkits (MCScan) from TBtools was used to analyze the gene replication events and synt relationships among Cucumber and Arabidopsis, Cucumber and Banana, Cucumber and Watermelon, and Cucumber and Soya bean. The orthologs for YABBY genes between Cucumber and Arabidopsis, Banana, Soybean and Watermelon were found using Dual Syntenny Plotter using the collinearity, gff and ctl file retrieved from MCScan in TBtools.

Result:

Identification of the YABBY in Cucumber*:* BLAST YABBY domain against the genome of *C. sativus* in the phytozome database led to the initial retrieval of 16 peptide sequences. The identical gene isoforms encoding the same peptide sequence were excluded. For further identification and characterization, eight non redundant YABBY protein sequences were isolated, which possessed C- the terminal YABBY domain and N-terminal zinc finger domain as their conserved domains. YABBY transcription factor's typical feature is its possession of these conserved domains. Within the highly conserved sequences of the cucumber C_2C_2 YABBY domain, 55 amino acids were found to be constitutive of the C_2C_2 domain. The amino acids include Cystine, Asparagine, Phenylalanine, Glycine, Leucine, Aspartic Acid, Serine, Threonine, Valine, Alanine, Proline, Tyrosine, Arginine, Histidine, Tryptophan, Methionine, and Lysine (as shown in fig.1).

Determination of physio-chemical properties of cucumber YABBY proteins: The number of amino acids ranged from 173 to 194; molecular weight ranged from 19.2 to 21.5 kDa in the peptide sequences encoding YABBY proteins. YABBY-1 was observed to be the smallest and YABBY-5 was observed to be the largest protein. The value of isoelectric points of the peptide sequences ranged from 7.04 to 9.51 (Table1).

J. Appl. Res Plant Sci. Vol. 3(2), 325-341, 2022 eISSN 2708-3004; pISSN: 2708-2997
www.joarps.org https://doi.org/10.38211/joarps.2022.3.2.40

Sequence logo:

Figure 1: The sequence logos are based on alignments of all Cucumber *YABBY* domains. *YABBY* domains are highly conserved across all 51 *YABBY* proteins in Cucumber. Multiple alignment analysis of 51 typical Cucumber *YABBY* domains was performed with ClustalW. The bit score indicates the information content for each position in the sequence. (Cys) in the *YABBY* domain are conserved and are present at the position no 1, 4, 12, 26 and 29. The zinc finger motif is also indicated as the green line

Gene name	**Accession no**		**Chromosome no**	**Chromosome location**	**direction**	**size**		**pI**	**Mw**
Name	**phytozome**	**V2**	**No.**	**(bp)**		**mRNA lengh**	**peptide**		**(KD)AA**
CsYABBY1	Cucsa.086660	Csa6G426940.1	00873	49827..56070	R	522	173	9.51	19.26185
CsYABBY2	Cucsa.132170	Csa2G006820.1	01037	199733..203179	R	579	192	8.88	21.52624
CsYABBY3	Cucsa.284430	Csa5G606780.1	02653	1208908..1210949	R	528	175	8.71	19.29361
CsYABBY4	Cucsa.286190	Csa5G600930.1	5	2307401..2308769	R	573	190	8.09	21.7921
CsYABBY5	Cucsa.295080	CsaUNG026620.1	02852	60287..63809	F	836	277	9.03	3.1795
CsYABBY6	Cucsa.302100	Csa5G160210.1	5	322496..324780	F	747	248	8.66	27.43895
CsYABBY7	Cucsa.320600	Csa3G033810.1	3	380494..384914	F	558	185	8.85	20.59038
CsYABBY8	Cucsa.394330	Csa2G348870.1	2	317693..319054	R	585	194	7.04	21.59577
CmYABBY1	CmaCh02G015570.1		02	8833805 ..8835686	R	1023	231	8.80	25.57087
CmYABBY2	CmaCh15G011490.1		15	7308199 ..7310013	R	1077	236	8.80	25.92929
CmYABBY3	CmaCh06G007630.1		06	3875334 ..3877797	F	1171	201	8.61	22.19604
CmYABBY4	CmaCh11G005360.1		11	2583470 ..2590191	F	912	226	8.06	25.34608

CmYABBY5	CmaCh05G012870.1		05	9820965.. 9824076	F	919	193	8.95	21.7156
CmYABBY6	CmaCh02G012970.1		02	7630829 ..7634766	F	940	179	9.3	19.97884
CmYABBY7	CmaCh18G010980.1		19	8991641 ..8995269	R	660	219	9.00	23.6531
CmYABBY8	CmaCh12G012020.1		12	9427635 ..9431144	F	1031	189	8.86	21.17989
CmYABBY9	CmaCh04G017410.1		04	8746724 ..8752040	R	1179	392	8.19	43.63639
CmYABBY10	CmaCh05G000220.1		05	105423 ..106388	R	576	191	6.37	21.36642
CmYABBY11	CmaCh04G017090.1		04	8594676 .. 8596684	F	537	178	8.81	19.47912
CmYABBY12	CmaCh18G010980.1		18	8991641 .. 8995269	R	660	219	9.00	23.65351
CmYABBY13	CmaCh20G001390.1		20	671323 ..674795	F	1264	201	9.69	22.21354
AtYABBY1	AT1G08465	AT1G08465.1	01	2675813..2679824	F	555	184	9.40	20.70052
AtYABBY2	AT1G23420	AT1G23420.1	01	8317297..8319491	F	789	231	5.93	25.95731
AtYABBY3	AT1G69180	AT1G69180.1	01	26007350..26009141	R	546	181	9.56	19.72248
AtYABBY4	AT2G26580	AT2G26580.2	2	11303455..11307010	R	495	164	9.47	18.50519
AtYABBY5	AT2G45190	AT2G45190.1	02	18628252..18630779	R	690	229	6.87	25.77942
AtYABBY6	AT4G00180	AT4G00180.1	04	72545..75576	R	723	240	8.67	26.33796
MaYABBY1	GSMUA_Achr11G03800_001	GSMUA_Achr11P03800_001	11	2745059..2747145	F	657	218	6.88	24.0
MaYABBY2	GSMUA_Achr11G17520_001	GSMUA_Achr11P17520_001	11	19080472..19082313	F	663	220	7.66	24.0
MaYABBY3	GSMUA_Achr1G04480_001	GSMUA_Achr1P04480_001	1	3755485..3760510	F	564	187	8.33	20.9
MaYABBY4	GSMUA_Achr1G24490_001	GSMUA_Achr1P24490_001	1	18879686..18883701	R	597	198	9.31	22.3
MaYABBY5	GSMUA_Achr1G27150_001	GSMUA_Achr1P27150_001	1	23732533..23735833	R	666	221	8.96	24.7
MaYABBY6	GSMUA_Achr2G12850_001	GSMUA_Achr2P12850_001	2	15161454..15167137	R	627	208	9.07	23.0
MaYABBY7	GSMUA_Achr3G15060_001	GSMUA_Achr3P15060_001	3	15562554..15565872	R	495	164	9.08	18.2
MaYABBY8	GSMUA_Achr3G25290_001	GSMUA_Achr3P25290_001	3	25524512..25526549	R	372	123	8.80	14.0
MaYABBY9	GSMUA_Achr3G25660_001	GSMUA_Achr3P25660_001	3	25766837..25768771	F	489	162	8.67	18.0
MaYABBY10	GSMUA_Achr3G30800_001	GSMUA_Achr3P30800_001	3	29409464..29411903	F	492	163	9.60	18.4

MaYABBY11	GSMUA_Achr4G01430_001	GSMUA_Achr4P01430_001	4	1200764..1203846	R	522	173	9.05	19.1
MaYABBY12	GSMUA_Achr4G22750_001	GSMUA_Achr4P22750_001	4	23037607..23044855	R	567	188	9.30	20.9
MaYABBY13	GSMUA_Achr4G31310_001	GSMUA_Achr4P31310_001	4	28558065..28561990	R	561	186	8.82	20.6
MaYABBY14	GSMUA_Achr5G08930_001	GSMUA_Achr5P08930_001	5	6501039..6504888	F	573	190	6.97	20.9
MaYABBY15	GSMUA_Achr5G23610_001	GSMUA_Achr5P23610_001	5	25301820..25303485	R	585	194	6.08	21.1
MaYABBY16	GSMUA_Achr6G24550_001	GSMUA_Achr6P24550_001	6	25378582..25382055	R	573	190	8.73	21.1
MaYABBY17	GSMUA_Achr6G31080_001	GSMUA_Achr6P31080_001	6	30976910..30982322	R	567	188	9.16	21.0
MaYABBY18	GSMUA_Achr7G01330_001	GSMUA_Achr7P01330_001	7	1036534..1042332	F	567	188	8.83	20.6
MaYABBY19	GSMUA_Achr7G07130_001	GSMUA_Achr7P07130_001	7	5220983..5222607	R	642	213	8.94	23.9
MaYABBY20	GSMUA_Achr7G07830_001	GSMUA_Achr7P07830_001	7	5813769..5818755	R	546	181	8.80	20.0
MaYABBY21	GSMUA_Achr8G04340_001	GSMUA_Achr8P04340_001	8	2905599..2907319	R	456	151	8.79	17.1
MaYABBY22	GSMUA_Achr8G04750_001	GSMUA_Achr8P04750_001	8	3154576..3156573	R	492	163	8.85	18.2
MaYABBY23	GSMUA_Achr8G11580_001	GSMUA_Achr8P11580_001	8	8361302..8363098	R	660	219	7.66	23.9
MaYABBY24	GSMUA_Achr8G26600_001	GSMUA_Achr8P26600_001	8	30103931..30107036	F	504	167	9.54	18.9
MaYABBY25	GSMUA_AchrUn_randomG19410_001	GSMUA_AchrUn_randomP19410_001		93420246..93422425	R	465	154	9.81	17.5

Gene structures analysis: To determine the evolutionary history and relations among genes or organisms, the orientation of introns and exons provides the grounds (Koralewski *et al*., 2011). The pattern of distribution and quantity of introns or exons serve as an evolutionary mark for the identified gene family. It was seen through the orientation of introns and exons that the pattern of the structure of the gene was consistent with phylogenetic analysis. The number of introns varied from one to nine in cucumber (Fig. 2). In group YAB2, containing *AtYAB2*, *CsYABBY*1, *CmYABBY6* and *CmYABBY13*, the number of introns varied from 5-7. In group YAB5 including *AtYAB5*, *CsYABBY*5, *CmYABBY4*, *CsYABBY*2, *CmYABBY5* and *CmYABBY8*, the number of introns varied from 6 to 8. In group AFO (YAB3) comprising of *MaYABBY1*, *MaYaBBY23*, *MaYABBY2*, *MaYABBY19*, *CsYABBY*7, *CmYABBY3*, *AtAFO*, *AtYAB3*, *CsYABBY*6, *CmYABBY1* and *CmYABBY2*, the number of introns varied from 5 to 7. The group CRC inclusive of *MaYABBY11*, *MaYABBY14*, *MaYABBY5*, *MaYABBY3*, *MaYABBY16*, *MaYABBY4*, *MaYABBY8*, *MaYABBY21*, *AtCRC*, *CmYABBY11*, *CsYABBY3*, *CmYABBY7* and *CmYABBY12*, the number of introns vary from 5 to 7. The group INO, including *MaYABBY15*, *AtINO*, *CsYABBY*4, *CsYABBY*8 and *CmYABBY10* contains 5-6 introns whereas *CmYABBY9* does not contain any introns. The number of introns in the members of different groups vary over a very narrow range (Fig. 2).

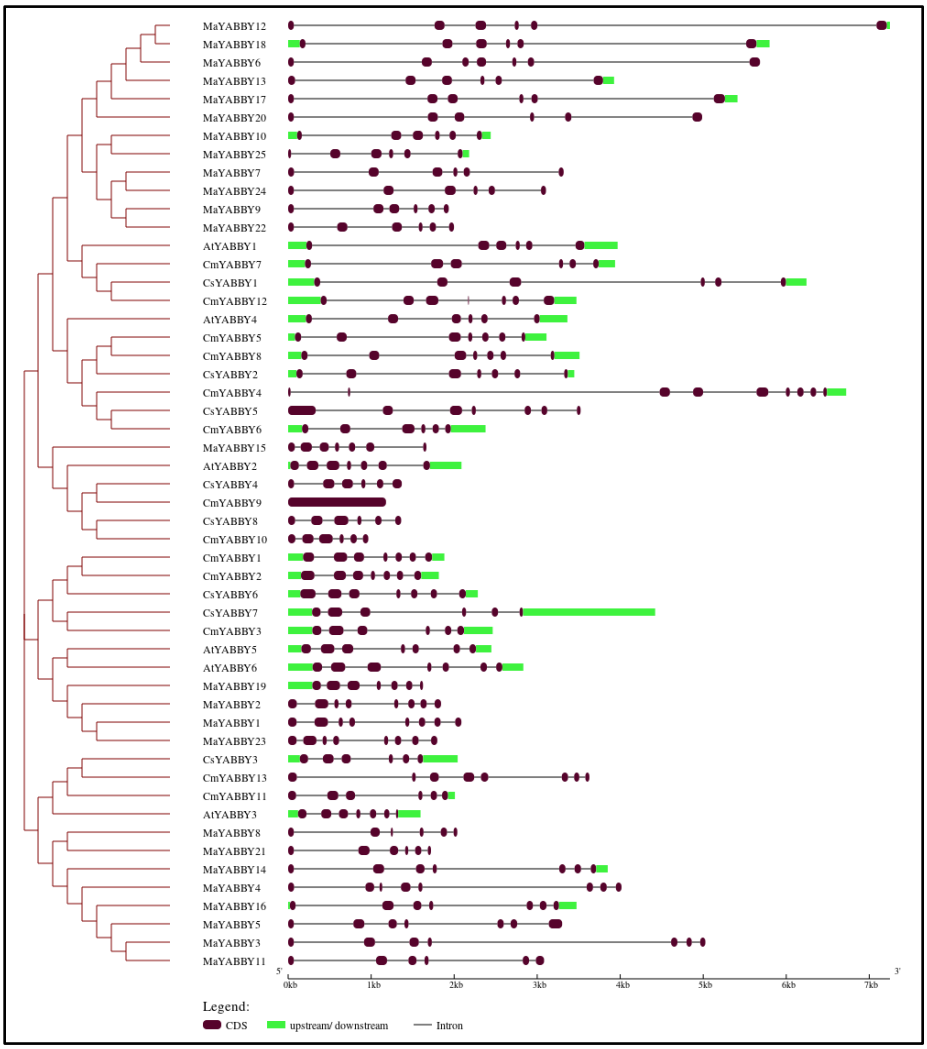

Figure 2: Gene structure intron exon display

Recognition of conserved motifs and domain: All members in group INO contained motifs 1 and 2. The majority of the members of the group CRC contained motifs 1, 2, 9 and 7. Only *AtCRC* and *MaYABBY4* lacked motif 9 and *AtCRC*, *CmYABBY11*, *CsYABBY*3, *CmYABBY7* and *CmYABBY12* lacked motif 7. Most members of group AFO/FIL/YAB3 contained motifs 1, 2, 3, 5, 6, 7 and 10. *CsYABBY*7 lacked motif 3 *CmYABBY3*

lacked motif 3 and 6, and *AtAFO* and *AtYAB3* lacked motif 5. In the family YAB5, motifs 1, 2, 3 and 4 were common in all the sequences, whereas, *AtYABBY5* lacked motif one and *CmYABBY8* lacked motif 3. In the family YAB2, motifs 2, 7 and 8 were common in all the sequences except *AtYAB2* lacked motif 8 and *CmYABBY6* had motif 3 (Fig. 3).

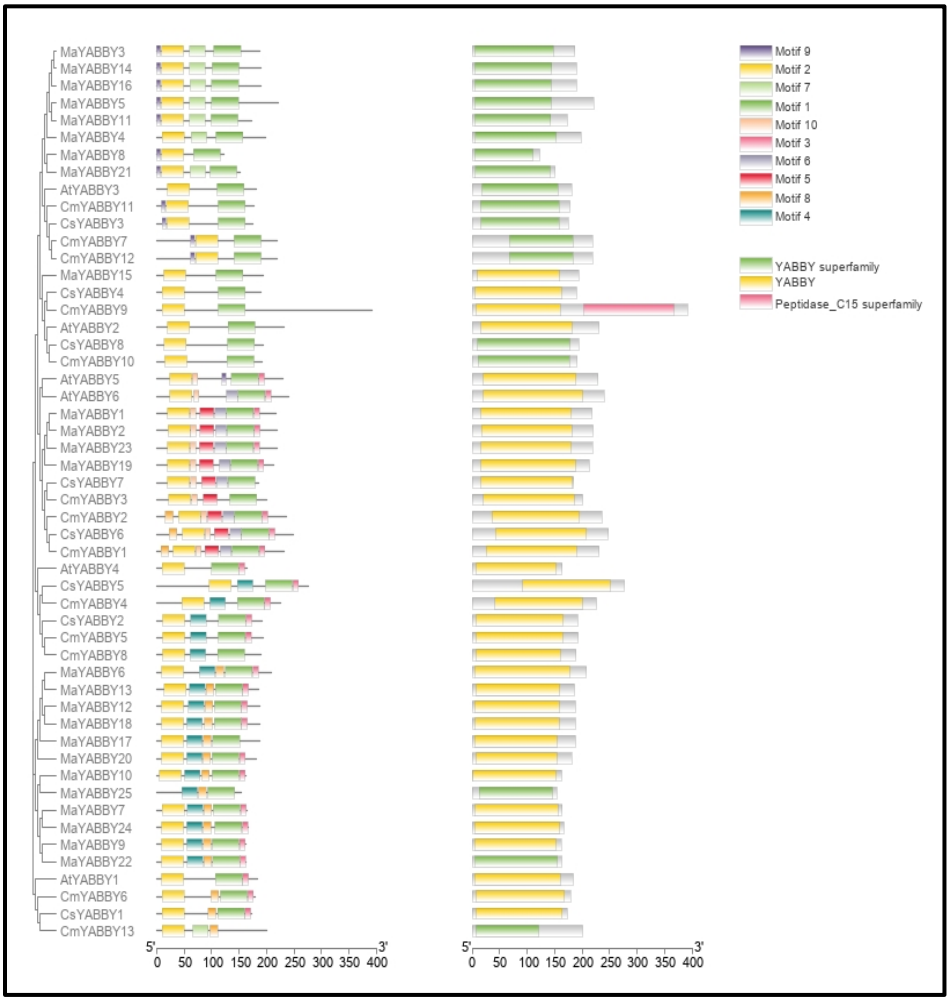

Figure 3: The distribution of 20 motifs on 51 *YABBY* proteins of Cucumber by using MEME version 4.9.0 and interlinking it with the phylogenetic tree to develop a good understanding of their association. The bars represent motifs with different color codes for different types of motifs.

Comparative phylogenetic relatedness of cucumber YABBY gene family with Arabidopsis, banana and *Cucurbita maxima*: Using Mega-X for aligning the complete length peptide sequences, a neighbor-joining tree was constructed to discover the evolutionary history between the YABBY transcription factors of *C. sativus*, *C. maxima*, *A. thaliana* and *M. acuminate*. A total of 54 YABBY proteins were depicted in cucumber, banana, *C. maxima* and *A. thaliana* and were distributed among five subgroups named CRC, INO, FIL/YAB3, YAB5,YAB2 (Table S1 and Fig. 4). Subgroup 1 named as CRC consisted of 13 proteins including *AtCRC*, *CsYABBY*3, *CmYABBY11*, *CmYABBY12*, *CmYABBY7*, *MaYABBY21*, *MaYABBY8*, *MaYABBY16*, *MaYABBY3*, *MaYABBY5*, *MaYABBY14* and *MaYABBY11*. The following subgroup was named INO and consists of 6 elements, including *AtINO*, *MaYABBY15*, *CsYABBY*4, *CmYABBY9*, *CsYABBY*8 and *CmYABBY10*. The third subgroup was called FIL/YAB3 and it included *AtYAB3*, *AtAFO*, *MaYABBY1*, *MaYABBY23*, *MaYABBY2*, *MaYABBY19*, *CsYABBY*7, *CmYABBY3*, *CsYABBY*6, *CmYABBY1*, and *CmYABBY2*. The other subgroup was YAB5 and included *AtYAB5*, *CsYABBY*5, *CmYABBY4*, *CsYABBY*2, *CmYABBY5*, and *CmYABBY8*. The last subgroup was YAB2 and consisted of *AtYAB2*, *CsYABBY*1, *CmYABBY6* and *CmYABBY13*. There were 12 sequences of *Musa acuminata; MaYABBY6*, *MaYABBY13*, *MaYABBY12*, *MaYABBY18*, *MaYABBY17*, *MaYABBY20*, *MaYABBY10*,

MaYABBY25, *MaYABBY22*, *MaYABBY9*, *MaYABBY7* and *MaYABBY24*, which did not fit in any group of YABBY, so, they were named as *MaYAB.* Proteins in a common clade usually seem to show similarity in structure and function. So, all the YABBY proteins in similar clades may have a similar structure and functions.

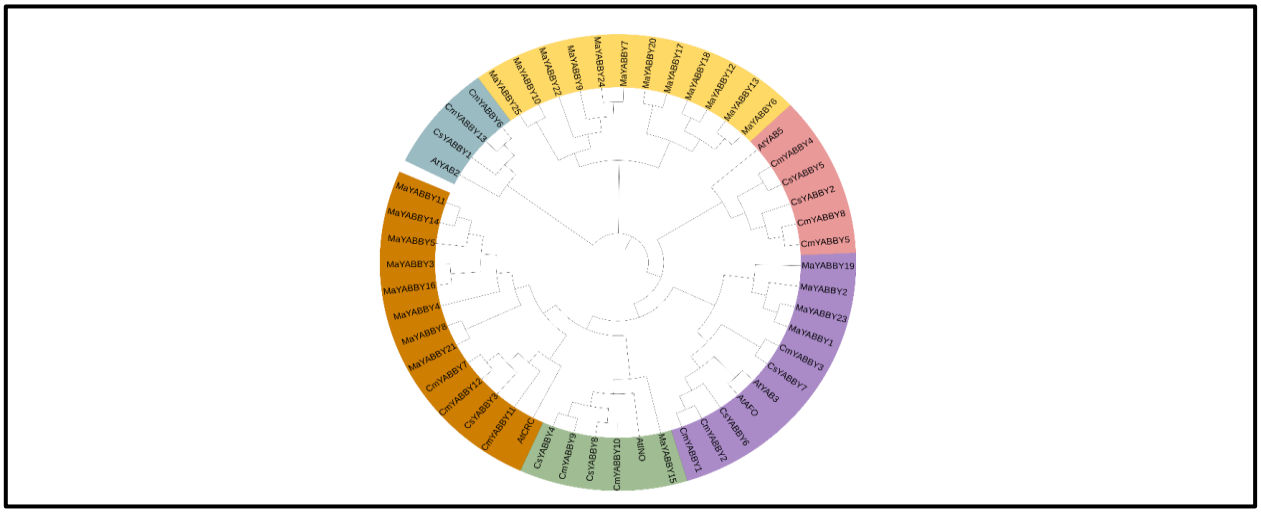

Figure 4: Phylogenetic Relationship between *CsYABBY* and *AtYABBY* proteins where *CsYABBY* proteins are marked with red stars. The evolutionary history was inferred using the UPGMA method with 1000 Bootstrap. This analysis involved 86 amino acid sequences. All ambiguous positions were removed for each sequence pair (pairwise deletion option). There was a total of 852 positions in the final dataset. Evolutionary analyses were conducted in MEGA X (P.H.A. *et al.*, 1973); (J., 1985); (E. *et al.*, 1965); (S. *et al.*, 2018))

Analysis of Cis-regulatory elements: The existence and orientation of many cis-regulatory elements (CREs) on the promoter region affect the expression of genes. An in-silico CREs analysis can be used to evaluate the putative functions of genes (Bulow *et al.*, 2016; Jones *et al.*, 2020). *CsYABBY*3 is a member of the CRC group, and it has an AAGA motif, AA*TATA box, CAAT box, MYB, TATA box, TCT motif, At rich repeats, Box 4, MBS, Myb, TATA, TC rich repeats, TCA elements, AT rich elements, BOX4, MBS, Myb, TATA, TC rich repeats, TCA elements, GCN4 motif, chs-CMa1a and circadian. Most of them are involved in the initiation of transcription, a fraction of others are involved in dealing with abiotic stress. *CsYABBY*4 and *CsYABBY*8 are members of the group INO. *CsYABBY*4 has AAGAA motif, ABRE, CAAT box, ERE, MYB, MYC, TATA box, Box 4, TC rich repeats, TGA elements, circadian, CARE, CAT box, LTR and Myb binding site. Most of them are involved in dealing with abiotic stress and have a developmental role. *CsYABBY*8 has 3 AF1 binding sites, AAGA motif, ABRE, ABRE3a, ABRE4, AT*TATA box, CAAT box, ERE, G box, MYB, Myb like sequences, MYC, Myc, O2 site, TATA box, ARE, BOX4, MBS, TATA, TC rich repeats, TGA elements, CAT box, GAP box and LS7. They are mostly involved in plants' abiotic stress tolerance. *CsYABBY*6 and *CsYABBY*7 are the members of group AFO. *CsYABBY*6 has an AAGA motif, ABRE, AT*TATA motif, CAAT box, G box, MRE, MYB, Myb like sequences, ARE, Box4, GT1 motifs, CCAAT box and WRE3. They are involved in abiotic stress responses, phytohormone responses, development and transcription initiation. *CsYABBY*7 possesses AT*TATA box, CAAT box, TATA box, WUN motif, CCAAT box and Myb recognition site. Most of them are involved in transcription initiation; rest are involved in wound responses, abiotic stresses, and development. *CsYABBY*2 and *CsYABBY*5 are members of group YAB5. *CsYABBY*2 possesses AAGAA motif, ERE, G box, MYB, STRE, TCT motif, A box, ARE, AT rich repeats, Box 4, CCGTCC motif, GT1 motif, MBS, Myb, TC rich repeats, TCA elements, TGA elements and WUN motif. They are involved in transcription initiation, development and abiotic stresses. *CsYABBY*5 has AAGA motif, CAAT box, ERE, MYB, MYC, ARE, BOX 4, Myb, TC rich repeats, TGA elements, circadian, CARE, CAT box, LTR, and myb binding sites. They have their involvement in low-temperature stress, phytohormone responses and development. *CsYABBY*1 belongs to the group YAB2. It possesses 3-AF1 binding site, AAGAA-motif, ABRE, ABRE3a, ABRE4, AE-box, AT-rich sequence, AT~TATA-box, CAAT-box, CGTCA-motif, ERE, G-box, GATA-motif, MRE, MYB, MYB-like sequence, MYC, Myc, O2-site, P-box, STRE, TATA-box, TCA, TCT-motif, TGA-box, TGACG-motif and as-1. They are involved in transcription initiation, phytohormone responses, development and response to abiotic stresses (Fig. 5)

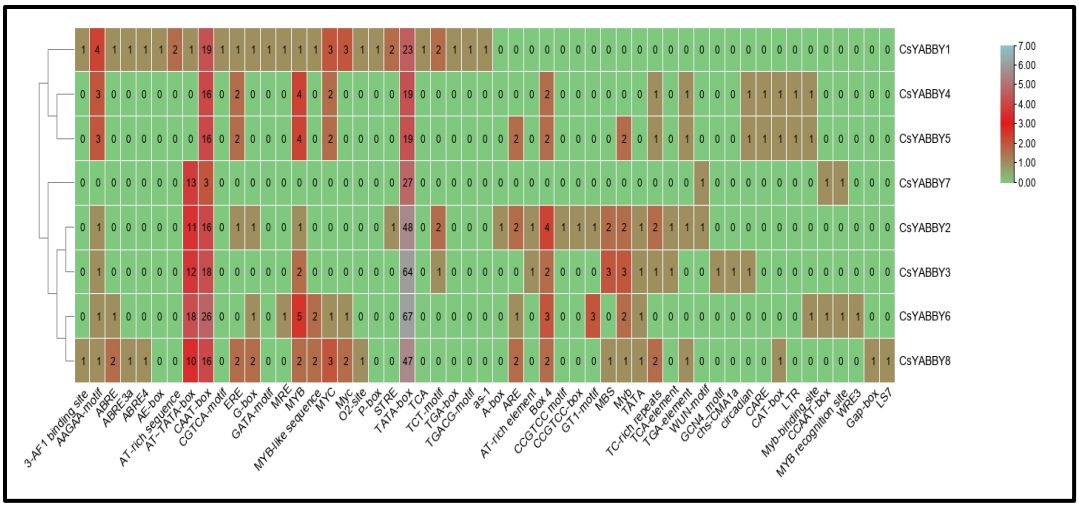

Figure 5: Different *cis*-acting elements in putative *CsYABBY* promoters which associated with abiotic stresses, hormone responses, growth and development. Color legends indicating the number of cis-elements found in each *CsYABBY* gene

Ks/Ka Analysis: Generally, the Ka/Ks<1 infers evolution as a purifying selection. Ka/Ks = 1 implies an evolution to be a neutral selection, whereas Ka/Ks > 1 is indicative of a positive selection. The ratio for Ka/Ks was found for all the cucumber pairs. It was seen that Ka/Ks ratio was less than 1 for all the pairs, which mean that they are the result of purification selection by environment (Fig. 6).

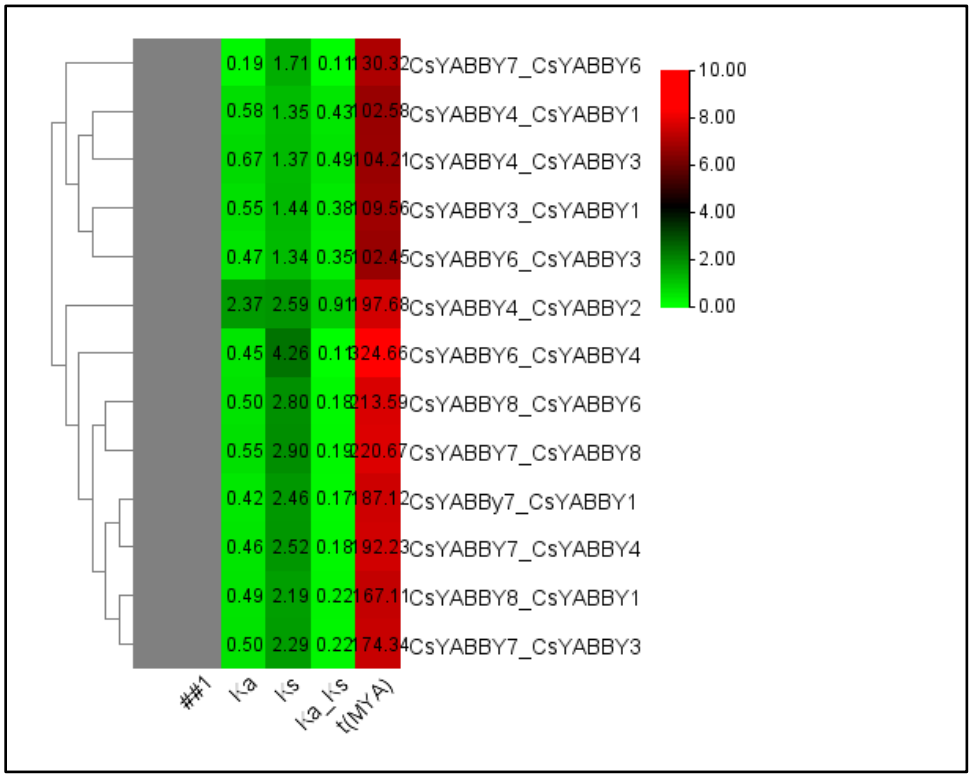

Figure 6: Time of gene duplication estimated for different paralogous pairs of Cucumber YABBY genes based on Ks and Ka values. Analysis was conducted using the Nei-Gojobori model. *Ka* represents the number of nonsynonymous substitutions per nonsynonymous site ,and *Ks* is the number of synonymous substitutions per synonymous site. While *Ka* /*Ks* represents the ratio of nonsynonymous (*Ka*) versus synonymous (*Ks*) mutations

Chromosomal Mapping: On chromosome 5, *CsYABBY*3 and *CsYABBY*4 are tandemly duplicated whereas *CsYABBY*1 on chromosome 6 and *CsYABBY*3 on chromosome 5 are in segmented duplication. *CsYABBY*6 on chromosome 5 and *CsYABBY*7 on chromosome 3 are in segmented duplication (Fig. 7).

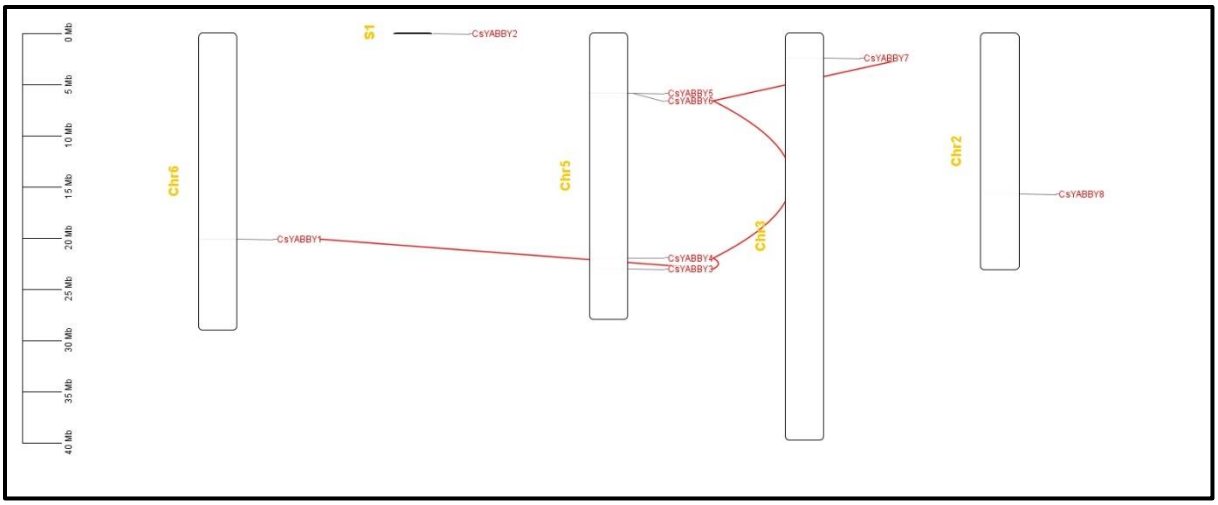

Figure 7: Distribution of *CsYABBY* genes on Cucumber chromosomes. *YABBY* genes present on the exact location within the same chromosome are colored differently than the other genes. Arctic blue color represents chromosomes having 1 *YABBY* gene, pink 2, Azure blue 3, dark blue 4, royal blue 5, and white color represents chromosomes with no *YABBY* genes in them, respectively. The scale represents a 10 Mb chromosomal distance. Genes on the scaffold are mapped imaginary due to the lack of complete scaffold length data

Synteny analysis: The synteny graph shows the syntenic relationships between the YABBY genes of the cucumber itself. It was seen that *CsYABBY*5 and *CsYABBY*8 on scaffold000044-7 are in segmented duplication with *CsYABBY*6 on chromosome 5. *CsYABBY*6 on chromosome 5 is in segmented duplication with *CsYABBY*7 and *CsYABBY*8 on chromosomes 2 and 3. Also, *CsYABBY*1 is in segmented duplication with *CsYABBY*8 on chromosome 2 and *CsYABBY*7 on chromosome 3, *CsYABBY*5 and *CsYABBY*3 on chromosome 5 and *CsYABBY*2 and *CsYABBY*5 on scaffold000044-7. *CsYABBY*3 and *CsYABBY*5 are in tandem duplication on chromosome 5. *CsYABBY*5 and *CsYABBY*2 are in tandem duplication on chromosome 5 (Fig. 8). Dual synteny analysis of cucumber was done with *C. maxima*, *G. Max*, *M. acuminata* and *A. thaliana*, *C. sativus* was shown to have more than two syntenic relationships with *C. maxima*, *G. max* and *A. thaliana*, proving their evolutionary relationship, implying that they share the same common ancestor throughout their evolutionary history. Whereas *C. sativus* shares only one syntenic pare with monocot *M. acuminata*, implying there's no common evolutionary history between these two.

Transcriptome Analysis: Among all the members of the *CsYABBY* family that were given the different conditions of temperature and photoperiodism, *CsYABBY*1, *CsYABBY*2, *CsYABBY*6, *CsYABBY*7 and *CsYABBY*8 had no difference in expression at 0 hours of treatment, downregulated after 3 hours of the treatment and downregulated more after the 6 hours of treatment. *CsYABBY*5 was over-expressed at 0 hours of the treatment, somewhat downregulated after the 3 hours of treatment whereas it was upregulated after the 6 hours of the treatment. *CsYABBY*3 had no difference in expression at all the points of treatments. The expression of *CsYABBY*4 was downregulated after 3 hours of the treatment and was upregulated the most after the 6 hours of treatment. We can conclude that *CsYABBY*4 has some role to play in its overexpression under heat stress (Fig. 9).

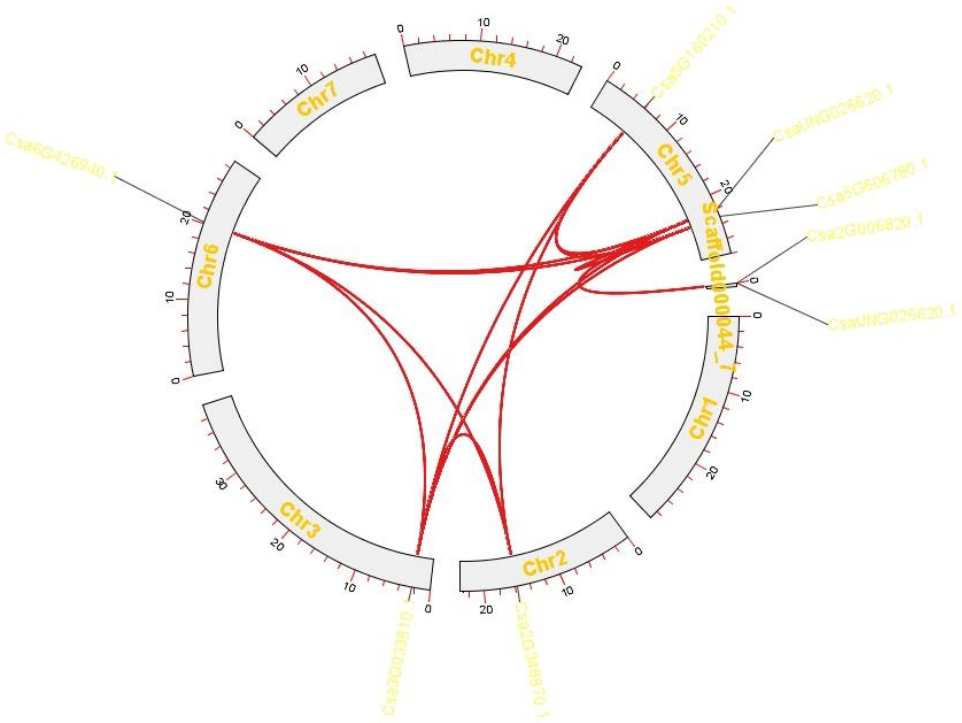

Figure 8: Genome-wide synteny analysis of Cucumber *YABBY* genes showing the dominance of segmental duplication and rare occurrence of tandem duplication

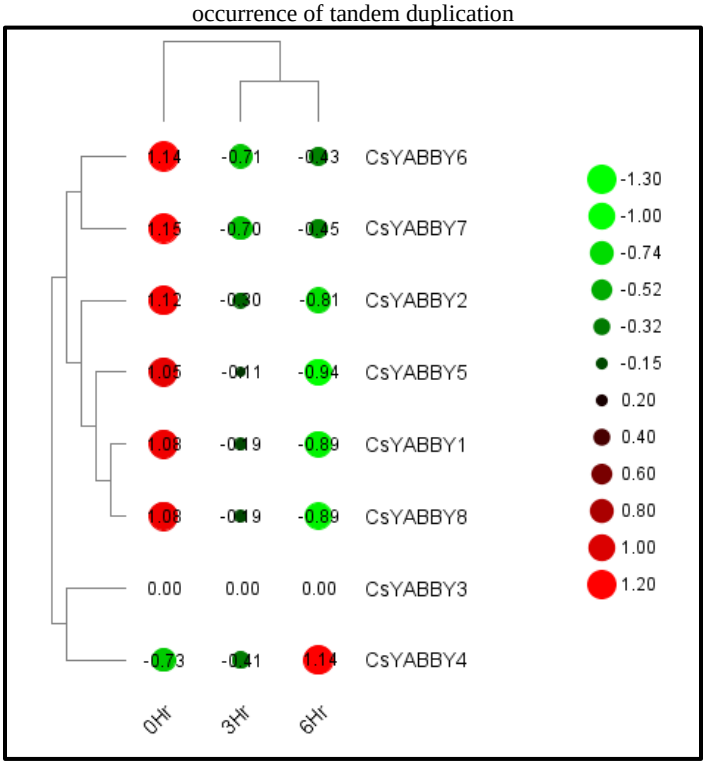

Figure 9: The heat map shows the expression profile of the Cucumber *YABBY* genes in different organs in the mature Cucumber tree. The x-axis represents the names of the six organs in the mature Cucumber tree, and the y-axis represents different *CsYABBY* genes. The expression levels of *CsYABBY* genes are revealed by different colors, which increase from blue to red

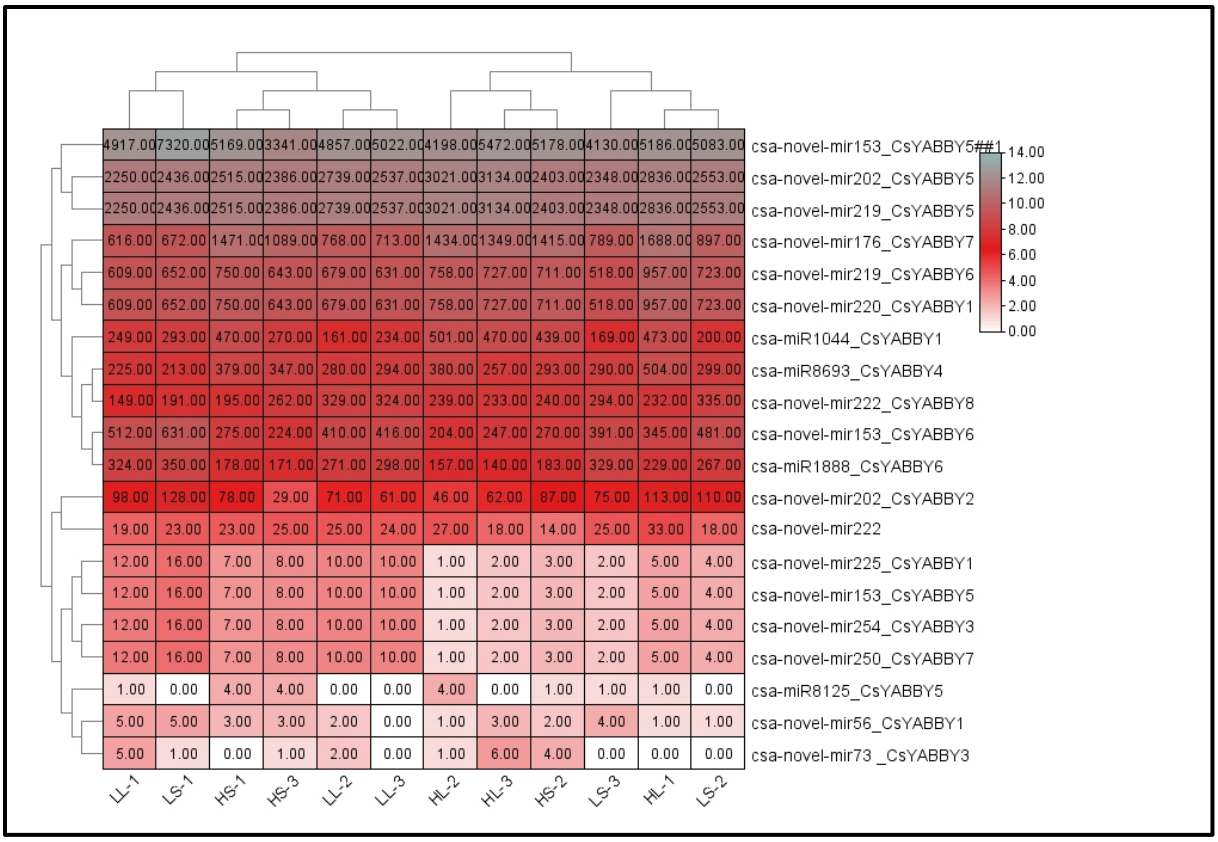

Figure 10: Putative mi-RNA analysis:

Putative miRNA Targets in *C. sativus*: The miRNA sequences in *C. sativus* were retrieved from the Plant MicroRNA Encyclopedia database. With the assistance of the psRNA target, the Micro RNA sequences that could target *CsYABBY* genes were figured out (https://plantgrn.noble.org/psRNATarget/analysis). As a result, 13 miRNAs were identified which targeted all 8 YABBY genes in cucumber. Csa-novel-mir153 targeted *CsYABBY*1, *CsYABBY*5 and *CsYABBY*6 which are involved in the sex ratio regulation of plants (Zhang *et al.*, 2018). *CsYABBY*2 was targeted by csa-novel-mir202. *CsYABBY*1 was also targeted by csa-novel-mir225. *CsYABBY*3 was targeted by csa-novel-mir254 and csa-novel-mir56. *CsYABBY*4 is targeted by csa-novel-mir8693. *CsYABBY*5 is targeted by csa-novel-mir-8125, csa-novel-mir202 and csa-novel-mir219. *CsYABBY*6 is targeted by csa-mir-1888 and csa-novel-mir219. *CsYABBY*7 is targeted by csa-novel-mir176 and csa-novel-mir250. *CsYABBY*8 is targeted by csa-novel-mir222. The miRNA's functions are not yet identified except for csa-novel-mir153 (Fig. 10)

Subcellular Localization: Subcellular localization was carried out to find the localization of different domains of YABBY transcription factor in *C. sativus*, *C. maxima*, *M. Acuminata* and *A. thaliana* in various cell organelles. It was found that in Nucleus, *CsYABBY*1, *CsYABBY*2, *CsYABBY*6, *CsYABBY*7, *CsYABBY*8, *CmYABBY5*, *CmYABBY6*, *CmYABBY8*, *CmYABBY11*, *AtYABBY3*, *AtYABBY5*, *AtYABBY6*, *MaYABBY1*, *MaYABBY2*, *MaYABBY3*, *MaYABBY4*, *MaYABBY5*, *MaYABBY11*, *MaYABBY14*, *MaYABBY16*, *MaYABBY19*, *MaYABBY22*, *MaYABBY23* and *MaYABBY25* are found in very high concentration whereas *CmYABBY3*, *CmYABBY10*, *CmYABBY13*, *MaYABBY7*, and *MaYABBY9* are found in relatively less quantities in the Nucleus. The cytoplasm contained *MaYABBY8* and *MaYABBY9*. In Chloroplast, *CsYABBY*5, *AtYABBY2*, *MaYABBY12*, *MaYABBY13*, *MaYABBY17*, *MaYABBY15*, *MaYABBY20*, *MaYABBY21*, *MaYABBY24* and *AtYABBY4* are present in comparatively high quantities whereas *MaYABBY6*, *MaYABBY7* and *MaYABBY8* are found in minute amounts in chloroplasts. Mitochondria contains the small amounts of *MaYABBY13*, *MaYABBY18*, *CsYABBY*5, *CmYABBY4*, *CmYABBY9*, *AtYABBY1*, *AtYABBY3*, *MaYABBY25*, *MaYABBY5*, *MaYABBY7*, *MaYABBY9* and *AtYABBY4*. In plastids, *MaYABBY25*, *MaYABBY23*, *MaYABBY22*, *MaYABBY18*, *MaYABBY17*, *MaYABBY15*, *MaYABBY1*, *MaYABBY9*, *MaYABBY6*, *MaYABBY2*, *AtYABBY1*, *CmYABBY13*, *CmYABBY12*, *CmYABBY10*, *CmYABBY9*, *CmYABBY4*, *CsYABBY4* and *CsYABBY*5 are present in minute

quantities. Cysk contains small amounts of *MaYABBY24*, *MaYABBY8*, *CmYABBY10* and *CsYABBY*8. Minute quantities of *MaYABBY18*, *MaYABBY15*, *MaYABBY13*, *MaYABBY10*, *MaYABBY6*, *CmYABBY9*, *CmYABBY4* and *CsYABBY*4 are present in vacuole. In peroxisomes, *CmYABBY1*, *CmYABBY2*, *CmYABBY6* are present in high quantities whereas *CmYABBY12* is present in a small amount. In ER, *CmYABBY9* and *CmYABBY4* are present in a minimant (Fig. 11).

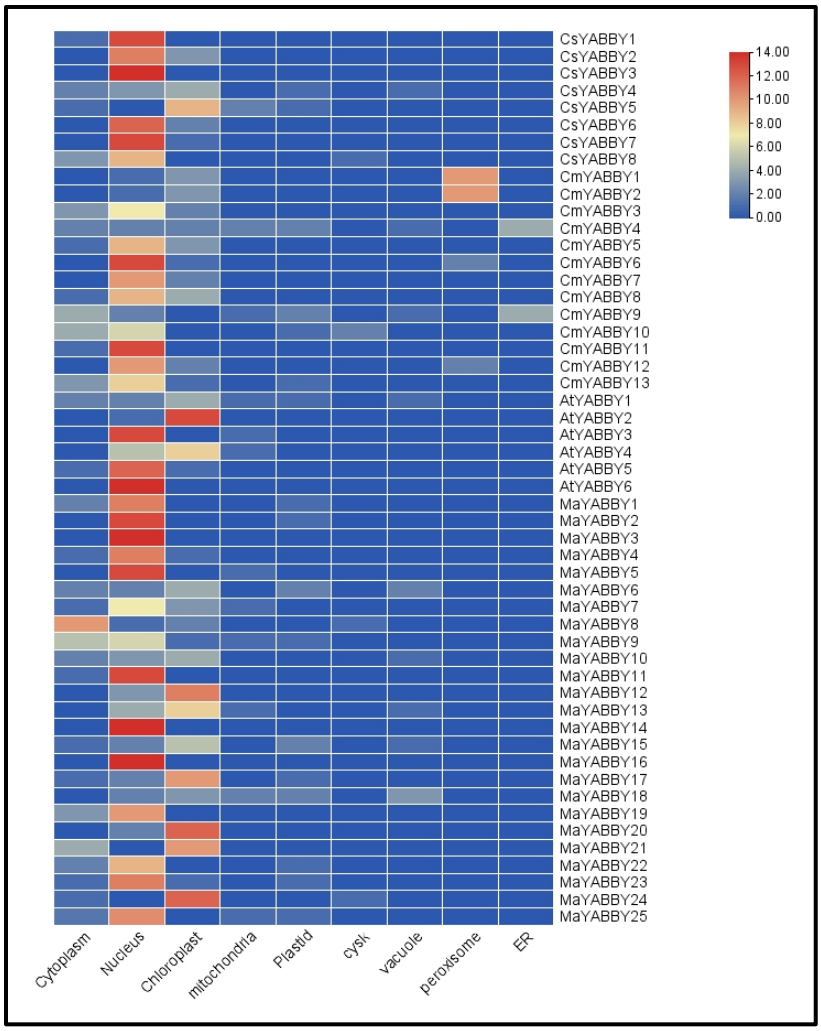

Figure 11: Subcellular Localization

Discussion

The development of abaxial polarity and the lateral organs' development have a vital involvement in plant-specific transcription factor YABBY. YABBY has also been identified as a major agent in the wake of responding to various abiotic stresses (S Zhang *et al.*, 2019b). Most of the YABBY's function is associated with the development of various organs, also the formation of seed integument and carpel (Bowman *et al.*, 1989). To understand the basic structure and function of the YABBY transcription factor in *C. sativus*, eight sequences of YABBY protein were identified in cucumber. For the comparison, sequences of YABBY proteins were also identified in *A. thaliana*. *A. thaliana* had 6 sequences of YABBY proteins, belonging to the groups CRC, INO, YAB2, FIL/AFO/YAB3, YAB5. Thirteen sequences of YABBY proteins were also retrieved in *C. maxima*, a member of the family Cucurbitaceae, to compare the analyses of YABBY protein in cucumber with a member of its own family and for its comparative analysis with monocots, 25 sequences of YABBY proteins were identified in *M. acuminata*. The alignment of the sequences illustrated a conserved C_2C_2 domain in all the YABBY sequences of all the species (Fig. 1). Using the same alignment, a phylogenetic tree was constructed to illustrate the evolutionary relationship between YABBY genes in *C. sativus* and other species. The phylogenetic tree was subdivided into five groups according to the groups of *A. thaliana* and all the sequences of *C. sativus* were in correspondence with the groups of *A. thaliana*.

*CsYABBY*3 was included in group CRC. *CsYABBY*4 and *CsYABBY*8 were the members of group INO. *CsYABBY*6 and *CsYABBY*7 were found to belong to the group AFO/FIL. *CsYABBY*2 and *CsYABBY*5 belonged to YAB5 and *CsYABBY*1 belonged to the group YAB2. The correspondence of these genes may imply similarity in their function (Wan *et al.*, 2018). The conserved domain analysis of the YABBY protein family confirmed the presence and conservation of the Zinc Finger Domain in all the sequences of *C. sativus.* Also, a graph illustrating the number of introns and exons was generated to further investigate the structure of YABBY genes. (Fig. 3). It was seen that majority of the members of the same group, including *C. sativus*, *C. maxima*, *A. thaliana* and *M. acuminata*, contained the same number of introns, with a variation among a very narrow range, suggesting that the structure of YABBY protein has been conserved throughout the species (Xia *et al.*, 2021). The dual analysis of Cucumber with *C. maxima*, *A. thaliana* and watermelon showed that *C. sativus* had syntenic relationships with all of them, so they might share some common evolutionary history (Yin *et al.*, 2022). It was seen that majority of the members of the same group of YABBY protein possessed the same number of motifs with a few being exceptions for having a few more motifs or missing a few motifs, which might reflect its structural similarity and variation in some specific sequences might deduce a great deal of mutation in their structure. MYB, EBOX, MYC, GT1 and ABRE elements were found as cis-regulatory elements in the promoter region of *C. sativus.* It was found that these cis-regulatory elements were involved in responding to various abiotic stresses like heat stress, drought stress, low-temperature stress and osmotic stress by the combination of these cis-regulatory elements with the corresponding transcription factors. Cis-regulatory elements were also involved in light response, wound healing, hormone responses and controlling of various aspects of flowering. This indicated that *CsYABBYs* have a role in response to abiotic stress, and also play a role in some light-driven mechanisms and the development of various plant organs (Zhao *et al.*, 2017). Transcription factors target the specific cis-regulatory elements and control the patterns of expression of genes. The cis-regulatory elements like ABRE, and ARE are involved in phytohormone responses, mainly responds to abscisic acid. Cis-regulatory elements like TGA motif are also involved in nitrogen assimilation in leaves (Araújo *et al.*, 2014). The binding sites 3-AF1, AAGAA-motif, ABRE, ABRE3a, ABRE4, AE-box, AT-rich sequence, AT~TATA-box, CAAT-box, CGTCA-motif, ERE, G-box, GATA-motif ,MRE,MYB, MYB-like sequence, MYC, Myc,O2-site, P-box, STRE, TATA-box, TCA, TCT-motif, TGA-box, TGACG-motif, as-1 have a role in various stresses like osmotic stress, drought stress by which the function of *CsYABBY* in various stresses are concluded (He *et al.*, 2012). The involvement of cis regulatory elements of YABBY proteins in various phytohormone responses and abiotic stresses may reveal that it had a role in the adaptation of plants to changing environments (Zhao *et al.*, 2017). Cis-regulatory elements lCGTCA-motif, O2-site, P-box, STRE, TATA-box, TCA, TCT-motif, TGA-box, TGACG-motif, as-1, A-box,ARE, AT-rich element are involved in the initiation of various kinds of transcription. This may suggest that in addition to developmental roles, YABBY protein may also have a resistant role against various abiotic stresses like NaCl stress, phytohormone stress, drought stress, heat stress (Ha *et al.*, 2010). It was seen in chromosomal mapping that mostly segmented duplication is found in *CsYABBYs.* analysis also reveals that most of the *CsYABBYs* are in segmented duplication. The expression of YABBY genes was particularly high in the nucleus, which is a strong evidence of it being a major developmental and functional gene*; the* expression of *CsYABBYs* was monitored under different conditions of heat stress, photoperiodism and *CsYABBY*4, belonging to INO was seen to be significantly overexpressed under heat stress. *CsYABBY*4 belonging to group INO which is involved in forming the outer integument of the seed might also suggest modification in the seed's structure under heat stress. So, further studies on the function of *CsYABBY* may prove to be groundbreaking in breeding and agricultural programs on making *C. sativus* heat tolerant or seedless or modification in its development (Dai *et al.*, 2007).

Conclusion:

During this research, a comprehensive analysis of *CsYABBY* TFs in the cucumber genome was done. A total of 54 YABBY proteins were depicted in cucumber, banana, *C. maxima* and *A. thaliana* and were distributed among six subgroups named 1, 2, 3, 4, 5 and 6. They are further divided into 5 subfamilies named FIL/YAB3, INO, YAB2, CRC, YAB5. In accordance with the presence of CsYABBYs in the same clade as A. thaliana, they are known to function in abaxial organ development, carpel and nectary development, cell division, abiotic stress, flower development and seed coat formation. *CsYABBYs* also have evolutionary relationships with *C. maxima*, *A. thaliana*, *G. max and C. lanatus* while it showed no evolutionary relation with *M. acuminata.* Within *C. sativus*, segmented duplication has been seen mostly; only a single pair of *CsYABBYs* was in tandem duplication. Gene structure of *CsYABBYs* was seen to be almost conserved during the course of evolution. They were shown to have a function in abiotic stress, wound healing, phytohormone responses and transcription initiation. *CsYABBY*4 was observed to have a significant overexpression during heat stress and long day conditions of photoperiodism. Future aspects of this research might contain the possibility of making the *C. sativus* plants heat

tolerant by mutating *CsYABBY*4, modifying its expression pattern to over-expression all the time in the conditions of heat stress.

Authors' Contribution:

MS, MM, MMH, ZK and MQ were involved in data analysis. MS and MM provided overall direction and experimental design. MQ and ZK wrote the manuscript.

Acknowledgements:

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

Conflict of Interest:

The authors have no potential conflict of interest.

References:

Araújo, W. L., Martins, A. O., Fernie, A. R., & Tohge, T. (2014). 2-oxoglutarate: Linking tca cycle function with amino acid, glucosinolate, flavonoid, alkaloid, and gibberellin biosynthesis. Frontiers in plant science, **5**, 552.

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

Bisognin, D. A. (2002). Origin and evolution of cultivated cucurbits. Ciência Rural, **32**, 715-723.

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. Methods Mol Biol, **1482**, 233-245. doi:10.1007/978-1-4939-6396-6_15

Cavagnaro, P. F., Senalik, D. A., Yang, L., Simon, P. W., Harkins, T. T., Kodira, C. D., . Weng, Y. (2010). Genome-wide characterization of simple sequence repeats in cucumber (cucumis sativus l.). BMC genomics, **11**(1), 1-18.

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. doi:https://doi.org/10.1016/j.molp.2020.06.009

Dai, M., Hu, Y., Zhao, Y., Liu, H., & Zhou, D.-X. (2007). A wuschel-like homeobox gene represses a yabby gene expression required for rice leaf development. Plant physiology, **144**(1), 380-390.

E., Z., & L., P. (1965). Evolutionary divergence and convergence in proteins.

Finn, R. D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R. Y., Eddy, S. R., . Mistry, J. (2014). Pfam: The protein families database. Nucleic acids research, **42**(D1), D222-D230.

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.

Goodstein, D., Batra, S., Carlson, J., Hayes, R., Phillips, J., Shu, S., Rokhsar, D. (2014). Phytozome comparative plant genomics portal.

Ha, C. M., Jun, J. H., & Fletcher, J. C. (2010). Control of arabidopsis leaf morphogenesis through regulation of the yabby and knox families of transcription factors. Genetics, **186**(1), 197-206.

He, Y., Li, W., Lv, J., Jia, Y., Wang, M., & Xia, G. (2012). Ectopic expression of a wheat myb transcription factor gene, tamyb73, improves salinity stress tolerance in arabidopsis thaliana. Journal of Experimental Botany, **63**(3), 1511-1522.

Higo, K., Ugawa, Y., Iwamoto, M., & Higo, H. (1998). Place: A database of plant cis-acting regulatory DNA elements. Nucleic Acids Res, **26**(1), 358-359. doi:10.1093/nar/26.1.358

Higo, K., Ugawa, Y., Iwamoto, M., & Korenaga, T. (1999). Plant cis-acting regulatory DNA elements (place) database: 1999. Nucleic Acids Res, **27**(1), 297-300. doi:10.1093/nar/27.1.297

Hu, B., Jin, J., Guo, A. Y., Zhang, H., Luo, J., & Gao, G. (2015). Gsds 2.0: An upgraded gene feature visualization server. Bioinformatics, **31**(8), 1296-1297. doi:10.1093/bioinformatics/btu817

Hudson, B. P., Martinez-Yamout, M. A., Dyson, H. J., & Wright, P. E. (2004). Recognition of the mrna au-rich element by the zinc finger domain of tis11d. Nature structural & molecular biology, **11**(3), 257-264.

J., F. (1985). Confidence limits on phylogenies: An approach using the bootstrap. Evolution, 783-791.

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. doi:10.1016/j.pbi.2019.12.008

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. doi:10.1093/molbev/msy096

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 acyabby4 involvement in salt stress. International journal of molecular sciences, **20**(23), 5863.

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. Nucleic Acids Res, **48**(D1), D265-D268. doi:10.1093/nar/gkz991

Mousavizadeh, S. J., Mashayekhi, K., Garmakhany, A. D., Ehteshamnia, A., & Jafari, S. (2010). Evaluation of some physical properties of cucumber (cucumis sativus l.). Nong Ye Ke Xue Yu Ji Shu, **4**(4), 107.

Mukherjee, P. K., Nema, N. K., Maity, N., & Sarkar, B. K. (2013). Phytochemical and therapeutic potential of cucumber. Fitoterapia, **84**, 227-236.

P.H.A., S., & R.R., S. (1973). Numerical taxonomy.

Ramirez-Tejero, J. A., Jimenez-Ruiz, J., Leyva-Perez, M. O., Barroso, J. B., & Luque, F. (2020). Gene expression pattern in olive tree organs (olea europaea l.). Genes (Basel), **11**(5). doi:10.3390/genes11050544

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.

S., K., G., S., M., L., C., K., & K., T. (2018). Mega x: Molecular evolutionary genetics analysis across computing platforms. Molecular biology and evolution, 1547-1549.

Samad, A. F. A. (2017). Microrna and transcription factor: Key players in plant regulatory network. Frontiers in Plant Science.

Sarojam, R., Sappl, P. G., Goldshmidt, A., Efroni, I., Floyd, S. K., Eshed, Y., & Bowman, J. L. (2010). Differentiating arabidopsis shoots from leaves by combined yabby activities. The Plant Cell, **22**(7), 2113-2130.

Tatlioglu, T. (1993). Cucumber: Cucumis sativus l *Genetic improvement of vegetable crops* (pp. 197-234): Elsevier.

Thompson, J. D., Gibson, T. J., & Higgins, D. G. (2003). Multiple sequence alignment using clustalw and clustalx. Current protocols in bioinformatics(1), 2.3. 1-2.3. 22.

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.

Wan, T., Liu, Z.-M., Li, L.-F., Leitch, A. R., Leitch, I. J., Lohaus, R., Liu, Y. (2018). A genome for gnetophytes and early evolution of seed plants. Nature Plants, **4**(2), 82-89.

Wang, Y., Li, J., & Paterson, A. H. (2013). Mcscanx-transposed: Detecting transposed gene duplications based on multiple colinearity scans. Bioinformatics, **29**(11), 1458-1460. doi:10.1093/bioinformatics/btt150

Xia, J., Wang, D., Peng, Y., Wang, W., Wang, Q., Xu, Y., . . . Xu, X. (2021). Genome-wide analysis of the yabby transcription factor family in rapeseed (brassica napus l.). Genes, **12**(7), 981.

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. (2019a). Genome-wide analysis of the yabby gene family in grapevine and functional characterization of vvyabby4. Frontiers in plant science, **10**, 1207.

Zhang, S., Wang, L., Sun, X., Li, Y., Yao, J., Nocker, S. v., & Wang, X. (2019b). Genome-wide analysis of the yabby gene family in grapevine and functional characterization of vvyabby4. Frontiers in plant science, 1207.

Zhang, T., Li, C., Li, D., Liu, Y., & Yang, X. (2020). Roles of yabby transcription factors in the modulation of morphogenesis, development, and phytohormone and stress responses in plants. Journal of Plant Research, **133**(6), 751-763.

Zhang, X., Lai, Y., Zhang, W., Ahmad, J., Qiu, Y., Zhang, X., . Wang, H. (2018). Micrornas and their targets in cucumber shoot apices in response to temperature and photoperiod. BMC genomics, **19**(1), 1-15.

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 gmyabby10 involvement in high salt and drought stresses. Plant Physiology and Biochemistry, **119**, 132-146.