



# *In silico* identification and annotation of drought responsive candidate genes in Solanaceous plants

<sup>1</sup>Mahammed Faizan\*, <sup>2</sup>Harish Babu, B. N., <sup>3</sup>Fakrudin, B., <sup>1</sup>Lakshmana, D and. <sup>1</sup>Rakshith, M.

<sup>1</sup>Department of Genetics and Plant Breeding, College of Horticulture, Mudigere (University of Agricultural and Horticultural Sciences, Shivamogga, Karnataka, India)

<sup>2</sup>Zonal Agricultural and Horticultural Research Station, Hiriyyur (University of Agricultural and Horticultural Sciences, Shivamogga, Karnataka, India)

<sup>3</sup>Department of Biotechnology and Crop improvement, PG centre, GKVK, Bangalore (University of Horticultural Sciences, Bagalkot, Karnataka, India,)

\*Corresponding author email- [faizanscf@gmail.com](mailto:faizanscf@gmail.com)

**Abstract:** An *in silico* analysis was undertaken to identify drought responsive candidate genes in solanaceous plants like potato, chilli, tomato, eggplant and *etc.*, using NCBI database. Drought tolerant plants counter the drought by producing various enzymes and transcription factors through genes associated with drought tolerance. The drought stress responsive genes were validated using *Arabidopsis thaliana* as reference species. The stress responsive unigenes were identified from STIFDB V.2.0 database and subjected for BLAST analysis to identify related homologous sequences from *Solanum* taxid using NCBI database. A total of 795 stress responsive unigenes were identified for drought stress among which 109 drought responsive unigenes were found in *solanaceous* taxid with sequence identities (ID) or homology of more than 80 per cent. Based on Bulk Data Retrieval (BDR) from The Arabidopsis Information Resource (TAIR), the Gene Ontology analysis for GO functional categories has revealed that, in cellular component category, genes assigned to the other intracellular components (20.37%) accounted for largest group followed by other cytoplasmic components (19.05%) and genes of chloroplast (11.77%). In biological process category, largest group of genes belong to other cellular processes (25.89%) followed by other metabolic processes (21.89%) and response to abiotic or biotic stimulus (12.44%). In Molecular functional categorization, 11.59% of genes have transferase activity; protein binding activity (10.90%); other enzymic activities (17.27 %) and other binding activities (17.95 %). The information on the drought responsive candidate genes can be used for identification and isolation of drought responsive genes for utilization in drought tolerance breeding in solanaceous crops.

**Index Terms -** *In silico*, drought, candidate genes, STIFDB V.2.0, NCBI data base

## I INTRODUCTION

Solanaceae is one of the important families in vegetable kingdom consisting of globally important vegetable crops like tomato (*Solanum lycopersicon*), potato (*Solanum tuberosum*), brinjal (*Solanum melongena*) and chilli (*Capsicum annum*). These vegetables have prominent role in Indian culinary as well as pharmaceutical industry. In the era of climate change, drought has become a major challenge in agriculture sector which not only affects yield but also leads to production of inferior quality produce. Inbuilt drought tolerance in crop plants is a durable and eco-friendly approach to counter the ill effects of drought. Breeding for drought tolerance has been considered as a major area of plant breeding in recent times all over the globe including development of drought tolerant cultivars in solanaceous vegetables.

Drought tolerant plants are known to counter the effect of drought by expressing various genes involved in different biochemical pathways that produce various enzymes and transcription factors *etc.*, associated with drought tolerance. The understanding of plant responses to drought stress at transcriptome level provides a foundation for the identification of stress response genes. One of the reliable approaches to identify drought responsive genes is mining of different databases for retrieving the already reported gene sequences which are known to have a role in drought tolerance mechanism. The identification and analysis of genes exhibiting better responses to drought through *in silico* approaches and categorization of the genes in various organisms under different conditions got importance due to growing data in the databases (Aceituno *et al.*, 2008). In the present study, an effort has been made through *in silico* analysis for identification of drought responsive candidate genes in solanaceous plants for subsequent utilization in drought tolerance breeding of solanaceous crops.

## II MATERIAL AND METHODS

### 3.1 MATERIAL

#### 3.1.1 reagents

Gene list data set from genome-wide mining experiments.

#### 3.1.2 equipment

Laptop or desktop computer with internet connection. High-speed internet connection is highly recommended.

#### 3.1.3 equipment setup

**Operating system requirements:** Windows XP or Windows 7 for PC users and MacOS 10.5 or upgraded versions for Mac users are required. A minimum of 2 GB RAM is recommended.

**Browser requirements:** It requires Microsoft Internet Explorer 8 or upgraded versions (recommended for PC users), Safari version 5 (recommended for Mac users), Firefox version 19 or Google Chrome version 26.

**Java requirements:** The latest version of Java is required (can be downloaded from <http://www.java.com/en/download/>); JavaScript, Java applets and cookies must be enabled in your browser; Java applet runtime parameters should be set to -ms128m -mx512m -Xss16m (Mi *et al.*, 2013).

### 3.2 METHODOLOGY

#### 3.2.1 Data Set

A total of 795 annotated drought stress responsive gene IDs of *Arabidopsis thaliana* identified from STIFDB V.2.0 (Stress Responsive Transcription Factor Database) ([http://caps.ncbs.res.in/cgi-bin/mini/databases/stifdb2/browse\\_by\\_gene.pl](http://caps.ncbs.res.in/cgi-bin/mini/databases/stifdb2/browse_by_gene.pl)) (Naika *et al.*, 2013) during December, 2018 and the FASTA sequences for such genes were extracted from NCBI (National Centre for Biotechnology Information) database.

#### 3.2.2 Basic Local Alignment Search Tool

The extracted FASTA (Fast Adaptive Shrinkage Threshold Algorithm) sequences of 795 drought responsive unigenes from NCBI were subjected to Basic Local Alignment Search Tool (BLAST) ([https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE\\_TYPE=BlastSearch&LINK\\_LOC=blasthome](https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome)) (Ye *et al.*, 2006) for organism *solanum* taxid with default parameter settings. Out of 795, a total of 109 stress responsive unigenes which were showing more than 80 per cent identities or sequence similarities were selected from *solanum* taxid. The detailed information these 109 stress responsive unigenes like gene ID, gene symbol, locus tag, type of gene, annotation or gene description, exon count, CDS (coding sequences) length along with their role as transcription factors were described in the table no. 1.

#### 3.2.3 Functional Categorization

These 109 stress responsive unigenes from *solanum* taxid were subjected to Bulk Data Retrieval (BDR) (<https://www.arabidopsis.org/tools/bulk/go/index.jsp>) from the database TAIR (The Arabidopsis Information Resource) (Rhee *et al.*, 2003) and the Gene Ontology functional categorization by annotation was carried out. The standard formula for Gene Ontology functional categorization by annotation is as follows:

$$\text{Gene Ontology functional categorization by annotation} = \frac{(\# \text{ of annotation to terms in this GOslim category})}{100 \times \# \text{ of total annotation to in this ontology}}$$

## III RESULTS

The mining of STIFDB V.2.0 database revealed 795 unigenes associated with drought tolerance. When these 795 unigenes were subjected for BLAST analysis to find out the probable drought responsive unigenes in *solanum* taxid, 109 out of 795 unigenes showing  $\geq 80$  per cent sequence similarity with the genes of *solanum* taxid were selected for further analysis.

The gene description and the probable function for each of the 109 drought responsive candidate genes are represented in the Table 1. All the genes were found to have protein coding which are involved in different cellular activities except two genes that were found to be associated with rRNA synthesis (AT2G01010 and AT2G01020).

The length of Coding Sequence (CDS) ranged from 164 to 4534 nucleotides among the 109 drought responsive candidate genes from *solanum* taxid. The lowest CDS length was recorded in the locus tag ID AT2G01020 (164 nucleotides) and the maximum CDS length of 4534 nucleotides was seen in the locus tag ID AT4G03430 (Table 1).

The exon count among the 109-drought tolerant unigenes from *solanum* taxid ranged from 1 to 22. Twenty-one out of 109 unigenes were represented by only one exon and the maximum number of exons were present in the unigene locus tag ID

AT5G65930 (22 exons). Further, when these 109 drought responsive candidate loci were analysed for their distribution on the *Arabidopsis* chromosomes, it was found that, maximum number of genes (35 loci out of 109) were distributed on chromosome number 1 compared to other 4 chromosomes. The chromosome number 2 and 4 contain 19 genes each and chromosome number 3 and 5 contain 18 genes each (Figure 1).

These 109 drought responsive candidate genes were also known to serve as transcription factors in regulation of different genes involved in various metabolic pathways. It is evident from the present analysis that, there are 19 different transcription factor activities shown by these 109 genes which are abbreviated as follows: ABRE, AuxRE, C/ABRE, CRT/DRE, G/ABRE, G box1, G box2, G box, GCC-box, HB, HSE, Myb\_box1, Myb\_box2, Myb\_box3, Myb\_box4, Myb\_box5, NAC, N box and W\_box (Table 1). Some of the important transcription factors include ABA-responsive element, basic domain leucine zipper, dehydration responsive element, ethylene responsive element binding factor, APETALA2, heat shock transcription factor etc.

All the 109 drought responsive unigenes were subjected to Bulk Data Retrieval (BDR) from TAIR Gene Ontology functional categorization by annotations (Swarbreck *et al.*, 2007). When these 109 unigenes were functionally categorized, they were grouped into three broad categories based on cellular components, biological processes and molecular function using gene ontology (GO) terms.

In cellular component category, the genes assigned to other intracellular components accounted for largest group (20.37%) followed by other cytoplasmic components (19.05%) and genes of chloroplast (11.77%). The smallest groups in cellular component category were belonging to the genes of ribosome (1.06%) and endoplasmic reticulum (0.53%) (figure 2A).

Upon categorization of these 109 genes based on their role in biological processes, maximum number of genes were belonging to those that were involved in other cellular processes (25.89%) followed by other metabolic processes (21.89%) and response to abiotic or biotic stimulus (12.44%). However, the smallest groups were represented by the genes involved in pollination (0.33%), unknown biological processes (0.33%) and DNA or RNA metabolism (0.33%) (Figure 2B).

The molecular functional categorization by annotation of these 109 genes has revealed that, these genes are known to play various functions in transferase activity (11.59%), protein binding (10.90%), DNA or RNA binding activity (9.32%), hydrolyse activity (7.05%), nucleotide binding activity (7.05%), kinase activity (6.14%), molecular transporter activity (3.63%), transcription factor activity (3.18%), nucleic acid binding activity (2.5%), other molecular functions (1.81%) and structural molecular activity (0.45%). However, about 17.95% of these genes have shown other binding activities; 17.27% have other enzymic activities (17.27%) and 1.13% have been recorded as having role in unknown molecular functions (Figure 2C).

#### IV DISCUSSION

Development of drought tolerant varieties is one of the biggest challenges of Solanaceous crop breeding programmes. An important approach for identification of drought tolerant genes is based on sequence similarity or homology using BLASTn tool in NCBI. A total of 109 genes were identified using BLASTn. Similarly, a number of candidate genes that were reported to respond to drought, salinity and cold stress at the transcriptional level have been described (Ingram and Bartels, 1996; Shinozaki and Yamaguchi-Shinozaki, 1996; Bray, 1997; Kim *et al.*, 2008; Appiano *et al.*, 2015). The functions of some gene products have been predicted from sequence homology or sequence similarities with known proteins and are thought to have a role in protecting the cells from water deficit (Ingram and Bartels, 1996; Bray, 1997). Genes induced during drought stress condition are known to function not only in protecting cell from water deficit by the production of important metabolic proteins through translation but also in the regulation of genes for signal transduction in water deficit responses. Based on these, genes which are responsible for stress tolerance are classified into broad two groups: first group includes genes involved in cellular protection including osmo-protectants, membrane stabilization, detoxification and transport proteins and the second group includes those genes that are serve as the transcription factors and signalling molecules (Vendruscolo *et al.*, 2007; Zhou *et al.*, 2016).

**Table 1. Details of 109 drought responsive candidate genes of *Solanum taxid* obtained through BLAST search in NCBI database**

Sl. No	Locus tag	Gene ID	Gene Symbol	Gene type	Gene Description	Exon count	CDS length	Transcription factors
1	AT1G35670	840471	CDPK2	protein coding	calcium-dependent protein kinase 2	7	1488	Myb_box3_MYB, Myb_box1_MYB, Myb_box5_MYB
2	AT1G64660	842774	MGL	protein coding	methionine gamma-lyase	3	1326	Myb_box3_MYB, Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, Myb_box2_MYB, Nac_box_NAC, G_box_bHLH
3	AT1G55490	841996	CPN60B	protein coding	chaperonin 60 beta	15	1803	Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, DREB_AP2_EREBP
4	AT1G63710	842675	CYP86A7	protein coding	cytochrome P450, family 86, subfamily A, polypeptide 7	1	1896*	Myb_box3_MYB, Myb_box1_MYB, Myb_box5_MYB, Nac_box_NAC, C_ABRE_bZIP
5	AT1G62750	842573	SCO1	protein coding	Translation elongation factor EFG/EF2 protein	4	2352	G_box2_Bzip, Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, W_box_WRKY, Myb_box2_MYB, AuxRE_ARF, DREB_AP2_EREBP
6	AT1G56600	842114	GolS2	protein coding	galactinol synthase 2	4	1008	HSE1_HSF, Myb_box5_MYB, G_box_bHLH, N_box_bHLH
7	AT1G54100	841849	ALDH7B4	protein coding	aldehyde dehydrogenase 7B4	14	1527	G_box2_bZIP, Myb_box3_MYB, Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, W_box_WRKY, Nac_box_NAC, G_box_bHLH
8	AT1G33590	840253	AT1G33590	protein coding	Leucine-rich repeat (LRR) family protein	1	1956	Myb_box3_MYB, Myb_box1_MYB, Myb_box5_MYB, W_box_WRKY, AuxRE_ARF
9	AT1G45249	841095	ABF2	protein coding	abscisic acid responsive elements-binding factor 2	5	1251	G_box2_bZIP, Myb_box1_MYB, Myb_box5_MYB, W_box_WRKY, Myb_box2_MYB, Nac_box_NAC
10	AT1G22370	838844	UGT85A5	protein coding	UDP-glucosyl transferase 85A5	2	1440	Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, W_box_WRKY, G_box_bHLH, G_ABRE_bZIP
11	AT1G37130	840630	NIA2	protein coding	nitrate reductase 2	3	2754	Myb_box3_MYB, Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, N_box_bHLH
12	AT1G52230	841653	PSAH2	protein coding	photosystem I subunit H2	3	438	Myb_box3_MYB, Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, AuxRE_ARF, G_box1_bZIP, G_box_bHLH, Myb_box4_MYB, G_ABRE_bZIP
13	AT1G67090	843029	RBCS1A	protein coding	ribulose bisphosphate carboxylase small chain 1A	3	543	Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, G_box1_bZIP, G_box_bHLH, Myb_box4_MYB, G_ABRE_bZIP
14	AT1G12900	837848	GAPA-2	protein coding	glyceraldehyde 3-phosphate dehydrogenase A subunit 2	5	1326	G_box2_bZIP, Myb_box3_MYB, Myb_box1_MYB, Myb_box5_MYB, W_box_WRKY, Myb_box2_MYB, Nac_box_NAC
15	AT1G07670	837280	ECA4	protein coding	endomembrane-type CA-ATPase 4	9	3186	HSE1_HSF, Myb_box5_MYB, W_box_WRKY, AuxRE_ARF, G_box_bHLH
16	AT1G32060	840098	PRK	protein coding	phosphoribulokinase	5	1188	G_box2_bZIP, Myb_box3_MYB, Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, DREB_AP2_EREBP, G_box_bHLH
17	AT1G06040	837113	STO	protein coding	B-box zinc finger family protein	2	747	Myb_box3_MYB, Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, W_box_WRKY, Myb_box2_MYB, HBE_HB
18	AT1G64680	842776	AT1G64680	protein coding	beta-carotene isomerase D27	4	753	G_box2_bZIP, Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, G_box_bHLH
19	AT1G80460	844385	NHO1	protein coding	Actin-like ATPase superfamily protein	4	1569	Myb_box3_MYB, Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, W_box_WRKY, Myb_box2_MYB, Nac_box_NAC, DREB_AP2_EREBP, G_box1_bZIP, G_box_bHLH
20	AT1G53210	841755	AT1G53210	protein coding	sodium/calcium exchanger family protein / calcium-binding EF hand family protein	7	1758	HSE1_HSF, Myb_box5_MYB, W_box_WRKY



21	AT1G51090	841532	AT1G51090	protein coding	Heavy metal transport/detoxification superfamily protein	3	516	HSE1_HSF, Myb_box5_MYB, G_box_bHLH
22	AT1G27200	839609	AT1G27200	protein coding	glycosyltransferase family protein (DUF23)	1	2431*	Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, W_box_WRKY, Myb_box2_MYB, AuxRE_ARF, DREB_AP2_EREBP
23	AT1G49240	841347	ACT8	protein coding	actin 8	5	1134	Myb_box3_MYB, Myb_box1_MYB, Myb_box5_MYB, AuxRE_ARF, DREB_AP2_EREBP, N_box_bHLH
24	AT1G10950	837638	TMN1	protein coding	transmembrane nine 1	12	1170	HSE1_HSF, Myb_box5_MYB, W_box_WRKY, Nac_box_NAC, G_box_bHLH, N_box_bHLH, GCC_box_AP2_EREBP, C_ABRE_bZIP
25	AT1G19400	838523	AT1G19400	protein coding	Erythronate-4-phosphate dehydrogenase family protein	3	912	Myb_box3_MYB, Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, AuxRE_ARF, Nac_box_NAC, DREB_AP2_EREBP
26	AT1G20980	838692	SPL14	protein coding	squamosa promoter binding protein-like 14	10	3108	G_box2_bZIP, Myb_box3_MYB, Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, W_box_WRKY
27	AT1G20620	838651	CAT3	protein coding	catalase 3	6	1479	Myb_box3_MYB, Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, W_box_WRKY, N_box_bHLH, HBE_HB
28	AT1G61890	842485	AT1G61890	protein coding	MATE efflux family protein	7	1506	Myb_box1_MYB, Myb_box5_MYB, W_box_WRKY, N_box_bHLH
29	AT1G68820	843214	AT1G68820	protein coding	Transmembrane Fragile-X-F-associated protein	13	1317	Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, W_box_WRKY, AuxRE_ARF, Nac_box_NAC, DREB_AP2_EREBP
30	AT1G61800	842477	GPT2	protein coding	glucose-6-phosphate/phosphate translocator 2	5	1167	Myb_box3_MYB, Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, AuxRE_ARF, Nac_box_NAC, Myb_box4_MYB
31	AT1G19170	838500	AT1G19170	protein coding	Pectin lyase-like superfamily protein	5	1521	G_box2_bZIP, Myb_box3_MYB, Myb_box1_MYB, Myb_box5_MYB, Nac_box_NAC, G_box_bHLH
32	AT1G04830	839404	AT1G04830	protein coding	Ypt/Rab-GAP domain of gyp1p superfamily protein	10	1380	AuxRE_ARF, HSE1_HSF, Myb_box5_MYB, Myb_box1_MYB, W_box_WRKY, G_box2_bZIP
33	AT1G17420	838314	LOX3	protein coding	lipoxygenase 3	8	2760	HSE1_HSF, Myb_box5_MYB, W_box_WRKY, Nac_box_NAC
34	AT1G31830	840073	PUT2	protein coding	Amino acid permease family protein	2	2373*	HSE1_HSF, Myb_box5_MYB, Myb_box1_MYB, W_box_WRKY, G_box_bHLH, G_box2_bZIP, N_box_bHLH, Nac_box_NAC
35	AT1G73500	843685	MKK9	protein coding	MAP kinase 9	1	1553*	AuxRE_ARF, HSE1_HSF, Myb_box5_MYB, Myb_box1_MYB, Myb_box2_MYB, G_box_bHLH, GCC_box_AP2_EREBP, Nac_box_NAC, Myb_box4_MYB
36	AT2G29550	817504	TUB7	protein coding	tubulin beta-7 chain	3	1350	Myb_box3_MYB, Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, W_box_WRKY, AuxRE_ARF, G_box_bHLH, G_ABRE_bZIP
37	AT2G29500	817499	AT2G29500	protein coding	HSP20-like chaperones superfamily protein	1	1056	HSE1_HSF, Myb_box5_MYB, W_box_WRKY, DREB_AP2_EREBP, N_box_bHLH
38	AT2G01020	3767992	AT2G01020	rRNA	—	1	164	Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, W_box_WRKY, Myb_box2_MYB, DREB_AP2_EREBP, N_box_bHLH, GCC_box_AP2_EREBP, C_ABRE_bZIP
39	AT2G37040	818280	PAL1	protein coding	PHE ammonia lyase 1	2	2178	G_box2_bZIP, Myb_box1_MYB, Myb_box5_MYB, W_box_WRKY, Myb_box2_MYB, DREB_AP2_EREBP, N_box_bHLH, C_ABRE_bZIP
40	AT2G05940	815147	RIPK	protein coding	Protein kinase superfamily protein	4	1389	G_box2_bZIP, Myb_box3_MYB, Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, W_box_WRKY, Nac_box_NAC
41	AT2G36530	818226	LOS2	protein coding	Enolase	12	1335	G_box2_bZIP, Myb_box1_MYB, Myb_box5_MYB, W_box_WRKY, Myb_box2_MYB
42	AT2G24360	816972	AT2G24360	protein coding	Protein kinase superfamily protein	3	1236	Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, Nac_box_NAC, G_box_bHLH, N_box_bHLH, GCC_box_AP2_EREBP
43	AT2G39730	818558	RCA	protein coding	rubisco activase	7	1425	G_box2_bZIP, Myb_box1_MYB, Myb_box5_MYB, W_box_WRKY, G_box1_bZIP, G_box_bHLH, Myb_box4_MYB
44	AT2G44840	819093	ERF13	protein coding	ethylene-responsive element binding factor 13	1	1113*	HSE1_HSF, Myb_box5_MYB, W_box_WRKY, DREB_AP2_EREBP

45	AT2G47190	819332	MYB2	protein coding	myb domain protein 2	3	822	Myb_box3_MYB, Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, W_box_WRKY
46	AT2G36780	818250	AT2G36780	protein coding	UDP-Glycosyltransferase superfamily protein	1	2028*	Myb_box3_MYB, Myb_box1_MYB, Myb_box5_MYB, AuxRE_ARF, Nac_box_NAC, G_box_bHLH
47	AT2G01010	3767991	AT2G01010	rRNA	—	1	1808	Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, W_box_WRKY, Myb_box2_MYB, AuxRE_ARF
48	AT2G37570	818333	SLT1	protein coding	HSP20-like chaperones superfamily protein	1	1038	Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, Myb_box2_MYB, AuxRE_ARF
49	AT2G26980	817240	CIPK3	protein coding	CBL-interacting protein kinase 3	16	1326	Myb_box3_MYB, Myb_box1_MYB, Myb_box5_MYB, W_box_WRKY, AuxRE_ARF, N_box_bHLH
50	AT2G33120	817874	SAR1	protein coding	synaptobrevin-related protein 1	5	666	HSE1_HSF, Myb_box5_MYB, Myb_box2_MYB, Nac_box_NAC, G_box_bHLH, Myb_box4_MYB
51	AT2G21330	816672	FBA1	protein coding	fructose-bisphosphate aldolase 1	5	1200	HSE1_HSF, Myb_box5_MYB, DREB_AP2_EREBP, G_box1_bZIP, G_box_bHLH, C_ABRE_bZIP, G_ABRE_bZIP
52	AT2G34850	818050	MEE25	protein coding	NAD(P)-binding Rossmann-fold superfamily protein	10	711	G_box2_bZIP, Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, W_box_WRKY, Myb_box2_MYB, AuxRE_ARF, Nac_box_NAC, Myb_box4_MYB
53	AT2G02760	814805	UBC2	protein coding	ubiquitinating-conjugating enzyme 2	6	459	Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, W_box_WRKY, Myb_box2_MYB, AuxRE_ARF, DREB_AP2_EREBP
54	AT2G23320	816864	WRKY15	protein coding	WRKY DNA-binding protein 15	3	954	Myb_box3_MYB, Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, W_box_WRKY, Myb_box2_MYB, DREB_AP2_EREBP, G_box_bHLH
55	AT3G09390	820098	MT2A	protein coding	metallothionein 2A	2	246	Myb_box3_MYB, Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, W_box_WRKY, DREB_AP2_EREBP, HBE_HB
56	AT3G48520	824011	CYP94B3	protein coding	cytochrome P450, family 94, subfamily B, polypeptide 3	1	1922*	HSE1_HSF, Myb_box5_MYB, W_box_WRKY
57	AT3G19290	821463	ABF4	protein coding	ABRE binding factor 4	3	1299	G_box2_bZIP, Myb_box3_MYB, Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, W_box_WRKY, Myb_box2_MYB, C_ABRE_bZIP
58	AT3G12580	820438	HSP70	protein coding	heat shock protein 70	2	1935	Myb_box3_MYB, Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, W_box_WRKY, Nac_box_NAC, DREB_AP2_EREBP
59	AT3G28910	822525	MYB30	protein coding	myb domain protein 30	3	972	G_box2_bZIP, Myb_box5_MYB, Nac_box_NAC
60	AT3G50000	824162	CKA2	protein coding	casein kinase II, alpha chain 2	10	1212	G_box2_bZIP, Myb_box3_MYB, Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, AuxRE_ARF, G_box_bHLH, GCC_box_AP2_EREBP
61	AT3G14440	820667	NCED3	protein coding	nine-cis-epoxycarotenoid dioxygenase 3	1	2730	Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, AuxRE_ARF, Nac_box_NAC, G_box_bHLH, N_box_bHLH, G_ABRE_bZIP, HBE_HB
62	AT3G57520	824919	SIP2	protein coding	seed imbibition 2	14	2322	Myb_box1_MYB, Myb_box5_MYB, W_box_WRKY, AuxRE_ARF, DREB_AP2_EREBP, G_box_bHLH, G_ABRE_bZIP
63	AT3G17390	821003	MTO3	protein coding	S-adenosylmethionine synthetase family protein	2	1905*	G_box2_bZIP, Myb_box3_MYB, Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, W_box_WRKY, Myb_box2_MYB, DREB_AP2_EREBP
64	AT3G48000	823955	ALDH2B4	protein coding	aldehyde dehydrogenase 2B4	11	1617	G_box2_bZIP, Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, W_box_WRKY, Nac_box_NAC, N_box_bHLH
65	AT3G45980	823741	HTB9	protein coding	Histone superfamily protein	1	727*	Myb_box3_MYB, Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, W_box_WRKY, AuxRE_ARF
66	AT3G43190	823393	SUS4	protein coding	sucrose synthase 4	13	2427	G_box2_bZIP, HSE1_HSF, Myb_box5_MYB, W_box_WRKY, Myb_box2_MYB, N_box_bHLH
67	AT3G55800	824746	SBPASE	protein coding	sedoheptulose-bisphosphatase	8	1182	Myb_box3_MYB, Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, W_box_WRKY, G_box_bHLH
68	AT3G45970	823740	EXLA1	protein coding	expansin-like A1	3	798	G_box2_bZIP, Myb_box5_MYB, AuxRE_ARF

69	AT3G53460	824514	CP29	protein coding	chloroplast RNA-binding protein 29	5	1029	HSE1_HSF, Myb_box5_MYB, W_box_WRKY, Myb_box2_MYB, AuxRE_ARF, Nac_box_NAC
70	AT3G47000	823853	AT3G47000	protein coding	Glycosyl hydrolase family protein	9	1827	G_box2_bZIP, Myb_box3_MYB, Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, W_box_WRKY, G_box_bHLH, N_box_bHLH
71	AT3G15730	820816	PLDAL- -PHA1	protein coding	phospholipase D alpha 1	4	2433	AuxRE_ARF, HSE1_HSF, Myb_box5_MYB, W_box_WRKY, Myb_box2_MYB, Nac_box_NAC
72	AT3G16720	820924	ATL2	protein coding	TOXICOS EN LEVADURA 2	1	1575	AuxRE_ARF, Myb_box3_MYB, HSE1_HSF, Myb_box5_MYB, Myb_box1_MYB, W_box_WRKY, Myb_box2_MYB, DREB_AP2_EREBP, N_box_bHLH, HBE_HB
73	AT4G37930	829949	SHM1	protein coding	serine transhydroxymethyltransferase 1	15	1554	Myb_box3_MYB, Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, W_box_WRKY, DREB_AP2_EREBP
74	AT4G19230	827663	CYP707A1	protein coding	cytochrome P450, family 707, subfamily A, polypeptide 1	1	1404	Myb_box3_MYB, Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, AuxRE_ARF, Nac_box_NAC, G_box_bHLH
75	AT4G03430	827925	EMB2770	protein coding	pre-mRNA splicing factor-like protein	2	4534	Myb_box1_MYB, Myb_box5_MYB, W_box_WRKY, DREB_AP2_EREBP, C_ABRE_bZIP
76	AT4G17500	827464	ERF-1	protein coding	ethylene responsive element binding factor 1	1	1189	Myb_box3_MYB, Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, Nac_box_NAC, G_box_bHLH, Myb_box4_MYB
77	AT4G34230	829572	CAD5	protein coding	cinnamyl alcohol dehydrogenase 5	5	1074	Myb_box3_MYB, Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, W_box_WRKY, Myb_box2_MYB, DREB_AP2_EREBP, G_box_bHLH, G_ABRE_bZIP
78	AT4G20830	827831	AT4G20830	protein coding	FAD-binding Berberine family protein	2	1713	HSE1_HSF, Myb_box5_MYB, W_box_WRKY, Nac_box_NAC, DREB_AP2_EREBP
79	AT4G09010	826480	TL29	protein coding	ascorbate peroxidase 4	10	1050	Myb_box3_MYB, Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, W_box_WRKY, AuxRE_ARF, Nac_box_NAC, DREB_AP2_EREBP, GCC_box_AP2_EREBP, HBE_HB
80	AT4G23050	828404	AT4G23050	protein coding	PAS domain-containing protein tyrosine kinase family protein	14	2208	Myb_box3_MYB, Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, W_box_WRKY, G_box_bHLH, Myb_box4_MYB, G_ABRE_bZIP
81	AT4G18170	827542	WRKY28	protein coding	WRKY DNA-binding protein 28	3	957	Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, W_box_WRKY, N_box_bHLH, Myb_box4_MYB, C_ABRE_bZIP
82	AT4G31780	829306	MGD1	protein coding	monogalactosyl diacylglycerol synthase 1	8	1101	HSE1_HSF, Myb_box5_MYB, W_box_WRKY, Myb_box2_MYB, Nac_box_NAC, DREB_AP2_EREBP, N_box_bHLH
83	AT4G05050	825847	UBQ11	protein coding	ubiquitin 11	1	1434*	Myb_box3_MYB, Myb_box1_MYB, Myb_box5_MYB, W_box_WRKY, AuxRE_ARF, DREB_AP2_EREBP, N_box_bHLH
84	AT4G39090	830064	RD19	protein coding	Papain family cysteine protease	3	1107	Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, AuxRE_ARF, DREB_AP2_EREBP, G_box_bHLH, N_box_bHLH
85	AT4G05100	825855	MYB74	protein coding	myb domain protein 74	3	975	G_box2_bZIP, Myb_box5_MYB, W_box_WRKY, AuxRE_ARF, DREB_AP2_EREBP, G_box_bHLH, N_box_bHLH
86	AT4G37590	829914	NPY5	protein coding	Phototropic-responsive NPH3 family protein	5	1743	HSE1_HSF, Myb_box5_MYB, W_box_WRKY, AuxRE_ARF, Nac_box_NAC, C_ABRE_bZIP
87	AT4G00360	828019	CYP86A2	protein coding	cytochrome P450, family 86, subfamily A, polypeptide 2	2	1662	G_box2_bZIP, Myb_box3_MYB, Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, W_box_WRKY, Myb_box2_MYB, AuxRE_ARF, Nac_box_NAC
88	AT4G05320	825880	UBQ10	protein coding	polyubiquitin 10	3	2332*	G_box2_bZIP, Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, DREB_AP2_EREBP, G_box_bHLH, N_box_bHLH, Myb_box4_MYB, HBE_HB
89	AT4G29190	829040	OZF2	protein coding	Zinc finger C-x8-C-x5-C-x3-H type family protein	1	1553*	Myb_box1_MYB, Myb_box5_MYB, DREB_AP2_EREBP, Myb_box4_MYB
90	AT4G39890	830148	RABH1c	protein coding	RAB GTPase homolog H1C	6	645	Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, Myb_box2_MYB, C_ABRE_bZIP
91	AT4G27260	828834	WES1	protein coding	Auxin-responsive GH3 family protein	3	1839	AuxRE_ARF, Myb_box3_MYB, Myb_box5_MYB, Myb_box1_MYB, W_box_WRKY, Myb_box2_MYB

92	AT5G48655	834923	AT5G48655	protein coding	RING/U-box superfamily protein	6	612	HSE1_HSF, Myb_box5_MYB, Nac_box_NAC, DREB_AP2_EREBP, G_box_bHLH, Myb_box4_MYB
93	AT5G45130	834549	RHA1	protein coding	Ras small GTP-binding family protein	7	603	G_box2_bZIP, Myb_box3_MYB, Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, W_box_WRKY, AuxRE_ARF, DREB_AP2_EREBP, N_box_bHLH, C_ABRE_bZIP
94	AT5G39610	833957	NAC6	protein coding	NAC domain containing protein 6	3	858	Myb_box3_MYB, Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, W_box_WRKY, Nac_box_NAC, HBE_HB
95	AT5G65930	836723	ZWI	protein coding	kinesin-like calmodulin-binding protein (ZWICHEL)	22	3783	Myb_box1_MYB, Myb_box5_MYB, W_box_WRKY, Nac_box_NAC
96	AT5G20000	832122	AT5G20000	protein coding	AAA-type ATPase family protein	9	1260	Myb_box3_MYB, Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, W_box_WRKY, G_box_bHLH, Myb_box4_MYB
97	AT5G47220	834768	ERF2	protein coding	Ethylene responsive element binding factor 2	1	1314	HSE1_HSF, Myb_box5_MYB, Myb_box1_MYB, W_box_WRKY, G_box_bHLH, G_box2_bZIP, Nac_box_NAC
98	AT5G47880	834839	ERF1-1	protein coding	Eukaryotic release factor 1-1	3	2543*	HSE1_HSF, Myb_box5_MYB, Nac_box_NAC
99	AT5G35410	833502	SOS2	protein coding	Protein kinase superfamily protein	13	1341	Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, W_box_WRKY, Myb_box2_MYB, DREB_AP2_EREBP, N_box_bHLH
100	AT5G49360	834996	BXL1	protein coding	beta-xylosidase 1	6	2325	G_box2_bZIP, Myb_box5_MYB, AuxRE_ARF, C_ABRE_bZIP
101	AT5G06370	830525	AT5G06370	protein coding	NC domain-containing protein-like protein	3	780	HSE1_HSF, Myb_box5_MYB, W_box_WRKY, AuxRE_ARF, DREB_AP2_EREBP
102	AT5G66530	836785	AT5G66530	protein coding	Galactose mutarotase-like superfamily protein	13	924	Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, W_box_WRKY, AuxRE_ARF
103	AT5G43760	834398	KCS20	protein coding	3-ketoacyl-CoA synthase 20	2	1590	Myb_box3_MYB, Myb_box1_MYB, Myb_box5_MYB, W_box_WRKY, G_box_bHLH, N_box_bHLH
104	AT5G15650	831419	RGP2	protein coding	Reversibly glycosylated polypeptide 2	4	1083	Myb_box1_MYB, Myb_box5_MYB, W_box_WRKY, DREB_AP2_EREBP, G_box_bHLH
105	AT5G66570	836789	PSBO1	protein coding	PS II oxygen-evolving complex 1	3	999	Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, W_box_WRKY, G_box_bHLH, N_box_bHLH, G_ABRE_bZIP
106	AT5G57550	835859	XTH25	protein coding	Xyloglucan endotransglucosylase/hydrolase 25	3	855	Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, W_box_WRKY, G_box_bHLH, Myb_box4_MYB
107	AT5G09440	830803	EXL4	protein coding	EXORDIUM like 4	1	1523	Myb_box3_MYB, Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, G_box1_bZIP, G_box_bHLH, Myb_box4_MYB, G_ABRE_bZIP
108	AT5G64740	836595	CESA6	protein coding	Cellulose synthase 6	13	3255	G_box2_bZIP, Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, W_box_WRKY, Myb_box2_MYB, AuxRE_ARF, Nac_box_NAC, G_box_bHLH
109	AT5G37540	833732	AT5G37540	protein coding	Eukaryotic aspartyl protease family protein	1	1791*	Myb_box1_MYB, HSE1_HSF, Myb_box5_MYB, W_box_WRKY, Myb_box2_MYB, AuxRE_ARF, DREB_AP2_EREBP, Myb_box4_MYB, C_ABRE_bZIP

\* span length in number of nucleotide



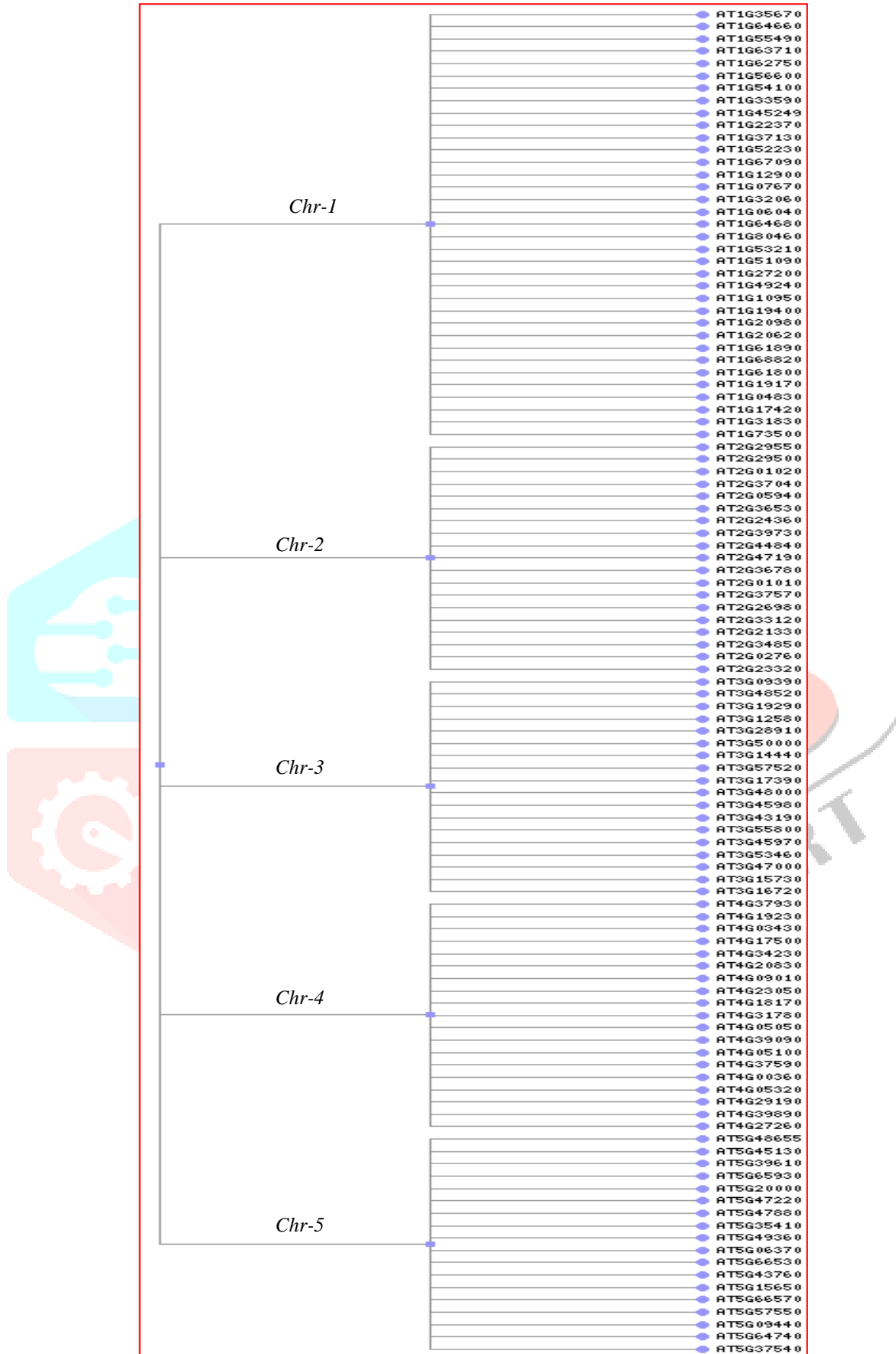
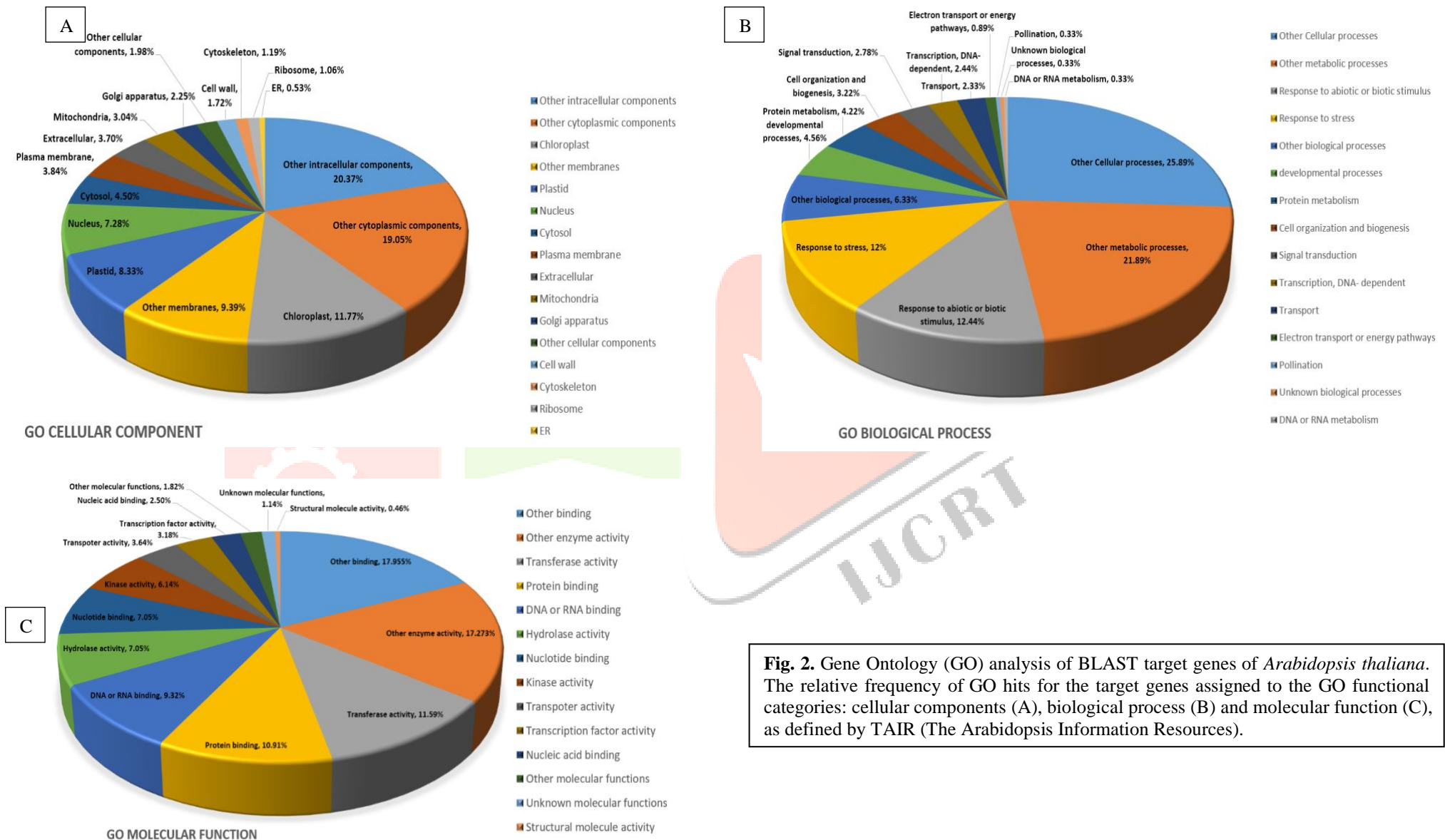


Fig.1. Distribution of drought responsive candidate genes from *solanum taxoides* on 5 chromosomes of *Arabidopsis thaliana*



**Fig. 2.** Gene Ontology (GO) analysis of BLAST target genes of *Arabidopsis thaliana*. The relative frequency of GO hits for the target genes assigned to the GO functional categories: cellular components (A), biological process (B) and molecular function (C), as defined by TAIR (The Arabidopsis Information Resources).

The distribution of 109 stress responsive candidate genes indicated that, maximum number of such stress responsive candidate genes were located on chromosome-1 followed by other 4 chromosomes wherein chromosome 2 and 4 contain 19 genes each, chromosomes 3 and 5 contain 18 genes each which clearly demonstrates that drought tolerance is polygenic trait controlled by few or more quantitative trait loci (QTLs) and such loci/genes are distributed all over the genome. The polygenic nature of drought tolerance controlled by one or more QTLs has been reported in many crops (Bac-Molenaar *et al.*, 2015; Rotino *et al.*, 2013; Bolger *et al.*, 2014).

In the present study, we are reporting 19 transcription factors (TFs) which are involved in stress responsive gene regulation. Among those, MYB/MYC, bZIP, AP2/EREBP, WRKY, DREB (AP2/ERF), HSF and NAC are well-known TF families that are known to play prominent role in the regulation of different genes associated with drought stress (Singh and Laxmi, 2015).

The present *in silico* analysis has revealed that, the genes like CDPK, galactinol synthase 2, lipoxygenase 3, cinnamyl alcohol dehydrogenase 5 and Cellulose synthase 6 are known to be associated with tolerance to stresses like drought. The involvement of such genes in drought tolerance was validated by pre-treatment of wheat seedlings with sodium hydrosulphide and it was found that, drought enhanced the differential expression or regulation of genes, pathways like SOD, transport, CDPK, ABA, Auxin, ribosome biogenesis (Li *et al.*, 2017). Further, the *in silico* analysis has identified that, chaperonin 60 beta, HSP70 and HSP20-like chaperones superfamily protein genes are known express during drought stress which has been experimentally substantiated and validated in the form of enhanced expression of Heat Shock Proteins (HSPs) in wheat (Liu *et al.*, 2017).

AP2/ERF family transcription factors are well recorded in many crop plant species for their role as mediators of both abiotic and biotic stress responses in plants (Licausi *et al.*, 2013). AP2/ERF transcription factors are classified into four sub-families in wheat, including DREB, ERF, AP2, and RAV (Zhuang *et al.*, 2011). It has been reported that, ERFs were rapidly upregulated in response to stresses (He *et al.*, 2011); Further, different kinds of heat-shock proteins (HSPs) and molecular chaperones have been reported to play key role in abiotic stress response (Wang *et al.*, 2004). A chloroplast enhancing stress tolerance protein (CEST) has been shown to overexpress during drought and was reported to have a role in drought tolerance in Arabidopsis (Yokotani *et al.*, 2011). Mitochondria have uncoupling proteins (UCPs) that uncouple electron transport from ATP synthesis. Higher levels of Arabidopsis thaliana uncoupling proteins (AtUCP1) improve tolerance to drought stress. The manipulation of the UCP proteins expression in mitochondria may leads to improved crop varieties with greater drought tolerance (Begcy *et al.*, 2011).

The Gene Ontology categorization of these 109 stress responsive candidate genes based on their role in cellular components has revealed that, 20.37 % of the genes were associated with functions in intracellular components which forms the largest group followed by other cytoplasmic components (19.05%) and genes of chloroplast (11.77%). Similar finding were reported by Vanisri *et al.* (2017) wherein they have indicated that, 57% of the genes were associated with the proteins of intracellular regions followed by plastids (15 %), nucleus and mitochondria (14%). Further, it was evident that, maximum of those genes were coding for DREB proteins which help in understanding the mechanism of drought responses.

The Gene Ontology analysis has revealed that, maximum GO slim terms were recorded for GO molecular function like transferase activity (11.59 %) followed by protein binding (10.90 %), DNA or RNA binding activity (9.32 %), hydrolyse activity (7.05 %) and nucleotide binding activity (7.05 %). Similarly, it has been reported that, proteins involved in various chemical activities like protein kinases, pyrophosphatases, reductases and hydrogenases play key roles in detecting and relaying abiotic stress signals particularly salt stress signals for the regulation of specific genes and thus mediate cellular responses to drought stress. Further, the nucleotide binding activity has closely related with regulation of drought stress modulated genes expression in various plant species (Hrabak *et al.*, 2003; Kalifa *et al.*, 2004a; Kalifa *et al.*, 2004b).

Upon Gene Ontology categorization based on their role in biological processes, maximum number of genes were shown to have role in other cellular processes (25.89%) followed by other metabolic processes (21.89%) and response to abiotic or biotic stimulus (12.44%). Similarly, Orellana *et al.* (2010) reported expressivity of salt and drought regulated genes like Photoassimilate-responsive protein (PAR1), Phosphoglycerate mutase family protein, TSW-12 protein – non-specific lipid protein and others involved in biological processes and all these have some or the other role in drought stress response (Chen *et al.*, 2002; Shanker *et al.*, 2014).

## V CONCLUSION

The major sources for data curation of the *in-silico* research literature are at present limited to Arabidopsis research articles. Present study revealed the activity of candidate genes during drought stress, which can be validated through molecular approach. On confirmation, this information on the drought responsive candidate genes can be used for identification and isolation of drought responsive genes for utilization in drought tolerance breeding in solanaceous crops.

## Abbreviation:

**TFs** Transcription Factors; **ABA** Abscisic Acid; **ABF** ABRE-Binding Factor; **ABRE** ABA-Responsive Element; **AREB** ABRE-Binding Protein; **bZIP** basic domain leucine zipper; **CBF** C-repeat-binding factor; **CRT** C-repeat; **DRE** Dehydration-Responsive Element; **DREB** DRE-BINDING PROTEIN; **ERF** Ethylene-Responsive Element Binding Factor; **bHLH** basic Helix-Loop-Helix; **AP2** APETALA2; **HD-Zip** Homeodomain-leucine zipper; **HSF** Heat Shock Transcription Factor; **G-Box motif** 5'-CCACGTGG-3'; **HBE** Haemoglobin Subunit Epsilon; **MYB** Gene identified from avian myeloblastosis virus; **NAC** Domain from first letters of each genes (NAM- No Apical Meristem; ATAF Arabidopsis Transcription Activation Factor; CUC Cup-Shaped Cotyledon); **WRKY** One of the largest families of transcriptional regulators in plants, play important roles in plant development and various stress responses (Wan *et al.*, 2018); **GO** Gene Ontology

## VI REFERENCES

1. ACEITUNO, F. F., MOSEYKO, N., RHEE, S. Y. AND GUTIÉRREZ, R. A., 2008, The rules of gene expression in plants: Organ identity and gene body methylation are key factors for regulation of gene expression in *Arabidopsis thaliana*. *BMC Genomics*, **9** (1): 438.
2. APPIANO, M., PAVAN, S., CATALANO, D., ZHENG, Z., BRACUTO, V., LOTTI, C., RICHARD G. F. V., RICCIARDI, L. AND BAI, Y., 2015, Identification of candidate MLO powdery mildew susceptibility genes in cultivated Solanaceae and functional characterization of tobacco NtMLO1. *Transgenic Res.*, **24**:847–858.
3. BAC-MOLENAAR, J. A., GRANIER, C., KEURENTJES, J. J. B. AND VREUGDENHIL, D., 2015, Genome-wide association mapping of time-dependent growth responses to moderate drought stress in Arabidopsis. *Plant, Cell & Environ.*, **39**(1), 88–102.
4. BRAY, E. A., 1997, Plant responses to water deficit. *Trends Plant Sci.*, **2**: 48-54.
5. BEGCY, K., MARIANO, E. D., MATTIELLO, L., NUNES, A. V., MAZZAFERA, P., MAIA, I. G. AND MENOSSI, M., 2011, An Arabidopsis Mitochondrial Uncoupling Protein Confers Tolerance to Drought and Salt Stress in Transgenic Tobacco Plants. *PLoS ONE* **6**: e23776.
6. BOLGER, A., SCOSSA, F., BOLGER, M. E., LANZ, C., MAUMUS, F., TOHGE, T.... AND FERNIE, A. R., 2014, The genome of the stress-tolerant wild tomato species *Solanum pennellii*. *Nature Genetics*, **46**(9): 1034–1038.
7. CHEN, W., PROVART, N. J., GLAZEBROOK, J., KATAGIRI, F., CHANG, H. S., EULGEM, T., MAUCH, F., LUAN, S., ZOU, G., WHITHAM, S. A., BUDWORTH, P. R., TAO, Y., XIE, Z., CHEN, X., LAM, S., KREPS, J. A., HARPER, J. F., SI-AMMOUR, A., MAUCH-MANI, B., HEINLEIN, M., KOBAYASHI, K., HOHN, T., DANGL, J. L., WANG, X. AND ZHU, T., 2002, Expression profile matrix of Arabidopsis transcription factor genes suggests their putative functions in response to environmental stresses. *Plant Cell*, **14**: 559-574.
8. HE, Y., LI, W., LV, J., JIA, Y., WANG, M. AND XIA, G., 2011, Ectopic expression of a wheat MYB transcription factor gene, TaMYB73, improves salinity stress tolerance in *Arabidopsis thaliana*. *J. Exp. Bot.*, **63**: 1511–1522.
9. HRABAK, E. E., CHAN, C. C., GRIBSKOV, M., HARPER, J. J., CHOI, J. J., HALFORD, N., KUDLA, J., LUAN, S., NIMMO, H. G. AND SUSSMAN, M. R., 2003, The Arabidopsis CDPK-SnRK superfamily of protein kinases. *Plant Physiol.*, **132**: 666–680.
10. INGRAM, J. AND BARTELS, D., 1996, The molecular basis of dehydration tolerance in plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, **47**: 377-403.
11. KALIFA, Y., GILAD, A., KONRAD, Z., ZACCAI, M., SCOLNIK P. A. AND BAR-ZVI, D., 2004a, the water- and salt-stress-regulated Asr1 (abscisic acid stress ripening) gene encodes a zinc dependent DNA-binding protein. *Biochem. J.*, **381**: 373-378.
12. KALIFA, Y., PERLSON, E., GILAD, A., KONRAD, Z., SCOLNIK, P. A., BAR-ZVI, D., 2004b, Overexpression of the water and salt stress-regulated Asr1 gene confers an increased salt tolerance. *Plant Cell Environ.*, **27**: 1459-1468.
13. KIM, H., BAEK, K., LEE, S., KIM, E. J., LEE, B., CHO, H., KIM, T. W., CHOI, D. AND HUR, C., 2008, Pepper EST database: comprehensive in silico tool for analyzing the chili pepper (*Capsicum annuum*) transcriptome. *BMC Plant Biol.*, **8**:101
14. LI, H., LI, M., WEI, X., ZHANG, X., XUE, R., ZHAO, Y. AND ZHAO, H., 2017, Transcriptome analysis of drought-responsive genes regulated by hydrogen sulfide in wheat (*Triticum aestivum* L.) leaves. *Mol. Genet. Genomics*, **292**: 1091–1110.
15. LIU, Z., QIN, J., TIAN, X., XU, S., WANG, Y., LI, H., 2017, Global profiling of alternative splicing landscape responsive to drought, heat and their combination in wheat (*Triticum aetivum* L.). *Plant Biotechnol. J.*, **16**(3): 714–726.
16. LICAUSI, F., OHME-TAKAGI, M. AND PERATA, P., 2013, APETALA2/ethylene responsive factor (AP2/ERF) transcription factors: mediators of stress responses and developmental programs. *New Phytol.* **199**: 639–649.
17. MI, H., MURUGANUJAN, A., CASAGRANDE, J. T. AND THOMAS, P. D., 2013, Large-scale gene function analysis with the PANTHER classification system. *Nature Protocols*, **8** (8): 1551–1566.



18. NAIKA, M., SHAMEER, K., MATHEW, O. K., GOWDA, R. AND SOWDHAMINI, R., 2013, STIFDB2: an updated version of plant stress-responsive transcription factor database with additional stress signals, stress-responsive transcription factor binding sites and stress-responsive genes in Arabidopsis and rice. *Plant and Cell Physiology.*, **54** (2): 1-15.
19. ORELLANA, S., YAÑEZ, M., ESPINOZA, A., VERDUGO, I., GONZÁLEZ, E., RUIZ-LARA, S. and CASARETTO, J. A., 2010, The transcription factor SIAREB1 confers drought, salt stress tolerance and regulates biotic and abiotic stress-related genes in tomato. *Plant, Cell & Environ.*, **33**(12): 2191–2208.
20. RHEE, S. Y., BEAVIS, W., TANYA, Z., BERARDINI., CHEN, G., DIXON, D., DOYLE, A., HERNANDEZ, M. G., HUALA, E., LANDER, G., MONTOYA, M., MILLER, N., LUKAS, A., MUELLER., MUNDODI, S., REISER, L., TACKLIND, J., WEEM, D. C., WU, Y., XU, I., YOO, D., YOON, J., AND ZHANG, P., 2003, The Arabidopsis Information Resource (TAIR): a model organism database providing a centralized, curated gateway to Arabidopsis biology, research materials and community. *Nucleic Acids Research*, **31**(1): 224–228.
21. ROTINO, G. L., SALA, T. AND TOPPINO, L. 2013, Eggplant: *Alien Gene Transfer in Crop Plants*, **2**: 381–409.
22. SWARBRECK, D., WILKS, C., LAMESCH, P., BERARDINI, T. Z., HERNANDEZ, M. G., FOERSTER, H., LI, D., MEYER, T., MULLER, R., PLOETZ, L., RADENBAUGH, A., SINGH, S., SWING, V., TISSIER, C., ZHANG, P. AND HUALA, E., 2007, The Arabidopsis Information Resource (TAIR): gene structure and function annotation. *Nucleic Acids Research*, **36**: D1009–D1014.
23. SHANKER, A. K., MAHESWARI, M., YADAV, S. K., DESAI, S., BHANU, D., ATTAL, N. B. AND VENKATESWARLU, B. 2014, Drought stress responses in crops. *Func. & Integrative Genomics*, **14**(1): 11–22.
24. SHINOZAKI, K. AND YAMAGUCHI-SHINOZAKI, K., 1996, Molecular responses to drought and cold stress. *Curr. Opin. Biotechnol.*, **7**: 161-167.
25. SINGH, D. AND LAXMI, A., 2015, Transcriptional regulation of drought response: A tortuous network of transcriptional factors. *Front. Plant Sci.*, **6**: 895.
26. VANISRI, S., R ERUVURI, D., AKULA, SAI, J. D., LANAH, S. AND MAMATHA, P., 2017, In silico identification and development of conserved primers from DREB gene in rice. *J Genetics, Genomics & Plant Bre.*, **1**(1): 5-9.
27. VENDRUSCOLO, E. C. G., SCHUSTER, I., PILEGGI, M., SCAPIM, C. C., MOLINARI, H. B. C., MARUR, C. C. AND VIEIRA, L. G. E., 2007, Stress-induced synthesis of proline confers tolerance to water deficit in transgenic wheat. *J. Plant Physiol.*, **164**: 1367–1376.
28. WAN, Y., MAO, M., WAN, D., YANG, Q., YANG, F., MANDLAA, LI, G. AND WANG, R., 2018, Identification of the WRKY gene family and functional analysis of two genes in Caragana intermedia. *BMC Plant Biol.*, **18**: 31.
29. WANG, W., VINO CUR, B., SHOSEYOVAND, O. AND ALTMAN, A., 2004, Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. *TRENDS in Plant Sci.*, **9**(5): 244-253.
30. YOKOTANI, N., HIGUCHI, M., KONDOU, Y., ICHIKAWA, T., IWABUCHI, M., HIROCHIKA, H., MATSUI, M. AND ODA, K., 2011, A novel chloroplast protein CEST induces tolerance to multiple environmental stresses and reduces photooxidative damage in transgenic Arabidopsis. *J. Exp. Bot.*, **6**: 557-569.
31. YE, J., MCGINNIS, S. AND MADDEN, T. L., 2006, BLAST: improvements for better sequence analysis. *Nucleic Acids Research*, **34**: W6–W9.
32. ZHUANG, J., CHEN, J. M., YAO, Q. H., XIONG, F., SUN, C. C. AND ZHOU, X. R., 2011, Discovery and expression profile analysis of AP2/ERF family genes from *Triticum aestivum*. *Mol. Biol. Rep.*, **38**: 745–753.
33. ZHOU, B., ZHANG, L., ULLAH, A., JIN, X., YANG, X. AND ZHANG, X., 2016, Identification of Multiple Stress Responsive Genes by Sequencing a Normalized cDNA Library from Sea-Land Cotton (*Gossypium barbadense* L.). *PLOS ONE*, **11**(3): 1-12