International Journal of Biological, Life and Agricultural Sciences
ISSN: 2415-6612
Vol:8, No:12, 2014

Genome-Wide Analysis of BES1/BZR1 Gene Family in Five Plant Species

Jafar Ahmadi, Zhohreh Asiaban, Sedigheh Fabriki Ourang

***Abstract*—**Brassinosteroids (BRs) regulate cell elongation, vascular differentiation, senescence, and stress responses. BRs signal through the BES1/BZR1 family of transcription factors, which regulate hundreds of target genes involved in this pathway. In this research a comprehensive genome-wide analysis was carried out in BES1/BZR1 gene family in *Arabidopsis thaliana*, *Cucumis sativus, Vitis vinifera*, *Glycin max* and *Brachypodium distachyon*. Specifications of the desired sequences, dot plot and hydropathy plot were analyzed in the protein and genome sequences of five plant species. The maximum amino acid length was attributed to protein sequence Brdic3g with 374aa and the minimum amino acid length was attributed to protein sequence Gm7g with 163aa. The maximum Instability index was attributed to protein sequence AT1G19350 equal with 79.99 and the minimum Instability index was attributed to protein sequence Gm5g equal with 33.22. Aliphatic index of these protein sequences ranged from 47.82 to 78.79 in *Arabidopsis thaliana*, 49.91 to 57.50 in *Vitis vinifera*, 55.09 to 82.43 in *Glycin max*, 54.09 to 54.28 in *Brachypodium distachyon* 55.36 to 56.83 in *Cucumis sativus*. Overall, data obtained from our investigation contributes a better understanding of the complexity of the BES1/BZR1 gene family and provides the first step towards directing future experimental designs to perform systematic analysis of the functions of the BES1/BZR1 gene family.

***Keywords*—**BES1/BZR1, Brassinosteroids, Phylogenetic analysis, Transcription factor.

I. Introduction

BRASSINOSTERIOIDS (BRs) play important roles in many plant growth and developmental processes, including germination, cell expansion and division, photo-morphogenesis, vascular differentiation, senescence and stress/disease resistance [1]-[3]. Mutants defective in *BR* biosynthesis or perception display dwarf phenotypes under both light and dark conditions [4]-[7]. Unlike animal steroid hormones that bind to nuclear receptors to directly regulate expression of target genes, *BRs* function through membrane localized receptor *BRI1* and several other components to regulate the protein levels and activities of BES1 and BZR1 family transcription factors. BES1 and BZR1 have a typical basic helix–loop–helix (bHLH) *DNA* binding domain and bind to E-box (CANNTG) and/or BRRE (BR) Response Element, CGTGT/CG) to regulate *BR* target gene expression. In addition, *BES1* interacts with other transcription factors such as BIM1 and *AtMYB30* as well as putative histone demethylases ELF6 and REF6. *BES1* and BZR1 were identified in genetic screens for positive regulators that function downstream of *BRI1*. The phosphorylation (PTM) states of *BES1* and BZR1 are currently considered to be reliable molecular gauges to assess the level of the *BR* signal output. In addition to BES1/BZR1 gene at biological process: regulation of transcription, DNA-dependent and traceable author statement, BES1 and BZR1 were identified in genetic screens for positive regulators that function downstream of BRI1. *Arabidopsis* IWS1 interacts with transcription factor *BES1* and is involved in brassinosteroid regulated gene expression. However, knowledge about the transcriptional mechanisms by which *BES1/BZR1* regulate gene expression is limited. BIN2 phosphorylated BES1 and BZR1 and negatively regulates their functions [8]-[11]. Other BES1 induced transcription factors (TFs) may function similarly to AtMYB30 to regulate other BR target genes. promoters together with Brassinosteroid (BR) homeostasis and signaling are crucial for normal growth and development of plants. BR signaling through cell-surface receptor kinases and intracellular components leads to dephosphorylation and accumulation of the nuclear protein BZR1. How BR signaling regulates gene expression, however, remains unknown. BZR1 is a transcriptional repressor that has a previously unknown DNA binding domain and binds directly to the promoters of feedback regulated BR biosynthetic genes [12].

Jafar Ahmadi is Associate Professor of Imam Khomeini International University, Qazvin, Iran (Corresponding author to provide phone: +982833901139; Fax: +982833780073; Address: Department of Plant Breeding, Engineering Faculty, Imam Khomeini International University, 34149-16818 Qazvin, Iran, e-mail: njahmadi910@yahoo.com).

Zhohreh Asiaban was M.Sc. Student of Imam Khomeini International University, Qazvin, Iran.

Sedigheh Fabriki Ourang is Assistant Professor of Imam Khomeini International University, Qazvin, Iran (e-mail: s.ourang910@gmail.cm).

II. Material and Methods

A. The Collection of Protein Sequences

In this study, seven genes in *Arabidopsis thaliana (At)*, five genes in *Cucumis sativus (Cs)*, four genes in *Brachypodium distachyon (Brdic)*, four genes in *Vitis vinifera (Vv)* and eight genes in *Glycin max (Gm)* in cooperation with their Cds sequences (mRNA)were identified and collected after searching in National Center for Biotechnology Information (www.ncbi.nlm.nih.gov). Genes loci were listed as AT1G19350, ARALYDRAFT, AT1G78700, AT4G18890, AT3G50750, AT4G36780, AT1G75080, Cs1g (101231601), Cs2g (101230113), Cs3g (101216171), Cs4g (101210002), Cs5g (101204637), Gm1g (100814517), Gm2g (100814217), Gm3g (100813985), Gm4g (100812095), Gm5g (100812053), Gm6g (100795898), Gm7g (100795837), Gm8g (100795369), Vv1g (100263555), Vv2g (100854662), Vv3g (100249673), Vv4g (100241756), Brdic1g (100842420), Brdic2g (100828296), Brdic3g (100825899), Brdic4g (100823855).

B. Dot Plot

Drawing dot plot (repeat within the sequences) was performed with help (arbl.cvmbs.colostate.edu/molkit/) to show the similarity of two DNA sequences by production of a similarity matrix displayed as a dot plot.

C. Hydrophobic Plot

Hydrophobic plot of protein sequences Obtained ofNcbi (http://www.ncbi.nlm.nih.gov/) database was drawn using kyte and Doolittle program (web.expasy.org/protscale/pscale/Hphob.Doolittle.html).

D. Specification of Desired Sequences

Specifications of gene and protein sequences obtained from five plant species analyzed using of programs protparam (http://web.expasy.org/protparam), protscale (http://web.expasy.org/protscale), Ncbi (http://www.ncbi.nlm.nih.gov/), gsds (http://gsds.cbi.pku.edu.cn/chinese.php).

III. Results and Discussion

Number of amino acids in selected BES1/BZR1 protein sequences generally ranged from 265aa to 357aa (At), 206aa to 371aa (Vv), 163aa to 334aa (Gm), 346aa to 374aa (Brdic), 319aa to 327aa (Cs). The maximum aa length is attributed to protein sequence Brdic3g with 374aa and the minimum aa length is attributed to protein sequence Gm7g with 163aa (Table I). The molecular weight of selected BES1/BZR1 protein sequences generally ranged from 28.77 kDa to 39.23 kDa (At), 21.57kDa to 39.97 kDa (Vv), 18.27 kDa to 35.52 kDa (Gm), 36.51 kDa to 38.61 kDa (Brdic), 34.63 kDa to 35.34 kDa (Cc). The maximum molecular weight is attributed to protein sequence Vv3g with 39.97kDa and the minimum molecular weight is attributed to protein sequence Gm7g with 18.27kDa (Table I). Theoretical PI of selected BES1/BZR1 protein sequences generally ranged from 6.84 to 10.74 (At), 8.53 to 9.47 (Vv), 8.49 to 9.37 (Gm), 7.70 to 9.51 (Brdic), 8.50 to 8.97 (Cc). The minimum Theoretical PI is attributed to protein sequence (AT1G78700) with some 6.84 and the maximum Theoretical PI is attributed to protein sequence (AT4G36780) with some 10.74 (Table I).

Total number of negatively charged residues (Asp+Glu) of selected BES1/BZR1 protein sequences generally ranged from 13 to 28(At), 16 to 31 (Vv), 17 to 30 (Gm), 28 to 34 (Brdic), 28 to 31 (Cc). In this study the minimum of negatively charged is attributed to protein sequence (AT4G36780) with 13 numbers and the maximum of negatively charged is attributed to protein sequence Brdic1g with 34 numbers. Total number of positively charged residues (Arg+Lys) of these deduced At, Vv, Gm, Brdic and Cs protein sequences generally ranged from 28 to 42 (At), 23 to 44 (Vv), 21 to 37 (Gm), 29 to 38 (Brdic), 31 to 37 (Cc). The minimum of positively charged is attributed to protein sequence Gm5g with 21 numbers and the maximum of positively charged is attributed to protein sequence Vv3g with 44 numbers (Table I). Instability index of selected BES1/BZR1 protein sequences generally respectively ranged from 49.70 to 79.99 (At), 55.31 to 70.80 (Vv), 33.22 to 73.90 (Gm), 51.87 to 68.64 (Brdic), 58.31 to 70.72 (Cc). The maximum Instability index is attributed to protein sequence (AT1G19350) with some 79.99 and the minimum Instability index is attributed to protein sequence Gm5g with some 33.22 (Table I). Aliphatic index of selected BES1/BZR1 protein sequences generally respectively ranged from 47.82to 78.79 (At), 49.91 to 57.50 (Vv), 55.09 to 82.43 (Gm), 54.09 to 54.28 (Brdic), 55.36 to 56.83 (Vv). The maximum aliphatic index is attributed to protein sequence (AT4G36780) with some 78.79 and the minimum aliphatic is attributed index to protein sequence (AT4G18890) with some 47.82 (Table I).

The dot plot diagram was presented in *Arabidopsis thaliana* Fig. 1. The main diagonal represents the sequence's alignment with itself; lines off the main diagonal represent similar or repetitive patterns within the sequence. The part in outside of the main diagonal indicates mutations in nucleotide sequence (Fig. 1). According to our study, all the *Arabidopsis*, *Vitis vinifera*, *Brachypodium distachyon*, *Glycin max* and *Cucumis sativus* genes were formed the main diagonal. In this five plant species other iterations in outside of the main diagonal have been formed. It is worthy that these iterations often occurred in N-terminal region of nucleotide sequences. Resulting, in all BES1/BZR1 plant genes the main diagonal was involved in transcriptional regulation mechanisms and production of brassinosteroids (Fig. 1).

Kyte-Doolittle is a widely applied scale for delineating hydrophobic character of a protein. Regions with values above zero are hydrophobic in character. To predict potential antigenic sites of globular proteins, which are likely to be rich in charged and polar residues, solvent accessible scales are developed which delineate hydrophobic and hydrophilic characteristics of amino acids. The protein under this study was exposed to Janin, Kyte & Doolittle, Abraham & Leo and Bull & Breese methods to predict its nature and prediction flexibility.All proteins of *Arabidopsis thaliana L., Brachypodium distachyon, Cucumissativus, Glycin max* and *Vitisvinifera* plant species were hydrophilic, of which the proteins AT4G18890 (NP-974145), Brdic3p (XP_003565213.1), Cu5p (XP_004172293.1), Gly5p (XP_003545501.1), Vv2p (XP_003632998.1) showed higher hydrophobic (Fig. 2). All proteins five plant species that quantity Kyte-Doolittle less than (1/6), are hydrophilic. There were not transmembrane region at none of protein sequence in gene family BES1/BZR1. In the other words in these gene families there were not transmembrane regions.

TABLE I
WHOLE-GENOME ANALYSIS OF THE BES1/BZR1 GENE FAMILY IN *ARABIDOPSIS THALIANA*, *CUCUMIS SATIVUS*, *VITIS VINIFERA*, *GLYCIN MAX* AND *BRACHYPODIUM*

Gene name	Length (aa)	MV(kDa)	IP	N	P	Intron	Ch.	Instability index	Aliphatic index
AT1G19350	335	36.48	9.24	28	34	1	I	78	50.27
AT1G19350	337	39.23	9.55	31	41	1	I	79.99	51.2
AT1G78701	333	36	9.24	28	34	1	UN	78.64	53.99
AT1G78700	325	34	6.84	28	28	1	V	62.2	52.92
AT4G18890	284	30.9	8.63	27	31	1	III	49.7	47.82
AT3G50750	276	30.11	9.48	22	32	1	IVX	76.76	61.45
AT4G36780	265	28.77	10.74	13	34	1	I	75.78	78.79
AT1G75080	336	36.48	9.19	28	35	1	I	75.02	52.08
Cs1g(101231601)	319	34.63	8.96	31	37	1	UN	70.72	55.36
Cs2g(101230113)	325	34.68	8.5	28	31	1	UN	58.31	56.83
Cs3g(101216171)	319	34.63	8.96	31	37	1	UN	70.72	55.36
Cs4g(101210002)	325	34.68	8.5	28	31	1	UN	58.31	56.83
Cs5g(101204637)	327	35.64	8.97	29	36	1	UN	61.76	53.79
Cs5g(101204637)	327	35.34	8.97	28	35	1	UN	61.76	53.79
Gm1g(100814517)	330	35.14	9.05	29	37	3	XII	53.71	55.09
Gm2g(100814217)	310	33.56	9.37	30	33	1	VII	73.9	62.71
Gm3g(100813985)	325	35.13	8.94	27	32	1	XII	58.58	55.88
Gm4g(100812095)	322	34.79	8.83	17	21	4	I	55.94	55.81
Gm5g(100812053)	169	19	8.87	29	37	1	XIV	32.22	82.43
Gm6g(100795898)	308	33.28	9.37	18	22	1	XIII	67.3	60.9
Gm7g(100795837)	163	18.27	8.93	28	35	1	XI	47.55	71.78
Gm8g(100795369)	334	35.52	8.99	28	33	1	XIII	56.6	56.47
Vv1g(100263555)	341	36.29	8.53	31	34	1	X	55.31	49.91
Vv2g(100854662)	206	21.57	9.47	31	34	1	X	62.51	56.8
Vv3g(100249673)	371	39.97	9.31	31	37	1	V	60.79	56.63
Vv4g(100241756)	316	33.93	9.04	34	38	1	XIX	70.8	57.5
Brdic1g(100842420)	346	36.51	8.37	29	33	1	I	66.14	54.28
Brdic2g(100828296)	355	37.6	8.64	26	37	1	II	68.64	52.79
Brdic3g(100825899)	374	38.61	9.51	26	37	1	III	51.87	52.09
Brdic4g(100823855)	355	37.48	7.11	28	29	1	IV	61.12	52.56

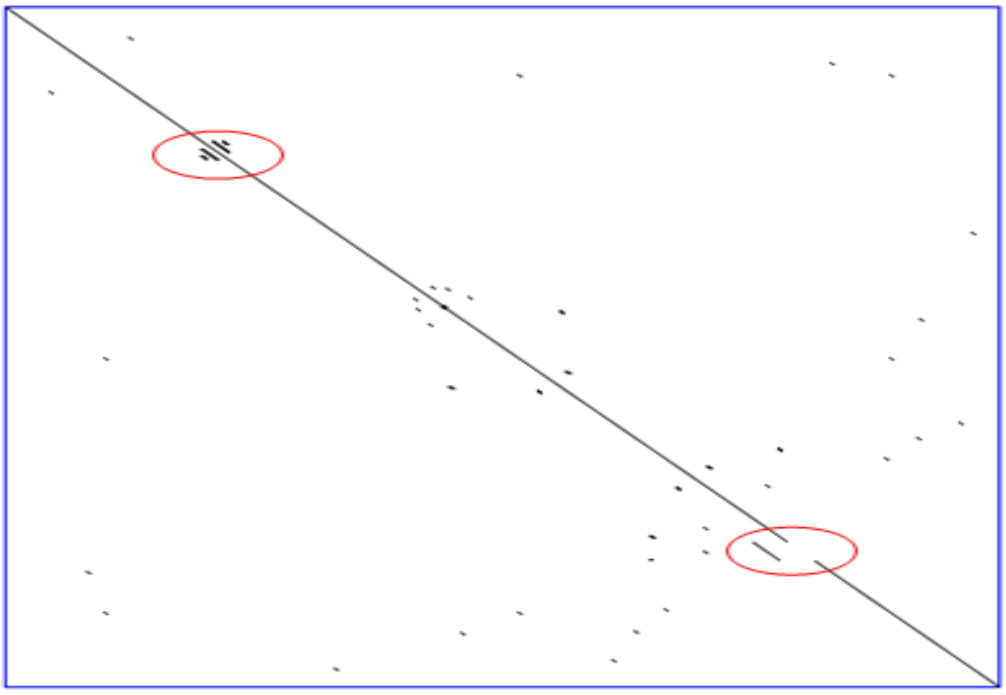

Fig. 1 DNA dot plot of an *Arabidopsis* transcription factor (GenBankNM_106517), showing regional self-similarity

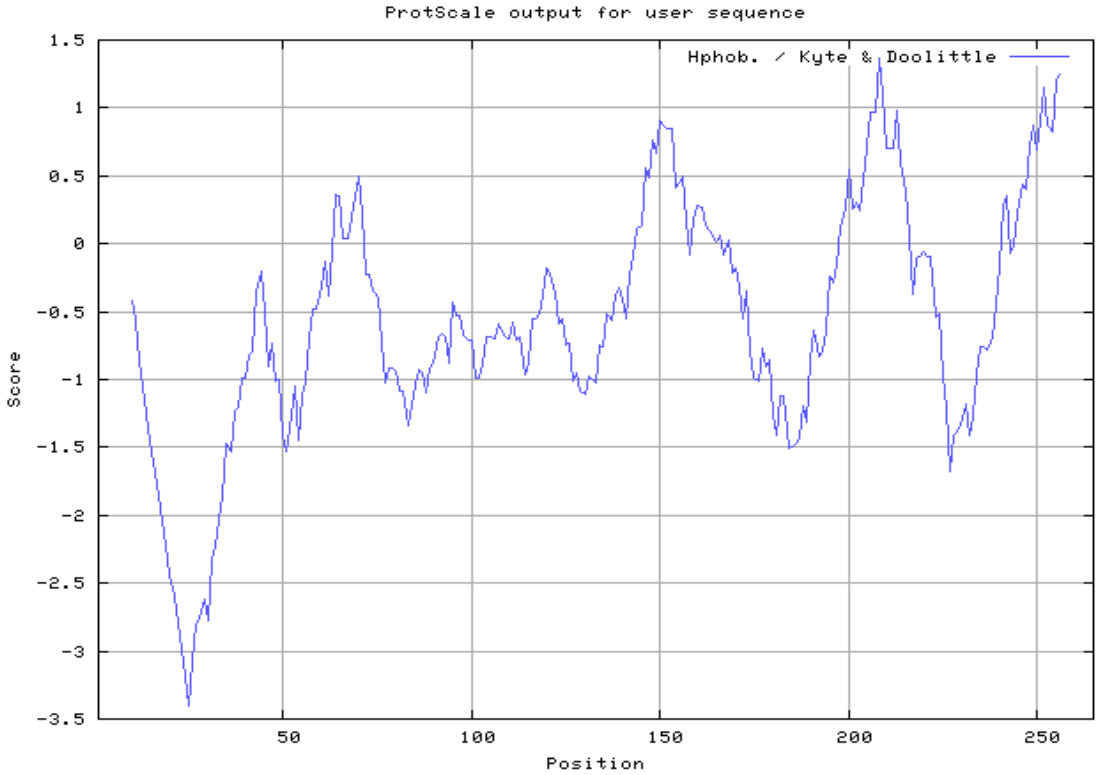

Fig. 2 Hydrophobicity plot of Kyte& Doolittle for the *Arabidopsis thaliana* protein (NP-974145).

ACKNOWLEDGMENT

This research was supported by a research grant in genomics laboratory at Imam Khomeini International University. We would like to thank the authority of Imam Khomeini International University (IKIU) for providing us with a good environment and facilities to complete this project

REFERENCES

[1] L., Li, X., Yu, A., Thompson, M., Guo, T., Yoshida, J. Chory, and Y. Yin. "Arabidopsis MYB30 is a direct target of BES1 and cooperates with BES1 to regulate brassinosteroid-induced gene expression" The plant journal. 2009, 58: 275-285.

[2] S. Clouse "Brassinosteroid Signal Transduction: From Receptor Kinase Activation to Transcriptional Networks Regulating Plant Development." The Plant Cell, 2011, 23: 1219-1230.I. Treich, B. R. Cairns, T. delos Santos, E. Brewster, M. Carlson, "SNF11, a new component of the yeast SNF-SWI complex that interacts with a conserved region of SNF2" Mol. Cell. Biol. 1995, 15: 4240–4248.

[3] Y., Yin, Z., Wang, S., Mora-Garcia, I., Li, S., Yoshida, and J. Chory, "BES1 Accumulates in the Nucleus in Response to Brassinosteroids to Regulate Gene Expression and Promote Stem Elongation" Cell, 2002, 109: 181-191.

[4] L., Li, H., Ye, H Guo, and Y. Yin. "Arabidopsis IWS1 interacts with transcription factor BES1 and is involved in plant steroid hormone brassinosteroid regulated gene expression" PINAS, 2010, 107:3918-3923.

[5] H., Ryu, H., Ch, K. Kim, and J. Hwang, "Phosphorylation Dependent Nucleocytoplasmic Shuttling of BES1 Is a Key Regulatory Event in Brassinosteroid Signaling"Mol. Cells, 2010, 29: 283-290.

[6] S. Kagale, and K. Rozwadowski, "Genome-Wide Analysis of Ethylene-Responsive Element Binding Factor-Associated Amphiphilic Repression Motif-Containing Transcriptional Regulators in Arabidopsis1"Plant Physiol. 2010, 152:1109-1134.

[7] Y., Lin, H., Jiang, Z., Chu, X., Tang, S. Zhu, and B. Cheng, "Genome-wide identification, classification and analysis of heat shock transcription factor family in maize" BMC Genomics, 2011:1-14.

[8] N. Saitou, and M. Nei, "The neighbor-joining method: a new method for reconstructing phylogenetic trees" Mol. Biol. Evol, 1987, 4: 406-425.

[9] J. Kyte, and R. Doolittle, "A simple method for displaying the hydropathic character of a protein" J. Mol. Biol. 1982, 157: 105-132.

[10] T.L. Bailey, C. Elkan "The value of prior knowledge in discovering motifs with MEME" Proc. in Conf. Intell. Syst. Mol. Biol., 2005, 3:21-29.

[11] K., Tamura, J., Dudley, M., Nei, S. Kumar, "MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0." Molecular Biology and Evolution, 2007, 24:1596-1599.

[12] P., Rushton, M., Bokowiec, S., Han, H., Zhang, J., Brannock, X., Chen, T., Laudeman, and M. Timko" Tobacco Transcription Factors: Novel Insights into Transcriptional Regulation in the Solanaceae1" Plant Physiology, 2008, 147: 280-295.