Plant Gene and Trait 2024, Vol.15, No.1, 8-14
http://genbreedpublisher.com/index.php/pgt

Research Report Open Access

Bioinformatics Identification and Expression Profiles of SBP Family Genes in Cucumber (*Cucumis sativus* L.)

Gao Dongju *, Zhang Qin *, Xu Taibai, Zhou Peng, Cheng Wenjing ✉, Zhang Weiwei ✉
Department of Plant Science and Technology, Shanghai Vocational College of Agriculture and Forestry, Shanghai, 201699, China
* These authors contributed equally to this work
✉ Co-Corresponding authors email: chengwj@shafc.edu.cn; zhangww@shafc.edu.cn
Plant Gene and Trait, 2024, Vol.15, No.1 doi: 10.5376/pgt.2024.15.0002
Received: 12 Dec., 2023
Accepted: 20 Jan., 2024
Published: 19 Feb., 2024

Preferred citation for this article:
Gao D.J., Zhang Q., Xu T.B., Zhou P., Cheng W.J., and Zhang W.W., 2024, Bioinformatics identification and expression profiles of SBP family genes in cucumber (*Cucumis sativus* L.), Plant Gene and Trait, 15(1): 8-14 (doi: 10.5376/pgt.2024.15.0002)

Abstract SQUAMOSA promoter binding protein (SBP), as a plant-specific transcription factor, plays an important role in plant growth development. In this study, 15 *SBP* genes were identified from the cucumber genome by bioinformatics methods, and the physicochemical property, gene structure, phylogeny, and expression of these genes in different tissues were analyzed. The results showed that 15 genes were distributed on 4 chromosomes, and divided into 6 groups. Genes in the same group had similar structure and conservative motifs. Expression analysis shows *CsSBP9*, *CsSBP12*, *CsSBP10*, *CsSBP3*, *CsSBP8* and *CsSBP7* are expressed in all tissues, and the other genes are expressed in specific tissues, suggesting that *SBP* genes play an important role in the growth development of cucumber at different stages. This study lays a foundation for the further identification of cucumber *SBP* gene function.
Keywords Cucumber; *SBP*; Gene structure; Gene expression

Transcription factors play a crucial role in plant growth and development. Currently, more than 60 transcription factors have been reported in plants, among which the SBP (SQUAMOSA Promoter Binding Protein) is a plant-specific transcription factor. Goldfish *SBP1* and *SBP2* were the first identified *SBP* genes, and they were named SBP (Klein et al., 1996) because they can bind to the promoter of the floral meristem identity gene SQUAMOSA. The SBP protein contains a conserved SBP domain of about 79 amino acids, generally consisting of two zinc finger structures (Zn1 and Zn2) and a conserved nuclear localization signal (NLS) (Cardon et al., 1999).

Many species have been identified for the SBP gene family, including 16 in Arabidopsis, 18 in rice, 20 in tea, and 32 in bamboo, respectively (Preston and Hileman, 2013; Pan et al., 2017; Wang et al., 2018). Further functional studies have shown that *SBP* genes play important roles in plant growth and development, hormone, and stress signal transduction. Arabidopsis *SPL9* and *SPL15* are involved in the transition from vegetative to reproductive growth (Schwarz et al., 2008); rice *OsSPL10* regulates the initiation of epidermal hair development (Lan et al., 2019); Arabidopsis *SPL8* participates in flower and root development by responding to gibberellin signaling and also affects seed production (Unteand, 2003; Zhang et al., 2007); Overexpression of *VpSBP16* in grape can enhance transgenic plants' tolerance to salt and drought stress (Hou et al., 2018).

Cucumber (*Cucumis sativus* L.) is one of the important vegetables in the world. It is a valuable model plant for studying sex differentiation for its abundant floral types. Although the *SBP* genes play important roles in plant growth and development, the identification of the SBP gene family in cucumber has not been reported. This study identified 15 *SBP* genes from the cucumber genome, and analyzed their chromosomal location, gene structure, conserved motifs, and evolutionary relationships. The expression profiles of the *SBP* genes in different cucumber organs were also analyzed, in order to provide references for further research on the functions of the cucumber *SBP* genes.

1 Results and Analysis

1.1 Identification of *SBP* genes in cucumber

Using the *Arabidopsis thaliana* SBP protein sequence as a reference, the cucumber genome database was searched with the help of bidirectional BLAST method to confirm the presence of structural domains in the obtained SBP protein sequences. Eventually, a total of 15 *SBP* genes were identified in cucumber and named *CsSBP1-CsSBP15* based on their chromosomal locations (Table 1). The longest ORF (Open reading frame) among them was *CsSBP12*, which was 3 096 bp in length, while the shortest gene was *CsSBP15*, which was 426 bp. The molecular weights ranged from 15.86 kDa (*CsSBP15*) to 114.64 kDa (*CsSBP12*), and the pI values ranged from 5.79 (*CsSBP9*) to 9.23 (*CsSBP11*).

Table 1 Information of SBP gene family in cucumber

Gene name	Gene ID	ORF (bp)	Isoelectric point (pI)	Molecular weight (MW)
CsSBP1	Csa1G001450	1140	6.54	42116.83
CsSBP2	Csa1G015680	1653	7.85	60182.15
CsSBP3	Csa1G039890	1647	8.17	60369.26
CsSBP4	Csa1G051590	945	8.82	34844.13
CsSBP5	Csa1G074980	489	6.09	18146.83
CsSBP6	Csa3G117960	1023	7.06	38272.58
CsSBP7	Csa3G151350	1746	6	64916.31
CsSBP8	Csa3G567830	1035	6.73	38621.64
CsSBP9	Csa3G664550	3042	5.79	111376.7
CsSBP10	Csa3G809420	1149	8.82	41157.74
CsSBP11	Csa4G631590	609	9.23	22405.77
CsSBP12	Csa4G664590	3096	8.74	114642.8
CsSBP13	Csa6G094760	987	8.82	36105.68
CsSBP14	Csa6G109120	894	8.81	32717.16
CsSBP15	Csa6G517960	426	6.29	15856.36

The 15 *SBP* genes were distributed on cucumber chromosomes 1, 3, 4, and 6, with the highest number of genes presented on chromosomes 1 and 3, with five genes each. Chromosome 6 contained three genes, while chromosome 4 contained two genes (Figure 1).

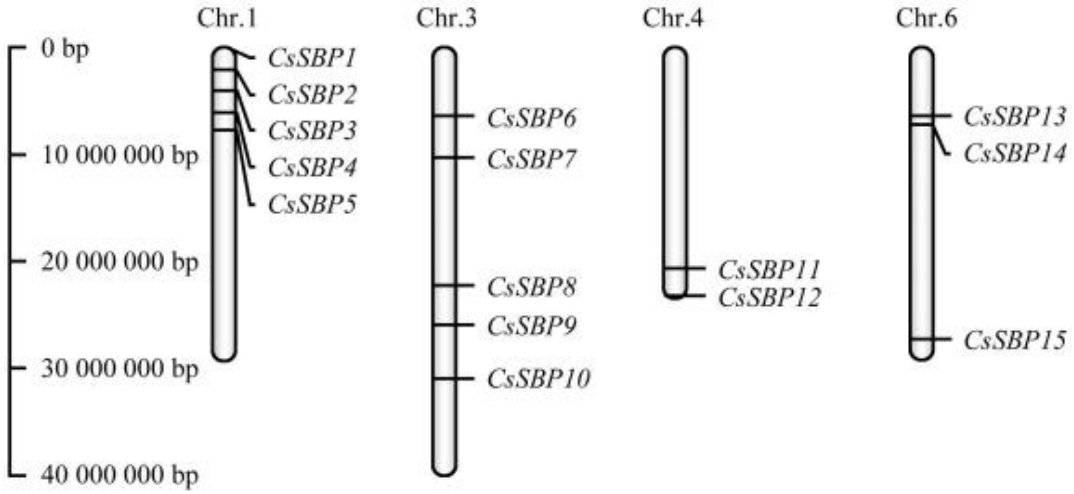

Figure 1 Gene locations of *SBP* genes in cucumber

Note: Scale bar on the left represents the length of the chromosome (bp)

1.2 Phylogenetic analysis of SBP protein

To further clarify the evolutionary relationships among the members of the cucumber SBP gene family, a phylogenetic tree was constructed using SBP family members from Arabidopsis and cucumber (Figure 2). The results showed that the phylogenetic tree could be mainly divided into six major groups (Class I, Class II, Class III, Class IV, Class V, and Class VI). Class II did not contain any cucumber *SBP* genes, while the other five groups included members from both species. Class I had the largest number of members, with six *CsSBP* members. Genes located on the same branch of the phylogenetic tree were closely related orthologous genes. Analysis showed that cucumber and Arabidopsis have direct orthologous genes, such as *CsSBP7* and *AT5G18830*, *CsSBP14* and *AT1G02085*. In addition, there were also a large number of paralogous genes within cucumber, such as *CsSBP2* and *CsSBP3*, *CsSBP1* and *CsSBP4*, and *CsSBP5* and *CsSBP11*, suggesting that *SBP* genes exist in cucumber in the form of a large number of homologous genes.

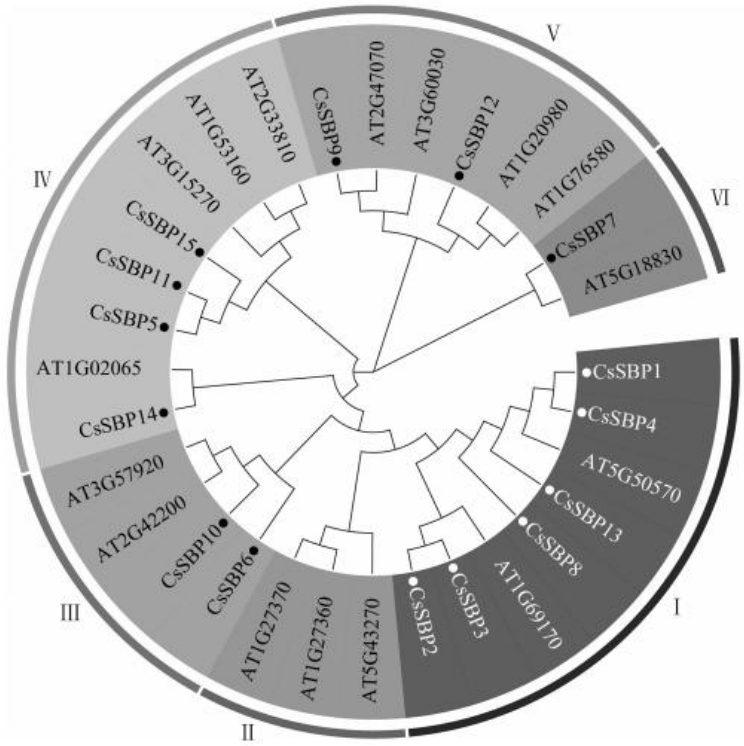

Figure 2 Phylogenetic analysis of *SBP*s in Arabidopsis and cucumber

1.3 Analysis of gene structure and conserved motifs of SBP proteins

To clarify the relationship between the structure, function, and phylogenetic history of the cucumber *SBP* gene family, we analyzed the conserved motifs and gene structures of the members of this gene family (Figure 3). Protein clustering showed that the 15 *SBP* genes were divided into six groups (I, II, III, IV, V, and VI) containing four, two, two, two, four, and one member, respectively. Conserved motif analysis showed that *CsSBP* had three conserved motifs, namely Motif-1, Motif-2, and Motif-3. Except for *CsSBP7*, all 14 other proteins contained these three motifs. Motif-1 was located at the position of the zinc finger Zn2, Motif-2 was located at the position of the zinc finger Zn1, and Motif-3 belonged to the nuclear localization signal domain. In addition, members of the same group had similar motif compositions. For example, members of Class I only contained Motif 1, 2, and 3; Motif-9 only appeared in Class II; members of Class IV, in addition to containing Motif-1, 2, and 3, also had Motif-6 and Motif-10.

Analysis of the *SBP* gene structure showed similar results, with *SBP* genes in the same group having roughly the same number of exons and introns. For example, most members of Class I contained two exons, members of Class II contained ten exons, and members of Class III, Class IV, and Class V contained three exons.

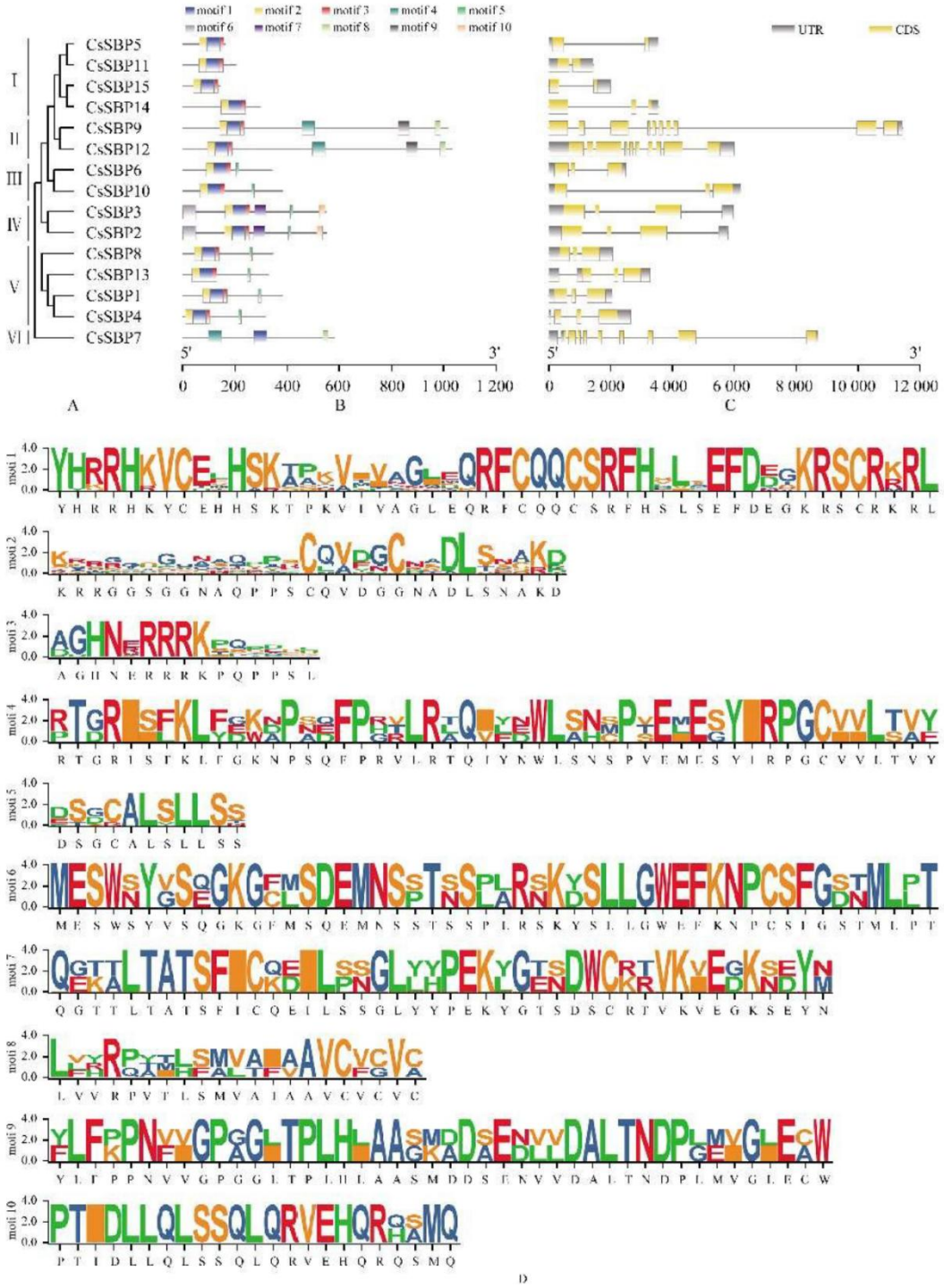

Figure 3 Phylogenetic tree, motif analysis and gene structure of cucumber *SBP* genes

Note: (A) Phylogenetic analysis of 15 cucumber SBP proteins; (B) Conserved motifs analysis; (C) Gene structure analysis; (D) Conserved motif sequence

1.4 Expression analysis of *SBP* genes in different tissues

Based on cucumber transcriptome data, the tissue-specific expression of the 15 *SBP* genes was analyzed in 23 tissues (Figure 4). The results showed that six genes (*CsSBP9*, *CsSBP12*, *CsSBP10*, *CsSBP3*, *CsSBP8* and *CsSBP7*) were expressed in all tissues, especially *CsSBP9* and *CsSBP12* from Class II, which had very high expression levels in all 23 tissues. Other genes showed relatively high expression levels in specific tissues, such as *CsSBP14*, which had higher expression levels in the ovary and fruit skin, and *CsSBP15*, which was mainly expressed in the reproductive organs and had the highest expression levels in leaves and petioles. These results

suggested that *SBP* genes with high expression levels in specific tissues may play important roles in the development of specific organs.

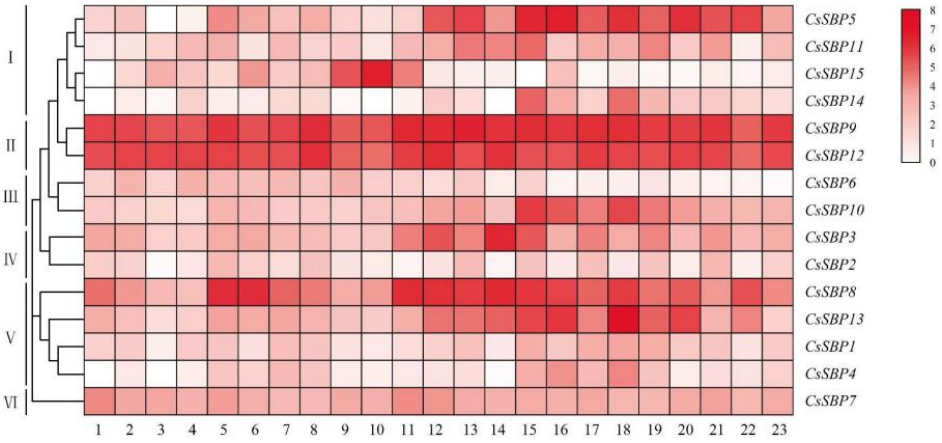

Figure 4 Expression analysis of cucumber SBP genes in different tissues

Note: 1: Roots (4 week seedlings); 2: Hypocotyls (4 week seedlings); 3: Cotyledons (4 week seedlings); 4: Euphylla (4 week seedlings); 5: Root; 6: Stem; 7: Spire; 8: Petiole (Spire); 9: Old leaves; 10: Petiole (Old leaf); 11: Tendril; 12: Female flowers; 13: Male flower bud; 14: Male flowers; 15: Unfertilized ovary; 16: Pericarp (Unfertilized ovary); 17: Pulp (Unfertilized ovary); 18: Pericarp (One week after pollination); 19: Pulp (One week after pollination); 20: Pericarp (Two weeks after pollination); 21: Pulp (Two weeks after pollination); 22: Pericarp (Three weeks after pollination); 23: Pulp (Three weeks after pollination)

2 Discussion

In this study, we identified and analyzed the cucumber SBP gene family members at the whole-genome level using bioinformatics tools, including their chromosome positions, phylogenetic relationships, conserved motifs, gene structures, and expression patterns.

We identified 15 cucumber *SBP* genes in this study, which is similar to the number of SBP members in Arabidopsis (16 members), despite the cucumber genome (367 Mb) being approximately three times larger than the Arabidopsis genome (125 Mb) (Huang et al., 2009). Whole-genome duplication (WGD) is common in angiosperms and can result in gene duplication and the potential for new gene functions. Studies have shown that Arabidopsis experienced three rounds of WGD, with the most recent two rounds (α and β) playing important roles in the rapid expansion of genes, following a whole-genome triplication (γ) event in the common ancestor of flowering plants (Cannon et al., 2004). However, cucumber lacks the two most recent WGD events (Huang et al., 2009), which may explain why the number of *SBP* genes identified in this study is not higher than that in Arabidopsis.

Phylogenetic analysis revealed that cucumber and Arabidopsis SBP members could be classified into six classes (Figure 4), with Class II lacking cucumber SBP members, suggesting that the genes in Class II of Arabidopsis may have undergone independent evolutionary events. Typically, genes of the same type have similar gene structures and motif compositions. Transcription factor domains and motifs are often related to protein interactions, transcriptional activity, and DNA binding (Liu et al., 1999). Eight cucumber members contained the conserved motif Motif-5 (ALSLLS), which corresponds to the target sequence of mRNA156 in Arabidopsis (Rhoades et al., 2002).

Expression analysis showed that *CsSBP9* was expressed in all tissues, and its Arabidopsis homolog *SPL14* (AT1G20980) was expressed in cotyledons, leaves, roots, and floral organs. *CsSBP14* was mainly expressed in reproductive organs, particularly in fruit flesh and skin, and its Arabidopsis homolog *SPL8* was also mainly expressed in inflorescences and siliques. This suggested that these genes have similar biological functions in different species (Stone et al., 2005) and may have undergone convergent evolution (Qian and Zhang, 2014). Our bioinformatics analysis of the cucumber SBP gene family provides a theoretical basis for future studies on the functions of SBP transcription factors.

3 Materials and Methods

3.1 Identification of *SBP* genes in cucumber

In this study, the cucumber SBP family genes were identified using a bidirectional BLAST approach. Firstly, the Arabidopsis SBP protein sequences were aligned to the cucumber genome using TBtools software (e-value, 1e-5) to search for members of the cucumber SBP family (Chen et al., 2020). The cucumber genome file was obtained from the cucumber genome database (ftp://cucurbitgenomics.org/pub/cucurbit/genome/cucumber/Chinese_long/v2/), while the Arabidopsis SBP protein sequences were downloaded from TAIR (https://www.arabidopsis.org/index.jsp). Subsequently, the cucumber SBP obtained from the previous step were confirmed using BLASTP (e-value, 1e-5) in NCBI (https://www.ncbi.nlm.nih.gov/). The SBP functional domains were analyzed using SMART (http://smart.embl.de/) to confirm that the selected proteins were cucumber SBP proteins. The isoelectric point and molecular weight of the cucumber SBP proteins were analyzed using the ProtParam platform (https://web.expasy.org/compute_pi/).

3.2 Chromosomal localization and phylogenetic analysis

TBtools was used to identify the location and distribution of cucumber *SBP* genes on chromosomes. Phylogenetic analysis was then performed using the SBP protein sequences of cucumber and *Arabidopsis thaliana*. The MEGA X software was used to construct a phylogenetic tree using the neighbor-joining (NJ) method with 1 000 bootstrap replicates. Beautification of the phylogenetic tree was performed using Evolview V3 (https://www.evolgenius.info//evolview/#login).

3.3 Analysis of gene structure and conserved protein motifs

TBtools was used to identify the gene structure of cucumber *SBP* genes, while MEME 5.0.5 (http://meme-suite.org/tools/meme) was used to identify the conserved protein motifs of cucumber SBP proteins.

3.4 Expression profiling analysis

To investigate the expression profile of cucumber *SBP* genes in different organs, transcriptome data of various cucumber tissues were obtained from the NCBI website (Accession number: SRP071224), and the analysis methods were followed as described by Wei et al. (2016). TBtools was used to generate a heatmap of the expression profiles of cucumber *SBP* genes.

Authors' contributions

GDJ, ZQ, and ZWW are the designers and conductors of this experiments. GDJ, ZQ, and XTB performed the data analysis and wrote the draft of the manuscript. ZP and CWJ participated in experimental design and data analysis. ZWW conceived and supervised the project, guided the experimental design, data analysis, manuscript writing, and revising. All authors read and approved the final manuscript.

Acknowledgments

This study was supported by the Shanghai Science and Technology Innovation Action Plan in the Agriculture Field (20392001300), the Shanghai Natural Science Foundation (20ZR1439600), the Young Talents Project of Shanghai Agricultural and Forestry Vocational College (A2-0273-20-01-16), and the Internal Project of Shanghai Agricultural and Forestry Vocational College (KY2-0000-20-01).

References

Cannon S.B., Mitra A., Baumgarten A., Young N.D., and May G., 2004, The roles of segmental and tandem gene duplication in the evolution of large gene families in Arabidopsis thaliana, BMC Plant Biol., 4(1): 10.

Cardon G., Höhmann S., Klein J., Nettesheim K., Saedler H., and Huijser P., 1999, Molecular characterisation of the Arabidopsis SBP-box genes, Gene, 237(1): 91-104.
https://doi.org/10.1016/S0378-1119(99)00308-X
PMid:10524240

Chen C., Chen H., Zhang Y., Thomas H.R., Frank M.H., He Y., and Xia R., 2020, TBtools: An integrative toolkit developed for interactive analyses of big biological data, Molecular Plant, 13(8): 1194-1202.
https://doi.org/10.1016/j.molp.2020.06.009
PMid:32585190

Hou H., Jia H., Yan Q., and Wang X., 2018, Overexpression of a SBP-Box Gene (VpSBP16) from Chinese Wild vitis species in Arabidopsis improves salinity and drought stress tolerance, International Journal of Molecular Sciences, 19(4): 940.
https://doi.org/10.3390/ijms19040940
PMid:29565279 PMCid:PMC5979544

Huang S.W., Li R.Q., and Vossen V.D.E.A., 2009, The genome of the cucumber, Cucumis sativus L., Nature Genetics, 41(12): 1275-1281.

Klein J., Saedler H., and Huijse P., 1996, A new family of DNA binding proteins includes putative transcriptional regulators of the Antirrhinum majus floral meristem identity gene SQUAMOSA, Molecular and General Genetics, 250(1): 7-16.
https://doi.org/10.1007/BF02191820
PMid:8569690

Lan T., Zheng Y., Su Z., Yu S., Song H., Zheng X., Lin G., and Wu W., 2019, OsSPL10, a SBP-Box Gene, Plays a dual Role in salt tolerance and trichome formation in rice (Oryza sativa L.), G3: Genes Genomes Genetics, 9(12): 4107-4114.
https://doi.org/10.1534/g3.119.400700
PMid:31611344 PMCid:PMC6893181

Liu L., White M.J., and MacRae T.H., 1999, Transcription factors and their genes in higher plants, European Journal of Biochemistry, 262(2): 247-257.
https://doi.org/10.1046/j.1432-1327.1999.00349.x
PMid:10336605

Pan F., Wang Y., Liu H., Wu M., Chu W., Chen D., and Xiang Y., 2017, Genome-wide identification and expression analysis of SBP-like transcription factor genes in Moso Bamboo (Phyllostachys edulis), BMC Genomics, 18(1): 486.
https://doi.org/10.1186/s12864-017-3882-4
PMid:28655295 PMCid:PMC5488377

Preston J.C., and Hileman L.C., 2013, Functional evolution in the plant Squamosa-Promoter binding Protein-Like (SPL) gene family, Frontiers in Plant Sci., 4: 80.
https://doi.org/10.3389/fpls.2013.00080

Qian W., and Zhang J., 2014, Genomic evidence for adaptation by gene duplication, Genome Res., 24(8): 1356-1362.
https://doi.org/10.1101/gr.172098.114
PMid:24904045 PMCid:PMC4120088

Rhoades M.W., Reinhart B.J., Lim L.P., Burge C.B., Bartel B., and Bartel, D. P., 2002, Prediction of plant microRNA targets, Cell, 110(4): 513-520.
https://doi.org/10.1016/S0092-8674(02)00863-2
PMid:12202040

Schwarz S., Grande A.V., Bujdoso N., Saedler H., and Huijser P., 2008, The microRNA regulated SBP-box genes SPL9 and SPL15 control shoot maturation in Arabidopsis, Plant Mol. Biol., 67(1-2): 183-195.
https://doi.org/10.1007/s11103-008-9310-z
PMid:18278578 PMCid:PMC2295252

Stone J.M., Liang X., Nekl E.R., and Stiers J.J., 2005, Arabidopsis AtSPL14, a plant-specific SBP-domain transcription factor, participates in plant development and sensitivity to fumonisin B1, The Plant Journal, 41(5): 744-754.
https://doi.org/10.1111/j.1365-313X.2005.02334.x
PMid:15703061

Unteand S.U., Anna-Marie S., Paolo P., Madhuri G., Dario L., Heinz S., and Peter H., 2003, SPL8, an SBP-Box gene that affects Pollen Sac Development in Arabidopsis, Plant Cell, 15(4): 1009-1019.
https://doi.org/10.1105/tpc.010678
PMid:12671094 PMCid:PMC152345

Wang P., Chen D., Zheng Y., Jin S., Yang J., and Ye N., 2018, Identification and expression analyses of SBP-Box genes reveal their involvement in abiotic stress and hormone response in tea plant (Camellia Sinensis), International Journal of Molecular Sciences, 19(11): 3404.
https://doi.org/10.3390/ijms19113404
PMid:30380795 PMCid:PMC6274802

Wei G., Tian P., Zhang F., Qin H., Miao H., Chen Q., Hu Z., Cao L., Wang M., Gu X., Huang S., Chen M., and Wang G., 2016, Integrative analyses of Nontargeted Volatile Profiling and transcriptome data provide mlecular Insight into VOC Diversity in Cucumber Plants (Cucumis sativus), Plant Physiol., 172(1): 603-618.
https://doi.org/10.1104/pp.16.01051
PMid:27457123 PMCid:PMC5074635

Zhang Y., Schwarz S., Saedler H., and Huijser P., 2007, SPL8, a local regulator in a subset of gibberellin-mediated developmental processes in Arabidopsis, Plant Molecular Biology, 63(3): 429-439.
https://doi.org/10.1007/s11103-006-9099-6
PMid:17093870