Genome-wide identification and phylogenetic analysis of the SBP-box gene family in melons

Y. Ma[1,2], J.W. Guo[1], R. Bade[1], Z.H. Men[2] and A. Hasi[1]

[1]Inner Mongolia Key Laboratory of Herbage & Endemic Crop Biotechnology,
College of Life Sciences, Inner Mongolia University, Hohhot, China
[2]Department of Biological Science and Technology,
Baotou Teacher's College, Baotou, China

Corresponding author: A. Hasi
E-mail: hasind@sina.com

ABSTRACT. The SBP-box gene family is specific to plants and encodes a class of zinc finger-containing transcription factors with a broad range of functions. Although SBP-box genes have been identified in numerous plants, including green algae, moss, silver birch, snapdragon, *Arabidopsis*, rice, and maize, there is little information concerning SBP-box genes, or the corresponding miR156/157, function in melon. Using the highly conserved sequence of the *Arabidopsis thaliana* SBP-box domain protein as a probe of information sequence, the genome-wide protein database of melon was explored to obtain 13 SBP-box protein sequences, which were further divided into 4 groups, based on phylogenetic analysis. A further analysis centered on the melon SBP-box genetic family's phylogenetic evolution, sequence similarities, gene structure, and miR156 target sequence was also conducted. Analysis of all the expression patterns of melon SBP-box family genes showed that the SBP-box genes were detected in 7 kinds of tissue, and fruit had the highest expression level. *CmSBP11* tends to present its specific expression in melon fruit and root. *CmSBP09* expression was the highest in flower. Overall, the molecular evolution and expression

pattern of the melon SBP-box gene family, revealed by these results, suggest its function differentiation that followed gene duplication.

Key words: Melon; SBP-box protein; Phylogenetic analysis; Gene family

INTRODUCTION

Transcription factors, which are proteins that bind the promoter region of DNA in a sequence-specific manner and regulate gene expression by activating or repressing the transcription of downstream target genes, are found in virtually all living organisms and play an essential role in regulatory networks of many important developmental processes (Riechmann et al., 2000; Hou et al., 2013). A large number of transcription factors are present in plants; they are usually classified into different families and subfamilies based on the sequence of DNA-binding domains (Riechmann et al., 2000), such as the WRKY, AP2/EREBP, MYB, HB, GRAS, NAC, MADS, and SBP transcription factor families. SBP (squamosa promoter binding protein)-box genes encode a family of plant-specific transcription factors (Klein et al., 1996; Cardon et al., 1999) that contain a highly conserved DNA-binding domain, termed the SBP domain. This domain is an assembly of approximately 76 amino acid residues that are involved in both DNA binding and nuclear localization and features 2 zinc-binding sites (Klein et al., 1996; Yamasaki et al., 2004); such plant-specific transcription factor gene families contain several members that are mainly involved in plant growth, development, and a variety of physiological and biochemical processes (Guo et al., 2008). SBP domain proteins were first isolated from *Antirrhinum majus* in an *in vitro* approach to identify regulators of the MADS-box gene SQUAMOSA by their ability to bind to its promoter (Klein et al., 1996). To date, these gene family members have been identified in *Arabidopsis* (Unte et al., 2003), rice (Shao et al., 1999), white birch (Lännenpää et al., 2004), and other plant species. The SBP-box genes play a crucial role in regulating plant growth and development, including regulating flower and fruit development, as well as other physiological processes (Lännenpää et al., 2004; Arazi et al., 2005; Kropat et al., 2005), GA-mediated developmental processes, and copper response process regulation (Zhang et al., 2007). A tomato SBP-box gene was found to be a critical factor in fruit ripening (Manning et al., 2006). Xie et al. (2006) identified 19 SBP-box genes from the rice genome and revealed their predominant expression in several organs. MicroRNAs (miRNAs) have been found to play a crucial role in the regulation of gene function in plants. miRNAs are small RNA molecules (20-24 nucleotides in length) that can cause the degradation of mRNAs or repress translation by binding to the transcripts of their target genes, of which approximately half encode transcription factors (Bartel, 2004). Eleven recognition sites of miRNA156 (miR156) were identified (Schwab et al., 2005; Wu and Poethig, 2006; Gandikota et al., 2007).

Melon (*Cucumis melo* L.) is a eudicot diploid plant species (2n = 2x = 24) of interest for its specific biological properties and for its economic importance (Garcia-Mas et al., 2012). Melon is an attractive model for studying valuable biological characteristics, including sex determination (Boualem et al., 2008; Martin et al., 2009), vascular biology (Haritatos et al., 1996; Gómez et al., 2005), and fruit ripening (Pech et al., 2008). It exhibits extreme diversity in fruit traits. The species is highly polymorphic: melon fruits vary markedly in fruit shape,

size, rind form, firmness, flesh color, ripening behavior, and flavor (Nuñez-Palenius et al., 2008). In addition, melon fruits have significant variations in ripening physiology and can be categorized as either climacteric or nonclimacteric based on their ripening-related respiration rate and ethylene evolution profiles (Giovannoni, 2007). Melon fruit has the potential to be an alternative model to tomato (*Solanum lycopersicum* L.) for fleshy fruit ripening (Giovannoni, 2007; Ezura and Owino, 2008). In June 2012, the accomplishments of the whole-genome sequencing of melon were published, and its genome, being relatively small (approximately 374 MB), statistically supports the analysis, identification, phylogenetic analysis, and functional deduction of its family genes within the scope of the whole genome. The International Cucurbit Genomics Initiative (ICuGI) database [available at http://www.icugi.org (accessed April 3, 2012)] currently includes circa 8,144,030 expressed sequence tags (ESTs) and 611,438 unigenes, nearly a third of which have been derived from fruit libraries (Portnoy et al., 2011). In addition, MELONOMICS (https://melonomics.net) is available for the genome database, which currently includes CM_ protein_v3.5 (accessed May 4, 2011), melon_genome_pseudo-molecules_v3.5 (accessed June 21, 2012), and CM_assembly_scaffold_and_unscaffold_contigs_v3.5 (accessed May 4, 2011). The melon genome sequences are excellent tools for elucidating the genomic structure, gene function, and evolution of species.

The SBP-box gene family of transcription factors is specific to plants and plays roles in plant growth and development. These genes have been identified in a number of plants, but they have not yet been identified in melons. This study performed predictive computer analysis for SBP-box proteins in melons. On the basis of the highly conserved sequences in the domain of the SBP-box family, genes in the SBP-box family in the melon genome were surveyed, resulting in the identification of 13 genes in this family. Phylogenetic analyses were performed, as well as exon/intron and protein motif structural analyses of the SBP-box family genes. In addition, the expression patterns of SBP-box family genes in different tissues or organs were analyzed. The data generated from this study will contribute to studies on the selection of appropriate candidate genes from the SBP-box family in melon for further functional characterization and understanding of the precise regulatory checkpoints that operate during developmental and stress responses.

MATERIAL AND METHODS

Database search for melon SBP-box genes

An *Arabidopsis* protein with a typical structure of the SBP domain was selected as the probe sequence in the National Center for Biotechnology Information (NCBI) database (ACCESSION: NM_180137), and Basic Local Alignment Search Tool (BLAST) (Altschul et al., 1990) of the Bioedit 7.0.9 tool (Hall, 1999) was used to search the similar protein sequences in the CM_protein_v3.5 database (https://melonomics.net/files/Genome), which subordinates to MELONOMICS, with the aim of identifying all the candidate proteins that contain the SBP domain in the melon genome. We also searched the database using amino acid sequences of the SBP domain of other members of the *Arabidopsis* SBP-box family as the query to confirm that the search was exhaustive. Sequences with an E-value $\leq 10^{-3}$ were regarded as candidate proteins with an SBP domain, while other parameters were the default values.

To further confirm the postulated SBP-box family genes, the amino acid sequences were then searched for the SBP-box domain using the Simple Modular Architecture Research Tool (SMART) (Schultz et al., 1998; Letunic et al., 2009).

Multiple sequence alignment, tree building, and conserved motif prediction

Multiple sequence alignment, using Clustal W (Larkin et al., 2007) with default parameters, predicted the melon SBP-box protein sequences. The DNAMAN Version 7.0 software was also used as a secondary method for aligning sequences and rechecking results. To compare the evolutionary relationships of SBP-box family members, multiple sequence alignment was applied, by way of Clustal W, on already obtained SBP-box protein sequences.

A phylogenetic tree was constructed with aligned SBP-box protein sequences of the domains using MEGA5.10 (Tamura et al., 2007) and the neighbor-joining (NJ) method (Saitou and Nei, 1987) with Poisson correction, pairwise deletion, and bootstrap (1000 replicates; random seeds) as parameters. Simultaneously, the maximum parsimony (MP) method of the MEGA5.10 software (Tamura et al., 2007) was employed to create a second phylogenetic tree with a bootstrap of 1000 replicates to validate the results from the NJ method. The subsequent tree file was visualized by MEGA5.10 (Tamura et al., 2007).

MEME Suite version 4.8.1 (Bailey et al., 2009) (http://meme.nbcr.net/meme/cgi-bin/meme.cgi) was used with the following parameters: optimum width 6-300 amino acids, any number of repetitions of a motif, and maximum number of motifs set to 5. The resulting motifs were checked in the NCBI and SMART databases for their significance.

Location of miR156 target sites

The SBP-box gene family is characterized remarkably by the regulation of miR156; as in the GRAS gene family, the expression of some genes is regulated by miR171 - namely, SBP-box gene family members of some of the genes expressed by miR156 regulation. It is pivotal to identify whether the gene sequences of these transcription factors contain miR156 target sites to understand their roles in regulatory functions.

On the basis of the GFDB database (http://gfdb.sdau.edu.cn/miRNA.asp), miR156 target sites (TGACAGAAGAGAGGGAGCAC) were found, alignment analysis was carried out using the DNAMAN7.0 software with SBP-box gene family nucleotide sequences, and reverse complementary sequences of miR156 in these genes were aligned. The resulting miR156 target site sequences with no more than 3 mismatching bases were considered as positive sequences.

Expression analysis of melon SBP-box genes

The expression characteristics of the SBP-box gene family were detected using ESTs. We used nucleotide sequences of the SBP-box, which are available from MELONOMICS, as probing sequences for the proteins in melons. A nucleotide BLAST analysis in the NCBI EST database (http://www.ncbi.nlm.nih. gov/blast.cgi) was completed. As a result, 1948 ESTs were obtained. Then, the sequences that had a maximum identity of >90 and <100% were screened (Zhang et al., 2012). Redundant sequences were removed from the data set, and 789 final sequences were obtained. The EST annotation information of the 789 obtained sequences was used to analyze the expression patterns in different tissues.

RESULTS

Identification of 13 SBP-box genes in melon

In order to identify the CmSBP-box genes in melon genomes, the SBP-box domain of a melon SBP-box amino acid sequence was used as a protein BLAST query sequence; 13 genes were identified as possibly encoding proteins containing the SBP-box domain (Table 1). In addition, other members of the *Arabidopsis* SBP-box family were used as the query, and the results were consistent with the initial search (ACCESSION: NM_180137), thereby certifying its reliability.

The same 13 sequences were also obtained from the MELONOMICS database using a hidden Markov model with PF03110 containing a typical SBP-box domain. The individual genes are listed in Table 1. Among them, the 13 that were predicted to encode proteins containing the SBP-box domain were thus assigned to the SBP-box family. On the basis of the amino acid identity of SBP-box domains, the 13 SBP-box family members were further classified into 4 groups.

SMART analysis indicated that the SBP-box domain of each of the 13 CmSBP-box genes was typical, thereby certifying their reliability.

Table 1. Characteristics of the melon SBP-box gene family.

Group/Gene name	Gene ID	Scaffold distribution	Melonomics (5'-3')	Length of ORF (bp)	No. of introns	No. of exons
(I)CmSBP01	MELO3C002048	scaffold00001	875601-880467	1647	3	4
(I)CmSBP02	MELO3C002370	scaffold00001	3021965-3023985	969	3	4
(I)CmSBP07	MELO3C014895	scaffold00023	3552716-3553996	1035	2	3
(I)CmSBP09	MELO3C017245	scaffold00030	1173804-1177486	1650	3	4
(I)CmSBP11	MELO3C022318	scaffold00052	816363-818219	957	3	4
(I)CmSBP13	MELO3C026196	scaffold00089	268868-269561	543	1	2
(II)CmSBP06	MELO3C012909	scaffold00018	2907768-2918681	3042	10	11
(II)CmSBP08	MELO3C017069	scaffold00029	4110679-4115739	3081	9	10
(II)CmSBP10	MELO3C021144	scaffold00047	37851-40769	894	2	3
(II)CmSBP12	MELO3C025597	scaffold00081	154445-155157	600	1	2
(III)CmSBP03	MELO3C005966	scaffold00006	281687-284009	1020	2	3
(III)CmSBP05	MELO3C009639	scaffold00011	3705169-3709768	1149	2	3
(IV)CmSBP04	MELO3C006597	scaffold00006	4445631-4448756	1443	6	7

Phylogenetic analysis of the SBP-box gene family in melon

To determine the evolutionary relationship among CmSBP-box proteins, an unrooted NJ (Saitou and Nei, 1987) phylogenetic tree was constructed with bootstrap analysis (1000 replicates) based on the multiple sequence alignments of the 13 CmSBP-box proteins (Figure 1).

The analysis revealed that the 13 CmSBP-box members were divided into 4 groups, designated I to IV (Table 1). The NJ tree reliability was certified by generating another phylogenetic tree using the MP method, whereupon it was found that nearly all the CmSBP-box members were placed within the same groups. The results show that members in group I contained 1 or 3 introns; members of group II contained 1, 2, 9, or 10 introns; members of group III contained 2 introns; and members of group IV contained 6 introns (Table 1).

Functional divergence in the SBP-box gene family was caused by different numbers and intervals of introns in the members of the 4 groups. A large difference in the number of introns within the group implies that members in the group are in the differentiation period of gene function. This result was consistent with those of Guo et al. (2008) who classified the SBP-box gene family in 9 representative green plants and, Hou et al. (2013), who analyzed grapes.

A phylogenetic tree was constructed using amino acid sequences of SBP-box domains. The results showed that the SBP-box gene family includes 5 pairs of duplicated genes - *CmSBP02* and *CmSBP13*, *CmSBP01* and *CmSBP09*, *CmSBP08* and *CmSBP06*, *CmSBP10* and *CmSBP12*, and *CmSBP03* and *CmSBP05* (Figure 1).

These 5 pairs of duplicated genes are likely to have different numbers of introns and be gene duplications of the same ancestral gene in evolution, which occurred prior to the divergence of the monocotyledonous and dicotyledonous plant species.

The two external nodes that are in the end of the same clades of the phylogenetic tree are likely to be the closest homologous gene pairs. Analysis of the phylogenetic tree showed that genes in the same clade were duplicated several times before the formation of melon species. The SBP-box gene occurred after the formation of repeated events. This shows that the ancestral gene of the SBP-box gene family originated with the divergence of other species.

In this study, a phylogenetic tree of 13 melon SBP-box protein sequences was constructed. The phylogenetic tree (Figure 1) shows that these sequences are divided into 4 subfamilies. The classification of SBP-box transcription factor proteins in *Arabidopsis* is consistent with that in melon, and the SBP-box gene family is highly conserved between *Arabidopsis* and melon.
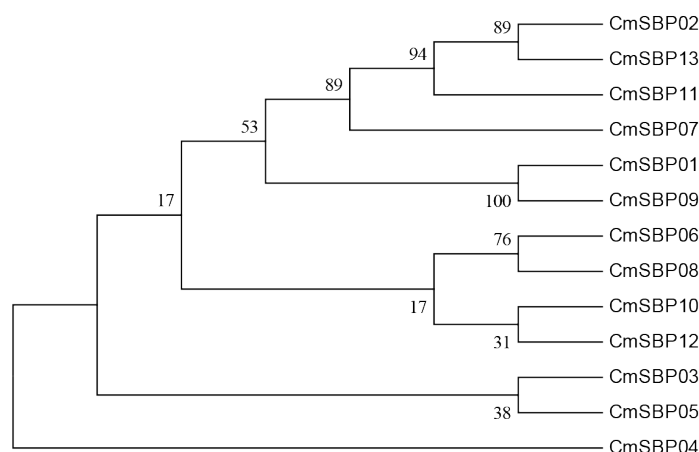

Figure 1. Phylogenetic analysis of melon SBP-box proteins.

Analysis of the melon SBP-box protein domain

For the convenience of this research, these 13 SBP genes were renamed *CmSBP01-CmSBP13* according to their Gene ID number. In order to compare the characteristics of these 13 CmSBP proteins, a multiple sequence alignment based on the CmSBP protein domain obtained through SMART and full-length SBP domain of amino acid sequences was performed with the DNAMAN version 7.0 software. The CmSBP proteins share 1 conserved SBP domain (Figure 2), 2 zinc finger protein class structures (Cys2HisCys), and a highly conserved bipartite nuclear localization signal (Figure 2). Melon SBP domains tend to be very identical, being highly conserved in certain areas in which the most conserved amino acid residues function with 2 zinc ions to coordinate cysteine and histidine residues. Similarly, the relatively conserved amino acid residues are involved in the nucleic localization of the basic

amino acid residues. In all SBP domains of the CmSBP proteins, there are no redundant amino acid residues. The sequence alignment analysis of the CmSBP protein amino acid sequences revealed that, except for the highly conserved CmSBP domain, the sequence variability of the N-terminus and C-terminus was relatively conspicuous (Figure 3); thus, the varied protein functional specificity of CmSBP can be determined.

By comparing the full-length sequence of CmSBP proteins, except for the conserved SBP domain, the most of the CmSBP amino acid sequences contain a conserved lysine near the upstream SBP domain.
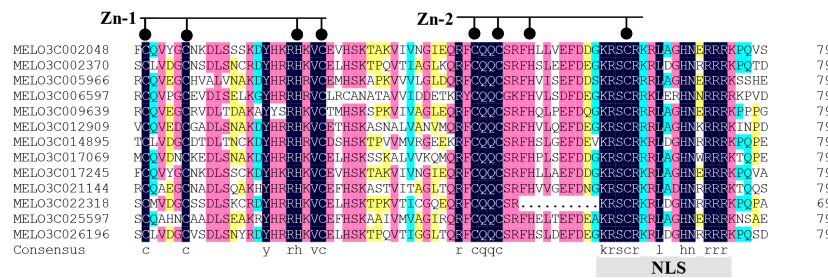
Figure 2. Analysis of the SBP-box domain in melon. Zn-1 and Zn-2 represent zinc finger protein class structures (Cys2HisCys), and NLS a highly conserved bipartite nuclear localization signal.
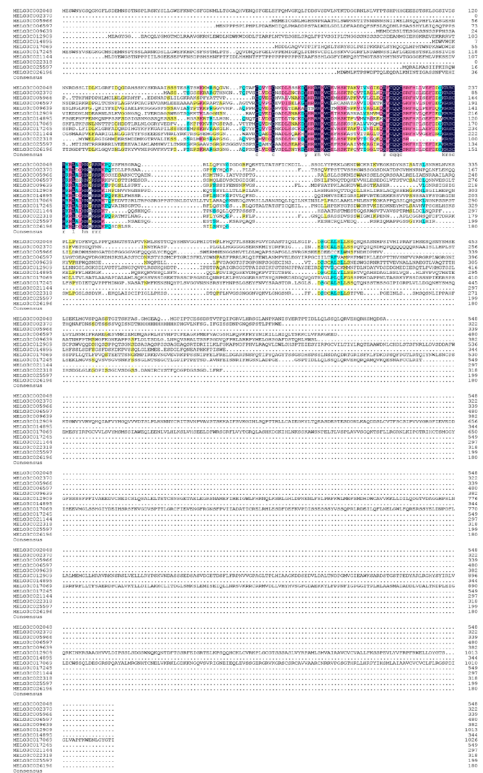

Figure 3. Full-length sequence analysis of melon SBP-box proteins.

Conserved motif analysis of melon SBP-box proteins

In order to verify the existence of a conserved motif in the CmSBP-box family, the CmSBP-box family was searched using the MEME online software, and 3 conserved motifs of the already-discovered SBP domain were detected (Figures 4 and 5). Very few biological functions of motifs are known. In all SBP-box family members, the SBP domain, composed of CM-1 and CM-2, has been widely identified. In addition, CmSBP01, 02, 05, 06, 09, and 11 contain CM-4 [DS(GD)CALSLLSS], which is the highly conserved bipartite nuclear localization signal (Figures 4 and 5). Also, CmSBP01 and 09 contain CM-5, while CmSBP04, 06, and 08 contain CM-4, and all 5 duplicated gene pairs share the motifs and are identical in type and quantity. The data above justify the prediction of the existence of the 5 duplicated gene pairs in SBP family proteins. Moreover, the intron analysis suggests the insertion of quantity-variable introns into the coded sequence of the SBP-box gene, verifying that the SBP-box protein was generated by gene duplication and differentiated by the insertion of introns.
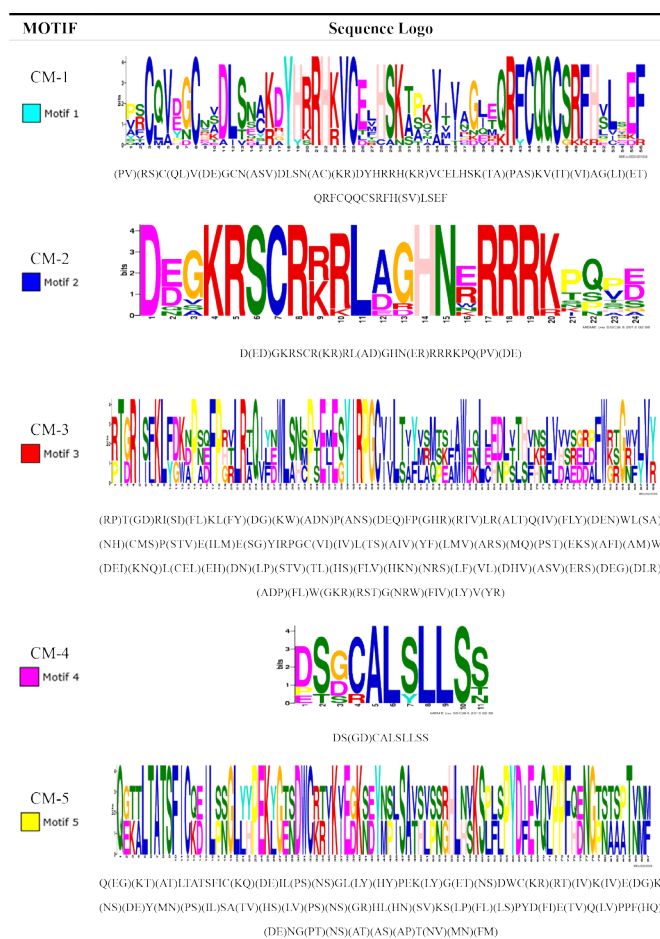

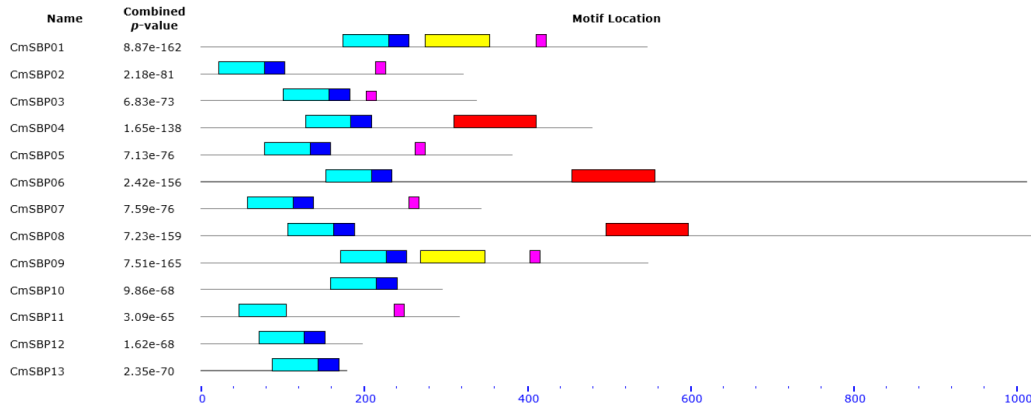**Figure 4.** Analysis of conserved motifs (CMs) of SBP-box family genes in melon.

Figure 5. Distribution of conserved motifs of SBP-box family genes in melon.

Localization of miR156 target sequences

Target sites of miR156 in all plants tend to be completely conserved, and varied miRNA target sites all encode a conserved peptide, ALSLLS, in exons of different genes (Guo et al., 2008). A comparison of the miR156 target sequence and the nucleotide sequence that corresponds with the 13-amino acid miR156 target sequence (Figures 3 and 6) showed that there were 6 miR156 recognition sites, accounting for 46.2% of the total family members. This result is consistent with that of related research in *Arabidopsis* and grapes, and it demonstrates the reliability and accuracy of the predictions of the SBP-box gene family. Although the genomes of *Arabidopsis thaliana*, melon, and grape are different in size, they share miR156 recognition sites, suggesting that SBP-box genes are extremely conserved in the evolution of various species, which might have been caused by the necessity and significance of SBP-box genes in plant functions.


```
CmSBP01    GIPDSGCALSLLSSQSQSASNHPSIVHIPRA      439
CmSBP02    .INESDCALSLLSSPQTQIQSDHQ.......      235
CmSBP05    SVTDSSCALSLLSNESWGSRNRTTTVEVN..      289
CmSBP07    ...DSGCALYLLSSRTVQAHSDAG.......      276
CmSBP09    GLSDSGCALSLLSSQTQNSSTHSSGIPIGRP      431
CmSBP11    .FFDSDCALSLLSTPVEPGEIN.........      256
Consensus     s cal lls
```

Figure 6. Analysis of the localization of miR156 target sequences.

Analysis of melon SBP-box gene expression

The 131,052 ESTs from *C. melo* have been deposited in the GenBank database (http://www.ncbi.nlm.nih.gov/nucest/?term=melon). These partial transcript sequences are valuable for gene expression pattern observations because the expressional characteristics and abundance of genes were revealed via these partial transcript sequences. To demonstrate the expression patterns of these SBP-box family genes across various tissues, SBP-box gene expression was analyzed in 7 tissues (fruit, root, flower, callus, leaf, cotyledon, and phloem) (Figure 7). Among them, fruit had the highest expression level, accounting for 29.5% of the

total expression. The root reached second place, with 22.8% of the total expression. This was followed by 16.7, 15.8, 7.6, 4.4, and 3.2% of the expression in flower, callus, leaf, cotyledon, and phloem, respectively. The percentage of SBP-box family gene expression in phloem (3.2%) was the lowest of all the tissues. To further study the SBP-box gene expression pattern of single genes among transcriptomes of different tissues, the expression profiles of 13 SBP-box family genes of *C. melo* were examined by analyzing EST counts (Figure 8). Redundant sequences were removed from the dataset, and 789 final sequences were obtained and used to determine the gene expression patterns in different tissues. *CmSBP11* tended to have specific expression in melon fruit and root with multiple copy forms. *CmSBP09* was expressed the highest in flower.
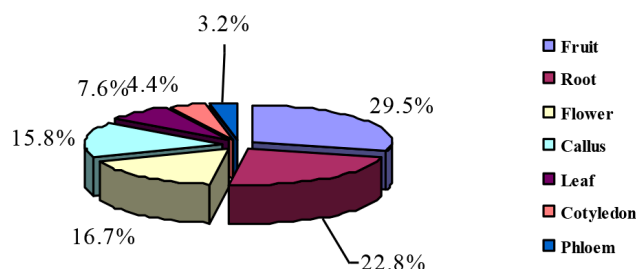

Figure 7. Expression of SBP-box family genes in different melon tissues and organs *in silico*.
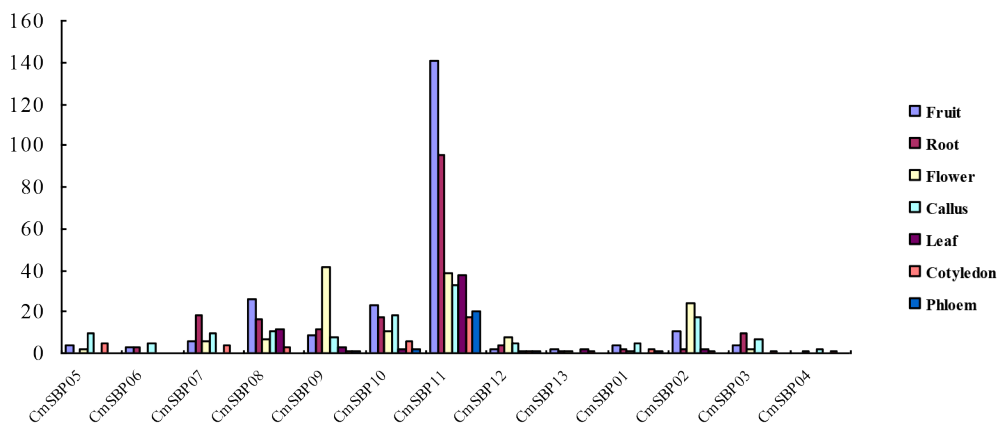

Figure 8. Expression patterns of melon SBP-box genes in various tissues. The expression profile was determined by analyzing the expressed sequence tag (EST) counts based on non-redundant ESTs of melon.

DISCUSSION

Thirteen SBP proteins exist in melon that could be divided into 4 groups. Among these groups, group I contains 6 genes and group II contains 2 genes. *A. thaliana* was found to contain 16 SBP-box genes (Riechmann et al., 2000), while strawberry was found to contain 16 SBP-box genes.

Gene duplication and differentiation have long been viewed as the major pathways of new gene origin and the differentiation of gene function. According to the classical theory,

when a gene duplicates, the initial function of duplicated genes tends to be completely redundant, but because of the relaxation of selection pressure, the two duplicate genes gradually and separately accumulate genetic variations to achieve changes in gene structure and function (Zhang, 2003).

Current views grant duplicated genes 3 different fates. First, one copy inherits the gene function of its ancestor, while the other one becomes a pseudogene. Second, one copy holds the gene function of its ancestor, while the other acquires a new function that is varied from the former. Third, 2 copies are sub-functionalized, and they together bear the gene function inherited from their ancestor (Moore and Purugganan, 2005). Among the 13 SBP-box genes that were identified, there are 5 duplicated gene pairs. Of these, *CmSBP01*, *CmSBP09*, *CmSBP02*, and *CmSBP13* shared expression characteristics. These pairs might have sub-functionalized so that they coordinately perform functions that once belonged to one of their ancestors, but the clarification of the primal and secondary status within those duplicated gene pairs needs further investigation.

Cardon et al. (1999) and Schmid et al. (2003) determined that the SBP-box genes (*AtSPL1*, *12*, *14*, and *17*) with relatively longer coding sequences in *Arabidopsis* were involved in constitutive expression, and the SBP-box genes with moderate or relatively short coding sequences were upregulated during the ripening of flowers (Birkenbihl et al., 2005). Thus, *CmSBP01* and *CmSBP09* might belong to SBP-box genes with longer coding sequences, and *CmSBP11*, *CmSBP10*, *CmSBP06*, *CmSBP03*, and *CmSBP12* pertain to SBP-box genes with moderate or short coding sequences.

ESTs are created by partially sequencing randomly isolated gene transcripts, and they have proved valuable in molecular biology. The identification of SBP-box family genes and analysis of gene expression patterns using ESTs are classic examples of their utility. In this study, to detect the expression patterns of the 13 genes in the SBP-box family of melon, 7 tissue types were analyzed according to EST annotations (Figure 8). Gene transcripts were the most abundant in fruit tissue, followed by root and flower tissue. Most of the SBP-box family genes were detected in all 7 tissue types. The *CmSBP11* gene was particularly interesting in that its level of expression was the highest among all the SBP-box family genes, and it was highly expressed in fruit tissue. Therefore, we speculate that the SBP-box family may play an important role mainly in fruit growth and development in melon. In addition, abundant ESTs were detected in the callus, supporting a role for SBP-box proteins in the stress response. The natural gene expression levels in various tissues were not caused by the frequencies of ESTs in GenBank. However, it is reasonable for researchers to outline the gene expression pattern of their interest. This study would be of interest to plant molecular biologists, particularly those interested in the SBP-box gene family. They could use this premise of this study to predict proteins in other types of plants.

ACKNOWLEDGMENTS

REFERENCES

Altschul SF, Gish W, Miller W, Myers EW, et al. (1990). Basic Local Alignment Search Tool. *J. Mol. Biol.* 215: 403-410.

Arazi T, Talmor-Neiman M, Stav R, Riese M, et al. (2005). Cloning and characterization of micro-RNAs from moss. *Plant J.* 43: 837-848.

Bailey TL, Boden M, Buske FA, Frith M, et al. (2009). MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.* 37: W202-W208.

Bartel DP (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116: 281-297.

Birkenbihl RP, Jach G, Saedler H and Huijser P (2005). Functional dissection of the plant-specific SBP-domain: overlap of the DNA-binding and nuclear localization domains. *J. Mol. Biol* 352: 585-596.

Boualem A, Fergany M, Fernandez R, Troadec C, et al. (2008). A conserved mutation in an ethylene biosynthesis enzyme leads to andromonoecy in melons. *Science* 321: 836-838.

Cardon G, Hohmann S, Klein J, Nettesheim K, et al. (1999). Molecular characterisation of the *Arabidopsis* SBP-box genes. *Gene* 237: 91-104.

Ezura H and Owino WO (2008). Melon, an alternative model plant for elucidating fruit ripening. *Plant Sci.* 175: 121-129.

Gandikota M, Birkenbihl RP, Hohmann S, Cardon GH, et al. (2007). The miRNA156/157 recognition element in the 3'UTR of the *Arabidopsis* SBP box gene SPL3 prevents early flowering by translational inhibition in seedlings. *Plant J.* 49: 683-693.

Garcia-Mas J, Benjak A, Sanseverino W, Bourgeois M, et al. (2012). The genome of melon (*Cucumis melo* L.). *Proc. Natl. Acad. Sci. U. S. A.* 109: 11872-11877.

Giovannoni JJ (2007). Fruit ripening mutants yield insights into ripening control. *Curr. Opin. Plant Biol.* 10: 283-289.

Gómez G, Torres H and Pallas V (2005). Identification of translocatable RNA-binding phloem proteins from melon, potential components of the long-distance RNA transport system. *Plant J.* 41: 107-116.

Guo AY, Zhu QH, Gu X, Ge S, et al. (2008). Genome-wide identification and evolutionary analysis of the plant specific SBP-box transcription factor family. *Gene* 418: 1-8.

Hall TA (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* 41: 95-98.

Haritatos E, Keller F and Turgeon R (1996). Raffinose oligosaccharide concentrations measured in individual cell and tissue types in *Cucumis melo* L. leaves: implications for phloem loading. *Planta* 198: 614-622.

Hou HM, Li J, Gao M, Singer SD, et al. (2013). Genomic organization, phylogenetic comparison and differential expression of the SBP-box family genes in grape. *PLoS One* 8: e59358.

Klein J, Saedler H and Huijser P (1996). A new family of DNA binding proteins includes putative transcriptional regulators of the *Antirrhinum majus* floral meristem identity gene *SQUAMOSA*. *Mol. Gen. Genet.* 250: 7-16.

Kropat J, Tottey S, Birkenbihl RP, Depege N, et al. (2005). A regulator of nutritional copper signaling in *Chlamydomonas* is an SBP domain protein that recognizes the GTAC core of copper response element. *Proc. Natl. Acad. Sci. U. S. A.* 102: 18730-18735.

Lännenpää M, Janonen I, Holtta-Vuori M, Gardemeister M, et al. (2004). A new SBP-box gene BpSPL1 in silver birch (*Betula pendula*). *Physiol. Plant* 120: 491-500.

Larkin MA, Blackshields G, Brown NP, Chenna R, et al. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* 23: 2947-2948.

Letunic I, Doerks T and Bork P (2009). SMART 6: recent updates and new developments. *Nucleic Acids Res.* 37: D229-D232.

Manning K, Tor M, Poole M, Hong Y, et al. (2006). A naturally occurring epigenetic mutation in a gene encoding an SBP-box transcription factor inhibits tomato fruit ripening. *Nat. Genet.* 38: 948-952.

Martin A, Troadec C, Boualem A, Rajab M, et al. (2009). A transposon-induced epigenetic change leads to sex determination in melon. *Nature* 461: 1135-1138.

Moore RC and Purugganan MD (2005). The evolutionary dynamics of plant duplicate genes. *Curr. Opin. Plant Biol.* 8: 122-128.

Nuñez-Palenius HG, Gomez-Lim M, Ochoa-Alejo N, Grumet R, et al. (2008). Melon fruits: genetic diversity, physiology, and biotechnology features. *Crit. Rev. Biotechnol.* 28: 13-55.

Pech JC, Bouzayen M and Latché A (2008). Climacteric fruit ripening: ethylene-dependent and independent regulation of ripening pathways in melon fruit. *Plant Sci.* 175: 114-120.

Portnoy V, Diber A, Pollock S, Karchi H, et al. (2011). Use of non-normalized, non-amplified cDNA for 454-based RNA sequencing of fleshy melon fruit. *Plant Genome* 4: 36-46.

Riechmann JL, Heard J, Martin G, Reuber L, et al. (2000). *Arabidopsis* transcription factors: genome-wide comparative

analysis among eukaryotes. *Science* 290: 2105-2110.

Saitou N and Nei M (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4: 406-425.

Schmid M, Uhlenhaut NH, Godard F, Demar M, et al. (2003). Dissection of floral induction pathways using global expression analysis. *Development* 130: 6001-6012.

Schultz J, Milpetz F, Bork P and Ponting CP (1998). SMART, a simple modular architecture research tool: identification of signaling domains. *Proc. Natl. Acad. Sci. U. S. A.* 95: 5857-5864.

Schwab R, Palatnik JF, Riester M, Schommer C, et al. (2005). Specific effects of microRNAs on the plant transcriptome. *Dev. Cell* 8: 517-527.

Shao CX, Takeda Y, Hatano S, Matsuoka M, et al. (1999). Rice genes encoding the SBP domain protein, which is a new type of transcription factor controlling plant development. *Rice Genet. Newsl.* 16: 114.

Tamura K, Dudley J, Nei M and Kumar S (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* 24: 1596-1599.

Unte US, Sorensen AM, Pesaresi P, Gandikota M, et al. (2003). SPL8, an SBP-box gene that affects pollen sac development in *Arabidopsis*. *Plant Cell* 15: 1009-1019.

Wu G and Poethig RS (2006). Temporal regulation of shoot development in *Arabidopsis thaliana* by miR156 and its target SPL3. *Development* 133: 3539-3547.

Xie K, Wu C and Xiong L (2006). Genomic organization, differential expression, and interaction of SQUAMOSA promoter-binding-like transcription factors and microRNA156 in rice. *Plant Physiol.* 142: 280-293.

Yamasaki K, Kigawa T, Inoue M, Tateno M, et al. (2004). A novel zinc-binding motif revealed by solution structures of DNA-binding domains of *Arabidopsis* SBP-family transcription factors. *J. Mol. Biol.* 337: 49-63.

Zhang CH, Shangguan LF, Ma RJ, Sun X, et al. (2012). Genome-wide analysis of the AP2/ERF superfamily in peach (*Prunus persica*). *Genet. Mol. Res.* 11: 4789-4809.

Zhang JZ (2003). Evolution by gene duplication: an update. *Trends Ecol. Evol.* 18: 292-298.

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 Mol. Biol.* 63: 429-439.