

Article

Genome-Wide Identification and Functional Exploration of SBP-Box Gene Family in Black Pepper (*Piper nigrum* L.)

Jing Li ^{1,2,3,4}, Rui Fan ^{1,2,3,4}, Baoduo Wu ^{1,2,3,4}, Xunzhi Ji ^{1,2,3,4} and Chaoyun Hao ^{1,2,3,4,*}

¹ Spice and Beverage Research Institute, Chinese Academy of Tropical Agricultural Sciences, Wanning 571533, China; xys8619@163.com (J.L.); tlfr83@163.com (R.F.); 0204114089@163.com (B.W.); 18876794730@163.com (X.J.)

² Key Laboratory of Genetic Resources Utilization of Spice and Beverage Crops, Ministry of Agriculture, Wanning 571533, China

³ Hainan Provincial Key Laboratory of Genetic Improvement and Quality Regulation for Tropical Spice and Beverage Crops, Wanning 571533, China

⁴ Academician Sim Soonliang of Hainan Province Research Station, Wanning 571533, China

* Correspondence: haochy79@163.com

Abstract: Black pepper (*Piper nigrum* L.), is dubbed “the King of Spices”. However, the lack of genic knowledge has limited the understanding of its physiological processes and hindered the development of its molecular breeding. The SBP-box gene family is an important family in plant development and integrates multiple physiological processes. Here, we made a genome-wide identification of the pepper SBP-box gene family to provide evolutionary and functional information about this conserved transcription factor. In total, 34 *SBP* genes were identified in pepper. All these pepper *SBP* genes were clustered into eight groups, and one pepper group was not found in *Arabidopsis thaliana*. Segment duplications played the most important role in the expansion process of pepper *SBP* genes, and all these duplications were subjected to purifying selection. Half of pepper *SBP* genes were found miR156 target sites, and 17 miR156s were predicted. The tissue expression analysis revealed the differential expression of pepper *SBP* genes. Eleven *SBP* genes were found in four co-expression networks, and the GO enrichment further provides a functional prediction for pepper *SBP* genes. This study lays a foundation for further studies of pepper and provides a valuable reference for functional mining of pepper *SBP* genes.

Keywords: *Piper nigrum*; SBP-box gene family; co-expression network; miR156



Citation: Li, J.; Fan, R.; Wu, B.; Ji, X.; Hao, C. Genome-Wide Identification and Functional Exploration of SBP-Box Gene Family in Black Pepper (*Piper nigrum* L.). *Genes* **2021**, *12*, 1740. <https://doi.org/10.3390/genes12111740>

Academic Editor: Jacqueline Batley

Received: 11 October 2021

Accepted: 28 October 2021

Published: 29 October 2021

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Transcription factors (TFs), which are proteins that regulate gene expression by activating or repressing the transcription of downstream target genes, play essential roles in the regulatory networks of many development processes in all living organisms [1,2]. Different families or subfamilies of TFs were found in plants according to their structure of DNA-binding domain, such as WRKY, MYB, MADS-box, DOF, and NAC transcription factor families. SQUAMOSA promoter-binding protein (briefly: SBP-box) or SBP-like (SPL) genes encode a type of TF family that is conserved in green plants. *SBP* genes were first found in *Antirrhinum majus* to regulate the expression of some MADS-box genes that play a critical role in floral development [3]. Since then, studies continue to reveal that *SBP* genes play critical roles in the regulation of plant development processes, signal transduction, defense processes from monocyte algae to flowering plants, and especially the outstanding roles in angiosperms flowering development and time-dependent phase transformation processes [4–7]. The SBP-box gene family play important roles in plant growth and development of the regulation network, and it is a key link to explore the molecular mechanism of plant growth and development [8]. In addition, the evolutionary patterns of the SBP-box gene family are highly consistent with the evolution of angiosperms, so it is also an important resource to reveal the plant lineage-specific evolution process.

The SBP domain, which contains about 75 amino acid residues (aa), is the typical structure for all SBP TFs. As studies revealed, the SBP domain is sufficient to bind to the GTAC core motif [4–7]. Generally, the SBP domain contains three characteristic structures: two zinc fingers and a nuclear localization signal (NLS). Mostly, the NLS overlaps with the second zinc finger [9]. In addition, some *SBP* genes are targets of miR156. The miR156 mature sequences can reduce the protein expression level at the transcriptional or translational stage by complementarily binding to their target mRNAs [10,11]. To angiosperms, a considerable part of *SBP* genes can be regulated by miR156s. It has been identified that 10 of 16 *AtSPL* genes are potential targets of miR156/157 (collectively known as miR156). Some *SBP* genes regulated by miR156s are involved in time-dependent or tissue-specific development processes. For example, the miR156 and SBP module play critical roles in the juvenile-to-adult phase transition of *A. thaliana* [12–14]. In *A. thaliana*, *AtSPL3/4/5*, which are direct upstream activators of *LEAFY*, *FRUITFUL*, and *APETALA1*, redundantly promote flowering [12]. They also integrate photoperiodic signals and developmental aging in the flowering locus T (FT)/flowering locus D (FD) module process [13]. *AtSPL9/15* are regarded as regulators of plastochron and branching [15,16]. *AtSPL7* is a regulator of copper homeostasis and responds to light and copper [17].

Pepper is a member of the family Piperaceae and is originally native to India. Pepper is known for its medicinal properties and unique aroma and flavor [18]. It is one of the most popular spices, thus also called “the King of Spices”. Different types of pepper products have been widely used in the medical, food, and cosmetics industries [19–21]. Due to the huge commercial value and increasing market demand, pepper has become increasingly attractive. Additionally, pepper is an early branch of angiosperms [22], and is likely a key link to explain the formation and expansion of angiosperms. The limited genic studies of pepper led to the lack of knowledge about the physiological processes such as growth and development, and stress response. At the same time, it also limits the development of worldwide pepper molecular breeding. To improve the yield, quality, and disease resistance, it is extremely necessary to perform more research on genes playing critical roles in multiple important physiological processes.

SBP genes have been identified in many species through the bioinformatic method, however, there has been no relevant reports in pepper. Fortunately, the high-precision pepper genome assembly allows more in-depth research to be conducted on the pepper SBP-box gene family [23]. In this study, we performed a genome-wide investigation of the SBP-box gene family in pepper. HMM and BLASTP searches were both adopted to predict *SBP* genes in a genome-wide way in pepper; the phylogenetic tree was reconstructed by both ML (maximum likelihood) and NJ (neighbor-joining) methods; motif analysis, and multiple sequence alignment were used to parse the protein structure; the duplications of pepper *SBP* genes were predicted and their selection pressure was calculated; miR156s were identified and the corresponding miR156 binding sites were predicted in pepper *SBP* genes; tissue expression pattern of pepper *SBP* genes was displayed; and, at last, functional analysis including construction of a co-expression network and a GO enrichment were carried out. Based on the above analysis, both evolutionary and functional information of the pepper SBP-box gene family was revealed, which is expected to be useful for more in-depth studies in the future.

2. Methods

2.1. Data Sources

The genomic and proteomic sequences of pepper were obtained from the assembly of our laboratory and have been deposited at the website (<http://cotton.hzau.edu.cn/EN/download.php> accessed on 30 February 2021). The *A. thaliana* genomic and proteomic sequences were obtained from TAIR database (accessed on 30 March 2021) [24]. The gene expression data for pepper was from previous experiment data of our laboratory and has been deposited in the SRA database under BioProject number PRJNA529760 (30 February 2021). The expression data of *A. thaliana* were obtained from TAIR database (30 March 2021).

2.2. Identification and Characterization of Pepper SBP Genes

Both HMM and BLASTP searches were performed to accurately predict the pepper SBP genes [25,26]. For HMM search, the SBP-specific HMM profile (PF03110) was obtained from the Pfam database [27], and the HMMER toolkit was adopted to find suspected SBP genes. In addition, the well-characterized *A. thaliana* SBP protein sequences were collected from PlantTFDB as queries for BLASTP searches ($e\text{-value} \leq 1 \times 10^{-10}$). The SBP structural integrity was confirmed using the Simple Modular Architecture Research Tool (SMART) [28]. The online toolkit ExPASy Proteomics Server was used to predict the physicochemical properties [29], including protein length, molecular weight (MW), and isoelectric point (pI).

2.3. Phylogenetic Analysis and Conserved Motif Prediction

Multiple sequence alignment of SBP protein sequences was performed using Multiple Sequence Comparison by Log-Expectation (MUSCLE) (v3.8) and ClustalX (v2.0) [30,31]. A neighbor-joining (NJ) tree was generated using MEGA 11 with 1000 bootstrap replicates [32]. To make a more reliable phylogenetic analysis, IQTREE was used to reconstruct the maximum likelihood (ML) tree [33]. The final phylogenetic relationship was based on the results obtained from the two methods, and the visualization was conducted by EvolView software [34]. The motifs of pepper SBP proteins were predicted using an online toolkit of Multiple Expectation Maximization for Motif Elucidation (MEME) [35]. All whole long SBP proteins were adopted as queries, and the parameters were set as follows: minimum width was 6, maximum width was 150, motif number was 10, and the minimum number of sites was 2.

2.4. Chromosomal Localization and Detection of Gene Duplication

According to the position on chromosomes, the location of each SBP gene was visualized. Both tandem duplications and segmental duplications here were predicted according to the method described in the Plant Genome Duplication Database [36]. MCScanX was used to detect the segment duplicate alignments [37], and the all-against-all BLASTP comparison ($e\text{-value} \leq 1 \times 10^{-10}$) was performed to supply the input data for MCScanX analysis. The segment duplications were manually confirmed from alignments produced from MCScanX analysis. Tandem duplications were accepted as those genes next to each other or separated by one unrelated gene.

2.5. Estimation of Synonymous (K_s) and Nonsynonymous (K_a) Substitutions per Site and Their Ratio (K_a/K_s)

The SBP genes that satisfied segment duplications were used to estimate K_a (the number of nonsynonymous substitutions per nonsynonymous site), K_s (the number of synonymous substitutions per synonymous site), and their ratio (K_a/K_s). Coding sequences from segmentally duplicated SBP genes were aligned using PRANK [38]. The estimation of K_a , K_s , and K_a/K_s were developed using KaKs_Calculator (v2.0) [39], and the MA (a model that averaged the result from many classical calculation models provided by KaKs_Calculator) model was adopted. The K_a/K_s value can reflect the selective pressure of duplicated gene pairs, and the K_s value can reflect the divergence time for duplication events.

2.6. MicroRNA and Target Prediction

The miR156/157 sequences were downloaded from the miRBase database [40]. Due to the high similarity of miR156 and miR157, they were collectively referred to as miR156 family. The prediction of pepper miR156 was performed according to the sequence similarity. The precursor sequences of *miR156* genes were used as queries to blastn against the pepper genome database, and the sequences were further confirmed using MFold to make sure a suitable RNA secondary structure with relative higher negative minimal free folding energy [41,42]. The psRNATarget program was used to predict target sequences

in pepper *SBP* genes, and the mature miR156 sequences downloaded from miRBase were used as queries [43].

2.7. Tissue Expression, WGCNA Analysis, GO Enrichment, and Co-Expression Network Construction

The expression data of pepper *SBP* genes in root, stem, leaf, flower, and fruit of four periods were used to construct tissue expression profile. The R package WGCNA was used to find the correlations between *SBP* genes and other genes [44]. BLASTP was made against the Uniport dataset, and DAVID was used to make GO enrichment analysis [45]. The co-expression network was constructed using Cytoscape [46].

3. Results

A genome-wide analysis result in the identification of 34 *SBP* genes, named *PnSBP1* to *PnSBP34* (Table 1), in the pepper genome. All identified *SBP* genes have been further confirmed using SMART to make sure complete *SBP* domain structures were obtained. Furthermore, some basic protein parameters including protein length, isoelectric point, and molecular weight were analyzed through the calculating of ExpASY. The protein length was ranging from 177aa to 1101aa, the molecular weight was from 19.80kDa to 120.90kDa, and the isoelectric point was from 6.30 to 10.28 (Table 1).

Table 1. The basic information of pepper *SBP* genes.

Names	Gene	Length (aa)	MV (kDa)	pI
<i>PnSBP8</i>	<i>Pn14.164</i>	337	36.7	8.72
<i>PnSBP13</i>	<i>Pn15.136</i>	354	38.2	8.62
<i>PnSBP4</i>	<i>Pn7.344</i>	365	40.3	8.72
<i>PnSBP28</i>	<i>Pn10.2075</i>	999	110.2	7.29
<i>PnSBP34</i>	<i>Pn5.1721</i>	787	88.1	6.40
<i>PnSBP33</i>	<i>Pn5.1651</i>	787	88.0	6.47
<i>PnSBP32</i>	<i>Pn18.304</i>	774	86.1	6.65
<i>PnSBP16</i>	<i>Pn17.1041</i>	269	30.7	9.61
<i>PnSBP9</i>	<i>Pn5.2134</i>	266	30.1	8.49
<i>PnSBP26</i>	<i>Pn28.549</i>	1038	114.4	6.30
<i>PnSBP27</i>	<i>Pn28.583</i>	1101	120.9	6.80
<i>PnSBP29</i>	<i>Pn12.1036</i>	1091	120.6	7.91
<i>PnSBP15</i>	<i>Pn1.1351</i>	203	23.2	10.28
<i>PnSBP1</i>	<i>Pn1.2504</i>	389	44.3	8.05
<i>PnSBP17</i>	<i>Pn11.1178</i>	368	40.0	8.44
<i>PnSBP18</i>	<i>Pn13.776</i>	396	43.8	9.62
<i>PnSBP22</i>	<i>Pn14.389</i>	344	38.7	7.63
<i>PnSBP12</i>	<i>Pn19.85</i>	289	30.9	8.94
<i>PnSBP5</i>	<i>Pn2.1033</i>	732	80.1	7.82
<i>PnSBP6</i>	<i>Pn2.1339</i>	476	52.3	8.73
<i>PnSBP7</i>	<i>Pn2.939</i>	713	78.9	8.66
<i>PnSBP14</i>	<i>Pn21.1417</i>	370	39.7	7.76
<i>PnSBP2</i>	<i>Pn3.1618</i>	581	64.4	8.54
<i>PnSBP10</i>	<i>Pn3.4074</i>	926	102.0	6.38
<i>PnSBP11</i>	<i>Pn3.426</i>	369	39.7	9.42
<i>PnSBP3</i>	<i>Pn4.2537</i>	323	35.9	8.43
<i>PnSBP23</i>	<i>Pn5.2416</i>	346	38.7	7.97
<i>PnSBP30</i>	<i>Pn6.1128</i>	177	19.8	9.47
<i>PnSBP24</i>	<i>Pn6.473</i>	274	29.8	8.75
<i>PnSBP20</i>	<i>Pn6.475</i>	327	35.5	8.43
<i>PnSBP21</i>	<i>Pn6.937</i>	327	35.6	8.37
<i>PnSBP25</i>	<i>Pn6.939</i>	274	29.8	8.75
<i>PnSBP19</i>	<i>Pn8.2451</i>	322	35.2	9.25
<i>PnSBP31</i>	<i>Pn8.2831</i>	177	19.9	9.18

3.1. Phylogenetic Analysis

To better understand the evolutionary trajectory of pepper *SBP* genes, a phylogenetic analysis of 34 pepper *SBP*s plus 16 *A. thaliana* *SPL*s was implemented (Figure 1). All these *SBP* genes were divided into nine groups according to the phylogenetic analysis, namely, group 1 to group 9. The bootstrap values were most concentrated from 80 to 100, which reflects a reliable phylogenetic relationship. Most groups were able to find both *A. thaliana* and pepper homologous genes, while group 4 was unique in *A. thaliana* and group 5 was specific to pepper. In addition, except group 8, all pepper groups have gene copies more than or equal to *A. thaliana*.

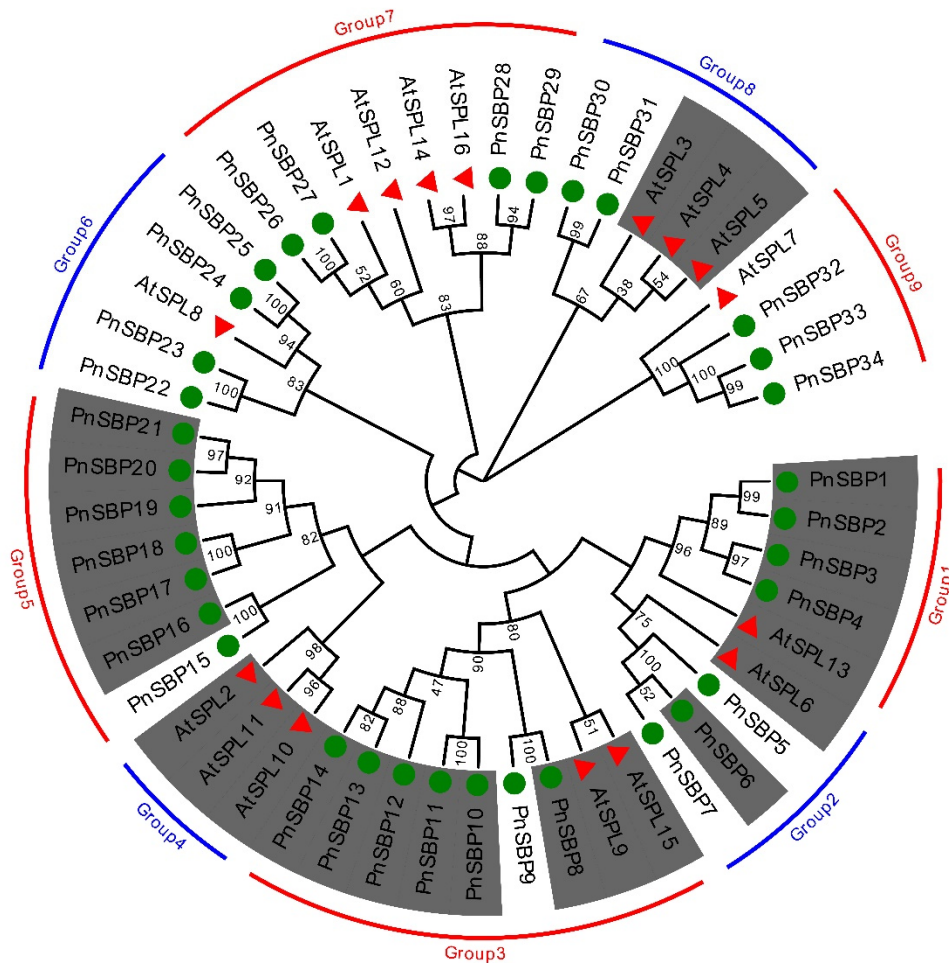


Figure 1. The phylogenetic tree of pepper *SBP* genes. The red triangles refer to *SBP* genes of *A. thaliana*; the green balls refer to pepper *SBP* genes. Each group is highlighted with a color stripe. The *SBP* genes with miR156 target site were marked with dark background color.

3.2. Multiple Sequence Alignment and Conserved Motif Analysis

A multiple sequence alignment was also conducted for the *SBP* domain structures of the pepper and *A. thaliana* *SBP* genes (Figure 2). Two Zn-fingers and one NLS were highly conserved in pepper *SBP* proteins, even so, we still retained five highly suspected sequences (PnSBP26, PnSBP27, PnSBP16, PnSBP3, PnSBP9) with defects in the three typical structures. The typical structure of the two Zn-fingers was Cys-Cys-His-Cys, and the second Zn-finger overlapped with the NLS. In addition, the sequences in the same group were more similar, and they often share group-specific nucleic acids.

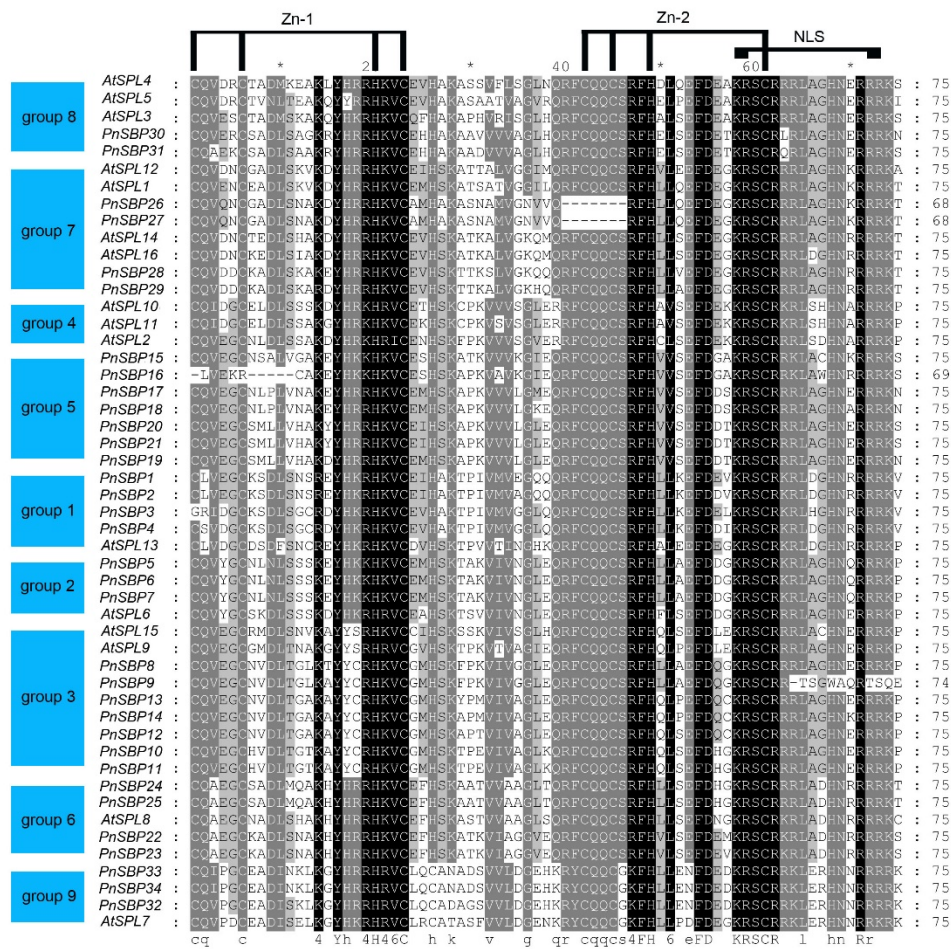


Figure 2. The multiple alignment of SBP-specific domain. Zn-1, Zn-2, and NLS are highlighted on the top.

We further predicted conserved motifs for all SBP proteins (Figure 3). A total of 10 motifs were found. Three motifs constitute the structure of SBP-specific domain (motif 1/3/9). The motif number was consistent with protein length; the relatively longer SBP proteins in group 7/9 were rich in motifs, sharply contrasting with the shorter SBP proteins in group 8. Some motifs were conserved across groups, such as motif 7/10; some motifs were group-specific, such as motif 5/6 were unique to group 7.

3.3. Chromosomal Localization and Gene Duplication Analysis

The chromosomal distribution of 34 SBP genes throughout the pepper genome was plotted. All SBP genes were spread unequally on 19 chromosomes (Figure 4). The gene duplications were recorded in a table (Table S1). After the MCScanX searching and manually screening, 28 pairs of duplicated SBP genes were found in accord with segment duplications. Most segment duplications were found between different chromosomes, while a smaller number of segment duplications were also found within the same chromosome. Compared to segment duplication, the tandem duplication had little effect on SBP gene expansion in pepper. Only two pairs of tandem duplications were discovered (*PnSBP24* and *PnSBP20*; *PnSBP25* and *PnSBP21*).

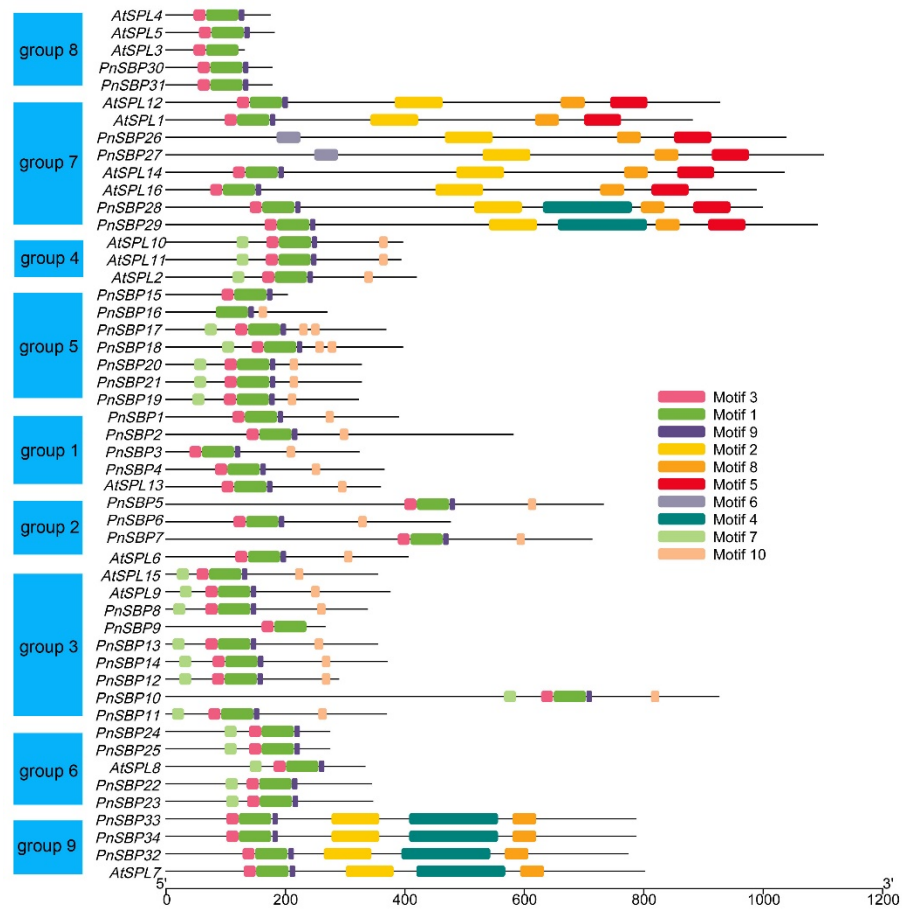


Figure 3. The conserved motifs in pepper *SBP* genes. The genes in the same group were drawn together, and the group number was marked to the left.

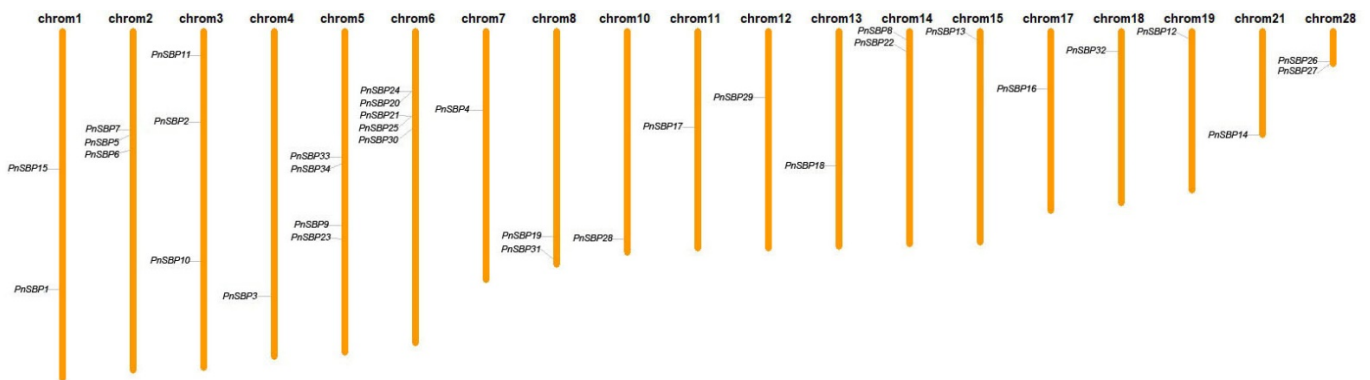


Figure 4. The gene distribution in the pepper genome.

The results of segment duplications were highly consistent with the phylogenetic grouping scheme that each pair of segmentally duplicated genes belong to the same evolutionary group. To further explore the evolutionary constraints on the pepper *SBP* genes, nonsynonymous substitution rate (Ka), synonymous substitution rate (Ks), and their ratio (Ka/Ks) were calculated for the segment duplication gene pairs. According to the Ka/Ks result (Figure 5), all *SBP* genes follow purifying selection ($Ka/Ks < 1$) which reflects they tend to rule out harmful mutations.

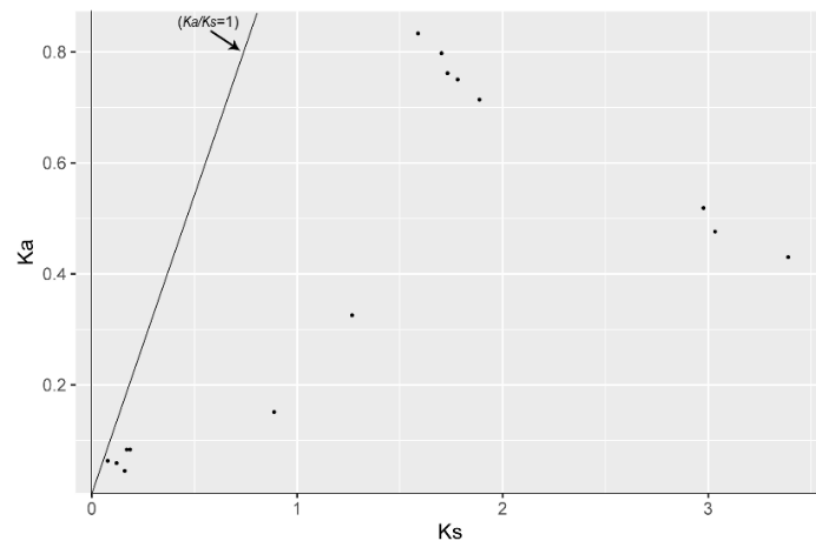


Figure 5. The K_a , K_s value of duplication gene pairs, and the $K_a/K_s = 1$ line was plotted in the graph.

3.4. Prediction of miR156 and Their Target Sites

The regulation of miR156 is a significant characteristic for *SBP* genes, and the miR156 sequences of many species from experimental identification or in silico prediction have been deposited on the miRBase database. Based on the similarity among homologous sequences, we performed a genome-wide search of miR156 precursor sequences in pepper. To further confirm the accuracy of the results, we predicted the RNA secondary structure for all predicted miR156 precursor sequences (Figure S1). Through the above analysis, 17 miR156s (*Pn-miR156a* to *Pn-miR156q*) were identified (Table S2). A phylogenetic analysis was also conducted on pepper miR156s. According to the phylogenetic tree, three groups were obtained from 17 *Pn-miR156s* (Figure 6), namely groups A to C. There are eight *Pn-miR156s* in group A, six in group B, and three in group C. The targets of miR156s were also predicted in pepper *SBP* genes. Seventeen pepper *SBP* genes were found in miR156 binding sequences (Table S3), leading to a 1:1 ratio of with and without miR156 target sites of pepper *SBP* genes. In general, the distribution of target sites in pepper and *A. thaliana* is similar, and the *A. thaliana* group with target site can also find targets in its pepper homologous group. However, the biggest difference was in group 8, where the target locus was present in *A. thaliana* but not in pepper. We do not exclude the possibility that the present genome assembly version may be inaccurate in predicting the target loci of 3' UTR regions (former studies have revealed that the target loci of homologous genes in group 8 are located in 3' UTR, while the target sites of other groups are generally located in CDS [6]). The pepper sequences with miR156 target sites were marked with background color in Figure 1. There are four groups in which miR156 target sites were found. Some sequences in group 2 and group 3 had lost target sites.

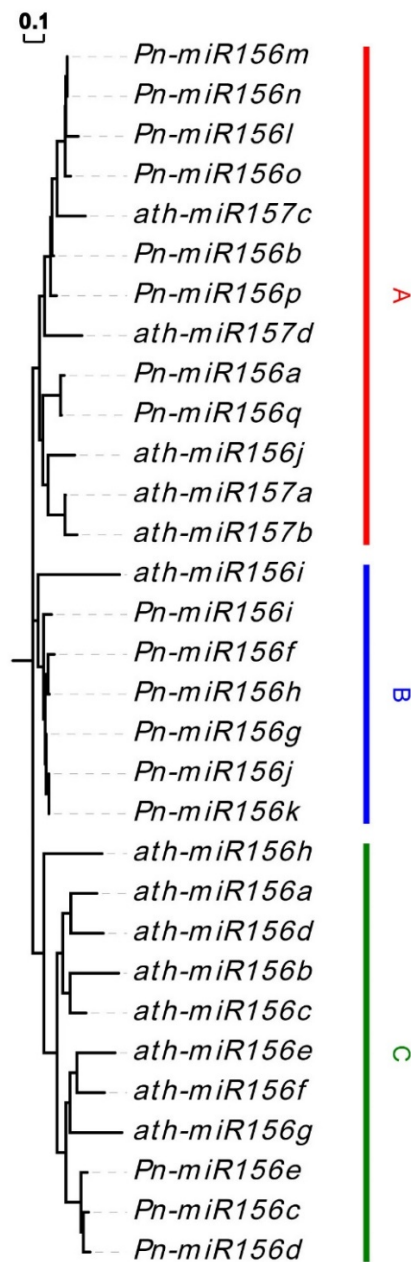


Figure 6. The phylogenetic tree of miR156s, three groups are highlighted in different color stripes. The *A. thaliana* sequences start with ath- and the pepper sequences start with Pn-. Three phylogenetic groups were marked by (A–C).

3.5. Tissue Expression and WGCNA Analysis

We extracted all the edges containing *SBP* genes from the network generated by WGCNA and then screened out the edges with the top 1000 correlation values (Figure 7). The 1000 edges contained 15 *SBP* genes and 684 associated genes. According to the expression values, a cluster analysis was conducted on the 699 genes (15 *SBP* genes plus 684 associated genes). Further, we made a GO enrichment analysis among the 699 genes. The bubble diagram below showed the top 20 enriched GO terms (Figure 8). In general, a lot of transcription factors related terms were obtained, and some terms related to flower development and stress responses were also obtained in BP (Biology Process) terms. To further reveal the functional relationships among the 699 genes, a co-expression network analysis was made. As shown in the graph (Figure 9), four networks (A, B, C, D) with a relatively larger number of associated genes were obtained. In the 4 networks, 11 *SBPs* were

dispersed alone or connected with several other *SBP* genes. Network A had the largest number of associated genes, with up to 6 *SBP* genes connected by multiple intermediate genes (the genes that connect more than one *SBP* gene). Among the 6 *SBP* genes in network A, one gene in group 1, two genes in group 2, two genes in group 3, and one gene in group 5. The other 3 networks were relatively smaller, and networks C and D found no connections with other *SBP* genes.

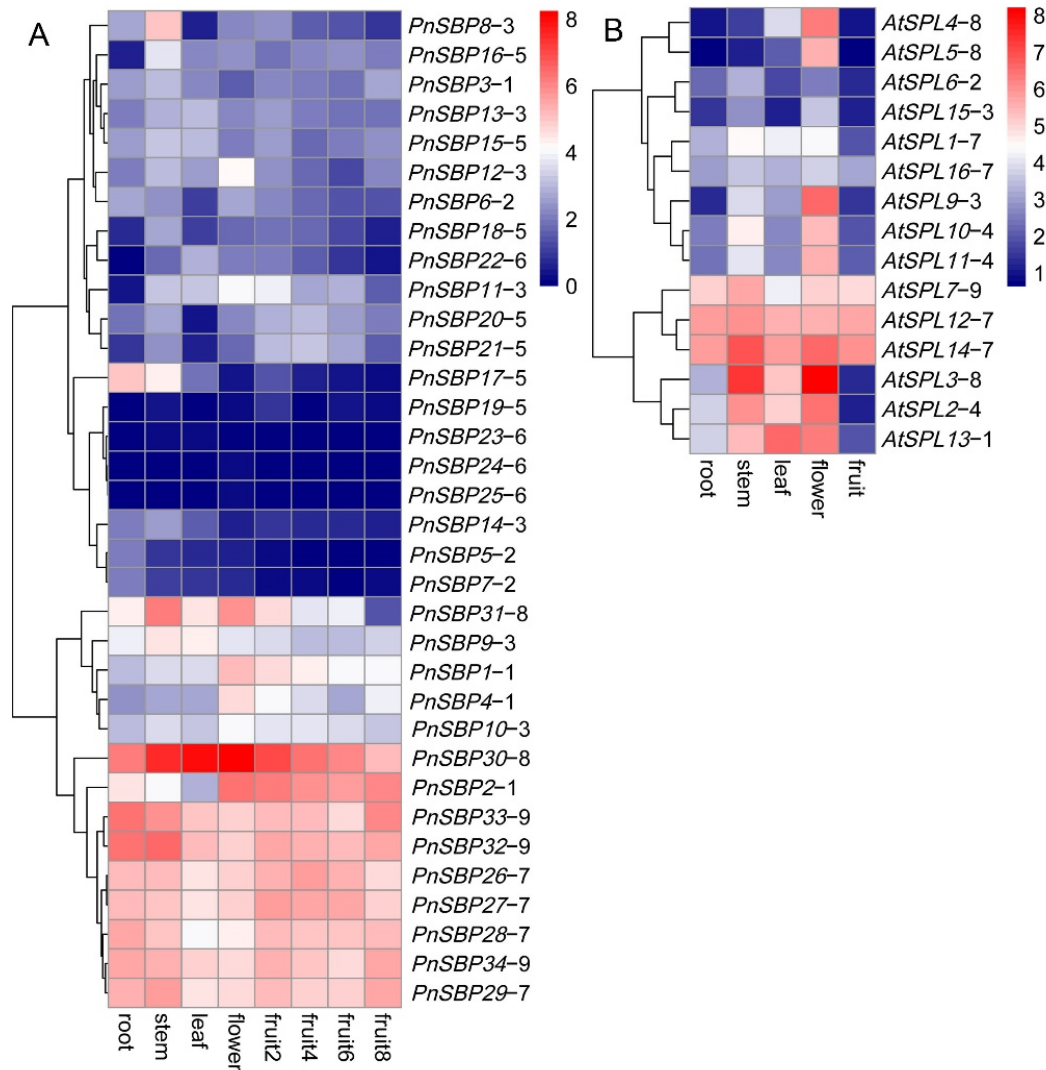


Figure 7. The heatmap of tissue expression of pepper *SBP* genes. The FPKM values have been log₂ transformed. The group number of each gene was marked after the gene name. (A) The expression patterns of pepper *SBP* genes. (B) The expression patterns of *A. thaliana* genes.

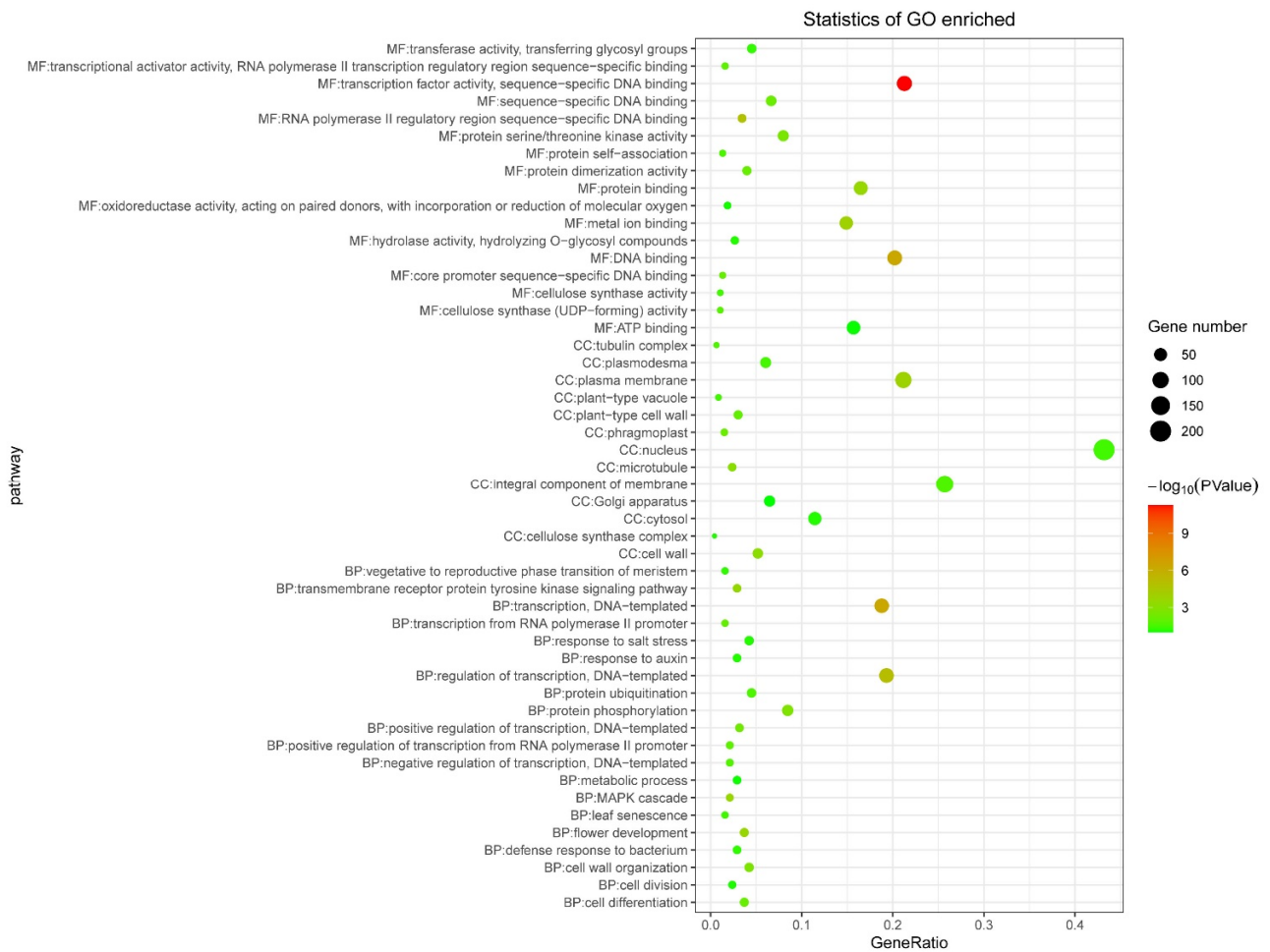


Figure 8. The bubble diagram of the top related genes in pepper *SBP* gene co-expression network. The GeneRatio refers to the ratio of genes enriched to all genes in a same pathway.

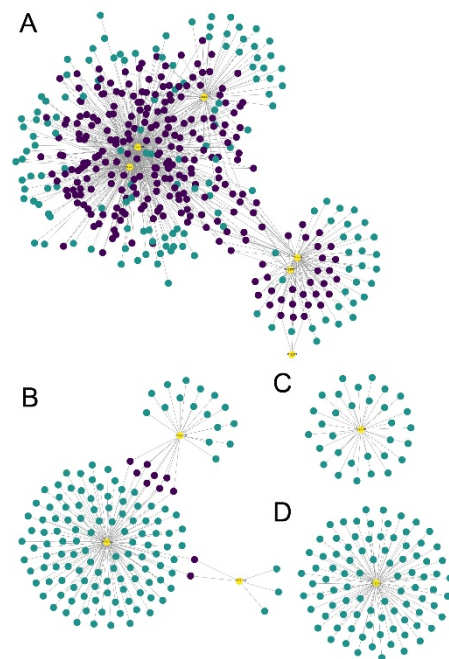


Figure 9. The co-expression networks of pepper *SBP* genes. Four networks were obtained and shown in (A–D).

4. Discussion

Pepper, as a natural spice crop, is in great demand not only in the food field, but also in the cosmetics and health industry [47,48]. However, genic mining for pepper is still limited, and more fundamental research is needed to provide data support in the pursuit of high yield, high disease resistance, high quality, and other excellent agronomic characters. The SBP-box gene family is important in regulating flower development, vegetative to reproductive phase transition, leaf development, plastochron, tillering/branching, panicle/tassel architecture, fruit ripening, fertility, signal transduction, and defense processes [4–7,12–17]. Particularly, *SBP* genes usually form an important regulatory module with miR156, which plays a hub-like role in plant growth and development of the regulatory network [8]. The module of SBP/miR156 integrates many endogenous and environmental cues, such as sugars, higher ambient temperature, some abiotic stresses, gibberellic acid, floral inductive photoperiod, and some hormones such as auxin and light to serve as a regulatory hub to determine the balance of two goals: to complete reproduction, or to grow more branches/leaves [8]. At present, little genic research has been done on pepper, and the knowledge about pepper physiological processes is very limited. Therefore, to improve crop agronomic traits and promote the development of molecular breeding, it is necessary to study the SBP-box gene family in pepper.

There are 34 *SBP* genes identified in pepper, twice as many as *A. thaliana*. Recent whole-genome duplication events are likely responsible for a large number of *SBP* genes in pepper [23]. This study did not find any segment duplications from earlier whole-genome duplication events, which is consistent with previous findings [23]. The pepper SBP-box gene family was amplified in a large scale after the recent duplication event, which resulted in genetic redundancy in some groups. Genes were doubled after whole-genome duplication, and gene redundancy led to new functionalization, sub-functionalization, and loss of function [49,50]. The tissue expression profile exhibited most of the pepper *SBP* genes which are probably functional virtually, which means that the regulation network of *SBP* genes in pepper is relatively complex. Compared with *A. thaliana*, some groups have a significant expansion in pepper; for example, group 1 and 6 both have only one copy in *A. thaliana*, but there are four copies in pepper; group 2 and group 9 both have one copy in *A. thaliana*, but there are three in pepper; group 3 has two copies in *A. thaliana*, but there are seven copies in pepper. More copies make the regulatory network of redundancy copies more complicated. For group 9, the three copies display similar expression profiles and they are all highly expressed with low tissue differences. Groups 1, 2, and 3 have both high expression copies and low ones. For group 6, two of its three members had no expression detected. These results indicate that sub-functionalization, neo-functionalization, or non-functionalization may have occurred among the duplicated copies, leading to functional differentiation among the redundant members. In addition, there may be significant functional differences between pepper and *A. thaliana* *SBP* genes. Group 4, which plays critical roles in *A. thaliana*, has no homologous gene in pepper. Conversely, group 5, which is missing in *A. thaliana*, has as many as seven copies in pepper, five copies of which have detected tissue differential expression patterns. Is the function of missing group 4 in pepper compensated by other genes? Have the genes in group 5 obtained new functions in pepper? Solving these problems is very interesting and necessary.

The regulation of miR156 is also a significant feature for pepper *SBP* genes, and the study of these genes may help to solve some long-standing problems in the pepper industry. Because *SBP* genes were found both with and without miR156 binding sites, it is probable that some genes have lost their target sites. Like other angiosperms [4–6], pepper has several groups with relatively longer protein lengths, no miR156 binding sites, and house-keeping-like expression patterns. These *SBP* genes are highly conserved in angiosperms and probably functionally conserved in land plants. Previous studies have shown that the *SBP* genes regulated by miR156 are often tissue-specific [6], which is in high accordance with our study. *AtSPL7* is a gene confirmed to this form, and has been identified as a regulator of copper homeostasis and responses to light and copper [17].

Three *AtSPL7* homologous genes were found in pepper, and the tissue expression patterns suggest that they may have similar functions to *AtSPL7*. In addition, *SBP* genes may be important regulators in pepper development. The genes in group 8 (*PnSBP30* and *PnSBP31*) are highly expressed in many tissues; their homologous genes in *A. thaliana* are key regulators in flower development and time-dependent pathways [13,14]. Pepper is considered to be a plant with flowering degeneration, and there are few studies on its flowering regulation mechanism. It is hoped that pepper can be harvested with the characteristics of consistent flowering, and flowering regulation genes may be the key to solving these problems. *AtSPL8* (the homologous gene of *PnSBP22*) has been reported to be involved in gynoceium and another development, male fertility, and GA-mediated development processes [51,52]. Thus, it can be seen, pepper *SBP* genes are likely functional like other angiosperms, and may be critical regulators in pepper growth and development.

According to previous studies, the *A. thaliana* *SBP* genes have evolved functional diversity and formed a complex regulatory network [53,54]. The pepper *SBP* genes also displayed obvious tissue differential expression patterns. This reflects that pepper *SBP* genes may play roles in a variety of developmental processes and physiological pathways. The pepper *SBP* genes have been functionally differentiated, 11 *SBP* genes clustered into four co-expression networks and no connection was found among the four networks. Since the co-expression networks of multiple *SBP* genes in Figure 9A are highly intersected and aggregated into a very large co-expression network, we speculate that these *SBP* genes may have functional correlations.

The study of pepper is of great significance to the evolution of angiosperms. Pepper has a special place in the evolutionary history of plants. According to APG IV, Piperaceae has a phylogenetic taxonomic position independent of dicotyledons and monocotyledons and belongs to a branch near the basal angiosperms. In contrast to the highly evolved floral organs of many angiosperms, Piperaceae is thought to be vestigial [23]. Angiosperms probably emerged in the early Cretaceous and expanded explosively, evolving into multiple taxa and diverse phenotypes. Former studies had revealed that the SBP-box gene family is a highly conserved gene family in plants, and its expansion pattern is highly consistent with the evolutionary process of angiosperms and is an important genic driving force for angiosperm evolution [6]. This study found that the SBP-box gene family of pepper is highly differentiated and results in a grouping scheme similar to that of other angiosperms. This result supports that the major branches of the SBP-box gene family have emerged and been preserved in the early angiosperms, similar to the previous study [6]. After the formation of angiosperms, the SBP-box gene family expanded in a lineage-specific manner. The pepper SBP-box gene family has undergone significant expansion, and specific groups, higher copy number, and a more complex regulatory mechanism of intra-group members have emerged compared with *A. thaliana*. These may be important genic drivers to promote the formation of unique physiological characteristics and competitive advantages of pepper. There are still many questions about the pepper SBP-box gene family to be answered, and further exploration is not only beneficial to improve the understanding of pepper, but also a key step to explore the evolution of angiosperms.

5. Conclusions

Our results provide a comprehensive phylogenetic classification of the pepper SBP-box gene family. Thirty-four *SBP* genes were identified in the pepper genome, and the phylogenetic analysis revealed that they can be classified into eight groups. The phylogenetic difference between *A. thaliana* and pepper was relatively significant, in that one *A. thaliana* group was missing in pepper and one group was pepper specific. Compared with *A. thaliana*, significant gene expansion was found in some groups, resulting in a large number of redundant genes in these groups. The distribution of genes regulated by miR156 in pepper was similar to that in *A. thaliana*, but some genes lost regulatory sites. The genes regulated by miR156 displayed significant tissue differential expression patterns, and a large number of redundancies also made the regulatory network of the pepper SBP-box

gene family more complex. The specific gene expansion and the redundancies functional changes provided the genic impetus for the evolution of pepper-specific physiological characteristics and competitive advantages. The present findings provide an improved understanding of the evolution and functions of the SBP-box gene family in pepper, and present a framework for further detailed analysis of this gene family.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/genes12111740/s1>, Figure S1: The prediction of RNA secondary structure of pepper miR156s, Table S1: The duplications of pepper SBP genes, Table S2: The information of predicted miR156 genes in pepper genome, Table S3: The distribution of miR156 binding sites in pepper SBP genes.

Author Contributions: C.H. conceived and supervised this study. C.H. and J.L. designed the experiments. J.L. carried out the experiments, and analyzed and interpreted the data. C.H., R.F., B.W. and X.J. participated in the discussion and provided valuable advice and practical contributions. C.H. and J.L. wrote the manuscript. All authors reviewed, edited, and approved the final manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Central Public-interest Scientific Institution Basal Research Fund for the Chinese Academy of Tropical Agricultural Sciences (NO. 1630142021015, No. 1630142020001, and No. 1630142017016), Natural Science Foundation of Hainan Province of China (2019CXTD417).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data used in this study are presented in the supplementary information.

Acknowledgments: We would like to thank Xiaowei Qin and Lisong Hu from Spice and Beverage Research Institute, Chinese Academy of Tropical Agricultural Science, for valuable comments on the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Van Lijsebettens, M.; Grasser, K.D. Transcript elongation factors: Shaping transcriptomes after transcript initiation. *Trends Plant Sci.* **2014**, *19*, 717–726. [[CrossRef](#)]
2. Yamasaki, K.; Kigawa, T.; Seki, M.; Shinozaki, K.; Yokoyama, S. DNA-binding domains of plant-specific transcription factors: Structure, function, and evolution. *Trends Plant Sci.* **2012**, *18*, 267–276. [[CrossRef](#)] [[PubMed](#)]
3. Klein, J.; Saedler, H.; Huijser, P. A new family of DNA binding proteins includes putative transcriptional regulators of the *Antirrhinum majus* floral meristem identity gene SQUAMOSA. *Mol. Gen. Genet.* **1996**, *250*, 7–16. [[PubMed](#)]
4. Riese, M.; Höhmann, S.; Saedler, H.; Münster, T.; Huijser, P. Comparative analysis of the SBP-box gene families in *P. patens* and seed plants. *Gene* **2007**, *401*, 28–37. [[CrossRef](#)] [[PubMed](#)]
5. Yang, Z.; Wang, X.; Gu, S.; Hu, Z.; Xu, H.; Xu, C. Comparative study of SBP-box gene family in *Arabidopsis* and rice. *Gene* **2008**, *407*, 1–11. [[CrossRef](#)]
6. Li, J.; Gao, X.; Zhang, X.; Liu, C. Dynamic expansion and functional evolutionary profiles of plant conservative gene family SBP-box in twenty two flowering plants and the origin of miR156. *Biomolecules* **2020**, *10*, 757. [[CrossRef](#)]
7. Zhang, H.X.; Jin, J.H.; He, Y.M.; Lu, B.Y.; Li, D.W.; Chai, W.G.; Khan, A.; Gong, Z.H. Genome-wide identification and analysis of the SBP-box family genes under *Phytophthora capsici* stress in pepper (*Capsicum annuum* L.). *Front. Plant Sci.* **2016**, *7*, 504. [[CrossRef](#)] [[PubMed](#)]
8. Wang, H.; Wang, H. The miR156/SPL module, a regulatory hub and versatile toolbox, gears up crops for enhanced agronomic traits. *Mol. Plant* **2015**, *8*, 677–688. [[CrossRef](#)] [[PubMed](#)]
9. Birkenbihl, R.P.; Jach, G.; Saedler, H.; Huijser, P. Functional dissection of the plant-specific SBP-domain: Overlap of the DNA-binding and nuclear localization domains. *J. Mol. Biol.* **2005**, *352*, 585–596. [[CrossRef](#)]
10. Rogers, K.; Chen, X. Biogenesis, turnover, and mode of action of plant microRNAs. *Plant Cell* **2013**, *25*, 2383–2399. [[CrossRef](#)]
11. Voinnet, O. Origin, biogenesis, and activity of plant microRNAs. *Cell* **2009**, *136*, 669–687. [[CrossRef](#)] [[PubMed](#)]
12. Yamaguchi, A.; Wu, M.F.; Yang, L.; Wu, G.; Poethig, R.S.; Wagner, D. The microRNA-regulated SBP-box transcription factor SPL3 is a direct upstream activator of leafy, FRUITFULL, and APETALA1. *Dev. Cell* **2009**, *17*, 268–278. [[CrossRef](#)]
13. Jung, J.H.; Lee, H.J.; Ryu, J.Y.; Park, C.M. SPL3/4/5 integrate developmental aging and photoperiodic signals into the FT-FD module in *Arabidopsis* flowering. *Mol. Plant* **2016**, *9*, 1647–1659. [[CrossRef](#)]
14. Cardon, G.H.; Höhmann, S.; Nettlesheim, K.; Saedler, H.; Huijser, P. Functional analysis of the *Arabidopsis thaliana* SBP-box gene SPL3: A novel gene involved in the floral transition. *Plant J.* **1997**, *12*, 367–377. [[CrossRef](#)] [[PubMed](#)]

15. Shikata, M.; Koyama, T.; Mitsuda, N.; Ohme-Takagi, M. *Arabidopsis* SBP-box genes *SPL10*, *SPL11* and *SPL2* control morphological change in association with shoot maturation in the reproductive phase. *Plant Cell Physiol.* **2009**, *50*, 2133–2145. [[CrossRef](#)]
16. Schwarz, S.; Grande, A.V.; Bujdosó, N.; Saedler, H.; Huijser, P. The microRNA regulated SBP-box genes *SPL9* and *SPL15* control shoot maturation in *Arabidopsis*. *Plant Mol. Biol.* **2008**, *67*, 183–195. [[CrossRef](#)]
17. Zhang, H.; Zhao, X.; Li, J.; Cai, H.; Deng, X.W.; Li, L. MicroRNA408 is critical for the *HY5-SPL7* gene network that mediates the coordinated response to light and copper. *Plant Cell* **2014**, *26*, 4933–4953. [[CrossRef](#)]
18. Salehi, B.; Zakaria, Z.A.; Gyawali, R.; Ibrahim, S.A.; Rajkovic, J.; Shinwari, Z.K.; Khan, T.; Sharifi-Rad, J.; Ozleyen, A.; Turkdonmez, E.; et al. Piper species: A comprehensive review on their phytochemistry, biological activities and applications. *Molecules* **2019**, *24*, 1364. [[CrossRef](#)]
19. Takooree, H.; Aumeeruddy, M.Z.; Rengasamy, K.R.R.; Venugopala, K.N.; Jeewon, R.; Zengin, G.A.-O.; Mahomoodally, M.F. A systematic review on black pepper (*Piper nigrum* L.): From folk uses to pharmacological applications. *Crit. Rev. Food Sci. Nutr.* **2019**, *59*, S210–S243. [[CrossRef](#)] [[PubMed](#)]
20. Butt, M.S.; Pasha, I.; Sultan, M.T.; Randhawa, M.A.; Saeed, F.; Ahmed, W. Black pepper and health claims: A comprehensive treatise. *Crit. Rev. Food Sci. Nutr.* **2013**, *53*, 875–886. [[CrossRef](#)]
21. Dosoky, N.S.; Satyal, P.; Barata, L.M.; da Silva, J.K.R.; Setzer, W.N. Volatiles of black pepper fruits (*Piper nigrum* L.). *Molecules* **2019**, *24*, 4244. [[CrossRef](#)]
22. Group, T.A.P.; Chase, M.W.; Christenhusz, M.J.M.; Fay, M.F.; Byng, J.W.; Judd, W.S.; Soltis, D.E.; Mabberley, D.J.; Sennikov, A.N.; Soltis, P.S.; et al. An update of the angiosperm phylogeny group classification for the orders and families of flowering plants: Apg IV. *Bot. J. Linn. Soc.* **2016**, *181*, 1–20.
23. Hu, L.; Xu, Z.A.-O.; Wang, M.A.-O.; Fan, R.; Yuan, D.A.-O.; Wu, B.; Wu, H.; Qin, X.; Yan, L.; Tan, L.; et al. The chromosome-scale reference genome of black pepper provides insight into piperine biosynthesis. *Nat. Commun.* **2019**, *10*, 4702. [[CrossRef](#)]
24. Lamesch, P.; Berardini, T.Z.; Li, D.; Swarbreck, D.; Wilks, C.; Sasidharan, R.; Muller, R.; Dreher, K.; Alexander, D.L.; Garcia-Hernandez, M.; et al. The *Arabidopsis* information resource (tair): Improved gene annotation and new tools. *Nucleic Acids Res.* **2011**, *40*, D1202–D1210. [[CrossRef](#)]
25. Ghahramani, Z. An introduction to hidden markov models and bayesian networks. In *Hidden markov models*. *World Sci.* **2001**, *45*, 9–41.
26. Camacho, C.; Coulouris, G.; Avagyan, V.; Ma, N.; Papadopoulos, J.; Bealer, K.; Madden, T.L. Blast+: Architecture and applications. *BMC Bioinform.* **2009**, *10*, 421. [[CrossRef](#)] [[PubMed](#)]
27. El-Gebali, S.; Mistry, J.; Bateman, A.; Eddy, S.R.; Luciani, A.; Potter, S.C.; Qureshi, M.; Richardson, L.J.; Salazar, G.A.; Smart, A.; et al. The pfam protein families database in 2019. *Nucleic Acids Res.* **2018**, *47*, D427–D432. [[CrossRef](#)] [[PubMed](#)]
28. Letunic, I.; Bork, P. 20 years of the smart protein domain annotation resource. *Nucleic Acids Res.* **2017**, *46*, D493–D496. [[CrossRef](#)]
29. Gasteiger, E.; Gattiker, A.; Hoogland, C.; Ivanyi, I.; Appel, R.D.; Bairoch, A. ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* **2003**, *31*, 3784–3788. [[CrossRef](#)]
30. Edgar, R.C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **2004**, *32*, 1792–1797. [[CrossRef](#)] [[PubMed](#)]
31. Larkin, M.A.; Blackshields, G.; Brown, N.P.; Chenna, R.; McGettigan, P.A.; McWilliam, H.; Valentin, F.; Wallace, I.M.; Wallace, I.M.; Wilm, A.; et al. Clustal W and clustal X version 2.0. *Bioinformatics* **2007**, *23*, 2947–2948. [[CrossRef](#)]
32. Koichiro, T.; Glen, S.; Sudhir, K. MEGA11: Molecular evolutionary genetics analysis Version 11. *Mol. Biol. Evol.* **2021**, *38*, 3022–3027.
33. Nguyen, L.T.; Schmidt, H.A.; von Haeseler, A.; Minh, B.Q. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **2015**, *32*, 268–274. [[CrossRef](#)] [[PubMed](#)]
34. Subramanian, B.; Gao, S.; Lercher, M.J.; Hu, S.; Chen, W.H. Evolvview v3: A webserver for visualization, annotation, and management of phylogenetic trees. *Nucleic Acids Res.* **2019**, *47*, W270–W275. [[CrossRef](#)] [[PubMed](#)]
35. Bailey, T.L.; Boden, M.; Buske, F.A.; Frith, M.; Grant, C.E.; Clementi, L.; Ren, J.; Li, W.W.; Noble, W.S. MEME suite: Tools for motif discovery and searching. *Nucleic Acids Res.* **2009**, *37*, W202–W208. [[CrossRef](#)]
36. Lee, T.H.; Tang, H.; Wang, X.; Paterson, A.H.; Paterson, A.H. PGDD: A database of gene and genome duplication in plants. *Nucleic Acids Res.* **2013**, *41*, D1152–D1158. [[CrossRef](#)]
37. Wang, Y.; Tang, H.; Debarry, J.D.; Tan, X.; Li, J.; Wang, X.; Lee, T.H.; Jin, H.; Marler, B.; Guo, H.; et al. Mcscanx: A toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* **2012**, *40*, e49. [[CrossRef](#)] [[PubMed](#)]
38. Löytynoja, A. Phylogeny-aware alignment with prank and pagan. *Methods Mol. Biol.* **2021**, *2231*, 17–37.
39. Wang, D.; Zhang, Y.; Zhang, Z.; Zhu, J.; Yu, J. Kaks_calculator 2.0: A toolkit incorporating γ -series methods and sliding window strategies. *Genom. Proteom. Bioinf.* **2010**, *8*, 77–80. [[CrossRef](#)]
40. Griffiths-Jones, S.; Grocock, R.J.; van Dongen, S.; Bateman, A.; Enright, A.J. miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res.* **2006**, *34*, D140–D144. [[CrossRef](#)]
41. Zuker, M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **2003**, *31*, 3406–3415. [[CrossRef](#)]
42. Patanun, O.; Lertpanyasampatha, M.; Sojikul, P.; Viboonjun, U.; Narangajavana, J. Computational identification of microRNAs and their targets in cassava (*Manihot esculenta* Crantz.). *Mol. Biotechnol.* **2013**, *53*, 257–269. [[CrossRef](#)]

43. Dai, X.; Zhuang, Z.; Zhao, P.X. psRNATarget: A plant small RNA target analysis server (2017 release). *Nucleic Acids Res.* **2018**, *46*, W49–W54. [[CrossRef](#)] [[PubMed](#)]
44. Langfelder, P.; Horvath, S. WGCNA: An R package for weighted correlation network analysis. *BMC Bioinform.* **2008**, *9*, 559. [[CrossRef](#)]
45. Huang, W.; Sherman, B.; Lempicki, R.A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **2009**, *4*, 44–57. [[CrossRef](#)] [[PubMed](#)]
46. Doncheva, N.A.-O.; Morris, J.A.-O.; Gorodkin, J.; Jensen, L.A.-O.X. Cytoscape stringapp: Network analysis and visualization of proteomics data. *J. Proteome Res.* **2019**, *18*, 623–632. [[CrossRef](#)]
47. Meghwal, M.; Goswami, T.K. *Piper nigrum* and piperine: An update. *Phytother. Res.* **2013**, *27*, 1121–1130. [[CrossRef](#)] [[PubMed](#)]
48. Smilkov, K.; Ackova, D.G.; Cvetkovski, A.; Ruskovska, T.; Vidovic, B.; Atalay, M. Piperine: Old spice and new nutraceutical? *Curr. Pharm. Des.* **2019**, *25*, 1729–1739. [[CrossRef](#)]
49. Panchy, N.A.-O.; Lehti-Shiu, M.A.-O.; Shiu, S.A.-O.X. Evolution of gene duplication in plants. *Plant Physiol.* **2016**, *171*, 2294–2316. [[CrossRef](#)]
50. Ren, R.; Wang, H.; Guo, C.; Zhang, N.; Zeng, L.; Chen, Y.; Ma, H.; Qi, J. Widespread whole genome duplications contribute to genome complexity and species diversity in angiosperms. *Mol. Plant* **2018**, *11*, 414–428. [[CrossRef](#)]
51. Xing, S.; Quodt, V.; Chandler, J.; Höhmann, S.; Berndtgen, R.; Huijser, P. SPL8 acts together with the brassinosteroid-signaling component bim1 in controlling *Arabidopsis thaliana* male fertility. *Plants* **2013**, *2*, 416–428. [[CrossRef](#)]
52. Zhang, Y.; Schwarz, S.; Saedler, H.; Huijser, P. SPL8, a local regulator in a subset of gibberellin-mediated developmental processes in *Arabidopsis*. *Plant Mol. Biol.* **2007**, *63*, 429–439. [[CrossRef](#)] [[PubMed](#)]
53. Zhang, S.D.; Ling, L.Z.; Yi, T.S. Evolution and divergence of SBP-box genes in land plants. *BMC Genom.* **2015**, *16*, 787. [[CrossRef](#)] [[PubMed](#)]
54. Wang, J.W.; Czech, B.; Weigel, D. miR156-regulated SPL transcription factors define an endogenous flowering pathway in *Arabidopsis thaliana*. *Cell* **2009**, *138*, 738–749. [[CrossRef](#)] [[PubMed](#)]