Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Genome-wide analysis of SPL/miR156 module and their expression analysis in vegetative and reproductive organs of oil palm (Elaeis guineensis)

Lixia Zhou (glzz_2009@163.com)
Chinese Academy of Tropical Agricultural Sciences
Rajesh Yarra
University of Wisconsin-Madison

Research Article

Keywords: EgSPLs, miR156 target sites, qRT-PCR, Inflorescences

Posted Date: April 17th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2795963/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Abstract

Background:

Oil palm (*Elaeis guineensis*, Jacq.) is an important oil-yielding plant, cultivated as source of oil. *SPL* (SQUAMOSA-promoter binding protein-like) gene family is one of the largest plant transcription factors and known to involve in the regulation of plant growth, development and stress responses. The genome-wide analysis of *SPL* gene members in diverse range of crops has been elucidated. However, none of the genome-wide studies on *SPL* gene family has been carried out in oil palm.

Results

In this research, a total of 24 Eg*SPL* genes were identified *via* genome wide approach. Phylogenetic analysis revealed that most of the *EgSPL*s are closely related to the *Arabidopsis* and Rice *SPL* gene members. *EgSPL* genes were mapped onto the only 9 chromosomes of oil palm genome. Motif analysis revealed conservation of SBP domain and the occurrence of 1-10 motifs in *EgSPL*gene members. Gene duplication analysis demonstrated the tandem duplication of *SPL*members in oil palm genome. Heatmap analysis indicated the significant expression of *SPL* genes in shoot and flower organs of oil palm plants. Among the identified *EgSPL* genes, a total 14 *EgSPL*s were shown to be targets of miR156. Real time PCR analysis of 14 SPL genes showed that the most of the *EgSPL* were highly expressed in female and male inflorescences of oil palm plants than vegetative tissues. Altogether, present study revealed the significant role of *EgSPL* genes in inflorescence development.

Conclusions

A total of 24 *EgSPL*s were identified and analyzed 14 *EgSPL*s (containing miR156 sites) genes expression in vegetative and reproductive tissues of oil palm. Moreover, detailed information on *SPL* gene structure, their miR156 target sites, motif composition, chromosomal location and phylogenetic analysis was also reported. Furthermore, the unique expression of *EgSPL*s (containing miR156 sites) in oil palm inflorescences was also revealed via qPCR analysis, predicting their putative role in male and female inflorescence development of oil palm.

Background

The African oil palm (*Elaeis guineensis*) belongs to the palm family (Arecaceae) and majorly cultivated as a source for palm oil. The palm oil is obtained from fruits of oil palm plants. The male and female inflorescences are formed separately in alternating cycle on the same plant [1, 2]. The inflorescences are continuously produced at the axil of each leaf along with the vegetative growth of oil palm plants [1, 2]. The development of inflorescences occur through different phases in a period of 2 to 3 years [3, 4]. The successful development of inflorescence through differential phases is the key step for the formation of oil palm fruit for better oil yield [4]. However, differential gene expression also influence the development

of oil palm inflorescences/flowers [4, 5].The differential expression of genes during the flowering are regulated at post-transcriptional level by various regulatory elements including microRNAs [4].

Plant transcription factors (TFs) play a vital role in regulating the growth and development [6]. Among the plant transcription factors, *SPL* (squamosa promoter binding protein, SBP-Box) gene family is an important plant specific transcription factor family known to regulate the growth and development of plants. The SPL gene was firstly identified in the cDNA library of *Antirrhinum majus* inflorescences as SPL proteins bind to the SQUAMOSA promoter of MADS-Box genes [7, 8]. The *SPL* proteins contain a highly conserved SBP domain with 70 amino acid residues, including two tandem zinc fingers (Cys-Cys-His-Cys and Cys-Cys-Cys-His), and C-terminus region with a nuclear localization signal (NLS) [9, 10]. The NLS overlaps with the zinc finger structure to direct the SBP proteins into nucleus for transcriptional regulation of downstream genes [11]. In recent years, various genome-wide studies identified the occurrence of *SPL* gene members in various plants including 16 in sweet cherry[12], 23 in alfalfa [13], 18 in foxtail millet [14], 37 in trifolium [15], 15 in jatropha [16], 17 in *Arabidopsis* [17, 18], 19 in rice [9], 56 in wheat [19], 24 in tartary buckwheat [20], 15 in Pomegranate [21], 15 in tomato [22, 28 in poplar [23], 27 in apple [24], 18 in grape [25], 17 in orchardgrass genome [26]. However, none of the genome wide studies are carried out for identifying the *SPL* genes in oil palm genome.

MicroRNAs (miRNAs) are endogenous non-coding RNAs and known to suppress the expression of target genes at the post-transcriptional level [27, 28]. Among all of the miRNAs, miR156 is highly conserved in plants and regulate the expression of *SPL* genes for transforming the plants from vegetative to reproductive phase [29, 30, 31]. The SPL genes that regulated by miR156 are categorized in to three groups including (i) SPL2, *SPL9, SPL10, SPL11, SPL13*, and *SPL15* (promoting juvenile-to-adult vegetative transition and vegetative-to-reproductive transition); (ii) *SPL3, SPL4*, and *SPL5* (promoting the floral meristem identify transition); and (iii) *SPL6* (function not yet known) [18]. Previous studies also have shown that various SPL genes have miR156 binding sites [30, 32]. Overexpression of miR156 in *Arabidopsis* down regulate the expression of SPL genes which are having miR156 target sites [33]. Various researchers reported that floral organ development is regulated by the *SPL*2 by activating ASYMMETRIC LEAVES 2 gene of *Arabidopsis thaliana* [34]. The expression of *LEAFY* (*LFY*), *FRUITFULL* (*FUL*) and *APETALA1* (*AP1*) transcription factors in floral meristems are activated by the *SPL3, SPL4* and *SPL5* genes [35, 36].

In this study, we identified 24 SPL genes in oil palm genome through bioinformatics approach with the available oil palm genome sequencing data. This is the first ample report on gene structure, conserved motif analysis, chromosomal distribution, phylogenetic analysis and duplication events of *EgSPL* genes in oil palm genome. Heat map analysis from available transcriptome data of oil palm genome revealed the significant expression of *EgSPL* genes in shoot and floral tissues of oil palm plants. A total of 14 *EgSPL* genes possess the miR156 target sites. In addition, real time PCR analysis of *EgSPL* genes in vegetative and reproductive tissues revealed their significant expression in male and female inflorescences tissues. Altogether, our study provides the involvement of *SPL* genes during the flower development in oil palm plants.

Results

SPL genes identification in *E. guineensis* genome

A total of 24 *SPL* genes were identified in oil palm genome through genome wide approach and they were named as *EgSPL*1-*EgSPL*24. The gene IDs for all the oil palm SPL members were provided in Supplementary Table 2. The sequence information of CDS and protein for all the oil palm SPL members were provided in Supplementary Table 3. The length of the CDS for *SPL* genes ranges from 537bp to 3282bp and the protein sequence length varies from 178 to 1093 amino acid residues (Supplementary Table 2).

EgSPL gene structural features and conserved motif analysis

To gain further insight into the structural features of oil palm *SPL* genes, we used Gene Structural DiSPLay Server tool 2.0 (http:// gsds. cbi. pku. edu. cn/) to analyze the exon/intron organization (Fig.1). Our analysis revealed the presence of varied number of exons (2-12) among the identified 24 SPL genes. Highest numbers of exons (12) are possessed by *EgSPL*21, whereas *EgSPL*1, *EgSPL*9, *EgSPL*19 and *EgSPL*24 contained the lowest number of exons (Fig.1). Moreover, each SPL gene contained at least one intron and varied among the *SPL* genes indicated the functional role of introns in evolution. Further motif analysis revealed the existence of 10 motifs in SPL proteins and varied from 4 to 10 in each SPL protein (Fig.2). Moreover, motifs (1& 2) related to SBP domains were found in all of the 24 oil palm SPL proteins (Fig.2). Presence of other motifs in addition to SBP domain motifs indicated the diverse functions of *SPL* genes.

miR156 target sites prediction in oil palm SPL genes

To identify the miR156 target sites in *EgSPL* genes, we searched the coding region and 3'-UTRs of all identified 24 *SPL* genes using psRNATarget tool. We found that a total of 14 *EgSPL* genes (*EgSPL*2, *EgSPL*4, *EgSPL*5, *EgSPL*7, *EgSPL*8, *EgSPL*10, *EgSPL*11, *EgSPL*12, *EgSPL*13, *EgSPL*15, *EgSPL* 16, *EgSPL*17, *EgSPL*18 & *EgSPL*22) have miR156 binding sites either in coding or in 3' UTR regions (Fig.3). The miR156 binding sites are present in the coding regions of 12 *EgSPL*s (*EgSPL*2, *EgSPL*5, *EgSPL*7, *EgSPL*8, *EgSPL*10, *EgSPL*11, *EgSPL*13, *EgSPL*15, *EgSPL* 16, *EgSPL*17, *EgSPL*18 & *EgSPL*22) members, where as 2 SPLs (*EgSPL*4 & *EgSPL12*) contained miR156 sites in their 3' UTR regions (Fig.3). Our results indicating the regulation of EgSPL genes by miR156 is confined to few genes among the identified 24 *EgSPL* genes.

Chromosomal distribution and SPL gene duplication in oil palm genome

The mapping of all identified 24 *EgSPL* gene across 16 chromosomes of oil palm genome was also studied. A total of 20 *EgSPL* members were mapped on the chromosomes (1, 2, 3, 4, 7, 8, 10, 11 & 14) and the remaining 4 *SPL* genes were not mapped any of the chromosomes. We did not find any SPL genes distributed on chromosome 5, 6, 9, 12, 13, 15 and 16. Chromosome 8 and chromosome 2 contained highest number of SPL genes (Fig.4). Chromosomes 1, 3,7,10 and 14 had only one SPL gene. These

results suggesting the uneven distribution of *EgSPL* genes on chromosomes of oil palm genome (Fig.4). Further, we used Circos algorithm to know the information on expansion of SPL gene duplications in oil palm genome. We found that a total of 16 pairs of SPL genes were duplicated among the 16 chromosomes of oil palm. The duplicated pairs were located on the different chromosomes of genome. Chromosomes 2 and 8 have the more number of duplicated pairs compared to other chromosomes (Fig.5). Most of the chromosomes did not contain any duplicated pairs of SPLs and duplication is existed only on the 7 number of chromosomes. Our findings indicating the expansion or duplication of SPLs exists through the only from the duplicated blocks of genome (Fig.5).

Evolutionary relationship of oil palm, rice and Arabidopsis *SPL* genes

We constructed the Maximum Likelihood tree for analyzing the evolutionary relation between oil palm SPL genes with the *SPL* genes of rice and Arabidopsis. Interestingly, oil palm *SPL* genes are distantly related with both rice Arabidopsis *SPL* members (Fig.6). However, some of them closely present in the clades where rice *SPL* members were grouped. Our results demonstrated that EgSPLs were separately evolved during evolution (Fig.6).

Expression profiles of *EgSPL*s in vegetative and reproductive tissues of oil palm

We examined the tissue specific expression of 24 EgSPLs in vegetative (leaf, root and shoot) and reproductive tissues (flower, fruit and mesocarp (15, 17, 21 and 23 weeks old)) of oil palm plants using the available transcriptome data (SRR851096, SRR851071, SRR851067, SRR851108, SRR851103, SRR190698, SRR190699, SRR190701 & SRR190702). A total of two genes (*EgSPL*1, *EgSPL*3) in flower, a total of four genes (*EgSPL*2, *EgSPL*13, *EgSPL*16, *EgSPL*22) in shoot and a total of two genes (*EgSPL*5, *EgSPL*6) in fruits are highly expressed compared to other *EgSPL* genes (Fig.7). However, majority of the *EgSPL*s were expressed in flower and shoot tissues. All the identified 24 SPL genes were downregulated in root tissues (Fig.7).

Realtime expression analysis of *EgSPL*s containing miR156 binding sites

Quantitative Realtime PCR experiment was performed to know the relative expression of 14 oil palm *SPL* genes (containing miR156 sites) in various vegetative and reproductive tissues (male and female inflorescences) of oil palm plants. Our results demonstrated that expression of *EgSPL*02 and *EgSPL*12 are specifically confined to male inflorescences, whereas *EgSPL*07, *EgSPL*08, *EgSPL*18, *EgSPL*22 are only confined to female inflorescences (Fig.8). A total of three genes (*EgSPL*05, *EgSPL*10, *EgSPL*16) were only highly expressed in both the male and female inflorescences (Fig.8). A total of 2 genes *EgSPL*13 & *EgSPL*17 were higly expressed in all the vegetative and reproductive tissues. Altogether, our results indicated the role of *SPL* genes in male and female inflorescences development in oil palm plants.

Discussion

SPLs are plant-specific transcription factors with highly conserved SBP domain and involved in regulating growth and development of plants. Various number of SPL genes were identified and characterized their expression in various crops including sweet cherry [12], alfalfa [13], Quinoa [37], Foxtail millet [14], wheat [38], pecan [39] , soybean[40], populas [41], Jatropha[16]. However, some of the studies were carried out to characterize SPL gene expression during reproductive tissue development in various crops including Pecan [39], Tartary Buckwheat [20], Rice [42], flowering cherry [43], Trifolium [15] and Petunia [44]. However, till date none of the information is available on identification and characterization of SPL genes in oil palm during inflorescences development. In this study, we identified a total of 24 SPL genes in oil palm genome through bioinformatics approach and dissected the expression of 14 SPLs in male and female inflorescences of oil palm plants through experimental approach. The number of identified SPLs (24) in this study are fewer than the number of SPLs identified in sweet cherry (16SPLs) [12], alfalfa (23*SPL*s) [13], foxtail millet (18 SPLs) [14], jatropha (15 *SPL*s) [16] and lesser than the number of SPLs identified in soybean(41 *SPL*s) [40], flowering cherry (32 *SPL*s) [43], wheat(56 *SPL*s) [38], populus (33 *SPL*s) [41], trifolium(37 *SPL*s) [15] genomes. Present information on oil palm *SPL*s would elucidate the evolutionary process of SPLs across various plants.

The structural organization across the identified oil palm *SPL* genes showed the variation in number of introns indicated the role of introns in *SPL* genes evolution. All the identified *SPL*s contained at least one intron and varied among them. Our results are consistent with the previous reports on identification of *SPL* genes in Jatropha [16], Soybean [40], Alfalfa [13] and sweet cherry [12]. The presence of SBP domain is the key feature in *SPL* gene family [9]. The motifs related to the SBP domain are found in all the identified *EgSPL* proteins. Conserved motif analysis revealed the occurrence of 10 motifs in oil palm SPL genes as reported in other plants including foxtail millet [14], Alfalfa [13] and Quinoa [37]. The post transcriptional regulation of SPLs by miR156 determines fine-tune functions of *SPL*s [45]. In our study, a total of 14 oil palm *SPL*s contained the miR156 sites mostly in the coding regions and lesser in UTR regions. Our results are consistent with the previous reports from Jatropha, Soybean, Medicago and mustard SPL genes [16, 40, 46, 47], indicating the conservation of *miR156*-mediated posttranscriptional regulation in plants. Phylogenetic analysis of 24 oil palm SPLs with *Arabidopsis* and rice SPLs suggested the diversification of oil palm SPL members among model-to-crop plants during evolution process. The uneven distribution of 24 SPL genes on 16 chromosomes of oil palm also coinciding with the previous reports on *SPL* gene family distribution on sweet cherry, Alfalfa and foxtail millet genomes [12,13,14].

SPL genes encode plant-specific transcription factors that play important roles in flower development, including vegetative to reproductive growth [48] . In our study, we analyzed the tissue specific expression of *EgSPL*s from the available transcriptome data and the majority of *EgSPL* genes expressed in flower and shoot. Our results are consistent with the previous reports on expression on *SPL* genes in blue berry [49]. Our qPCR data on expression of 14 *SPL* genes that contain miR156 sites, revealed that confined expression of some *SPL* genes in male and female inflorescences. The expression of miR156 targeted *EgSPL*02 and *EgSPL*12 were unique to male inflorescences, miR156 targeted *EgSPL*07, *EgSPL*08, *EgSPL*18, E*gSPL*22 were unique to female inflorescences and miR156 targeted *EgSPL*05, *EgSPL*10, *EgSPL*16 were only highly expressed in both the male and female inflorescences, suggesting the

involvement of *EgSPL* genes in inflorescences development. Our results are consistent with the previously reported the involvement of miR156 targeted *SPL* genes in inflorescences [40, 44].

Materials And Methods

Identification of SPL gene family in oil palm genome

The known SPL protein sequences of Arabidopsis (16) and rice (19) were used as queries to retrieve the putative oil palm SPL family members against oil palm (*E. guineensis*) genome database(http://palmx plore. mpob. gov. my/ palmX plore/). Then, all the retrieved putative oil palm SPL family proteins were queried against the CDD (https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi) and Pfam databases to confirm the occurrence of conserved SBP domain. The length of each putative SPL gene coding sequence was also determined via Blastn search against *E. guineensis* genome database.

Oil palm SPL gene structure, conserved motif analysis and miR156 target site prediction

The intron/exon structure analysis of 24 number of oil palm SPL gene structure was analyzed using Gene Structure DiSPLay Server (http:// gsds. cbi. pku. edu. cn/). Further, MEME tool was used to analyze the conserved motif analysis of all identified oil palm SPL protein members (http:// meme- suite. org/tools/ meme). The psRNATarget tool (http://plantgrn.noble.org/psRNATarget/?function=3) was used to predict the miR156 target site in oil palm full length SPL gene sequences

Phylogenetic analysis, Duplication and Chromosomal distribution of oil palm SPL genes

We generated the phylogenetic tree of oil palm *SPL* genes with the other known SPL genes of *Arabidopsis* and Rice with the help of MEGA 7.0 tool by Maximum Likelihood method, with a bootstrap value of 1000 replications [50]. We also investigated the SPL gene member's duplication events across the oil palm genome via MCScanX tool with default parameters [51]. Further, we also mapped the distribution of 24 oil palm SPL genes on 16 chromosomes of oil palm from the oil palm genome data base using TB tools software [52].

Plant material and RNA isolation

The healthy African oil palm (*Elaeis guineensis*) plants were grown in the field station of Coconut Research Institute, Chinese Academy of Tropical Agricultural Sciences, Wenchang, China and are all grown under institutional regulatory issues. The plant material used in this research work is collected by the corresponding author. To investigate the SPL gene expression analysis in different developmental stages, the samples including root, stem, leaf, male inflorescence and female inflorescence were collected separately from the 6 year old oil palm plants and quickly frozen in liquid nitrogen and stored at -80°C for further experiments. Total RNA from root, stem, leaf, male inflorescence and female inflorescence was isolated by following the method as described previosuly [53]. The yield, integrity and purity of extracted

RNA samples were quantified on Nanodrop and also electrophoresed on 1% agarose gel. The genomic DNA contamination was removed by treating the RNA samples with DNase I.

In silico expression analysis of oil palm *SPL* genes

To examine the *SPL* gene expression in various tissues of oil palm, the normalized RPKM values were retrieved from the available transcriptome data of six different tissues including leaf, root, flower, shoot and mesocarp (15, 17, 21, and 23 weeks). The heatmap was generated to analyze the SPL gene expression in various tissues of oil palm using pheatmap tool (https://cran.r-project.org/web/packages/pheatmap/index.html).

qRT-PCR analysis of *SPL* genes in vegetative and reproductive tissues

To investigate the real time expression of SPL genes in various vegetative and reproductive tissues, the real time PCR expression analysis was performed using Mastercycler ep realplex4

Machine. The cDNA synthesis was carried out with the extracted RNA by using MightyScript first-strand cDNA synthesis kit by following the manufacturer's instructions. The 2 × SYBR Green qPCR ProMix was used to carryout for qRT-PCR reactions with Mastercycler. The oil palm *SPL* gene-specific primers (Supplementary Table 1) were designed by using QuantPrime qPCR primer designing tool (https://quantprime.mpimp-golm.mpg.de/). Three biological and technical repeats were performed to know the reliability of gene expression studies. The oil palm *Actin1* gene was used as a house keeping gene to check the relative expression of SPLs by employing the $2^{-\Delta\Delta Ct}$ method. The statistical significance was determined at $p < 0.05$ and $p < 0.01$ using ANOVA.

Conclusions

To the best of our knowledge, this is the first report on genome-wide analysis of *SPL* genes in oil palm. In current investigation, a total of 24 *EgSPL*s were identified and analyzed 14 *EgSPL*s (containing miR156 sites) genes expression in vegetative and reproductive tissues of oil palm. Moreover, detailed information on *SPL* gene structure, their miR156 target sites, motif composition, chromosomal location and phylogenetic analysis was also reported. Furthermore, the unique expression of *EgSPL*s (containing miR156 sites) in oil palm inflorescences was also revealed via qPCR analysis, predicting their putative role in male and female inflorescence development of oil palm.

Declarations

Acknowledgements

We thank Dr. Rui Li for his help throughout this study.

Authors Contributions

LZ assessed the conceptualization of research. LZ conceived and designed the methodology of this study. RY performed the bioinformatics analysis. LZ analyzed the data. LZ and RY prepared and wrote original manuscript draft. LZ supervised the research. All authors have read and approved the final manuscript.

Funding:

This research work is financially supported by Hainan Provincial Natural Science Foundation of China (323MS073)

Availability of data and materials

The transcriptome data used in this study are available in the NCBI Sequence Read Archive (SRA) with accession numbers(SRR851096, SRR851071, SRR851067, SRR851108, SRR851103, SRR190698, SRR190699, SRR190701 & SRR190702). All data sets analyzed during this research are provided as Supplementary files and also included in the article. The data sets generated during this study are also available from the corresponding author on reasonable request.

Ethics approval and consent to participate.

All methods in this research were carried out in accordance with relevant guidelines and regulations of Coconut Research Institute, Wenchang, China.

Consent for publication

Not applicable

Competing interests

The authors declare that they have no competing interest.

References

1. Adam H, Collin M, Richaud F, Beulé T, Cros D, Omoré A, Nodichao L, Nouy B, Tregear JW. Environmental regulation of sex determination in oil palm: current knowledge and insights from other species. Ann Bot. 2011;108(8):1529-37.
2. Adam H, Jouannic S, Escoute J, Duval Y, Verdeil JL, Tregear JW. Reproductive developmental complexity in the African oil palm (*Elaeis guineensis*, Arecaceae). Am J Bot. 2005;92(11):1836-52.
3. Corley RHV, Gray BS. Growth and morphology. In: Corley RHV, Hardon JJ, Wood BJ (eds) Developments in crop. Elsevier, Amsterdam, 1976.
4. Ho H, Gudimella R, Ong-Abdullah M. Expression of microRNAs during female inflorescence development in African oil palm (*Elaeis uineensis* Jacq.). Tree Genetics & Genomes. 2017;13:35.

5. Beulé T, Camps C, Debiesse S, Tranchant C, Dussert S, Sabau X, Jaligot E, Syed Alwee SSR, Tregear JW. Transcriptome analysis reveals differentially expressed genes associated with the mantled homeotic flowering abnormality in oil palm (*Elaeis guineensis*). Tree Genet Genomes. 2011;7:169–182.
6. Gonzalez DH. Introduction to transcription factor structure and function. In Plant Transcription Factors; Elsevier: Amsterdam, The Netherlands. 2016;3–11.
7. Huijser P, Klein J, Lönnig WE, Meijer H, Saedler H, Sommer H. Bracteomania, an inflorescence anomaly, is caused by the loss of function of the MADS-box gene squamosa in Antirrhinum majus. EMBO J. 1992;11(4):1239–49.
8. Klein J, Saedler HP. A new family of DNA binding proteins includes putative transcriptional regulators of the ANTIRRHINUM MAJUS floral meristem identity gene SQUAMOSA. Mol Gen Genet MGG. 1996;250(1):7–16.
9. 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(1–2):1–11.
10. Yamasaki K, Kigawa T, Inoue M, Tateno M, Yamasaki T, Yabuki T, Aoki M, Seki E, Matsuda T, Nunokawa E. A novel zinc-binding motif revealed by solution structures of DNA-binding domains of Arabidopsis SBP-family transcription factors. J Mol Biol. 2004;337(1):49–63.
11. Birkenbihl RP, 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(3):585–96.
12. Sun Y, Wang Y, Xiao Y, Zhang X, Du B, Turupu M, Wang C, Yao Q, Gai S, Huang J, Tong S, Li T. Genome-Wide Identification of the SQUAMOSA Promoter-Binding Protein-like (SPL) Transcription Factor Family in Sweet Cherry Fruit. Int J Mol Sci. 2023;24(3):2880.
13. Wang Y, Ruan Q, Zhu X, Wang B, Wei B, Wei X. Identification of Alfalfa SPL gene family and expression analysis under biotic and abiotic stresses. Sci Rep. 2023;13(1):84
14. Lai D, Fan Y, Xue G, He A, Yang H, He C, Li Y, Ruan J, Yan J, Cheng J. Genome-wide identification and characterization of the SPL gene family and its expression in the various developmental stages and stress conditions in foxtail millet (*Setaria italica*). BMC Genomics. 2022;23(1):389.
15. Ma J, Nie G, Yang Z, Ma S, Fan J, Hu R, Wu F, Zhang X. Genome-Wide Identification, Characterization, and Expression Profiling Analysis of SPL Gene Family during the Inflorescence Development in Trifolium repens. Genes (Basel). 2022;13(5):900.
16. Yu N, Yang JC, Yin GT, Li RS, Zou WT. Genome-wide characterization of the SPL gene family involved in the age development of Jatropha curcas. BMC Genomics. 2020;21:368.
17. Cardon G, Hohmann S, Klein J, Nettesheim K, Saedler H, Huijser P. Molecular characterisation of the Arabidopsis SBP-box genes. Gene. 1999;237(1):91–104.
18. Xu M, Hu T, Zhao J, Park MY, Earley KW, Wu G. Developmental functions of miR156-regulated SQUAMOSA PROMOTER BINDING PROTEIN-LIKE (SPL) genes in *Arabidopsis thaliana*. PLoS Genet. 2016;12:e1006263.

19. Zhu T, Liu Y, Ma L, Wang XY, Zhang DZ, Han YC, Ding Q, Ma LJ. Genome-wide identification, phylogeny and expression analysis of the SPL gene family in wheat. BMC Plant Biol. 2020;20:420.
20. Liu M, Sun W, Ma Z, Huang L, Wu Q, Tang ZZ, Bu TL, Li CL, Chen H. Genome-wide identification of the SPL gene family in Tartary Buckwheat (*Fagopyrum tataricum*) and expression analysis during fruit development stages. BMC Plant Biol. 2019;19:299.
21. Li B, Zhao Y, Wang S, Zhang X, Wang Y, Shen Y, Yuan Z. Genome-wide identification, gene cloning, subcellular location and expression analysis of SPL gene family in *P. granatum* L. BMC Plant Biol. 2021;21(1):400.
22. Salinas M, Xing S, Höhmann S, Berndtgen R, Huijser P. Genomic organization, phylogenetic comparison and differential expression of the SBP-box family of transcription factors in tomato. Planta. 2012;235(6):1171–84.
23. Li C, Lu S. Molecular characterization of the SPL gene family in Populus trichocarpa. BMC Plant Biol. 2014;14(1):131.
24. Li J, Hou H, Li X, Xiang J, Yin X, Gao H, Zheng Y, Bassett CL, Wang XP. Genome-wide identification and analysis of the SBP-box family genes in apple (Malus×domestica Borkh.). Plant Physiol Biochem. 2013;70:100–14.
25. Hou H, Li J, Gao M, Singer S, Wang H, Mao L. Genomic organization, phylogenetic comparison and differential expression of the SBP-box family genes in grape. PLoS One. 2013;8(3):e59358.
26. Feng G, Han J, Yang Z, Liu Q, Shuai Y, Xu X, Nie G, Huang L, Liu W, Zhang X. Genome-wide identification, phylogenetic analysis, and expression analysis of the SPL gene family in orchardgrass (*Dactylis glomerata* L.). Genomics. 2021;113(4):2413-2425.
27. Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell. 2004;116(2):281–97.
28. Breakfield NW, Corcoran DL, Petricka JJ, Shen J, Sae-Seaw J, Rubio-Somoza I. High-resolution experimental and computational profiling of tissue specific known and novel miRNAs in *Arabidopsis*. Genome Res. 2012;22(1):163–76.
29. Wang Y, Liu W, Wang X, Yang R, Wu Z, Wang H. MiR156 regulates anthocyanin biosynthesis through SPL targets and other microRNAs in poplar. Horticult Res. 2020;7(1):6786–7.
30. Ye B, Zhang K, Wang J. The role of miR156 in rejuvenation in *Arabidopsis thaliana*. J Integr Plant Biol. 2020; 62(05):550–5.
31. Teotia S, Tang G. To bloom or not to bloom: role of microRNAs in plant flowering. Mol Plant. 2015;8(3):359–77.
32. Munusamy P, Zolotarov Y, Meteignier LV, Moffett P, Strömvik MV. novo computational identification of stress-related sequencemotifs and microRNA target sites in untranslated regions of a plant translatome. Sci Rep. 2017;7(1):43861.
33. Schwab R, Palatnik JF, Riester M, Schommer C, Schmid M, Weigel D. Specific effects of microRNAs on the plant transcriptome. Dev Cell. 2005;8(4):517–27.

34. Wang ZS, Wang Y, Kohalmi SE, Amyot L, Hannoufa A. SQUAMOSA promoter binding PROTEIN-LIKE 2 controls floral organ development and plant fertility by activating ASYMMETRIC LEAVES 2 in *Arabidopsis thaliana*. Plant Mol Biol. 2016; 9(7):1–14.
35. Kim JJ, Lee JH, Kim W, Jung HS, Huijser P, Ahn JH. The miR156-SPL3 module regulates ambient temperature-responsive flowering via FT in Arabidopsis thaliana. Plant Physiol. 2012;159(1):461–78.
36. Jung JH, Seo PJ, Kang SK, Park CM. miR172 signals are incorporated into the miR156 signaling pathway at the SPL3/4/5 genes in Arabidopsis developmental transitions. Plant Mol Biol. 2011;76(1/2):35–45.
37. Ren Y, Ma R, Fan Y, Zhao B, Cheng P, Fan Y, Wang B. Genome-wide identification and expression analysis of the SPL transcription factor family and its response to abiotic stress in Quinoa (*Chenopodium quinoa*). BMC Genomics. 2022;23(1):773.
38. Guo F, Lu Q, Cang J. Genome-wide identification and expression profiling of the SPL family genes in wheat. Botany. 2021;99:185–198.
39. Wang M, Mo Z, Lin R, Zhu C. Characterization and expression analysis of the SPL gene family during floral development and abiotic stress in pecan (*Carya illinoinensis*). Peer J. 2021;9:e12490.
40. Tripathi RK, Goel R, Kumari S, Dahuja A. Genomic organization, phylogenetic comparison, and expression profiles of the SPL family genes and their regulation in soybean. Dev. Genes Evol. 2017;227:101–119.
41. Qin S, Bao LH, He ZG, Li CL, La HG, Zhao LF. Identification and regulatory network analysis of SPL family transcription factors in Populus Euphratica Oliv. Heteromorphic Leaves. Sci. Rep. 2022;12(1):2856
42. Giaume F, Fornara F. SPL transcription factors prevent inflorescence reversion in rice. Mol Plant. 2021;14(7):1041-1043.
43. Gao L, Lyu T, Lyu Y. Genome-Wide Analysis of the SPLGene Family and Expression Analysisduring Flowering Induction in Prunus yedoensis 'Somei-yoshino'. Int. J.Mol. Sci. 2022;23:10052.
44. Zhou Q, Zhang S, Chen F, Liu B, Wu L, Li F, Zhang J, Bao M, Liu G. Genome-wide identification and characterization of the SBP-box gene family in Petunia. BMC Genomics. 2018;19(1):193.
45. Rhoades MW, Reinhart BJ, Lim LP, Burge CB, Bartel B, Bartel DP. Prediction of plant microRNA targets. Cell. 2002;110:513–520.
46. Wang H, Lu Z, Xu Y, Kong LC, Shi JJ, Liu YF, Fu CX, Wang XS, Wang ZY, Zhou CE, Han L. Genome-wide characterization of SPL family in Medicago truncatula reveals the novel roles of miR156/SPL module in spiky pod development. BMC Genomics. 2019;20:552
47. Gao J, Peng H, Chen F, Liu Y, Chen B, Li W. Genome-wide identification and characterization, phylogenetic comparison and expression profiles of SPL transcription factor family in B. juncea (Cruciferae). PLoS ONE. 2019;14(11):e0224704.
48. Chen X, Zhang Z, Liu D, Zhang K, Li A, Mao L. SQUAMOSA promoter-binding protein-like transcription factors: star players for plant growth and development. J Integr Plant Biol. 2010;52(11):946-51.

49. Feng X, Zhou BJ, Wu XL, Wu HL, Zhang SL, Jiang Y, Wang YP, Zhang YQ, Cao M, Guo BS, Su SC, Hou ZX. Molecular characterization of SPL gene family during flower morphogenesis and regulation in blueberry. BMC Plant Biol. 2023;23:40.
50. Kumar S, Stecher G, Tamura K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol. 2016;33:1870–4.
51. Wang Y, Tang H, DeBarry JD, Tan X, Li J, Wang X, Lee T, Jin H, Marler B, Guo H, Kissinger JC, Paterson AH. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucl Acids Res. 2012;40(7):e49.
52. Chen C, Xia R, Chen H, He Y. TBtools, a toolkit for biologists integrating various HTS-data handling tools with a user-friendly interface. BioRxiv. 2018;[Preprint] https:// doi. org/ 10. 1101/ 289660.
53. Badai SS, Rasid OA, Parveez GKA, Masani MYA. A rapid RNA extraction method from oil palm tissues suitable for reverse transcription quantitative real-time PCR (RT-qPCR). 3 Biotech. 2020;10(12):530.

Figures

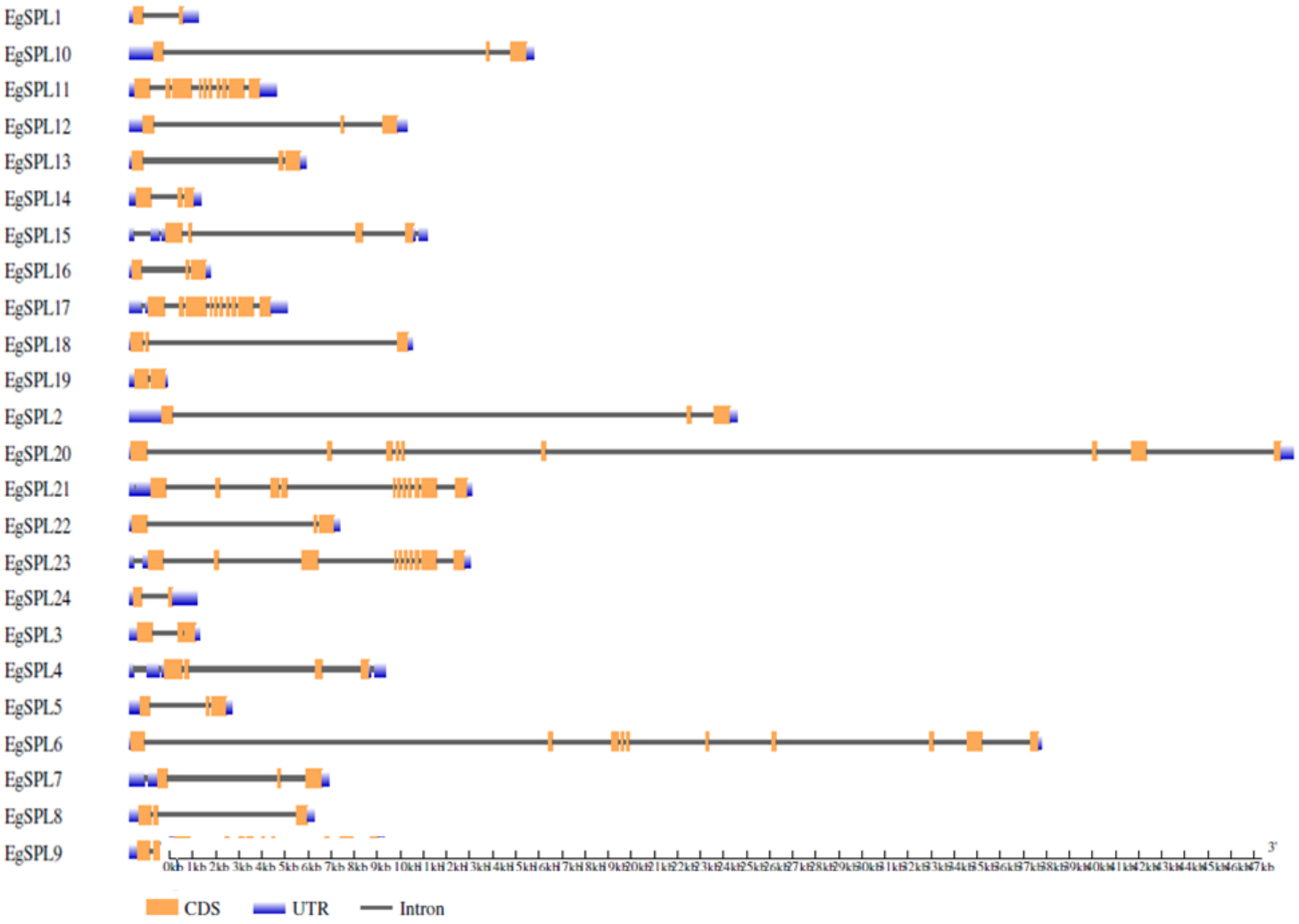


Figure 1

Intron-exon organization of 24 *EgSPL* genes. Coding sequences (CDS), untranslated regions (3' & 5'UTRs), intron r*egi*ons are represented by orange, blue and black color blocks respectively in the schematic presentation.

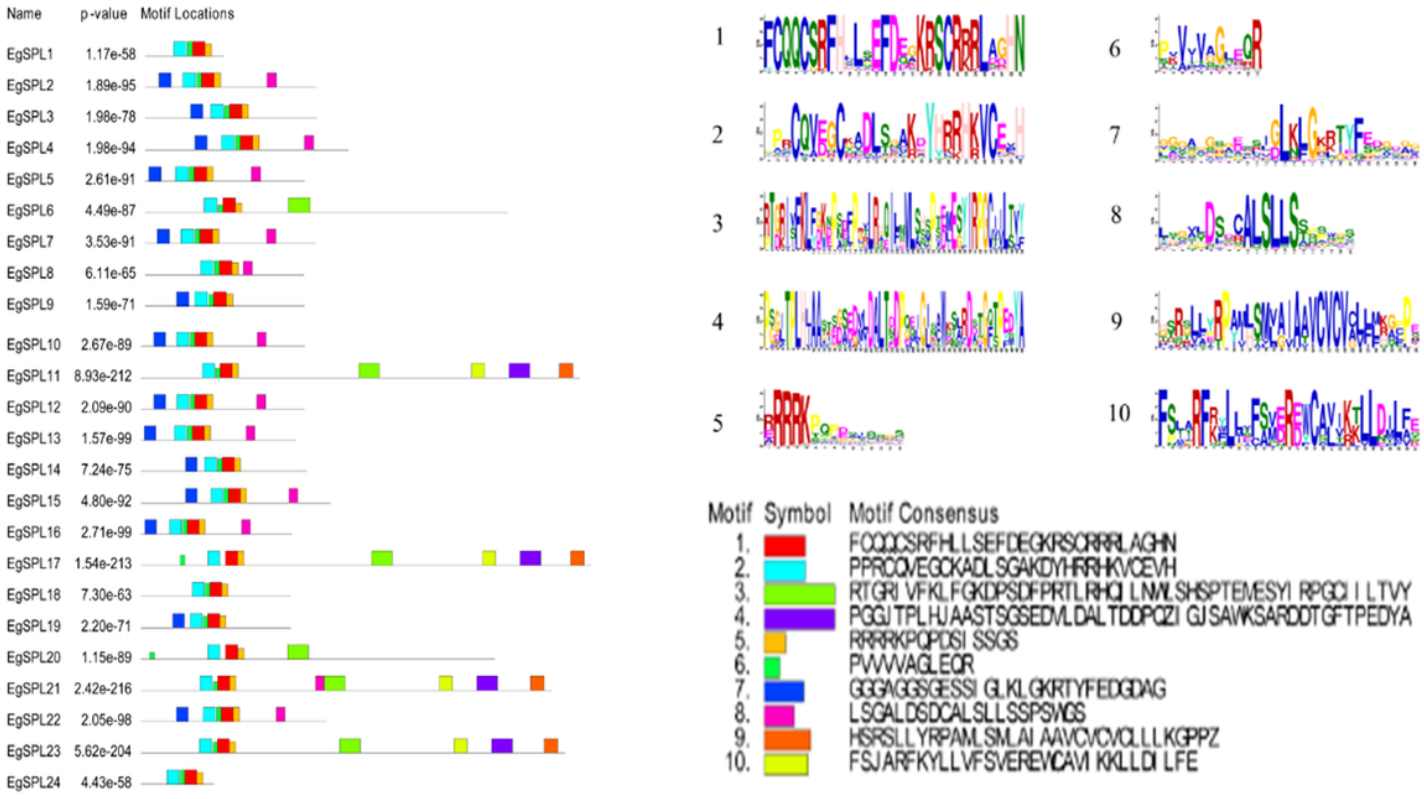


Figure 2

Motif analysis of *EgSPL* proteins using the MEME tool. Presence of 1-10 motifs in *EgSPL* proteins. Each motif is represented with different color. The abundance of each amino acid in their motifs is represented by sequence logo of each motif.

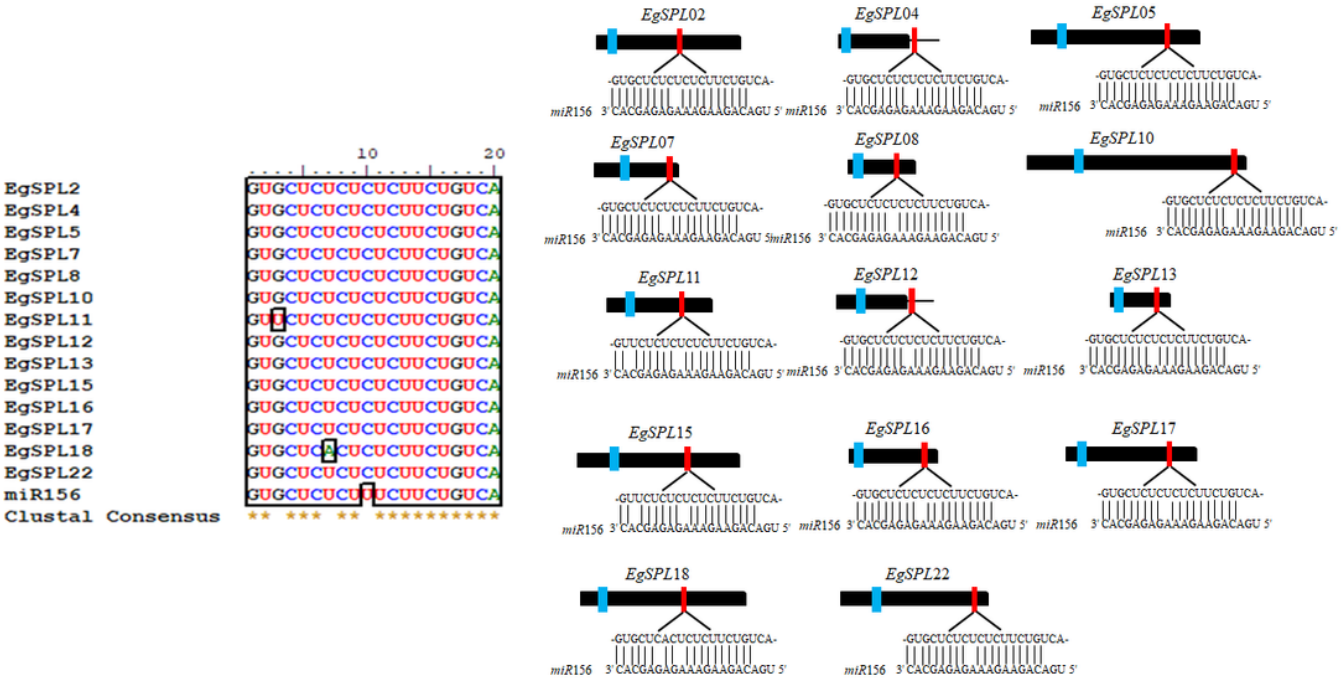


Figure 3

Prediction of miR156 target sequences of 14 *EgSPLs* through psRNATarget tool. Multiple alignment of miR156 complementary sequences with their targets (left). Right side diagram represents the gene structure of 14 *EgSPLs*. The black boxes represent CDS regions the black lines represent 3' and 5' UTR regions the blue boxes represent the SBP domains the red boxes represent *miR*156 target sites of *EgSPL* transcripts In the expanded regions the sequence direction of *EgSPL* is from 5' to 3' and the *miR*156 sequence direction is from 3' to 5'

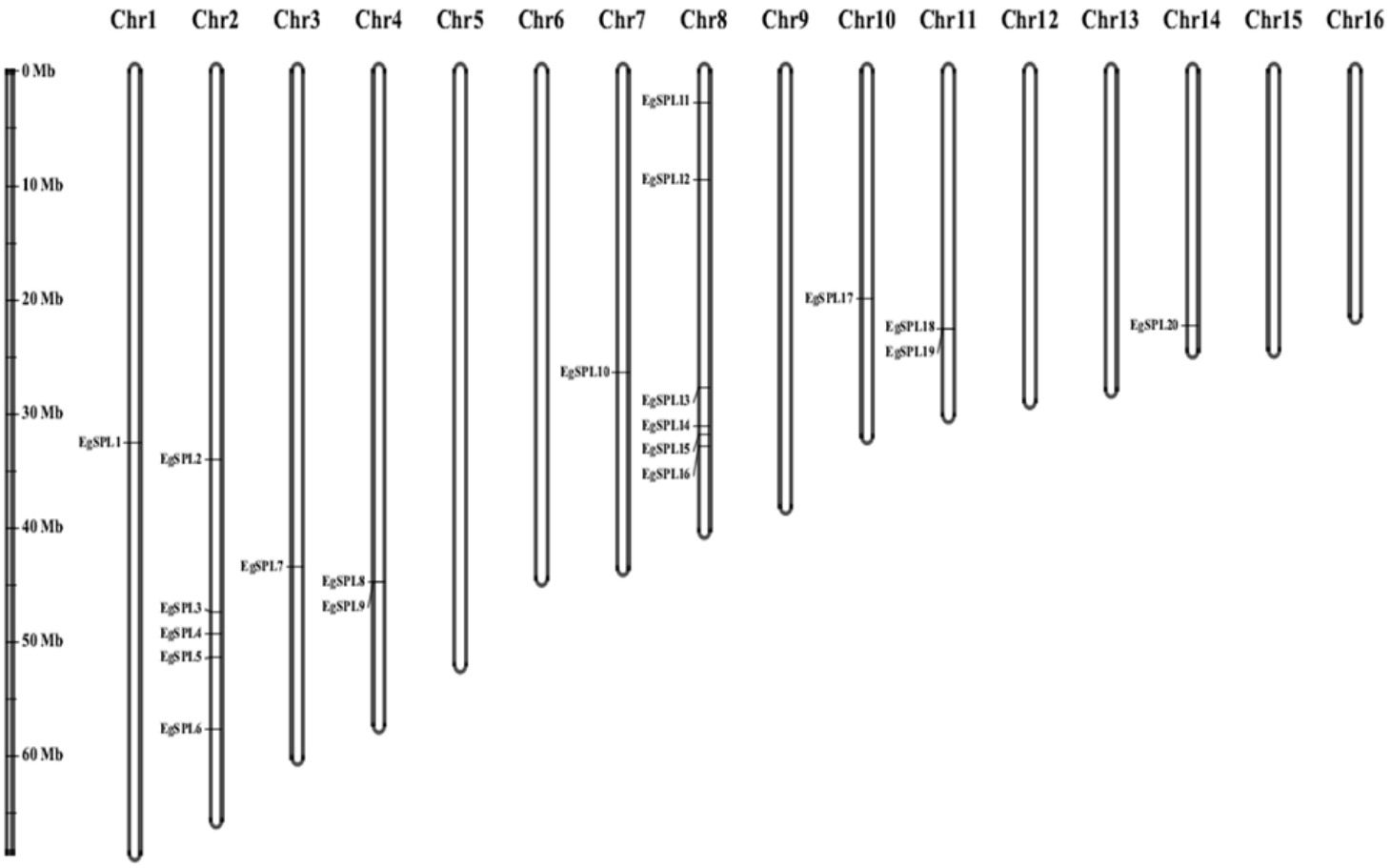


Figure 4

Uneven distribution of *EgSPL* genes on 16 chromosomes of oil palm genome. Chromosome numbers from 1-16 are marked on the top portion of each chromosome. The length of the oil palm chromosomes is represented with the vertical greyscale on the left side

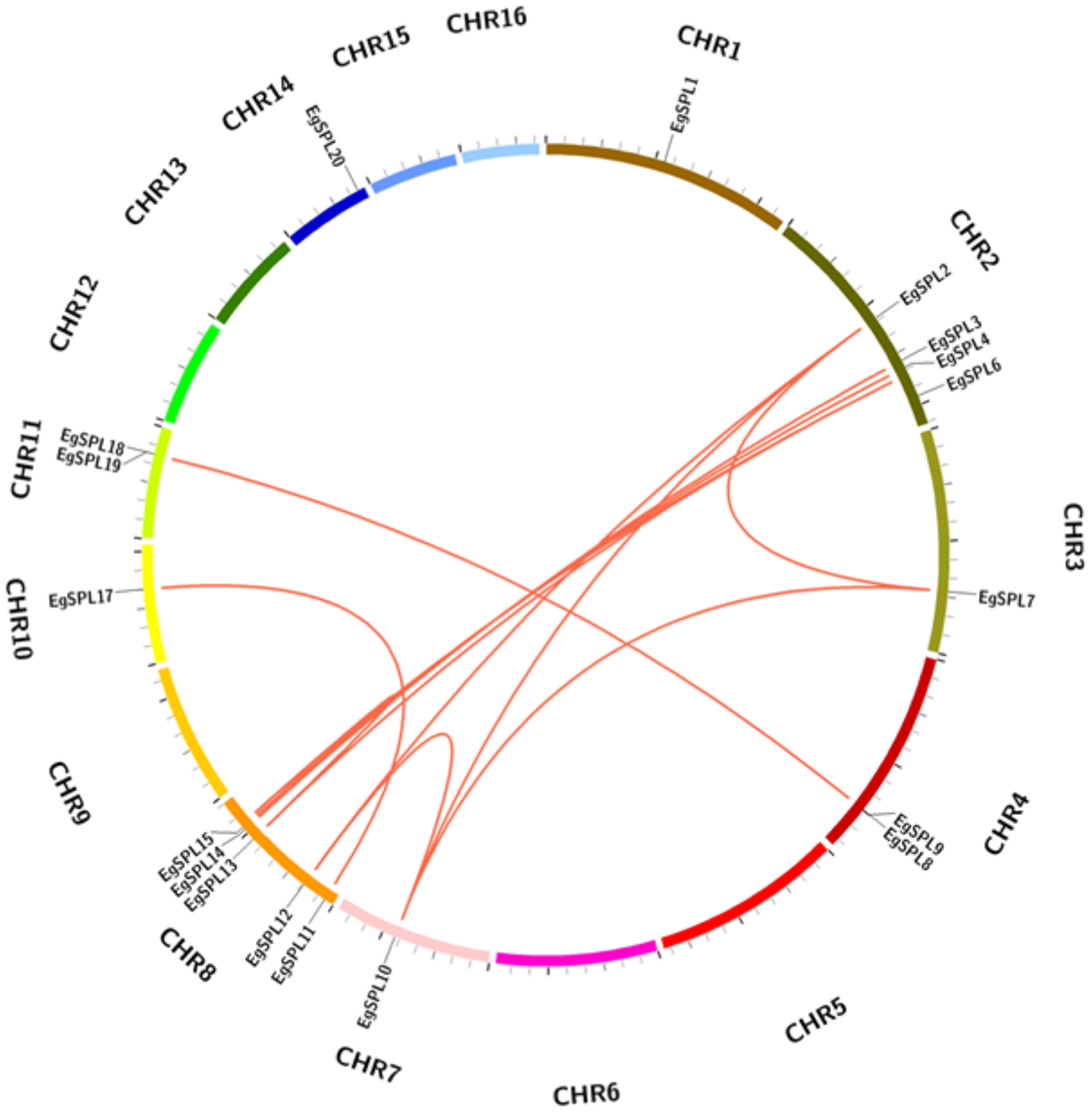


Figure 5

EgSPL gene duplication events in oil palm genome. The duplicated gene pairs are linked by the red lines inside the circle view as revealed by MC ScanX tool. Each chromosomal block was represented with a different color

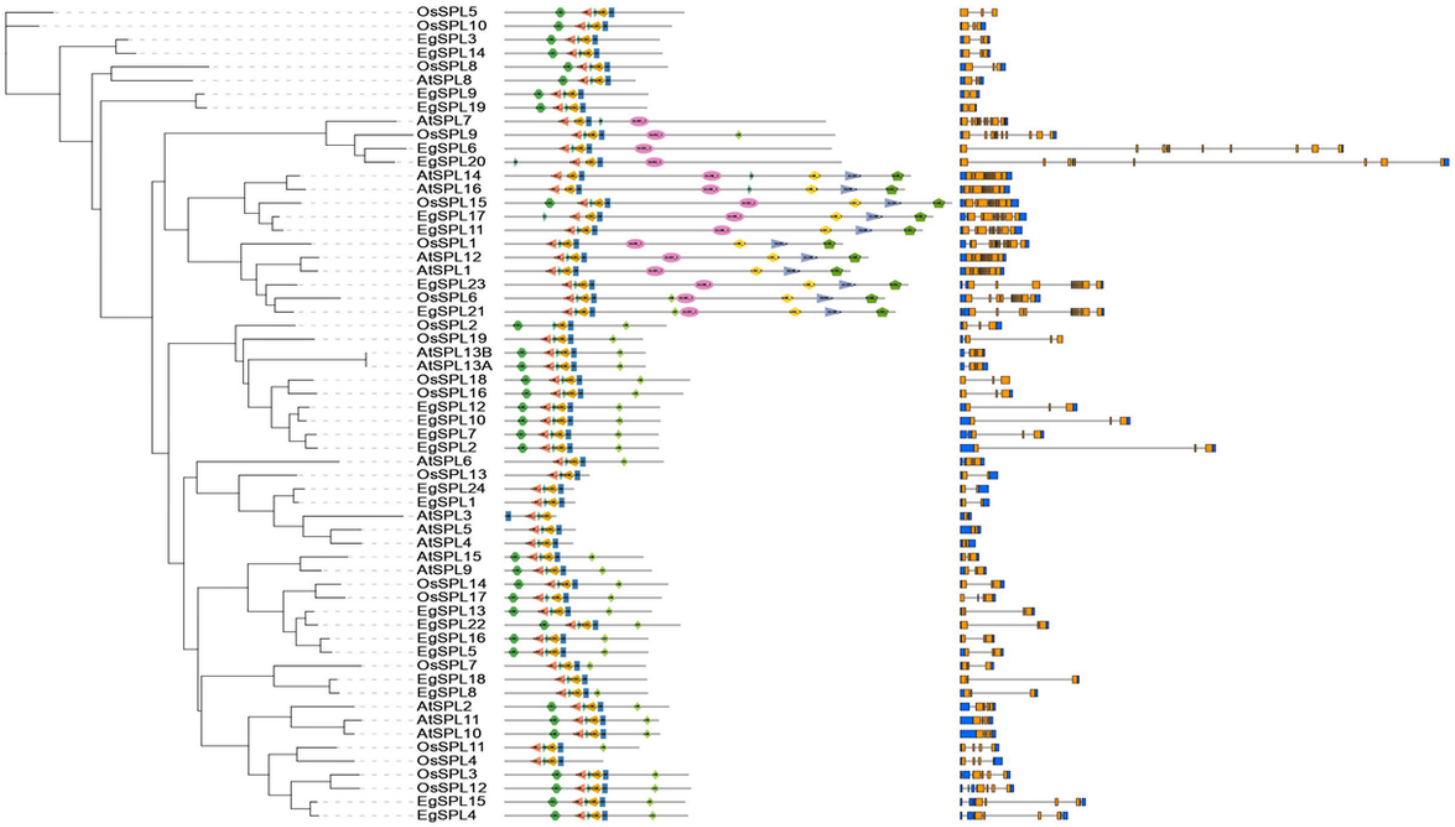


Figure 6

Phylogenetic analysis of *EgSPL* gene family with rice and *Arabidopsis SPL* family genes. Domain composition (middle), gene structural organization (right), of oil palm, rice, and *Arabidopsis SPL* members were also represented in the illustration.

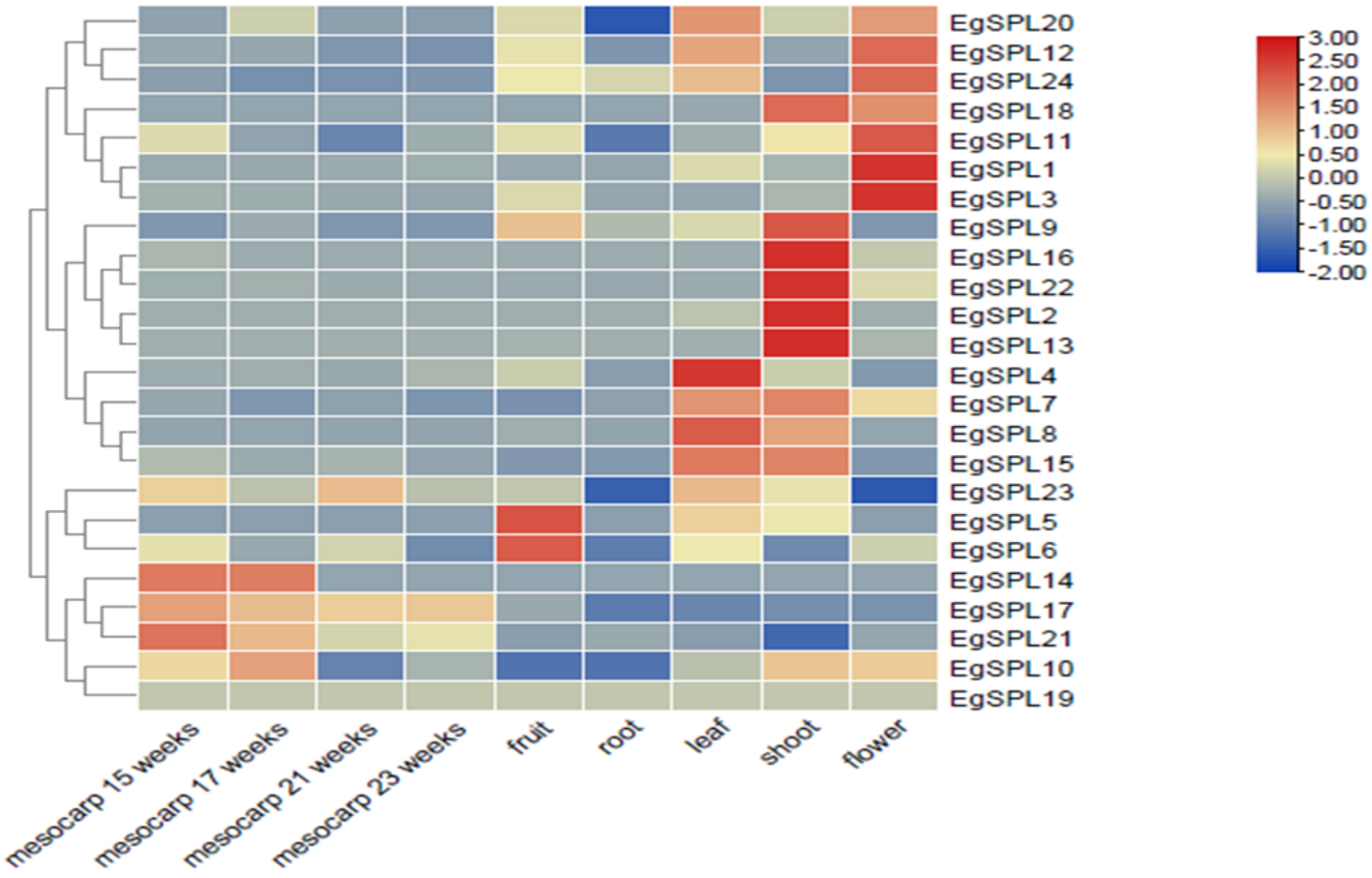


Figure 7

Tissue specific expression profile of *EgSPL*genes in vegetative (leaf, root, and shoot) and reproductive tissues (flower, mesocarp and fruit) based on the available transcriptome data.) of oil palm. SRR190698 represents transcriptome of mesocarp (15 weeks); SRR190699 represents transcriptome mesocarp (17 weeks); SRR190701 represents transcriptome mesocarp (21 weeks); SRR190702 represents transcriptome mesocarp (23 weeks); SRR851108 represents transcriptome of flower; SRR851067 represents transcriptome of fruit; SRR851096 represents transcriptome of the leaf; SRR851071 represents transcriptome root; SRR851103 represents transcriptome of the shoot.

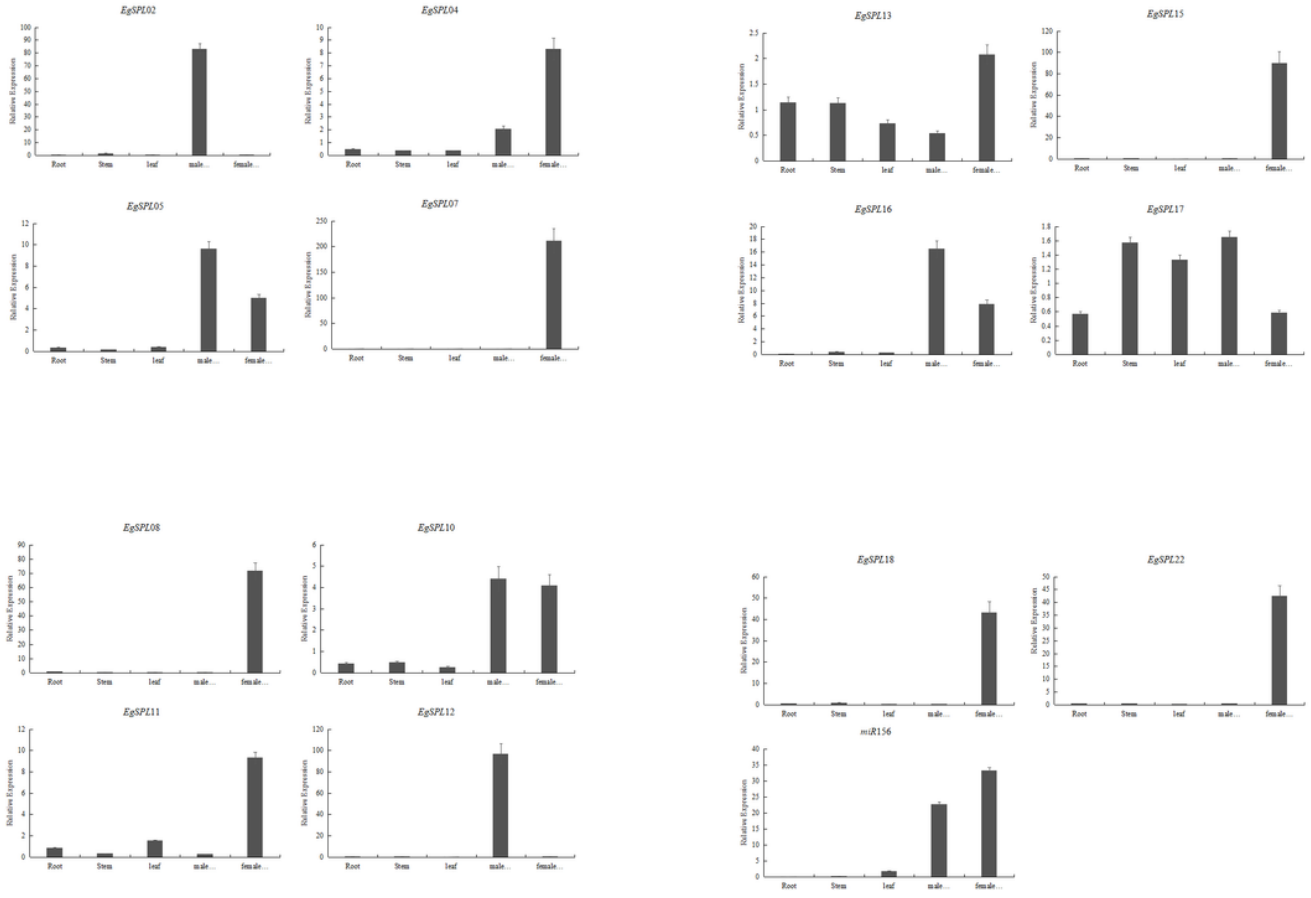


Figure 8

Relative expression of 14 *EgSPL* transcription factors (containing miR156 sites) in vegetative and reproductive tissues. Data represents the mean ± SE of three replicates. Asterisks represent a significant difference at $p \leq 0.05$

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryTable1.xls
- SupplementaryTable2.xls
- SupplementaryTable3.xls