The Journal of Animal & Plant Sciences, 30(1): 2020, Page: 133-146
ISSN: 1018-7081

GENOMIC ANALYSIS AND EXPRESSION INVESTIGATION OF STEROLEOSIN GENE FAMILY IN *ARABIDOPSIS THALIANA*

U. Aziz, N. Saleem, T. Tang, and M. Zhang

Northwest Agriculture and Forestry University, College of Agronomy, Taicheng Road, 3#, Yangling, Shaanxi, China, 712100
Corresponding author's email: zhangm@nwsuaf.edu.cn

ABSTRACT

Despite having growing evidence discussing the role of steroleosins in crop plants, their systemic analysis has not been worked extensively.–Therefore, bioinformatics tools were employed to study Arabidopsis (*Arabidopsis thaliana*) steroleosin in comparison with reported *Brassica napus* and pine steroleosins. Based on molecular weight, eight steroleosins in Arabidopsis were divided into two groups, L-steroleosin (alkalescence) and H-steroleosin (acidity except At5g50770). Phylogeny analysis suggested that H-steroleosins in Arabidopsis are primitive.–Moreover, phylogeny, protein alignment, motif analysis and gene structure analysis revealed that the reported steroleosins in *B. napus* are closely related to H-steroleosins. Similarly, Gene structure analysis revealed an insertion of 10 residues at the C-terminal of H-steroleosin which makes it relatively larger as compared to L-steroleosins. According to the Protein-protein interactions, Arabidopsis steroleosins are interrelated with genes involved in fatty acid biosynthesis and stress tolerance. Duplication studies suggest that among the steroleosin genes in Arabidopsis, segmental duplications are newly developed. Expression analysis suggests that H-steroleosins mainly express in seed tissue, while L-steroleosins diverge in two directions: expression or repression in all tissues. The upstream sequence analysis suggests that Arabidopsis steroleosins are involved in lipid metabolism and the brassinosteroid signaling pathways. The roles of steroleosins for stress responses are indicated by their induced expression under stress and abscisic acid treatments. Therefore, further investigations are required to evaluate the functions of steroleosins in plants.

Keywords: Steroleosins classification; Expression analysis; Potential functions, Evolution; Arabidopsis.

INTRODUCTION

For most eukaryotic organisms, the important source of energy is stored in the form of triacylglycerols (TAGs) and sterols (Murphy, 2012). On a per unit volume basis, the complete oxidation of TAGs provides energy twice as much as protein or carbohydrate hydrolysis do. Triacylglycerols (TAGs) in seeds are broken down during germination in order to offer energy as well as substrates for seedlings growth (Huang, 1996; Pritchard *et al.*, 2002; Murphy, 2012). The hydrophobic TAGs are enclosed by a phospholipid monolayer to form a small globular structure known as lipid droplets (LDs, also known as lipid bodies, oil bodies, spherosomes, lipid bodies and oleosomes). The phospholipid monolayer is embedded with different kinds of proteins, termed as oleosins, caleosins and steroleosins respectively (Murphy, 2012).

Steroleosins are the minor proteins related with LDs in plants. It is also known as sterol dehydrogenases, mainly because they share significant sequence similarity with the LD-associated hydroxysteroid dehydrogenases (HSDs) in mammals (Lin *et al.*, 2002) and a member of Arabidopsis steroleosins (At5g50600) shows *in vitro* activity of HSD (Baud *et al.*, 2009). Differently from the oleosin and caleosin, steroleosin has two main structural domains: a N-terminal hydrophobic domain which is required for LD association and a C-terminal domain which exhibits NADP(H) binding sub-domain and sterol binding sub-domain. Moreover, N-terminal hydrophobic domain contains conserved proline-knob instead of the proline-knot, which is proposed to have functions associated with the LD surface (Lin *et al.*, 2002; Chapman *et al.*, 2012).

Isoforms of steroleosin with distinctive sterol-binding sites are present in sesame oil bodies, which are regulated by distinct sterols to carry out numerous metabolic reactions in seeds, specially related to the degradation or formation of the oil bodies (Chen *et al.*, 1998; Pasaribu *et al.*, 2016). Sop2 in sesame was named as steroleosin because of its structural properties i.e. N-terminal hydrophobic domain consisting of sterol binding dehydrogenase on the surface of the lipid droplets. Steroleosins have also been discovered in numerous plant species such as pine tree (*Pinus massoniana*) (Pasaribu *et al.,* 2016). More recently, steroleosin B in sesame was identified as sesamin binding protein, putatively sesamin in plants is reported to have numerous functions related to maturation of seeds and germination (Tera *et al.*, 2019). Molecular study revealed that *AtHSD1* (At5g50600 and At5g50700) is the direct target of LEAFY COTYLEDON2 (LEC2), which can bind to the promoter of *AtHSD1* (Baud *et al.,* 2009). Down-

regulation of HSDs in Arabidopsis can affect the plant growth and normal development (Li *et al.,* 2007).

Arabidopsis (*Arabidopsis thaliana*) is an important model organism due to its small genomic size (125 MB) and short life span (Meinke *et al.*, 1998). To date, eight steroleosin genes have been discovered in the Arabidopsis genome (Lin *et al.*, 2002). Steroleosin is the third oil droplet protein with relatively high expression in the plant seeds. It is believed; steroleosin has key roles in brassinosteroid metabolism and plant signaling (Li *et al.*, 2007; van der Schoot *et al.*, 2011; Ferrer *et al.*, 2017). Brassinosteroids in plants and mammals have number of regulatory roles like stress response, reproduction, growth, development and so on (Clouse and Sasse, 1998; Fujioka and Yokota, 2003; Zhou *et al.*, 2015). Whether steroleosin has other functions, like regulation of oil contents in seeds and roles in brassinosteroid signaling pathways, still needs to be elucidated. Cis-regulatory elements are responsible for the transcriptional regulation of neighboring genes in order to carry out various biological functions. So far, less attention is paid to carry out promoter sequence analysis of steroleosin. Additionally, little study focuses on their structural features, properties and phylogenesis.

The Arabidopsis genomic sequence and various online gene expression databases offer an opportunity to analyze and predict gene functions by using different bioinformatics tools. In this study, the genomic wide analysis of steroleosin genes in Arabidopsis is presented. According to molecular weight, multiple sequence alignment and motif analysis, the eight steroleosin genes were classified into two groups, H-steroleosin and L-steroleosin. Moreover, the properties, evolutionary aspects, and structure of steroleosin genes were also investigated. Lastly, possible biological functions of steroleosin genes were predicted on the basis of promoter sequence analysis and their expression pattern.

MATERIALS AND METHODS

Identification of steroleosin genes in Arabidopsis: To find out steroleosin genes in Arabidopsis, HMM (Hidden Markov Model) and exact name searches were carried out in Arabidopsis genome. The knownArabidopsis steroleosin "At5g50700 was used as a query in the Arabidopsis information database (TAIR10.0, http://www.arabidopsis.org/) (Huala *et al.*, 2001). The raw HMM of steroleosin domain PF00106 was found from the Pfam (http://pfam.sanger.ac.uk/) (version. 31.0). The protein sequences of the steroleosin genes were retrieved from the TAIR database. To verify the existence of steroleosin domain (PF00106) all the putative amino acid sequences were analyzed by the Interproscan (http://www.ebi.ac.uk/Tools/pfa/iprscan/) (version.67) (Hunter *et al.*, 2012).

Physical and chemical properties of steroleosin genes: The ExPASY (http://web.expasy.org/compute_pi/) (Gasteiger *et al.*, 2003) was used to analyze the characteristic features of steroleosin proteins. Putative phosphorylation sites (Ser, Thr and Tyr) were predicted through the NetPhos (http://www.cbs.dtu.dk/services/NetPhos/) (version. 3.1) (Blom *et al.*, 1999). The hydropathic plot was drawn in the ProtScale (http://web.expasy.org/protscale/) with the Kyte and Doolittle method (Kyte and Doolittle, 1982) and the default setting. To find out the conserved domains within the steroleosin sequences, the MEME (http://meme.nbcr.net/meme/cgi-bin/meme.cgi) (Bailey *et al.*, 2006) was used for discovering 10 different motifs. The annotation of motifs was performed through the Pfam (https://pfam.xfam.org/) (Finn *et al.*, 2015).

Multiple sequence alignment and polygenetic analysis: The ClustalX (version. 2.1) (Larkin *et al.*, 2007) was used for multiple protein sequence alignment of eight Arabidopsis steroleosins, one pine (*P. massoniana*) steroleosin and five *Brassica napus* steroleosins. The protein sequence of pine steroleosin was retrieved from gene bank with the accession No. KT731102 and accession numbers for *B. napus* steroleosins, BnSLO1-1, BnSLO1-2, BnSLO1-3, BnSLO2-1 and BnSLO2-2 were ACG69522, ACG69523, ACG69524, ACG69525 and ACG69526 respectively. Output and shading of sequence alignment were carried out by the Boxshade (http://www.ch.embnet.org/software/BOX_form.html) (version. 3.21) (Gasteiger *et al.*, 2003).

To find out evolutionary relationship between different steroleosins, both maximum likelihood (ML) and neighbor joining (NJ) trees were populated. For NJ tree MEGA (version. 6.06) (Tamura *et al.*, 2013) was used by using the Poisson model, gaps were deleted pairwise, and bootstrap replication was set to 1000. The PhyML (http://www.atgcmontpellier.fr/phyml/) (version.3.0) was employed for the ML tree. The LG model was used with substitution rate of 6, SPR mode and bootstrap replication was 500.

Chromosomal Localization and duplication events of steroleosin: By using locus names on the TAIR chromosome map tool, the localization of genes were mapped on five chromosomes of Arabidopsis. Two genes were considered as tandem duplications if they were separated by utmost of five genes (Ma and Zhao, 2010). Segmental duplication events were searched according to the previous study (Maher *et al.*, 2006).

Upstream sequence element analysis: For the prediction of cis regulatory elements, about 2000 bp upstream sequences of eight steroleosin genes were retrieved from the Arabidopsis TAIR website as their promoters and analyzed by the Promoter Scan

(https://www-bimas.cit.nih.gov/molbio/proscan/) (version. 2) and the PlantCare (Lescot *et al.*, 2002).

Structural analysis of steroleosin genes in *Arabidopsis thaliana*: The gene structural and intron/exon analysis were carried out by using the Gene Structure Display Server (GSDS) (version. 2.0) (Hu *et al.,* 2014).

Protein-Protein interaction of steroleosins in Arabidopsis thaliana: Protein-protein interactions of Arabidopsis thaliana steroleosins were analyzed by using the STRING database (http://string.embl.de) (version. 10).

Expression Analysis of Steroleosin genes: Expression patterns of steroleosin group in different Arabidopsis tissues were determined by using RNA-seq data. The NCBI SRA database was used to obtain RNA-seq data of bud, flower, inflorescence, leaf, root, silique, seed and germinating seed. Gene sequences of different steroleosins were used for the BLASTN analysis in the relevant databases. Parameters were as follows: identity percentage was set to 100%, lengths were according to the sequence length in the database. The calibration of the data was performed as fragments per kilo base of exon/per million fragments mapped (FPKM).

Because RNA-seq data under different stress conditions was scarce, microarray data was used instead. The AtGenExpression visualization tool (AVT, http://jsp.weigelworld.org/expviz/expviz.jsp) was used to perform microarrays analysis and visualization by using locus search with mean normalized value.

The heatmaps were produced with the R script, an online tool (https://www.hiv.lanl.gov/content/sequence/HEATMAP/heatmap.html) by using the log2-transformed-calibrated values of expression data (Text S1).

RESULTS

Identification, classification and characteristics of steroleosins in Arabidopsis: Eight Arabidopsis steroleosin proteins were identified by TAIR and HMM searches and were named by referring to previous studies (Lin *et al.*, 2002; Li *et al.*, 2007; Baud *et al.*, 2009) with little modification. Among the identified steroleosins two i.e., At5g50600 and At5g50590 are completely same to At5g50700 and At5g50690 respectively, and were named as *AtHSD1a/AtHSD1b* (At5g50600/At5g50700) and *AtHSD4a/AtHSD4b* (At5g50590/At5g50690). The size of deduced amino acids, gene open reading frame (ORF), molecular weight (Mw) and isoelectric point (pI) value of each steroleosin are presented in table 1. The sizes, Mws, and pI values of *AtHSD1a* (At5g50600) and *AtHSD4a* (At5g50590) were found as same as *AtHSD1b* (At5g50700) and *AtHSD4b* (At5g50690) respectively. According to the molecular weight, the eight steroleosins were divided into two groups: high molecular weight and low molecular weight steroleosins, termed as H-steroleosin and L-steroleosin.

Protein alignment of eight steroleosins in Arabidopsis, one steroleosin in pine and five steroleosins in *B. napus* was performed to investigate structural features (Fig.1). Amino acids of H-steroleosin *AtHSD1a* (At5g50600) and *AtHSD1b* (At5g50700) are completely same, while L-steroleosin *AtHSD4a* (At5g50590) is fully identical to *AtHSD4b* (At5g50690). The analysis reveals that the insertion of 10 residues at the C-terminal is the main difference between H-steroleosin and L-steroleosin and this insertion makes the sterol binding region relatively larger. It is noteworthy that both the reported pine and *B. napus* steroleosin also exhibit the extra insertion (Fig.1).

The NADP(H) binding domain is quite conserved among all of the steroleosin proteins (Fig.1). The N-terminal hydrophobic domain is essential for the association with the oil droplet surface (Bhatla *et al.*, 2009; Pasaribu *et al.*, 2016). Our analysis also reveals that the proline knob motif is highly conserved in all steroleosin proteins (Fig.1). To further investigate the stability of the steroleosin proteins, hydrophobic plots were generated and the change of stability for amino acid mutation was predicted (Fig.S1). The result suggests that hydrophobic region of H-steroleosin is smaller than that of L-steroleosin, which led to decreased in stability (Fig.S1).

The putative phosphorylation sites, Ser, Thr and Tyr, were predicted and are marked on the Fig.1. According to *AtHSD5* (At4g10020), Y67, S233, Y279 sites are conserved in H-steroleosin and T53, S96, S107, T119, T165, S201, S246 and S253, according to *AtHSD4a* (At5g50590), are conserved in L-steroleosin.

Phylogenetic analysis of steroleosins: Neighbor joining (NJ) and maximum likelihood (Ml) trees were constructed to determine the evolutionary relationship of steroleosins (Fig.2). It was observed that bootstrap values of both methods were higher than 80%, which suggests the higher reliability of the trees. Moreover, both methods led to almost similar topological structures of both trees. The phylogenetic trees were divided into two groups and both pine and *B. napus* steroleosins were found closely related to H-steroleosin (Fig.2), indicating that H-steroleosins may be more primitive and L-steroleosins may be evolved from H-steroleosins. Moreover, BnSLO1-1, 2 and 3 are closely related to *AtHSD1a* (At5g50600) and *AtHSD1b* (At5g50700). Similarly, BnSLO2-1 and 2 were found closely related to *AtHSD5* (At4g10020). It is noteworthy that the results of polygenetic tree are matched to the two groups divided by the molecular weight.

Motif analysis: In order to find out different motifs in Arabidopsis and *B. napus*, ten motifs were identified and

analyzed. It was observed that all steroleosins in Arabidopsis and *B. napus* exhibit motif 1-4 and motif 6 (Fig.3A). Motif 5 on the other hand, was absent only in *AtHSD2* (At3g47350). Similarly, motif 9 was present in all steroleosins except *AtHSD6* (At5g50770). It is noteworthy that motif 7 and Motif 8 were present only in H-steroleosins (Fig. 3A). Moreover, motif 10 was present only in *AtHSD1a* (At5g50600), *AtHSD1b* (At5g50700) and BnSLO1-1 and 2. The annotation of the motifs revealed that two of them were the parts of the steroleosin whereas the annotation for the others was not available (Table 2).

Steroleosin genes structure analysis: The intron/exon analysis of eight steroleosins in Arabidopsis and five steroleosins in *B. napus* were performed (Fig.3B). The intron and exon regions of the Arabidopsis and *B. napus* steroleosin genes were conserved except *AtHSD2* (At3g47350). Six exons and five introns were found in seven Arabidopsis and five *B. napus* steroleosins with a 0, 1, 0, 0 and 0 intron phase, while in *AtHSD2* (At3g47350), seven exon and six intron were identified with a 0, 1, 0, 0, 0 and 1 intron phase. Moreover, the last exon in all four H-steroleosins of Arabidopsis and five steroleosisns in *B. napus* is the largest one, which is due to the extra insertion. Similar to the same protein sequences, sizes of introns and exons in *AtHSD4a* (At5g50590) and *AtHSD1a* (At5g50600) are also completely as same as *AtHSD4b* (At5g50690) and *AtHSD1b* (At5g50700) respectively. The first intron phase of BnSLO2-1 and 2 were relatively larger as compared to the other steroleosin genes. Moreover, compared to other steroleosin genes, *AtHSD2* (At3g47350) shows the conserved intron phase except that there is an extra last intron and the last exon is divided into two exons (Fig.3B). Combining the above evolution analysis, the results indicate that *AtHSD2* (At3g47350) may evolve from *AtHSD3* (At3g47360) by insertion of an intron in the last exon.

Protein-protein interaction of Arabidopsis steroleosins: In order to understand the possible roles of the steroleosins in Arabidopsis, the protein-protein interaction maps of six Arabidopsis steroleosins were drawn by using STRING software (Fig. 4). It was revealed that all the steroleosin genes were related to the genes involved in fatty acid metabolism. Our analysis suggested that *AtHSD1a* (At5g50600), and *AtHSD3* (At3g47360) were interacted with oleosin gene family (Fig. 4A and C). The oleosin gene family was reported to have role in seed lipid accumulations and their stability (Lu *et al.*, 2018). Similarly both *AtHSD1a* (At5g50600) and *AtHSD3* (At3g47360) are related to the LACS5 (At4g11030) and LACS8 (At2g04350), which are AMP-dependent synthetase and ligase family protein and are responsible for long chain fatty acid-CoA biosynthesis (Clark *et al.*, 2018). Furthermore, it was also found that *AtHSD3* (At3g47360) and *AtHSD4b* (At5g50690) were regulated by the Glycine Rich protein family (GRP) (Fig. 4C and F). Our study revealed that *AtHSD4a* (At5g50590) and *AtHSD4b* (At5g50690) were interrelated with each other (Fig. 4D and F). Similarly, *AtHSD2* (At3g47350) has interactions with BEN1 protein which has roles in brassinosteroids biosynthesis (Oklestkova *et al.*, 2015) (Fig. 4B). *AtHSD5* (At4g10020) was found to be interacted by KCS3 and KAS1 gene family which are reported to have important biological functions in fatty acid biosynthesis (Fig. 4E).

Chromosomal localization and duplications of steroleosin: The localization of steroleosin genes on five different chromosomes of Arabidopsis was investigated. As presented in Fig. 5, there is no steroleosin gene located on the chromosome 1 and 2. However, five steroleosin genes are located on the chromosome 5, whereas two are on chromosome 3 and only one was located on the chromosome 4.

The investigations on the duplication events suggest that *AtHSD2* (At3g47350) and *AtHSD3* (At3g47360) are tandemly duplicated, which are also closely related to each other on the proteins level (Fig.5, Table 3). Similarly, tandem duplication was also found among two pairs *AtHSD1a/AtHSD4a* and *AtHSD1b/AtHSD4b* (At5g50600/At5g50590 andAt5g50700/At5g50690) of steroleosin on the chromosome 5 (Fig.5, Table 3). On the other hand, segmental duplications analysis reveals that *AtHSD1a* (At5g50600) and *AtHSD4b* (At5g50590) are the segments of *AtHSD1b* (At5g50700) and *AtHSD4b* (At5g50690) respectively with 100% identity at protein level (Fig.5, Table 4), indicating that the segmental duplication occurred not long time ago.

Pormoter analysis of steroleosin genes in Arabidopsis: To determine the potential function of steroleosins in Arabidopsis, promoter sequence analysis was performed. Up to 2000 bp upstream sequences of steroleosins were retrieved and subjected to the Promoter Scan 2 and the Plant Care online tools.

According to the results from the promoter scan 2 (Text S2), the RY repeat element existed in *AtHSD2* (At3g47350), *AtHSD1a* (At5g50600) and *AtHSD1b* (At5g50700) (Text S2). In contrast, the Plant care analysis suggests, RY repeat element is present only in *AtHSD1a* (At5g50600) and its homologue partner *AtHSD1b* (At5g50700). Similarly Skn-1 motif (Text S2) is found in all Arabidopsis steroleosin genes except *AtHSD3* (At3g47360). Similarly, according to promoter scan 2 analyses MYB and E-Box elements exists in all of the steroleosin genes of Arabidopsis. In addition to this, many cis elements related to environmental stress, light responsiveness, and phytohormone signaling were also identified in different steroleosin genes.

Expression analysis: To determine the expression patterns of steroleosin genes in different tissues of Arabidopsis, the RNA-seq data was retrieved from SRA database and subjected to generate heatmap. As shown in Fig.5A, three H-steroleosin genes, including *AtHSD1a* (At5g0600), *AtHSD1b* (At5g50700) and *AtHSD5* (At4g10020), exhibit relatively higher expression levels in seeds, siliques and germinating seeds. Whereas, other H-steroleson *AtHSD6* (At5g50770) shows as similar expression patterns as other three H-sterolesoins except relatively low expression in germinating seeds. On the contrary, four L-steroleosin genes show two types of expression patterns: *AtHSD2* (At3g47350) and *AtHSD3* (At347360) express at a relative high levels in all tissues, while *AtHSD4a* (At5g50590) and *AtHSD4b* (At5g50690) express at very low levels in all tissues. These results indicate that H-steroleosin genes mainly express in seed related tissues and L-steroleosin genes evolve to two directions: express in all tissue or in no tissue. The two L-steroleosins *AtHSD4a* (At5g50590) and *AtHSD4b* (At5g50690) may not play function or just do minor function in all Arabidopsis tissues.

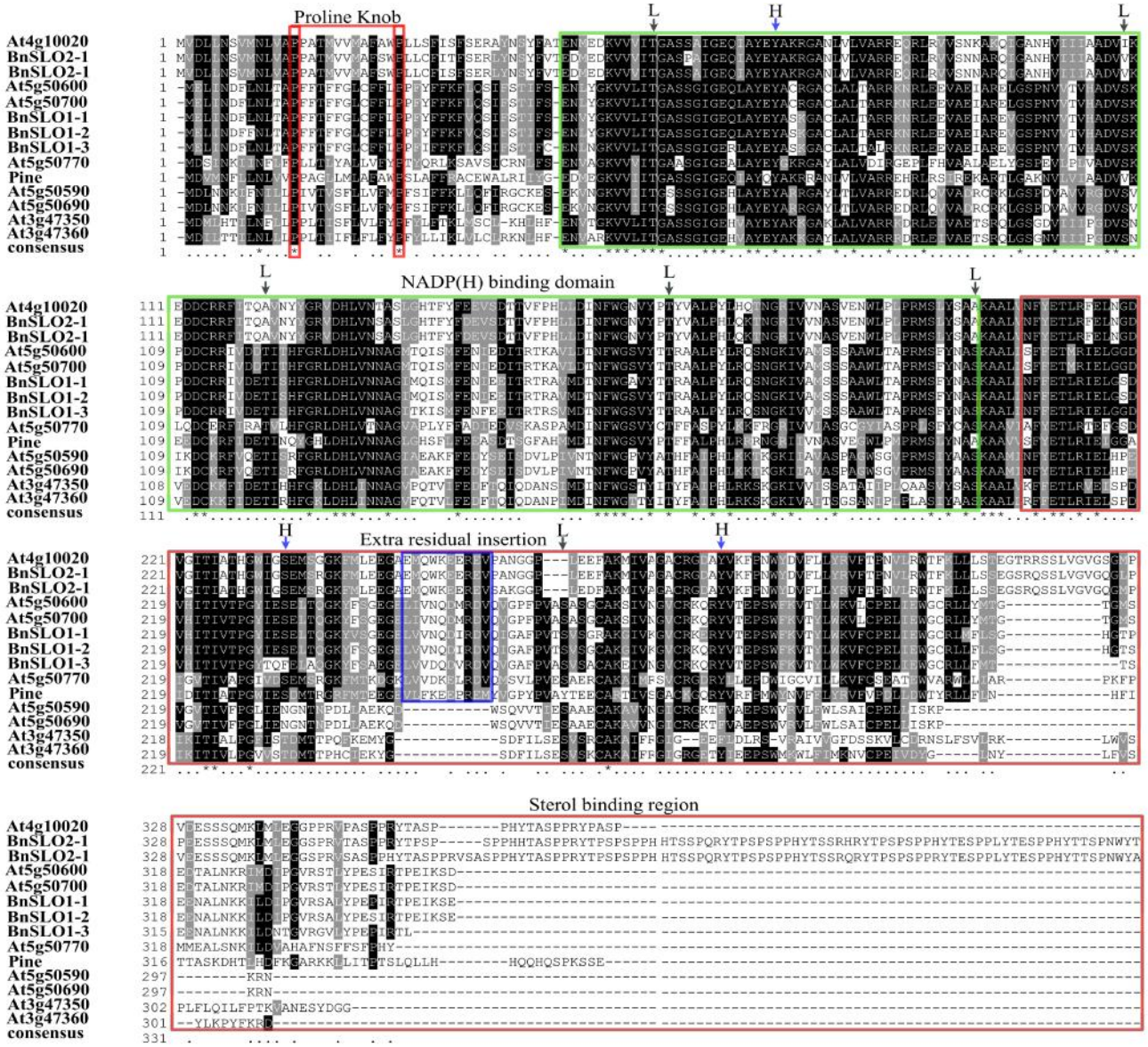

Fig. 1 Multiple sequence alignment of the eight steroleosins in Arabidopsis and one steroleosin in pine and five steroleosins in *Brassica napus*. Extra residual insertion is boxed with blue line. NADP(H) binding motif is boxed with green line. Sterol binding region is boxed with red line. Proline-knob domain is boxed with red line. Putative phosphorylation sites of L-steroleosinare marked by upper black arrow with ‘‘L’’ and H-steroleosin by blue arrow with ‘‘H’’.

Expression analysis under different stress conditions: In order to examine possible roles of steroleosins, expression patterns of steroleosin genes under different stress and hormonal conditions were investigated by mining microarray data. Among the eight different steroleosin genes, the microarray data was found for *AtHSD5* (At4g10020)*, AtHSD1a* (At5g50600)*, AtHSD4a* (At5g50590) and *AtHSD6* (At5g50770). Expression analysis under different hormonal conditions showed either in seeds or seedling steroleosins has inducive response to abscisic acid (Fig.5B), while other hormones showed very little or negligible effects on the steroleosin genes (data not shown). Our analysis suggests, among all steroleosins *AtHSD1* (At5g50600) was induced in seedlings under ABA stress conditions. On the contrary, all steroleosins in seed tissues were sensitive to different concentrations of ABA (Fig.5B). The result of various abiotic stresses shows that *AtHSD6* (At5g50770) and *AtHSD1a* (At5g50600) were induced in the aerial parts under both osmotic and heat stress condition (Fig.S2). Moreover, *AtHSD6* (At5g50770) was also induced in aerial parts under oxidative stress (Fig. S2). Similarly, A*tHSD1a* (At5g50600) was induced in roots under both osmotic and heat stress condition, and *AtHSD4a* (At5g50590) was induced in cell culture and roots under heat and genotoxic stresses (Fig.S2). These results suggest that these steroleosins may play some roles under numerous stress conditions.

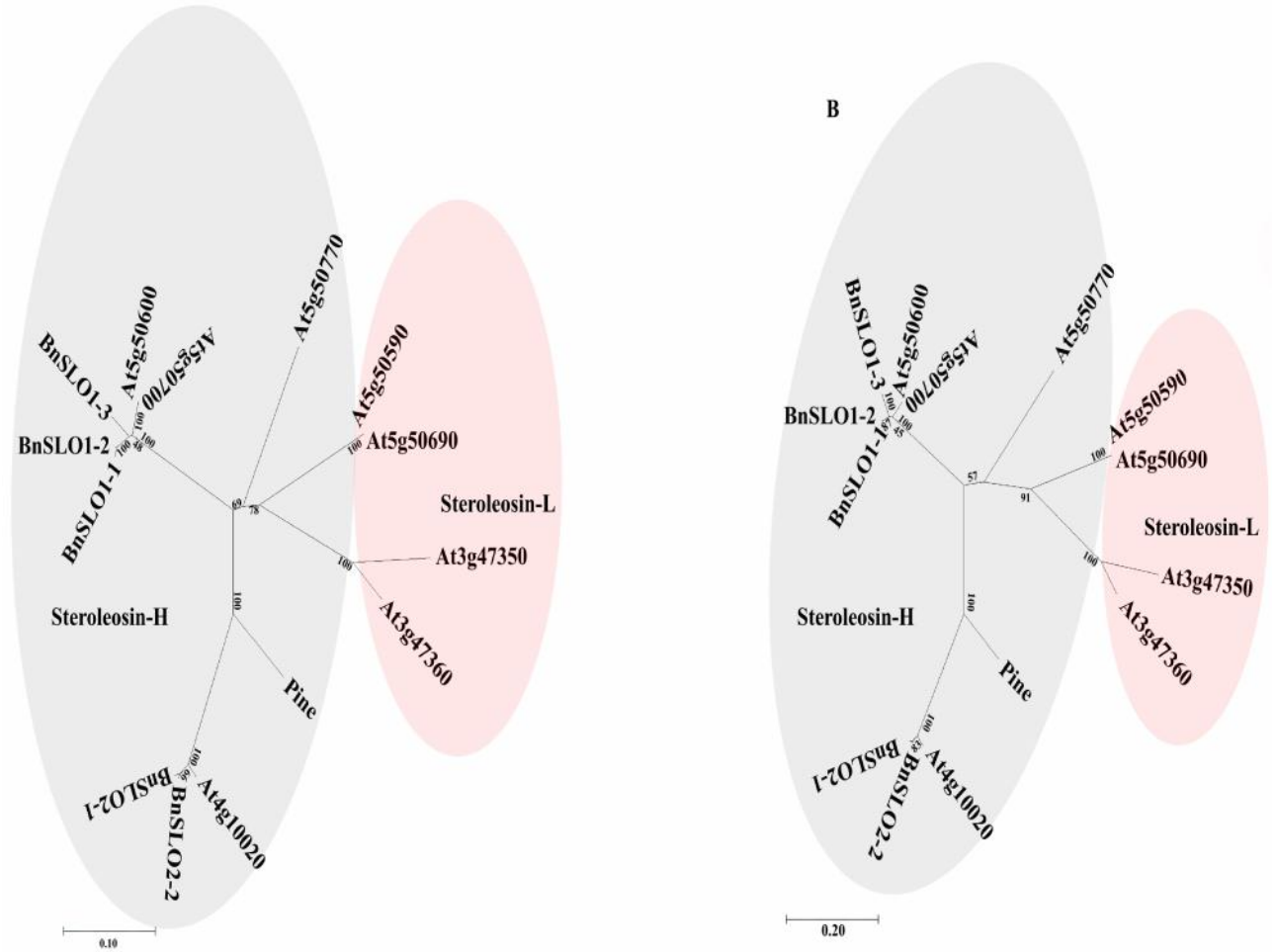

Fig. 2 Phylogenetic trees of eight steroleosins in Arabidopsis and one steroleosin in pine by maximum likelihood method in PhyML 3.0 (A) and neighbor-joining method in MEGA 6.06 (B).

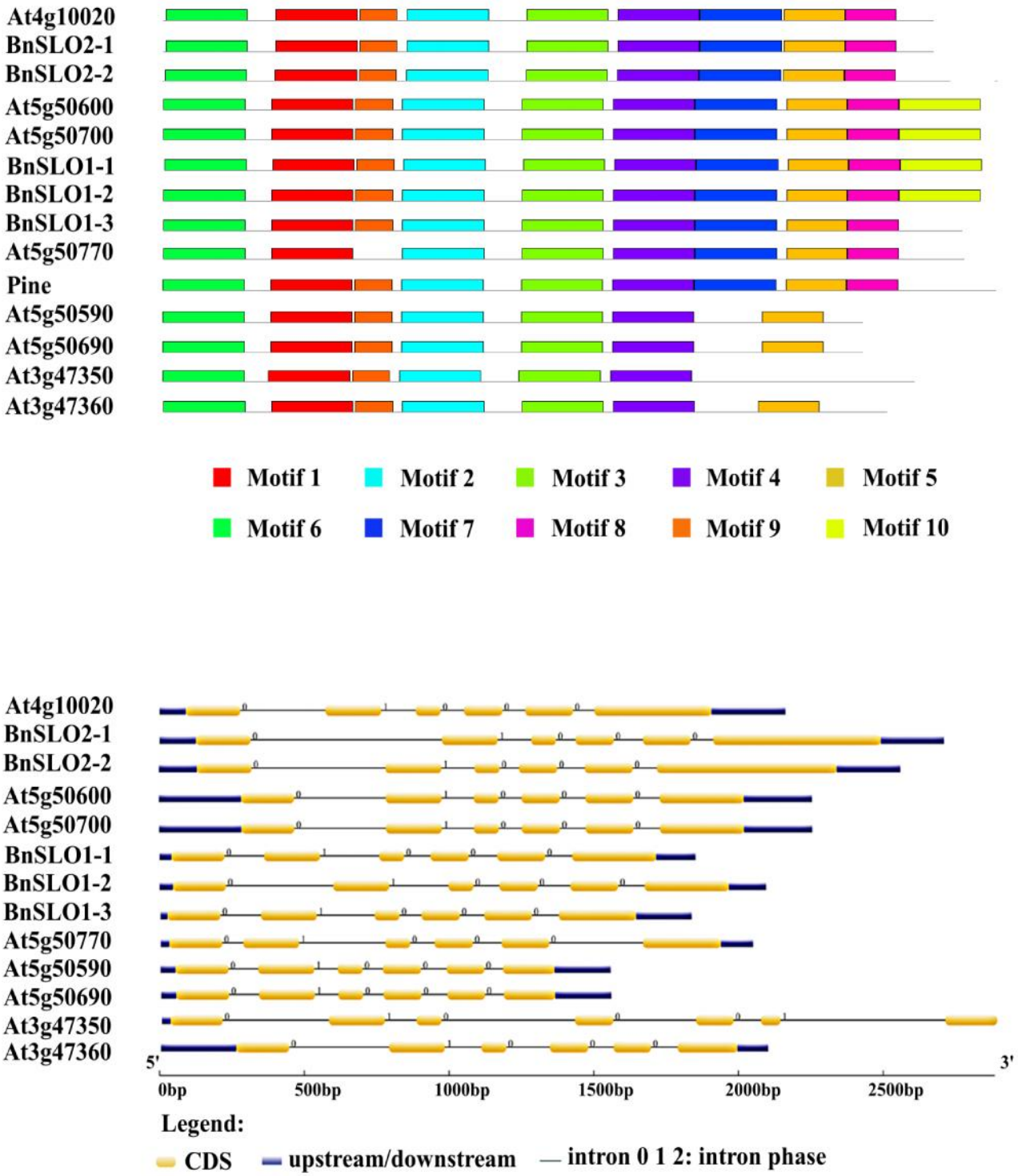

Fig. 3. Distribution of motifs in eight *Arabidopsis thaliana* one pine and five *Brassica napus* steroleosins (A). Analysis of gene structure and intron phase prediction of steroleosin genes in *Arabidopsis thaliana* and *Brassica napus* (B).

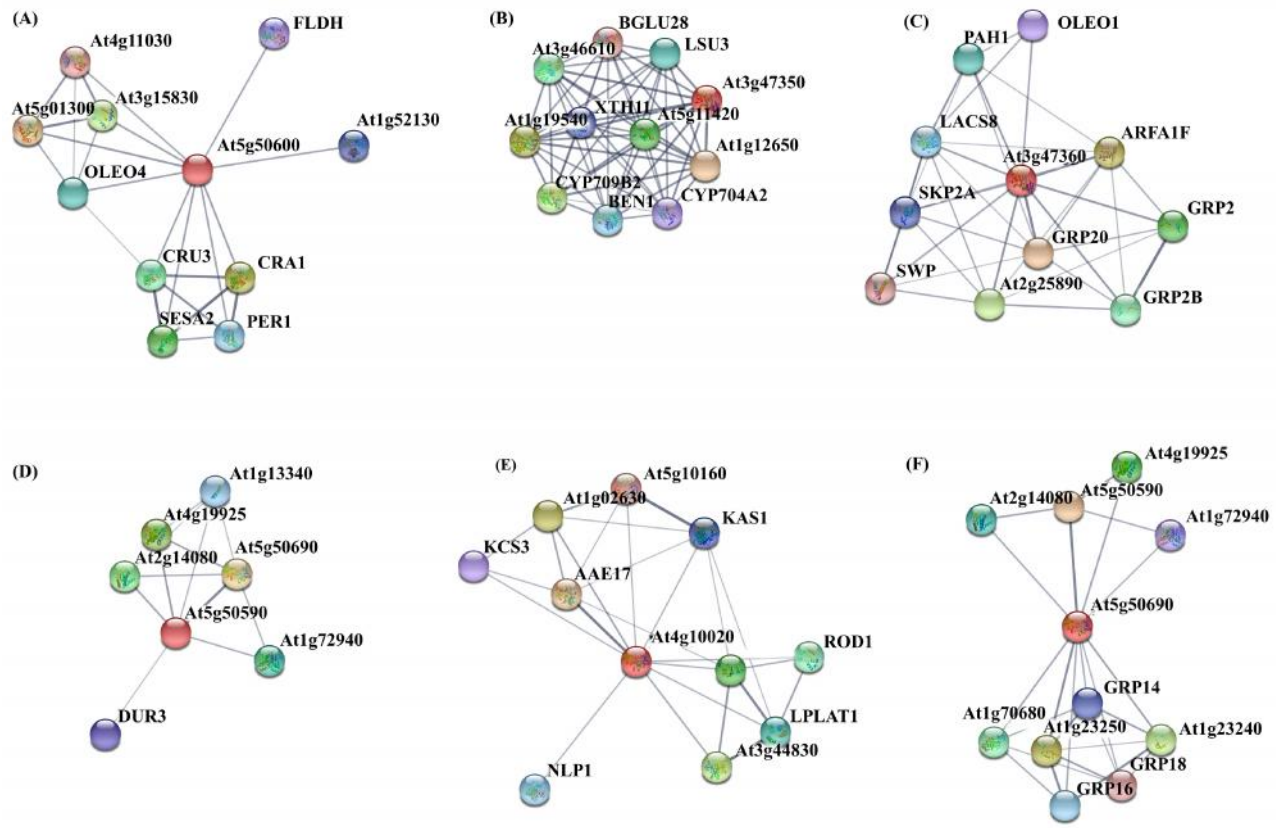

Fig 4. Protein-protein interaction of steroleosin genes in Arabidopsis. A, B, C, D, E and F represents the protein-protein interactions of At5g50600, At3g47350, At3g47360, At5g50590, At4g10020 and At5g50690, respectively.

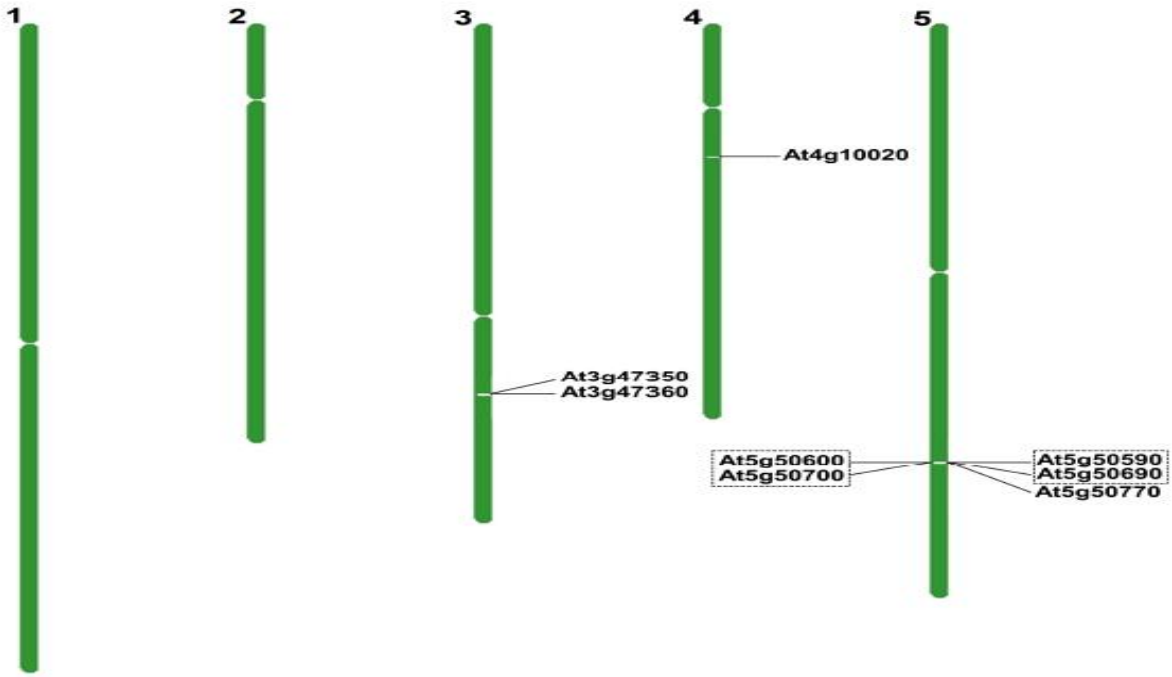

Fig. 5 Chromosomal localization of steroleosin genes in Arabidopsis. The number indicated at the top represents the chromosome number. The segmentally duplicated genes are indicated by dotted box.

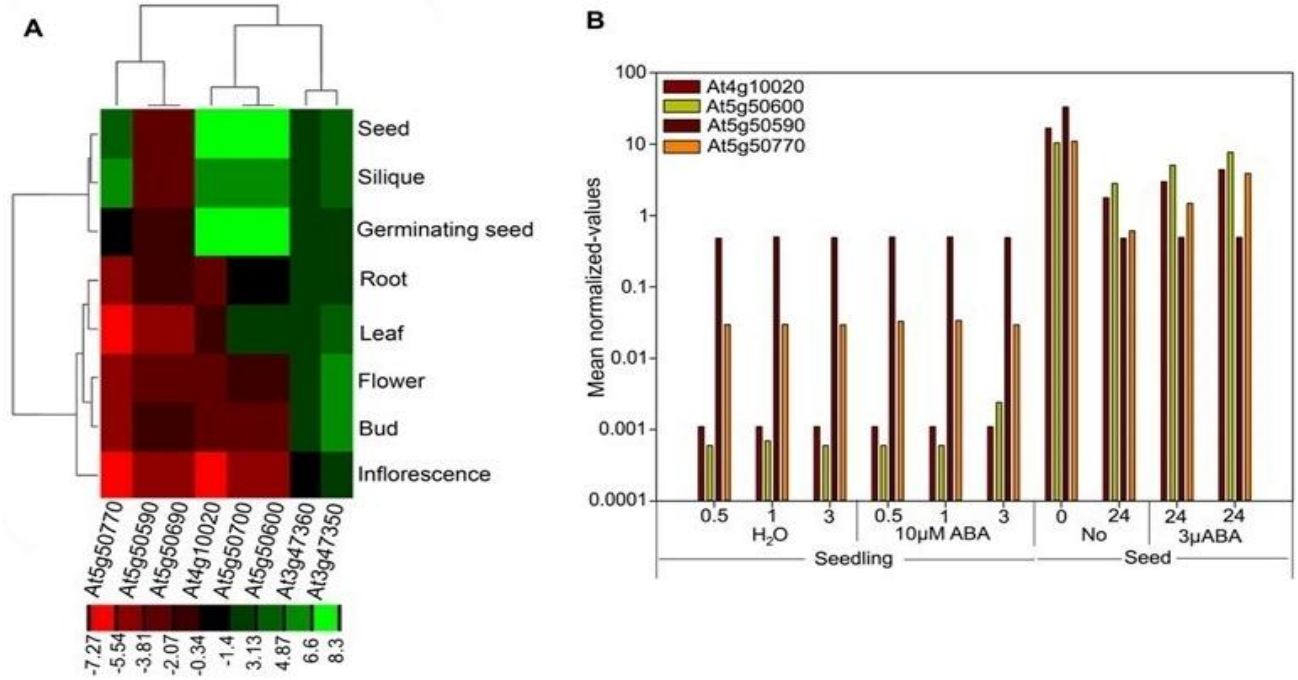

Fig. 6 Expression profiles of Arabidopsis steroleosin genes using RNA-seq data. (A) Expression patterns in different development tissues. The heat map was generated by using log2-transformed-calibrated values. The color scale representing average log signal values is shown at the bottom. (B) Expression patterns under various hormones (ABA, abscisic acid).

Table 1. Classification and characteristics of steroleosin genes in Arabidopsis.

Gene name	TAIR locus ID	Size (ORF, bp)	Size (AA)	pI	Mw
		Steroleosin H			
AtHSD1a	At5g50600	1050	349	5.91	39086.9
AtHSD1b	At5g50700	1050	349	5.91	39086.93
AtHSD5	At4g10020	1170	389	6.41	43441.75
AtHSD6	At5g50770	1029	342	8	37993.42
		Steroleosin L			
AtHSD2	At3g47350	966	321	7.69	35715.64
AtHSD3	At3g47360	930	309	8.66	34947.04
AtHSD4a	At5g50590	900	299	9.07	33247.78
AtHSD4b	At5g50690	900	299	9.07	33247.78

Table 2. Motif sequences and annotation.

Motif no.	Width	Sequence	Annotation
1	35	GKVVLITGASSGIGEQJAYEYAKRGACLALVARRK	Short chain dehydrogenase
2	35	HADVSKPDDCRRFIDETISHFGRLDHLVNNAGIGQ	Enoyl-(Acyl carrier protein) reductase
3	35	MDINFWGSVYTTYFALPYLRKSNGKIVVMASSAAW	NA
4	35	RMSFYNASKAALLNFYETLRIELGGDVGITIVTPG	NA
5	26	CAKAIVNGVCRGQRYVTEPSWYKVTY	NA
6	35	MDLJNDILNLLAPPFTFFGLCFFLPPFYFFKFLQS	NA
7	18	YIESELTQGKYMSEEGELVVNQEERDVQVGAFPVE	NA
8	22	LWKVFCPELJEWGCRLLFLSGG	NA
9	16	RLEVVAEIARZLGSPN	NA
10	35	GTSEDTALNKKILDIPGVRSALYPESIRTPEIKSD	NA

Table 3. Tandem duplicated steroleosin genes in Arabidopsis.

Duplicated Steroleosin 1	Location	Duplicated Steroleosin 2	Location	Identity (%)
At3g47350	17446818-17449703	At3g47360	17451112- 17452843	70.46%
At5g50700	20623024 -20625034	At5g50690	20621330-20622638	42.12%
At5g50600	20589702-20591689	At5g50590	20587988- 20589296	42.12%

Table 4. Segmental duplication in steroleosin genes.

Duplicated Steroleosin 1	Duplicated Steroleosin 2	Flanking genes	Identity (%)
At5g50600	At5g50700	10	100%
At5g50590	At5g50690	10	100%

DISCUSSION

Types of steroleosins in Arabidopsis thaliana: On the Mw basis, eight steroleosin genes were divided into two groups, H-steroleosin and L-steroleosin (Table 1). Similar classification system was previously used for both oleosin and caleosin, respectively (Wu *et al.*, 2010; Shen *et al.*, 2014). The results of phylogenetic tree, multiple sequence alignment and motif analysis also support this classification (Fig.2, Fig.1 and Fig.3A). Furthermore, comparisons were made between Arabidopsis, pine and *B. napus* steroleosins which further suggested that H-steroleosins of Arabidopsis were more closely related to both the reported pine and *B. napus* steroleosins. Similarly, phylogenetic analysis suggests that BnSLO1-1, 2 and 3 were related to *AtHSD1a* (At5g50600) and *AtHSD1b* (At5g50700) whereas BnSLO1-1 and 2 were related to *AtHSD5* (At4g10020), these results were similar to that of previous studies (Jolivet *et al.*, 2009). Unlike oleosin and caleosin, steroleosin N-terminal hydrophobic domain lacks a proline knot, instead it contains the conserved proline residues termed as "proline knob" (Fig.1) which is presumed to have similar functions like proline knot in oleosin and caleosin (Chapman *et al.*, 2012; Tzen, 2012). Similarly, modifications of N-terminal in steroleosin may be related to the function and stability of this protein.

Protein alignment shows the insertion of 10 extra residuals at the C-terminal of H-steroleosin, which is the main difference between two classes of steroleosins (Fig.1). The comparison between the two groups further suggests that the pI values of H-steroleosin were lower to 6.5 except *AtHSD6* (At5g50770) whereas, these values were found higher than 7 in L-steroleosin (Table 1), which suggests that two steroleosin groups may function or regulate under different microenvironments. Moreover, two groups have numerous different conserved phosphorylation sites (Ser, Thr and Tyr) (Fig.1). In the previous studies, structural analysis revealed that the C-terminal of different steroleosin genes varied in length (Lin *et al.*, 2002; Laibach *et al.*, 2015; Huang, 2018). Whether, the two steroleosin groups have diversified functions, is still needed to be elucidated.

Evolution study of steroleosins in Arabidopsis: Pine is relatively old gymnosperm and pine steroleosin is among the H-steroleosin group in phylogenetic tree (Fig.2), suggesting that H-steroleosin group is more primitive than L-steroleosin. Moreover, the phylogenetic analysis shows that *AtHSD6* (At5g50770) is between the two groups of steroleosins. Similarly, molecular weight and pI value are also in the middle of both groups, suggesting that *AtHSD6* (At5g50770) is probably in the evolutionary process from higher to lower molecular weight steroleosins.

Segmental duplication are the DNA segments, which are almost identical (90-100%) and exist on more than one site in the genome (Khaja *et al.*, 2006). Within Arabidopsis, *AtHSD1a* (At5g5g50600) and *AtHSD4a* (At5g50590) were found segments of *AtHSD1b* (At5g50700) and *AtHSD4b* (At5g50690), respectively (Fig.5, Table 4), which is due to the 33-kb duplication events on chromosome 5 (Baud *et al.*, 2009). Moreover, both pairs show 100% sequence similarity at protein levels (Fig. 5, Table 4), suggests that segmental duplications are newly developed in steroleosin genes. However, among the tandemly duplicated gene pairs in steroleosin, there exists variation at the protein identity levels (Table 3). Collectively, these findings suggest that tandem duplications of Arabidopsis steroleosin were much earlier as compared to those of segmental duplication events. The evolutionary processes in the gene family can also be assessed by intron/exon modifications, such as deletion or insertion of intron and exon (Xu *et al.*, 2012). Our results suggest that all the steroleosin in Arabidopsis exhibit conserved introns and exon with an exception that is *AtHSD2* (At3g47350) has one extra exon and intron (Fig. 3B).

Potential functions of steroleosins in *Arabidopsis thaliana*: The protein-protein interaction of steroleosins genes in Arabidopsis suggests that all the steroleosins in Arabidopsis are related to genes involved in Fatty acid biosynthesis (Fig. 4). There are several reports on

steroleosins which suggests that steroleosins have roles in the brassinosteroids signaling and pathways (Du *et al.*, 2019). Our analysis suggested that *AtHSD2* (At3g47350) is interacted with BEN1 which has a key role in brassinosteroid metabolism (Kekez *et al.*, 2018; Zheng *et al.*, 2018). *AtHSD3* (At3g47360) and *AtHSD4b* (At5g50690) were found to be interacted with several genes of the GRP family (Fig. 4C and F), which have important roles related to plant growth, development and response to abiotic stress (Yang *et al.*, 2019).

Promoter sequence analysis showed that RY-repeat elements were found in *AtHSD3* (At3g47360), *AtHSD5* (At4g10020), *AtHSD1a* (At5g50600) and *AtHSD1b* (At5g50700) genes (Text S2). RY element is known to be bound by the transcription factors with B3 domain (Ezcurra *et al.*, 2000; Jing *et al.*, 2019). Numerous transcription factors such as LEC2, FUS3 and ABI3 exhibit B3 domain (Freitas *et al.*, 2019), which are known to regulate lipid metabolism (McGlew *et al.*, 2015). The SKn-1 motif , which is also known to have roles in lipid metabolism(Pang *et al.*, 2014; Dowen, 2019), exists on promoters of most steroleosin genes. Collectively these results suggest that steroleosin genes may play roles in oil droplet accumulation or degradations. The three steroleosin genes *AtHSD5* (At4g10020), *AtHSD1a* (At5g50600) and *AtHSD1b* (At5g50700) are highly expressed in the seeds (Fig.5A), which are the main sites for TAG assembly and are believed to provide energy for germinating seeds. Therefore, the high expression of three steroleosins in seeds further suggests their roles in triacylglycerol metabolism.

Promoter scan 2 analysis also showed MYB element in various steroleosin genes (Text S2). This cis-element is responsible for brassinosteroids signaling pathway. GATA and GT1 motifs are the light-response promoter elements and GATA has also roles in BR-signaling pathways. The GATA element functions together with the G-box or GT1 motifs for their response to light signal which involve different photoreceptors (Luo *et al.*, 2010; Liu *et al.*, 2019). According to the Promoter scan 2 analyses; these two elements are present in all steroleosin genes (Text S2). By the Plant care analysis numerous stress and signal related cis-elements, such as MBS, TC-rich repeats and LTR, were also found in different steroleosin genes (Text S2). These stress signal responsive cis elements have also been reported to be related with BRZ transcription factors in *Zea mays* (Manoli *et al.*, 2018), which collectively suggest that steroleosin genes may play roles to different brassinosteroids.

Previous studies on steroleosins in relation to abscisic acid suggest steroleosins are sensitive to different ABA concentrations. Similarly, ABA and brassinosteroids are antagonistic to each other(Xu *et al.*, 2012; Ohri *et al.*, 2019). From the results of plant care and promoter scan 2 analysis ABRE element, responsible for abscisic acid responsiveness (Finn *et al.*, 2015) was found in H-steroleosins (Table S3). Expression analysis under different hormonal stress also suggests that steroleosins are sensitive to different ABA concentrations (Fig. 5B). Collectively, these results are also in support that steroleosins may have roles in ABA insensitivity and brasinosteroids signaling pathways.

In conclusion, eight steroleosins in Arabidopsis are divided into two groups with different molecular weights, motifs, properties and expression patterns. H-steroleosins are more primitive and larger than L-steroleosins, and mainly express in seeds and siliques, suggesting that they may have roles in storage lipids. L-steroleosins may evolve from H-steroleosins and have diversified their expression patterns, either expressing or not expressing in all tissues, indicating their diversified physiologic roles in plant development. Additionally, the existences of sterol binding domains and cis-elements, related to brassinosteroid signaling pathway, also suggest that steroleosins may play their roles through the regulation of brassinosteroids and the binding of sterols.

Acknowledgements: This work was supported by the National Natural Science Foundation of China (31270295) and the Program of Introducing Talents of Innovative Discipline to Universities (Project 111) from the State Administration of Foreign Experts Affairs (#B18042) “Crop breeding for disease resistance and genetic improvement”.

REFERENCES

Bailey, T.L., N. Williams, C. Misleh and W.W. Li (2006). Meme: Discovering and analyzing DNA and protein sequence motifs. Nucleic Acids research. 34(Web Server issue): W369-W373.

Baud, S., N.R. Dichow, Z. Kelemen, S. d'Andrea, A. To, N. Berger, M. Canonge, J. Kronenberger, D. Viterbo, B. Dubreucq, L. Lepiniec, T. Chardot and M. Miquel (2009). Regulation of hsd1 in seeds of arabidopsis thaliana. Plant and Cell Physiology. 50(8): 1463-1478.

Bhatla, S.C., S. Vandana and V. Kaushik (2009). Recent developments in the localization of oil body-associated signaling molecules during lipolysis in oilseeds. Plant Signaling and Behavior. 4(3): 176-182.

Blom, N., S. Gammeltoft and S. Brunak (1999). Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. J. Molecular Biology. 294(5): 1351-1362.

Chapman, K.D., J.M. Dyer and R.T. Mullen (2012). Biogenesis and functions of lipid droplets in plants: Thematic review series: Lipid droplet

synthesis and metabolism: From yeast to man. J. Lipid Research. 53(2): 215-226.

Chen, E.C., S.S. Tai, C.-C. Peng and J.T. Tzen (1998). Identification of three novel unique proteins in seed oil bodies of sesame. Plant and Cell Physiology. 39(9): 935-941.

Clark, L., D. Leatherby, E. Krilich, A.J. Ropelewski and J. Perozich (2018). In silico analysis of class i adenylate-forming enzymes reveals family and group-specific conservations. PloS One. 13(9): e0203218.

Clouse, S.D. and J.M. Sasse (1998). Brassinosteroids: Essential regulators of plant growth and development. Annual Review of Plant Physiology and Plant Molecular Biology. 49: 427-451.

Dowen, R.H. (2019). Ceh-60/pbx and unc-62/meis coordinate a metabolic switch that supports reproduction in c. Elegans. Developmental Cell. 49(2): 235-250. e237.

Du, C., A. Liu, L. Niu, D. Cao, H. Liu, X. Wu and W. Wang (2019). Proteomic identification of lipid-bodies-associated proteins in maize seeds. Acta Physiologiae Plantarum. 41(5): 70.

Ezcurra, I., P. Wycliffe, L. Nehlin, M. Ellerstrom and L. Rask (2000). Transactivation of the brassica napus napin promoter by abi3 requires interaction of the conserved b2 and b3 domains of abi3 with different cis-elements: B2 mediates activation through an abre, whereas b3 interacts with an ry/g-box. The Plant J. 24(1): 57-66.

Ferrer, A., T. Altabella, M. Arró and A. Boronat (2017). Emerging roles for conjugated sterols in plants. Progress in Lipid Research. 67: 27-37.

Finn, R.D., P. Coggill, R.Y. Eberhardt, S.R. Eddy, J. Mistry, A.L. Mitchell, S.C. Potter, M. Punta, M. Qureshi and A. Sangrador-Vegas (2015). The pfam protein families database: Towards a more sustainable future. Nucleic Acids Research. 44(D1): D279-D285.

Freitas, N.C., H.G. Barreto, L.F. Torres, L.L. Freire, L.A.Z. Rodrigues, L.E.C. Diniz, L.A. Beijo and L.V. Paiva (2019). In silico and in vivo analysis of abi3 and val2 genes during somatic embryogenesis of coffea arabica: Competence acquisition and developmental marker genes. Plant Cell, Tissue and Organ Culture. 137(3): 599-611.

Fujioka, S. and T. Yokota (2003). Biosynthesis and metabolism of brassinosteroids. Annual Review of Plant Biology. 54(1): 137-164.

Gasteiger, E., A. Gattiker, C. Hoogland, I. Ivanyi, R.D. Appel and A. Bairoch (2003). Expasy: The proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Research. 31(13): 3784-3788.

Hu, B., J. Jin, A.Y. Guo, H. Zhang, J. Luo and G. Gao (2014). Gsds 2.0: An upgraded gene feature visualization server. Bioinformatics. 31(8): 1296-1297.

Huala, E., A.W. Dickerman, M. Garcia-Hernandez, D. Weems, L. Reiser, F. LaFond, D. Hanley, D. Kiphart, M. Zhuang, W. Huang, L.A. Mueller, D. Bhattacharyya, D. Bhaya, B.W. Sobral, W. Beavis, D.W. Meinke, C.D. Town, C. Somerville and S.Y. Rhee (2001). The arabidopsis information resource (tair): A comprehensive database and web-based information retrieval, analysis, and visualization system for a model plant. Nucleic Acids Research. 29(1): 102-105.

Huang, A.H.C (1996). Oleosins and oil bodies in seeds and other organs. Plant Physiology. 110(4): 1055-1061.

Huang, A.H.C. (2018). Plant lipid droplets and their associated proteins: Potential for rapid advances. Plant Physiology. 176(3): 1894-1918.

Hunter, S., P. Jones, A. Mitchell, R. Apweiler, T.K. Attwood, A. Bateman, T. Bernard, D. Binns, P. Bork, S. Burge, E. de Castro, P. Coggill, M. Corbett, U. Das, L. Daugherty, L. Duquenne, R.D. Finn, M. Fraser, J. Gough, D. Haft, N. Hulo, D. Kahn, E. Kelly, I. Letunic, D. Lonsdale, R. Lopez, M. Madera, J. Maslen, C. McAnulla, J. McDowall, C. McMenamin, H. Mi, P. Mutowo-Muellenet, N. Mulder, D. Natale, C. Orengo, S. Pesseat, M. Punta, A.F. Quinn, C. Rivoire, A. Sangrador-Vegas, J.D. Selengut, C.J. Sigrist, M. Scheremetjew, J. Tate, M. Thimmajanarthanan, P.D. Thomas, C.H. Wu, C. Yeats and S.Y. Yong (2012). Interpro in 2011: New developments in the family and domain prediction database. Nucleic Acids Research. 40(Database issue): D306-D312.

Jing, Y., Q. Guo and R. Lin (2019). The b3-domain transcription factor val1 regulates the floral transition by repressing flowering locus T. Plant Physiology. 00642.

Jolivet, P., C. Boulard, A. Bellamy, C. Larre, M. Barre, H. Rogniaux, S. d'Andrea, T. Chardot and N. Nesi (2009). Protein composition of oil bodies from mature brassica napus seeds. Proteomics. 9(12): 3268-3284.

Kekez, M., V. Zanki, I. Kekez, D. Matković-Čalogović and J. Rokov-Plavec (2018). Where translation meets plant steroid metabolism: Interaction of seryl-trna synthetase and ben1 protein from *Arabidopsis thaliana*. In FEBS3+ From molecules to living systems.

Khaja, R., J.R. MacDonald, J. Zhang and S.W. Scherer (2006). Methods for identifying and mapping recent segmental and gene duplications in

eukaryotic genomes. Methods in Molecular Biology. 338: 9-20.

Kyte, J. and R.F. Doolittle (1982). A simple method for displaying the hydropathic character of a protein. J. Molecular Biology. 157(1): 105-132.

Laibach, N., J. Post, R.M. Twyman, C.S. Gronover and D. Prüfer (2015). The characteristics and potential applications of structural lipid droplet proteins in plants. J. Biotechnology. 201: 15-27.

Larkin, M.A., G. Blackshields, N.P. Brown, R. Chenna, P.A. McGettigan, H. McWilliam, F. Valentin, I.M. Wallace, A. Wilm, R. Lopez, J.D. Thompson, T.J. Gibson and D.G. Higgins (2007). Clustal w and clustal x version 2.0. Bioinformatics. 23(21): 2947-2948.

Lescot, M., P. Dehais, G. Thijs, K. Marchal, Y. Moreau, Y. Van de Peer, P. Rouze and S. Rombauts (2002). Plantcare, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. Nucleic Acids Research. 30(1): 325-327.

Li, F., T. Asami, X. Wu, E.W. Tsang and A.J. Cutler (2007). A putative hydroxysteroid dehydrogenase involved in regulating plant growth and development. Plant Physiology. 145(1): 87-97.

Lin, L.J., S.S. Tai, C.C. Peng and J.T. Tzen (2002). Steroleosin, a sterol-binding dehydrogenase in seed oil bodies. Plant Physiology. 128(4): 1200-1211.

Liu, Y., B. Patra, S. Pattanaik, Y. Wang and L. Yuan (2019). Gata and phytochrome interacting factor transcription factors regulate light-induced vindoline biosynthesis in Catharanthus roseus. Plant Physiology. 180(3): 1336-1350.

Lu, Y., M. Chi, L. Li, H. Li, M. Noman, Y. Yang, K. Ji, X. Lan, W. Qiang and L. Du, (2018). Genome-wide identification, expression profiling, and functional validation of oleosin gene family in carthamus tinctorius L. Frontiers in Plant Science. 9.

Luo, X.M., W.H. Lin, S. Zhu, J.Y. Zhu, Y. Sun, X.Y. Fan, M. Cheng, Y. Hao, E. Oh, M. Tian, L. Liu, M. Zhang, Q. Xie, K. Chong and Z.Y. Wang (2010). Integration of light- and brassinosteroid-signaling pathways by a gata transcription factor in arabidopsis. Developmental Cell. 19(6): 872-883.

Ma, H. and J. Zhao (2010). Genome-wide identification, classification, and expression analysis of the arabinogalactan protein gene family in rice (oryza sativa l.). J. Experimental Botany. 61(10): 2647-2668.

Maher, C., L. Stein and D. Ware (2006). Evolution of arabidopsis microrna families through duplication events. Genome Research. 16(4): 510-519.

Manoli, A., S. Trevisan, S. Quaggiotti and S. Varotto (2018). Identification and characterization of the bzr transcription factor family and its expression in response to abiotic stresses in zea mays l. Plant Growth Regulation. 84(3): 423-436.

McGlew, K., V. Shaw, M. Zhang, R.J. Kim, W. Yang, B. Shorrosh, M.C. Suh and J. Ohlrogge (2015). An annotated database of arabidopsis mutants of acyl lipid metabolism. Plant Cell Reports. 34(4): 519-532.

Meinke, D.W., J.M. Cherry, C. Dean, S.D. Rounsley and M. Koornneef (1998). Arabidopsis thaliana: A model plant for genome analysis. Science. 282(5389): 662-682.

Murphy, D.J. (2012). The dynamic roles of intracellular lipid droplets: From archaea to mammals. Protoplasma. 249(3): 541-585.

Ohri, P., R. Bhardwaj, R. Kaur, S. Jasrotia, R.D. Parihar, A. Khajuria and N. Sharma (2019). Emerging trends on crosstalk of brs with other phytohormones. In: Brassinosteroids: Plant growth and development. Springer: pp: 425-441.

Oklestkova, J., L. Rárová, M. Kvasnica and M. Strnad (2015). Brassinosteroids: Synthesis and biological activities. Phytochemistry reviews. 14(6): 1053-1072.

Pang, S., D.A. Lynn, J.Y. Lo, J. Paek and S.P. Curran (2014). Skn-1 and nrf2 couples proline catabolism with lipid metabolism during nutrient deprivation. Nature Communications. 5: 5048.

Pasaribu, B., T.Y. Chung, C.S. Chen, P.L. Jiang and J.T.C. Tzen (2016). Identification of steroleosin in oil bodies of pine megagametophytes. Plant Physiology and Biochemistry. 101: 173-181.

Pritchard, S.L., W.L. Charlton, A. Baker and I.A. Graham (2002). Germination and storage reserve mobilization are regulated independently in arabidopsis. The Plant J. 31(5): 639-647.

Shen, Y., J. Xie, R.D. Liu, X.F. Ni, X.H. Wang, Z.X. Li and M. Zhang (2014). Genomic analysis and expression investigation of caleosin gene family in arabidopsis. Biochemical and Biophysical Research Communications. 448(4): 365-371.

Tamura, K., G. Stecher, D. Peterson, A. Filipski and S. Kumar (2013). Mega6: Molecular evolutionary genetics analysis version 6.0. Molecular Biology and Evolution. 30(12): 2725-2729.

Tera, M., T. Koyama, J. Murata, A. Furukawa, S. Mori, T. Azuma, T. Watanabe, K. Hori, A. Okazawa, Y. Kabe, M. Suematsu, H. Satake, E. Ono and M. Horikawa (2019). Identification of a binding protein for sesamin and characterization of its roles in plant growth. Scientific Reports. 9(1): 8631.

Tzen, J.T. (2012). Integral proteins in plant oil bodies. ISRN Botany. 2012.

van der Schoot, C., L.K. Paul, S.B. Paul and P.L. Rinne (2011). Plant lipid bodies and cell-cell signaling: A new role for an old organelle? Plant Signaling and Behavior. 6(11): 1732-1738.

Wu, Y.Y., Y.R. Chou, C.S. Wang, T.H. Tseng, L.J. Chen and J.T. Tzen (2010). Different effects on triacylglycerol packaging to oil bodies in transgenic rice seeds by specifically eliminating one of their two oleosin isoforms. Plant Physiology and Biochemistry. 48(2-3): 81-89.

Xu, G., C. Guo, H. Shan and H. Kong (2012). Divergence of duplicate genes in exon–intron structure. Proceedings of the National Academy of Sciences. 109(4): 1187-1192.

Yang, W., M. Yu, C. Zou, C. Lu, D. Yu, H. Cheng, P. Jiang, X. Feng, Y. Zhang and Q. Wang (2019). Genome-wide comparative analysis of rna-binding glycine-rich protein family genes between *Gossypium arboreum* and *Gossypium raimondii*. PloS One. 14(6): e0218938.

Zheng, L., C. Zhao, J. Mao, C. Song, J. Ma, D. Zhang, M. Han and N. An (2018). Genome-wide identification and expression analysis of brassinosteroid biosynthesis and metabolism genes regulating apple tree shoot and lateral root growth. J. Plant Physiology. 231: 68-85.

Zhou, Y., X. Xia, G. Yu, J. Wang, J. Wu, M. Wang, Y. Yang, K. Shi, Y. Yu and Z. Chen (2015). Brassinosteroids play a critical role in the regulation of pesticide metabolism in crop plants. Scientific Reports. 5: 9018.