DENDROBIOLOGY

2024, vol. 91, 42–55

https://doi.org/10.12657/denbio.091.004

*Xiaoyun Yan, Ru Yi, Wenquan Bao, Sa Rula, Yu-e Bai, Junxing Chen, Yue Lin, Haicuang Huang, Daorentuya Han, Yanhong He**

Identification and analysis of the SAURs in *Prunus sibirica*

Received: 12 July 2023; Accepted: 7 February 2024

Abstract: In the present study, we identified the members of the small auxin-up RNA (*SAUR*) gene family in *Prunus sibirica* and analyzed their localization, phylogeny, duplication, *cis*-elements in the promoter, and expression patterns. In total, 57 *PsSAURs* were identified, which were randomly distributed along eight chromosomes. Among them, eight and ten pairs of segmentally duplicated genes and tandem duplication genes were found in 13 and 21 *PsSAURs*, respectively. Phylogenetic analysis indicated that the *PsSAURs* were divided into five groups (Group A–E). Light-responsive, methyl jasmonate, abscisic acid, salicylic acid, and gibberellin, as well as low-temperature responsiveness, and defense and stress responsiveness were identified by analyzing *PsSAURs* promoter sequences. The collinearity analysis of *P. sibirica SAURs* and *Prunus mume* and *Prunus persica SAURs* family genes detected 35 and 59 pairs of gene pairs, respectively, and the Ks values of all collinearity gene pairs were almost less than 1. Expression pattern analysis showed that *PsSAURs* had different tissue and stage expression patterns. However, research on *SAURs* in non-model plants remains limited, and studies on the expression and function of *SAURs* are lacking. This study provides a foundation for further investigations into the functional analyses of *SAURs* in *P. sibirica*.

Keywords: gene expression, gene family, auxin, bioinformatics

Addresses: X. Yan, R. Yi, W. Bao, R. Sa, Y. Bai, J. Chen, Y. Lin, D. Han, Y. He, College of Forestry, Inner Mongolia Agricultural University, Hohhot, 010018, e-mail: hyh20012008@imau.edu.cn
H. Huang, Inner Mongolia Academy of Forestry, Hohhot, China, 010010
* corresponding author

Introduction

Auxin plays an important role in plant growth, development, and stress response. Early auxin-induced genes are divided into three major categories: Auxin/indole-3-acetic acid (*Aux/IAA*), Gretchen Hagen 3 (*GH3*), and small auxin upregulated RNA (*SAUR*) (Hagen & Guilfoyle, 2002). *SAUR* was first identified in the hypocotyl elongation zone of soybean (*Glycine max* L.) (McClure & Guilfoyle, 1987). Subsequently, *SAURs* were identified in *Arabidopsis* (Hagen & Guilfoyle, 2002), corn (*Zea mays* L.) (Chen et al., 2014), wheat (*Triticum aestivum* L.) (Liu et al., 2022), apple (*Malus pumila* Mill.) (Wang, Lu et al., 2020), grape (*Vitis vinifera* L.) (Li et al., 2018), and citrus (*Citrus reticulata* Blanco) (Wang, Lu et al., 2020). *SAURs* are major response genes in the auxin pathway that can respond to the auxin family of genes within 2–5 minutes (Wang, Yu, et al., 2020). Currently, functional studies on *SAUR* in plant growth and development

have been reported. For example, *AtSAUR63* promotes the elongation of cells in *Arabidopsis* (Franco et al., 1990), *AtSAUR19–24* plays a major regulatory role in hypocotyl elongation and increases the leaf size in *Arabidopsis* (Spartz et al., 2012), and *AtSAUR36* is a positive regulator that mediates auxin-induced leaf senescence (Hou et al., 2013). The *SAURs* not only participate in the regulation of plant growth and development but also play an important role in the response to environmental stress. For example, *AtSAUR41* plays an important role in abiotic stress resistance, which could increase the resistance to salt stress in transgenic plants (Qiu et al., 2020).

Prunus sibirica is an ecological and economic tree species in Asia, which is broadly distributed in northern China, eastern Mongolia, eastern Siberia and the coastal areas of Russia (Niu et al., 2015). *Prunus sibirica* kernel is rich in nutrition, fat and protein, and it is widely used in industry, food, cosmetics, medicine and other aspects. It can be fried and salted, and can also be used to prepare a variety of food and beverages (Wang et al., 2014). In addition, *P. sibirica* is drought- and cold-resistant and can adapt to various harsh environmental conditions. It is a pioneer tree species for afforestation in arid and semi-arid mountainous areas (Wang et al., 2019). Therefore, *P. sibirica* has a very broad development and utilization potential.

Previous studies on *P. sibirica* mainly focused on the physiological characteristics of kernels; however, research on the *SAUR* gene family is lacking thus far (Wang et al., 2022). The completion of genome sequencing for *P. sibirica* provides valuable data for systematic analysis of the *PsSAUR* family (Huang et al., 2023). This study aims to use bioinformatics to identify *PsSAUR* family members and conduct a systematic analysis at the whole-genome level. The results of the study will provide a reference for studying the function of *PsSAUR* and provide technical support for the subsequent improvement of *P. sibirica* varieties and germplasm resources.

Materials and methods

Identification of PsSAURs: To identify *SAURs* in the *P. sibirica* genome, 78 *SAURs* in *Arabidopsis* (Phytozome 12 database) were used as reference sequences to search against the *P. sibirica* genome (https://www.rosaceae.org, tfGDR1049) by using BLAST (a minimum amino acid identity of ≥ 50, E-value < 1e-5). The *SAURs* structural domain sequence PF02519 was downloaded from the Pfam database (http://pfam.xfam.org/), and the Hidden Markov Model was used to screen candidate sequences containing the SAURs structural domain. In addition, Conserved Domain Search (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) and the SMART website (http://smart.embl-heidelberg.de/) were used to confirm the presence of conserved structural domains. ExPASy ProtParam (https://web.expasy.org/protparam/) was used to predict the physical and chemical parameters (Gasteiger et al., 2005), including the number of aa, Mw, PIs, grand average of hydropathicity, aliphatic index, and instability index. CELLO (http://cello.life.nctu.edu.tw/) was used to predict the subcellular location of the *SAUR*.

PsSAUR phylogenetic tree construction: A phylogenetic tree was established using the identified full-length protein sequences of SAURs from *P. sibirica* and *Arabidopsis thaliana*. The sequence alignment of 135 SAURs was established using Clustal X (Larkin et al., 2007). The phylogenetic tree was constructed in MEGA X (Kumar et al., 2018) based on the maximum likelihood method, the JTT+F model, and bootstrap values were calculated for 1000 replicates. Evolview was used to display the phylogenetic tree (Subramanian et al., 2019).

Structure, conserved motifs, and gene duplication analysis of PsSAURs: GSDS 2.0 (http://gsds.cbi.pku.edu.cn/index.php) (Hu et al., 2015) was used to analyze the exon and intron composition of *PsSAURs*. MEME (http://meme-suite.org/tools/meme) (Bailey & Elkan, 1994) was used to predict the conserved domains of *PsSAURs* with five different sequence motifs and a sequence motif width of 6–50 aa. Tbtools (Chen et al., 2020) was used to perform collinearity and gene duplication analysis of *PsSAURs*.

Gene duplication and evolutionary analysis of PsSAURs: The chromosomal location of *PsSAURs* was obtained from *P. sibirica* genome annotation files. BLASTP was used to search collinearity for each *PsSAUR* (e-value < $1e^{-5}$, the top 5 matches). Then, the replication events were examined using MCScanX in TBtools (Chen et al., 2020).

Analysis of cis-acting elements in PsSAUR promoters: To analyze the *cis*-acting elements in promoter sequences, the sequences 2,000 bp upstream of the coding regions of the identified *PsSAURs* were extracted. PlantCARE (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) (Lescot et al., 2002) database was used for predicting the *cis*-acting elements in the promoters, and the results were visualized using TBtools (Chen et al., 2020).

Analysis of synteny and Ka/Ks ratios: The chromosome location of *PsSAUR* genes were obtained by referring to the genome annotation file and visualized in Chromosome-Basic Circos by TBtools (Chen et al., 2020). To evaluate the synteny and gene duplication events of *SAUR* genes in *P. sibirica* and its related species (*Prunus mume* and *Prunus persica*), the genome and genome annotation files of two species were downloaded from the GDR database (https://www.rosaceae.org/). Synteny analysis was performed

using One Step MCScanX in TBtools, and the results were visualized using Circos and Dual Synteny Plot (Chen et al., 2020). The coding sequences (CDSs) and protein sequences of the gene pairs were compared, with Ka/Ks values = 1, >1 and <1 representing neutral, positive and negative selection, respectively, and non-synonymous substitution rates (Ka), synonymous substitution rates (Ks) and Ka/Ks ratios were calculated in TBtools (Chen et al., 2020).

Expression Analysis of PsSAURs: RNA-seq data, including pistil, stamen, petal, sepal, and the kernels (S1-S6) at six development stages of *P. sibirica*. At least two biological replicates for each sample were selected. FPKM (fragments per kilobase of exon per million fragments mapped) of each gene was calculated to present the expression level of *PsSAURs*. The expression patterns were based on the transformed data of log2 values and min-max normalization by Heat map in TBtools (Chen et al., 2020).

Results

Whole-genome identification and characteristics of PsSAURs: A total of 57 *PsSAURs* were identified in the *P. sibirica* genome, which were randomly distributed on eight chromosomes (Fig. 1, Table S1). Among them, the highest number (24) of *PsSAURs* was on chromosome 8, with several genes forming clusters, and the other seven chromosomes each containing 1–10 *PsSAURs*. *PsSAURs* were named *PsSAUR1-PsSAUR57* according to their position on the chromosome. The length *PsSAURs* were ranged from 70 to 236 amino acids (aa), the molecular weight (Mw) and putative isoelectric points (PI) ranging from 8.05 (*PsSAUR53*) to 27.03 (*PsSAUR17*) kDa and 5.11 (*PsSAUR33*) to 10.46 (*PsSAUR19*), respectively (Table 1). A total of forty-nine and 8 *PsSAURs* were hydrophilic and hydrophobic proteins. Protein instability index analysis showed that 17 *PsSAURs* were stable with an instability index of less than 40. In addition, the protein subcellular localization prediction showed that *PsSAURs* were primarily located in the cytoplasm, nucleus, and mitochondria.

Phylogenetics of PsSAURs: To study the phylogenetic relationships among *PsSAURs*, 78 AtSAUR protein sequences and 57 PsSAUR protein sequences were used to construct a phylogenetic tree (Fig. 2). The *PsSAURs* were divided into five groups (Groups A-E). Among them, Group E had the largest number of *PsSAURs,* which contained 24 *PsSAUR* members, followed by Group A, which contained 14 *PsSAUR* members. Groups B, C, and D consisted of 9, 7, and 3 *PsSAUR* members, respectively. Most of *PsSAURs* with similar locations were distributed in the same group as follows: *PsSAUR24*, *PsSAUR25*, *PsSAUR26*, *PsSAUR27*, *PsSAUR28*, and *PsSAUR29* from Group A were located on chromosome 7; *PsSAUR6*, *PsSAUR7*, *PsSAUR8*, *PsSAUR9*, and *PsSAUR11* from Group B were located on chromosome 2; *PsSAUR34*, *PsSAUR35*, *PsSAUR36*, *PsSAUR37*, *PsSAUR38*, *PsSAUR39*, *PsSAUR40*, *PsSAUR41*, *PsSAUR42*, *PsSAUR43*, *PsSAUR44*, *PsSAUR45*, *PsSAUR46*, *PsSAUR47*, *PsSAUR48*, *PsSAUR51*, *PsSAUR52*, *PsSAUR53*, *PsSAUR54*, *PsSAUR55*, and *PsSAUR56* from Group E were located on chromosome 8.

Gene structure and conserved motifs in PsSAURs: The exon-intron structure of *PsSAURs* and their conserved motifs were analyzed to understand their sequence characteristics (Fig. 3). In total, 73.68% (42/57) of *PsSAURs* contained no introns, ten *PsSAURs* contained one intron, only four, including *PsSAUR13*, *PsSAUR15*, *PsSAUR17*, and *PsSAUR23*, contained two introns, and *PsSAUR49*

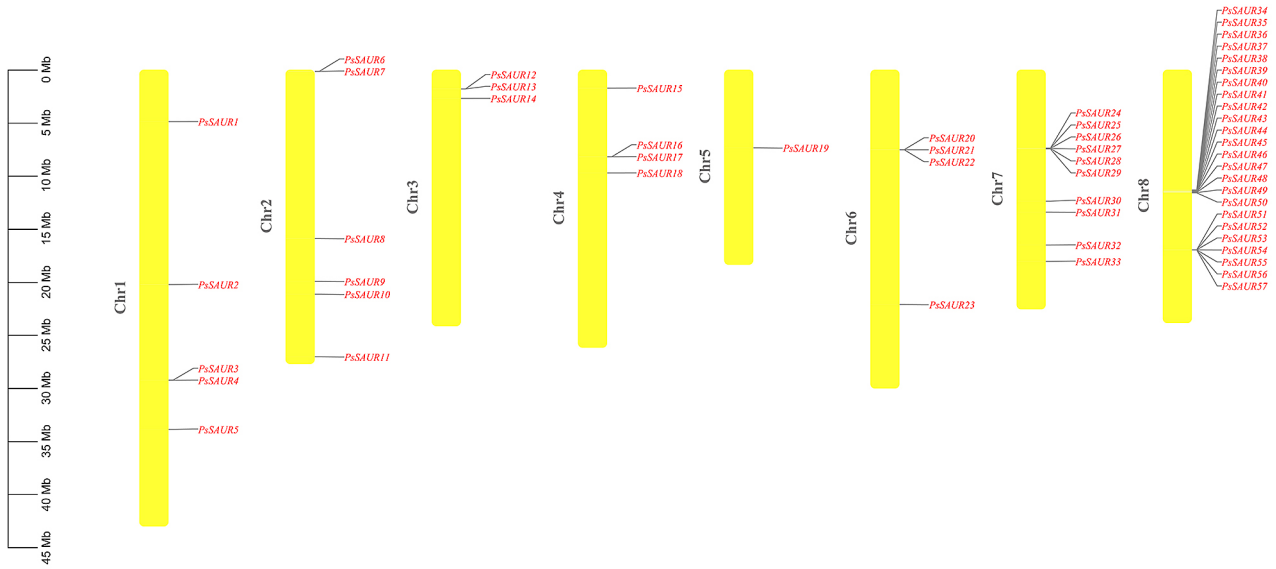

Fig. 1. Chromosome distribution of *PsSAURs* in *P. sibirica*

Table 1. Information related to PsSAURs in *P. sibirica*

Gene name	Accession	Chromosome location	Strand	Number of amino acids/aa	Molecular weight (kD)	Theoretical pI	Instability index	Aliphatic index	Grand average of hydropathicity (GRAVY)	Predicted subcellular localization
PsSAUR1	PaF106G0100000780.01	Chr1: 4851261-4851947	+	166	18.68785	9.9	59.13	90.42	−0.005	PlasmaMembrane/Mitochondrial
PsSAUR2	PaF106G0100002700.01	Chr1: 20203785-20204276	+	163	18.46055	9.37	56.43	69.26	−0.452	Nuclear
PsSAUR3	PaF106G0100004332.01	Chr1: 29206302-29206619	+	105	12.08995	7.83	61.56	87.24	−0.355	Mitochondrial
PsSAUR4	PaF106G0100004336.01	Chr1: 29228506-29229489	-	150	17.0335	9.79	62.65	82.53	−0.28	Mitochondrial/Nuclear
PsSAUR5	PaF106G0100005296.01	Chr1: 33845860-33846426	+	188	21.21526	8.74	56.44	74.15	−0.414	Nuclear
PsSAUR6	PaF106G0200006771.01	Chr2: 110636-111793	+	133	14.77948	5.17	44.96	57.22	−0.714	Nuclear
PsSAUR7	PaF106G0200006772.01	Chr2: 116535-116960	+	141	15.90597	6.9	40.75	73.97	−0.671	Nuclear
PsSAUR8	PaF106G0200008604.01	Chr2: 15878329-15879323	+	172	19.42461	10.45	29.56	71.4	−0.49	Mitochondrial
PsSAUR9	PaF106G0200009291.01	Chr2: 19915629-19916397	+	124	14.43145	7.89	27.07	75.32	−0.673	Nuclear
PsSAUR10	PaF106G0200009494.01	Chr2: 21116519-21116899	+	126	14.75434	8.75	34.76	88.17	−0.107	Extracellular
PsSAUR11	PaF106G0200010737.01	Chr2: 27010514-27011408	+	132	15.73591	7.12	29.14	70.76	−0.742	Nuclear
PsSAUR12	PaF106G0300014330.01	Chr3: 1769202-1769762	+	186	21.06221	9.22	40.42	78.06	−0.409	Nuclear
PsSAUR13	PaF106G0300014329.01	Chr3: 1770375-1771793	-	177	20.16526	6.79	60.88	86.5	−0.288	Cytoplasmic/Mitochondrial/Nuclear
PsSAUR14	PaF106G0300014203.01	Chr3: 2666558-2667170	+	117	13.1601	6.06	37.09	96.5	−0.099	Mitochondrial/chloroplast/Nuclear
PsSAUR15	PaF106G0400018086.01	Chr4: 1700338-1701561	+	117	12.85694	5.7	55.59	90	0.104	Plasma Membrane/chloroplast
PsSAUR16	PaF106G0400016888.01	Chr4: 8162844-8163143	+	99	11.12376	6.02	42.29	90.51	−0.153	Extracellular/Mitochondrial/Nuclear
PsSAUR17	PaF106G0400016886.01	Chr4: 8164068-8166793	-	236	27.02566	7.65	40.32	75.59	−0.398	Nuclear
PsSAUR18	PaF106G0400016640.01	Chr4: 9687821-9688240	-	139	15.99332	9.38	80.52	95.25	−0.408	Mitochondrial/Nuclear
PsSAUR19	PaF106G0500019384.01	Chr5: 7321462-7322312	+	166	19.46583	10.46	33.43	77.47	−0.627	Mitochondrial

Gene name	Accession	Chromosome location	Strand	Number of amino acids/aa	Molecular weight (kD)	Theoretical pI	Instability index	Aliphatic index	Grand average of hydropathicity (GRAVY)	Predicted subcellular localization
PsSAUR20	PaF106G0600022654.01	Chr6: 7496390-7497064	+	128	14.52475	6.9	43.44	88.28	−0.501	Nuclear
PsSAUR21	PaF106G0600022655.01	Chr6: 7500637-7501071	+	144	15.97596	5.65	51.05	68.4	−0.583	Nuclear
PsSAUR22	PaF106G0600022656.01	Chr6: 7505887-7506315	+	142	16.00309	6.9	41.36	73.45	−0.677	Nuclear
PsSAUR23	PaF106G0600024333.01	Chr6: 22092133-22092647	+	170	19.37449	8.76	54.84	85.35	−0.106	Extracellular
PsSAUR24	PaF106G0700028558.01	Chr7: 7344400-7344855	-	151	17.02373	9.55	63.49	92.98	−0.142	Plasma Membrane
PsSAUR25	PaF106G0700028555.01	Chr7: 7361195-7361650	-	151	17.03777	9.55	69.5	90.4	−0.165	Plasma Membrane/Nuclear
PsSAUR26	PaF106G0700028553.01	Chr7: 7377830-7378284	-	135	15.23753	7.81	73.42	87.41	−0.11	Extracellular/Plasma Membrane
PsSAUR27	PaF106G0700028551.01	Chr7: 7383579-7383929	-	116	12.98013	5.7	51.53	99.14	0.241	Extracellular/Plasma Membrane
PsSAUR28	PaF106G0700028552.01	Chr7: 7383579-7384040	-	153	17.24314	9.58	52.95	91.76	−0.092	Extracellular/Plasma Membrane/Nuclear
PsSAUR29	PaF106G0700028547.01	Chr7: 7429270-7429728	-	152	17.14089	9.3	59.34	92.43	−0.148	Plasma Membrane/Nuclear
PsSAUR30	PaF106G0700027893.01	Chr7: 12353407-12353796	-	129	14.30748	6.83	39.07	76.28	−0.188	Chloroplast
PsSAUR31	PaF106G0700027703.01	Chr7: 13383893-13384288	-	125	14.04426	8.78	53.76	87.36	−0.3	Nuclear
PsSAUR32	PaF106G0700027134.01	Chr7: 16479544-16480270	-	106	12.08287	8.6	53.09	77.26	−0.446	Mitochondrial
PsSAUR33	PaF106G0700026812.01	Chr7: 18024373-18025218	+	125	14.02486	5.11	59.04	82.56	−0.371	Mitochondrial/Nuclear
PsSAUR34	PaF106G0800031625.01	Chr8: 11282676-11284223	+	158	17.77833	8.61	50.16	78.29	−0.266	Extracellular/Nuclear
PsSAUR35	PaF106G0800031622.01	Chr8: 11300070-11300537	+	155	17.7714	6.88	45.6	84.26	0.059	Plasma Membrane
PsSAUR36	PaF106G0800031620.01	Chr8: 11310028-11310553	+	109	12.31222	6.07	36.93	89.36	−0.143	Extracellular/Mitochondrial
PsSAUR37	PaF106G0800031618.01	Chr8: 11335417-11335943	+	145	16.30579	8.47	36.42	87.38	−0.254	Extracellular/Mitochondrial
PsSAUR38	PaF106G0800031616.01	Chr8: 11349134-11349529	+	97	10.62313	6.71	36.13	76.29	−0.228	chloroplast/Nuclear

Gene name	Accession	Chromosome location	Strand	Number of amino acids/aa	Molecular weight (kD)	Theoretical pI	Instability index	Aliphatic index	Grand average of hydropathicity (GRAVY)	Predicted subcellular localization
PsSAUR39	PaF106G0800031609.01	Chr8: 11404246-11406526	-	120	13.21033	6.71	39.09	102.33	0.235	Plasma Membrane
PsSAUR40	PaF106G0800031605.01	Chr8: 11432471-11433969	+	158	17.76854	9.2	31.21	83.86	−0.046	Plasma Membrane
PsSAUR41	PaF106G0800031604.01	Chr8: 11436254-11436994	+	91	10.21372	6.55	39.31	85.71	−0.202	Mitochondrial
PsSAUR42	PaF106G0800031608.01	Chr8: 11404246-11438414	-	186	20.61977	6.42	31.52	90.65	−0.068	Cytoplasmic
PsSAUR43	PaF106G0800031599.01	Chr8: 11491197-11498191	-	186	20.49061	6.9	32.59	88.55	−0.062	Cytoplasmic/Mitochondrial
PsSAUR44	PaF106G0800031596.01	Chr8: 11506866-11507486	-	156	17.59013	5.94	49.6	90.58	−0.315	Mitochondrial/Nuclear
PsSAUR45	PaF106G0800031595.01	Chr8: 11508605-11508910	+	101	11.33027	8.71	40.39	94.65	0.051	Plasma Membrane/Mitochondrial/Chloroplast
PsSAUR46	PaF106G0800031594.01	Chr8: 11513969-11514385	+	101	11.25811	8.93	65.03	90.79	−0.023	Mitochondrial/Chloroplast
PsSAUR47	PaF106G0800031591.01	Chr8: 11521494-11522770	+	101	11.31822	9.36	64.39	88.91	−0.027	Mitochondrial/Chloroplast
PsSAUR48	PaF106G0800031590.01	Chr8: 11524532-11524816	+	94	10.64937	8.85	53.16	99.47	0.055	Mitochondrial/Chloroplast
PsSAUR49	PaF106G0800031588.01	Chr8: 11542778-11543299	+	171	19.52165	9.64	58.7	77.49	−0.587	Nuclear
PsSAUR50	PaF106G0800031587.01	Chr8: 11565588-11566287	-	150	16.86448	9.21	66.27	93.6	−0.157	Nuclear
PsSAUR51	PaF106G0800030701.01	Chr8: 16931768-16932043	-	91	10.14858	8.66	58.21	76.15	−0.297	Mitochondrial/Nuclear
PsSAUR52	PaF106G0800030700.01	Chr8: 16933274-16933676	+	94	10.49326	8.54	64.09	88.09	0.121	Plasma Membrane/Mitochondrial
PsSAUR53	PaF106G0800030699.01	Chr8: 16934478-16934909	-	70	8.05333	7.99	50.22	83.57	−0.056	Extracellular/Mitochondrial
PsSAUR54	PaF106G0800030698.01	Chr8: 16936378-16936848	+	94	10.59748	8.52	61.73	89.15	0.122	Extracellular/Mitochondrial
PsSAUR55	PaF106G0800030697.01	Chr8: 16938188-16938499	+	103	11.42508	7.79	16.37	81.46	−0.237	Mitochondrial
PsSAUR56	PaF106G0800030696.01	Chr8: 16940616-16940909	+	97	11.22181	7.06	38.9	75.36	−0.497	Cytoplasmic/Mitochondrial/Nuclear
PsSAUR57	PaF106G0800030695.01	Chr8: 16942827-16943126	+	99	11.17022	9.87	45.24	83.54	−0.242	Mitochondrial

contained three introns. MEME was used to analyze the conserved motifs of 57 PsSAUR proteins. A total of five conserved motifs were obtained (Fig. 3), including motif1, motif2, motif3, motif4, and motif5, at lengths of 32, 21, 11, 15, and 21 aa, respectively. Sixteen *PsSAURs* contained motif1, motif2, motif3, and motif4; 14 contained motif1, motif2, motif3, and motif5; 19 contained motif1, motif2, and motif3. However, two *PsSAURs* (*PsSAUR36*, *PsSAUR37*) contained motif1, motif2 and motif4; two (*PsSAUR28*, *PsSAUR29*) contained motif1, motif3, and motif5; *PsSAUR33* contained motif2, motif3, and motif5. Generally, some *PsSAURs* in the same group had similar conserved structures. For example, *PsSAURs* in Group D contained motif1, motif2, motif3, and motif5, and most *PsSAURs* in Group E contained motif1, motif2, motif3, and motif4.

Gene duplication in PsSAURs: The gene duplication of *PsSAURs* was analyzed, and even segmental duplications and 10 tandem duplications of *PsSAURs*

Table 2. Gene duplication of PsSAUR in *P. sibirica*

Type	Chromosome	Gene ID
Segmental duplication	Chr1&Chr7	*PsSAUR3/PsSAUR32*
	Chr1&Chr8	*PsSAUR3/PsSAUR46*
	Chr1&Chr8	*PsSAUR4/PsSAUR50*
	Chr2&Chr6	*PsSAUR6/PsSAUR20*
	Chr2&Chr2	*PsSAUR9/PsSAUR11*
	Chr2&Chr6	*PsSAUR10/PsSAUR23*
	Chr3&Chr7	*PsSAUR14/PsSAUR33*
Tandem duplication	Chr1	*PsSAUR6/PsSAUR7*
	Chr3	*PsSAUR12/PsSAUR13*
	Chr4	*PsSAUR16/PsSAUR17*
	Chr6	*PsSAUR20/PsSAUR21/ PsSAUR22*
	Chr7	*PsSAUR27/PsSAUR28*
	Chr8	*PsSAUR39/PsSAUR42*
	Chr8	*PsSAUR40/PsSAUR41*
	Chr8	*PsSAUR45/PsSAUR46*
	Chr8	*PsSAUR47/PsSAUR48*
	Chr8	*PsSAUR55/PsSAUR56*

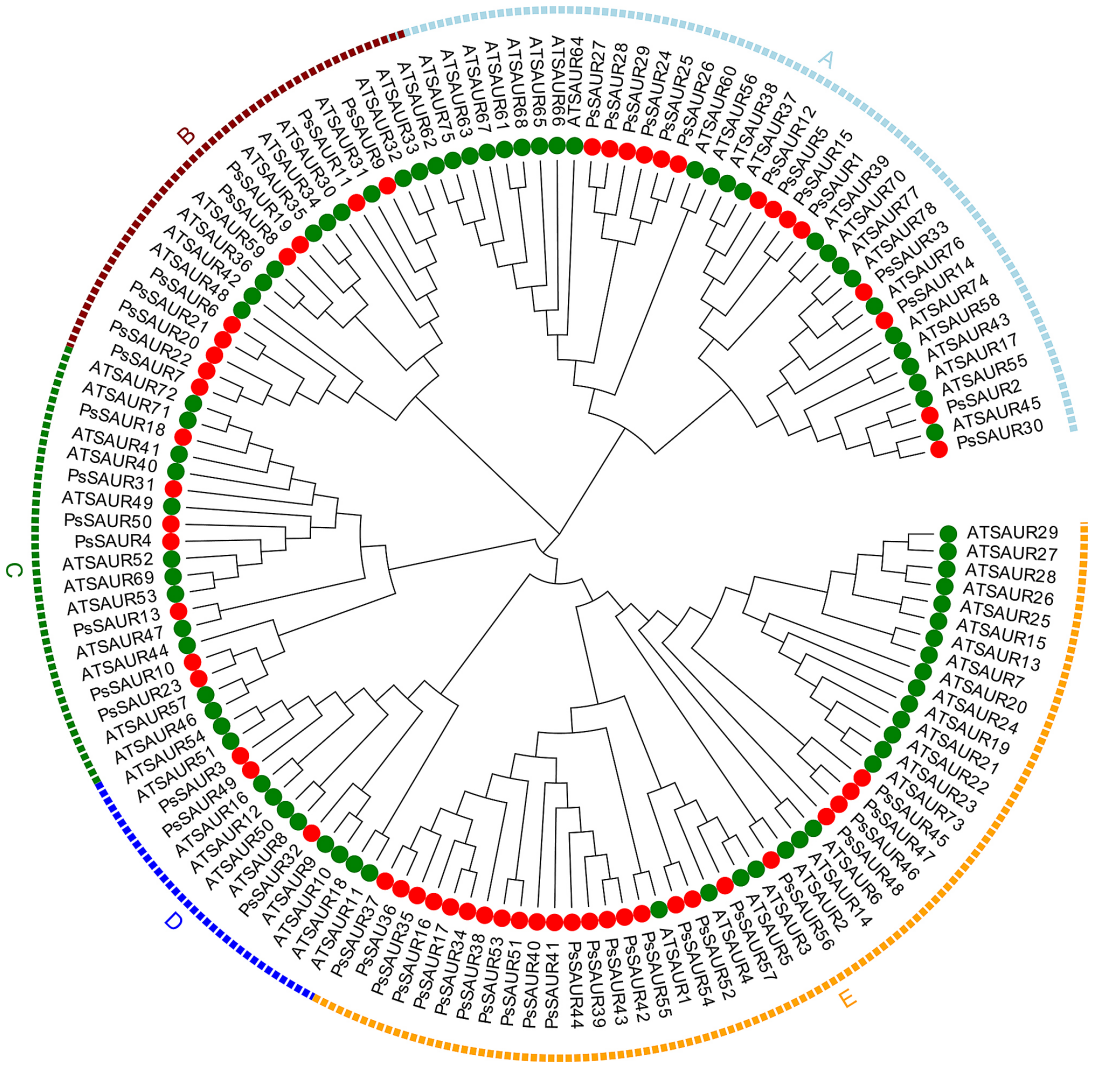

Fig. 2. Phylogenetic tree of SAUR proteins from *P. sibirica* (Ps) and *Arabidopsis* (At) and Arabidopsis. AtSAUR protein and PsSAUR protein were represented by red and green circles, respectively. Group A-E are represented by light blue, red, green, blue and yellow, respectively

were obtained (Table 2). These included 21 *PsSAURs* on chromosomes 1, 3, 4, 6, 7, and 8. Taken together with the phylogenetic relationship of *PsSAURs*, the result showed that most of the gene pairs for segmental duplication (6/7) and tandem duplication (9/10) were located in the same group, including the segmental duplication *PsSAUR3/PsSAUR46* and the tandem duplication *PsSAUR12/PsSAUR13*.

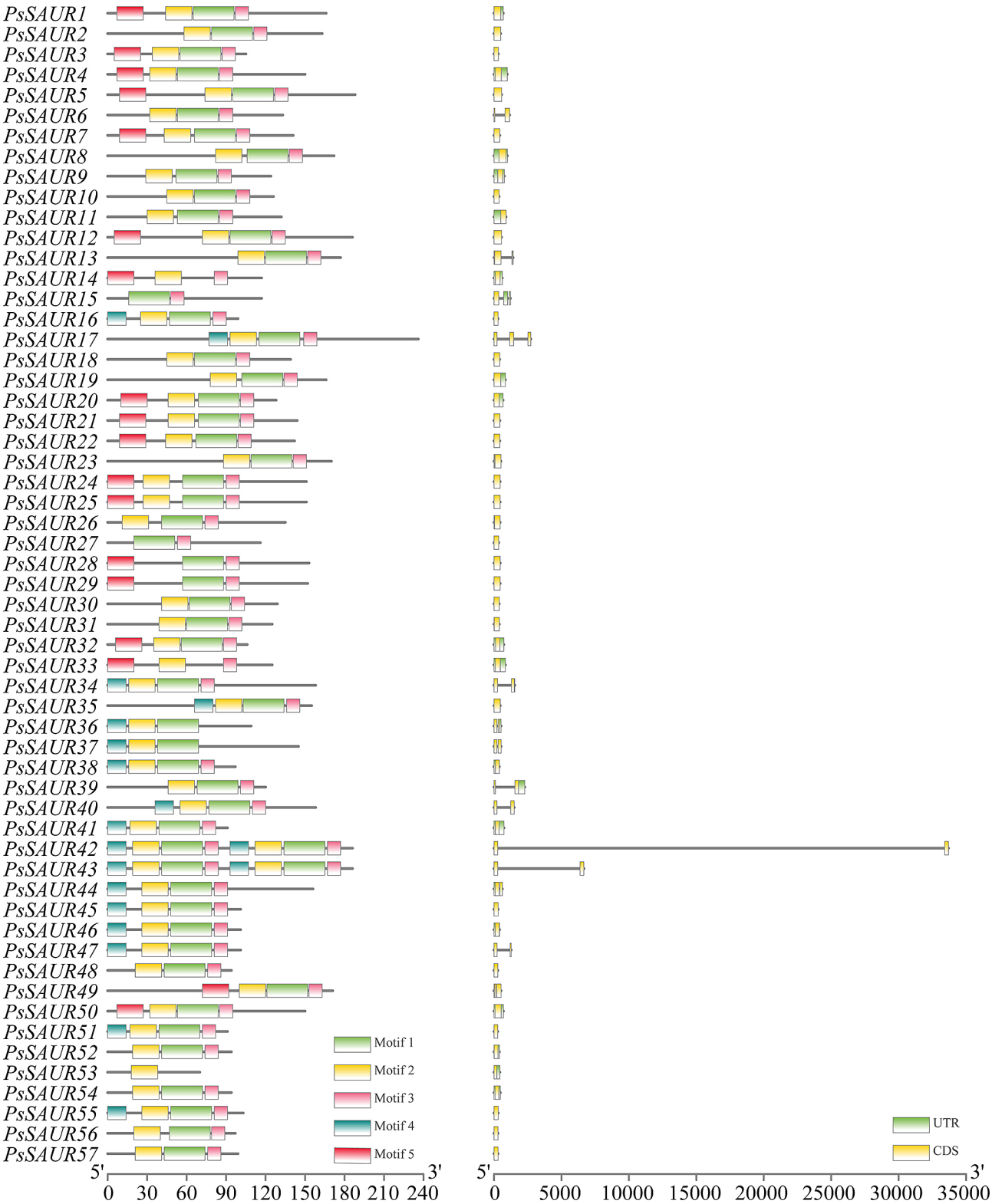

Fig. 3. Architecture of conserved motifs and gene structure of *PsSAUR*. Different numbers 1-5 in the five conserved motifs of PsSAUR proteins are represented by other colour boxes. In the structure of the *PsSAUR* gene, the yellow box represents the coding region, and the green box represents the untranslated region

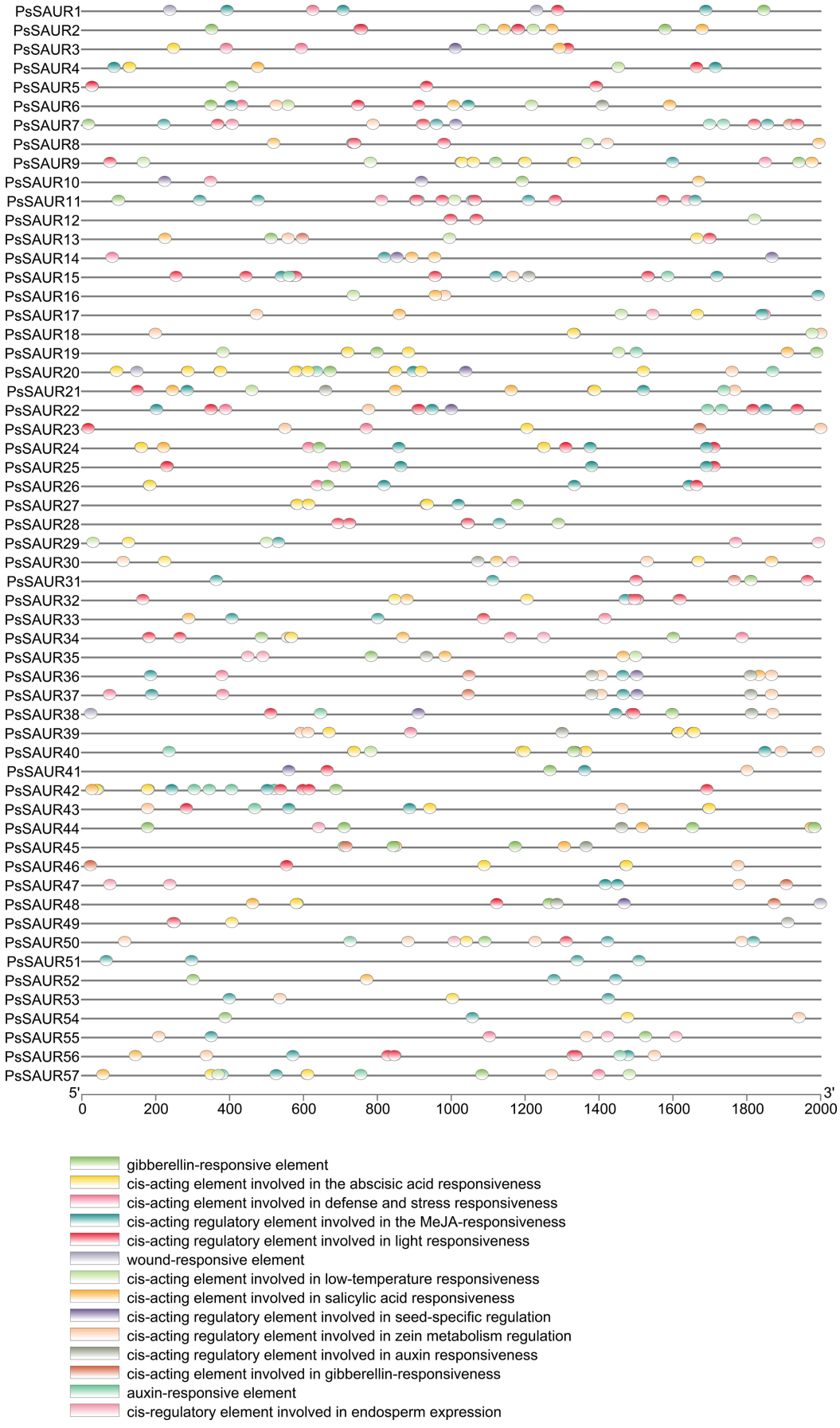

Fig. 4. Distribution of major *cis*-elements on the 2000 bp region of promoter of *PsSAURs*

Promoter analysis of PsSAURs: The *cis*-acting elements of promoter regions (from the start codon located 2000 bp upstream) were analyzed (Fig. 4, Table S2). The results revealed that the promoters of the 15 *PsSAURs* contained the auxin signalling transduction related *cis*-elements, including *PsSAUR6, PsSAUR15, PsSAUR21, PsSAUR30, PsSAUR35, PsSAUR36, PsSAUR36, PsSAUR37, PsSAUR37, PsSAUR38, PsSAUR39, PsSAUR44, PsSAUR45, PsSAUR48,* and *PsSAUR49*. *cis*-responsive elements were also involved in defense and stress response, low-temperature response, drought induction, hormonal response, and others, and primarily included 160 methyl jasmonate (MeJA), 15 auxin, 139 abscisic acid (ABA), 38 salicylic acid, and 39 gibberellin (GA) responsive elements. In addition, PsSAUR promoters also contained 24 low-temperature responsiveness elements, 26 defense and stress responsiveness elements, and 12 endosperm expression *cis*-responsive elements.

Collinearity analysis in PsSAURs: The synteny relationship between SAURs family genes in *P. mume* and the closely related *P. mume* and *P. persica* were analyzed. We detected 35 and 59 gene pairs (Fig. 5a, Table S3). The average Ks values between *P. sibirica* and the closely related *P. mume* and *P. persica* were

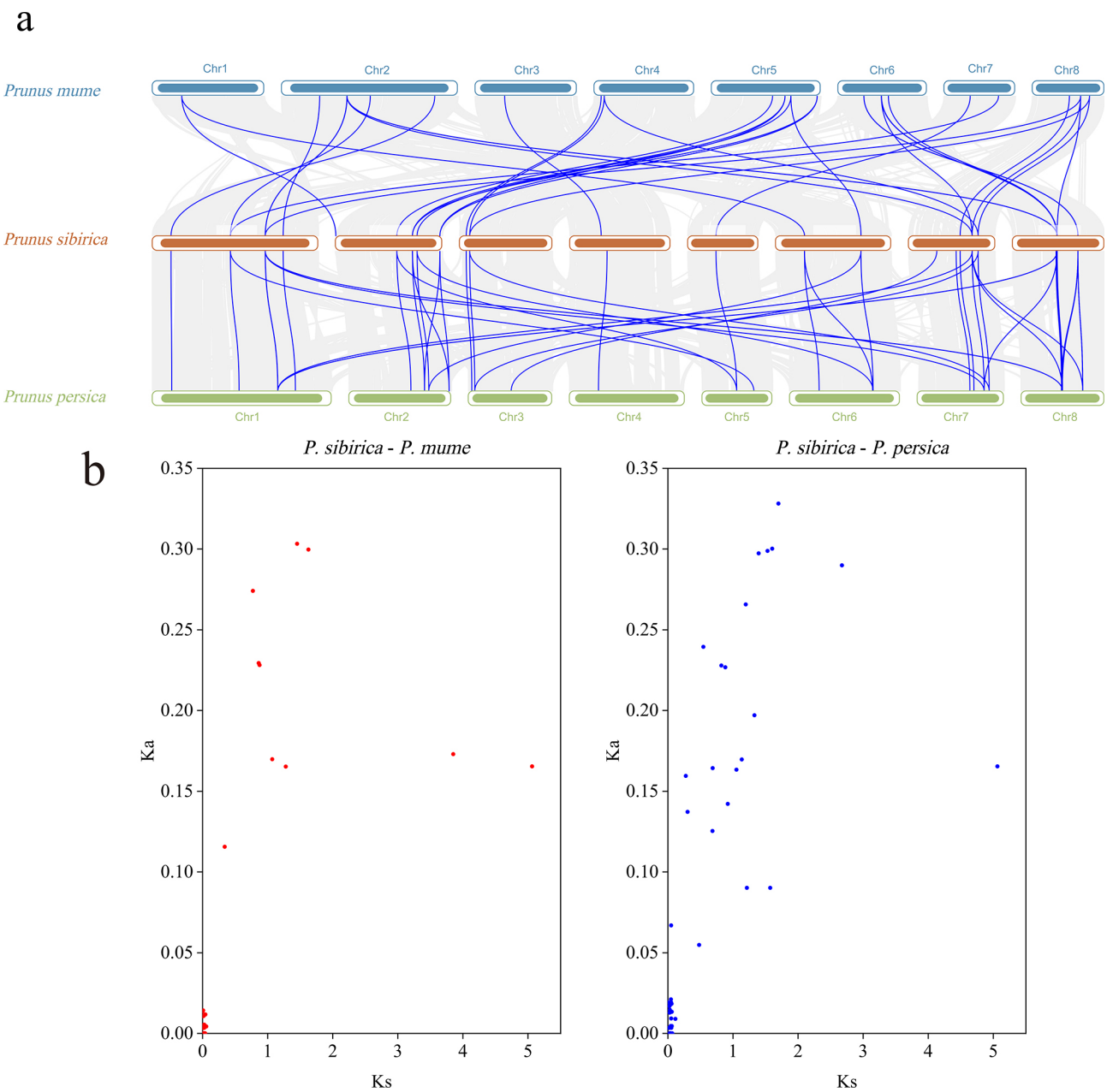

Fig. 5. Synteny analysis of *PsSAUR* genes between *P. sibirica* and two other plant species (*P. mume* and *P. persica*). (a) Grey lines in the background and blue lines between different species indicate the collinear blocks and syntenic *SAUR* pairs between *P. sibirica* and other species, respectively. (b) The scatter plot shows the Ka and Ks distributions of gene pairs between *P. sibirica* and the other plant species

0.57 and 0.74, respectively (Fig. 5b). These results suggest that the *PsSAUR* gene family is more closely related to *P. persica* than to *P. mume*. The Ks values of all collinear gene pairs are almost less than 1, confirming that the *SAURs* gene family has undergone strong negative selection (Table S4).

Expression patterns of PsSAURs in different floral organs: The expression patterns of *PsSAURs* in the pistil, stamen, petal, and sepal were analyzed (Fig. 6, Table S5). A total of 34 *PsSAURs* had an average expression level of fragments per kilobase million FPKM > 1, which were used in subsequent analyses. *PsSAUR4*, *PsSAUR8*, *PsSAUR19*, and *PsSAUR22* were highly expressed in the pistil; *PsSAUR13* was highly expressed in the stamen; *PsSAUR30*, *PsSAUR32*, *PsSAUR33*, *PsSAUR38*, *PsSAUR41*, *PsSAUR42*, *PsSAUR43*, *PsSAUR44*, *PsSAUR45*, *PsSAUR46*, *PsSAUR48*, *PsSAUR51*, and *PsSAUR52* were highly expressed in the petal; *PsSAUR9* and *PsSAUR11* were highly expressed in the sepal.

Expression patterns of PsSAURs in kernels at different developmental stages: To analyze the function of *PsSAURs* in kernel development, the expression patterns of *PsSAURs* in the kernel during six different developmental stages (S1-S6) were analyzed (Fig. 6, Table S6). The results showed that the average expression level of 17 *PsSAURs* revealed an FPKM > 1, and the expression patterns of *PsSAURs* in the kernel differed during different developmental

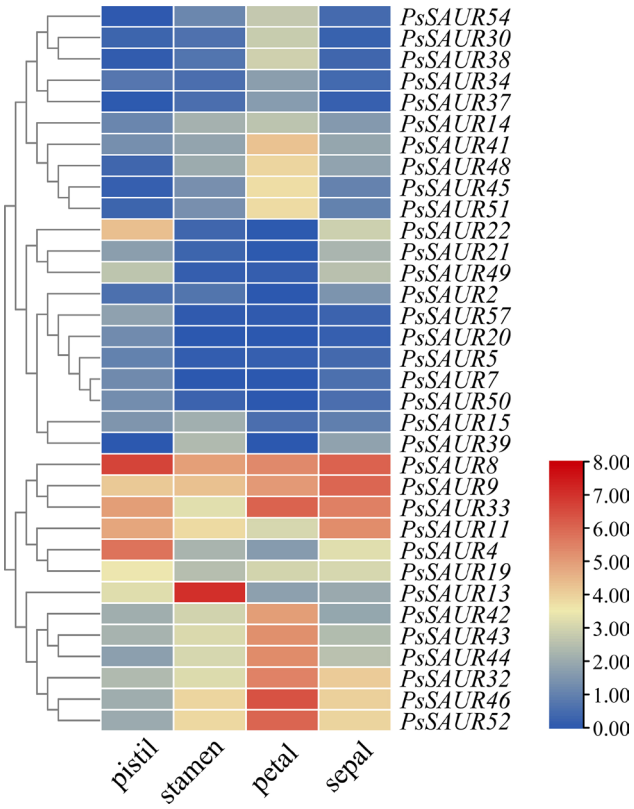

Fig. 6. Expression profile of *PsSAURs* in different floral organs

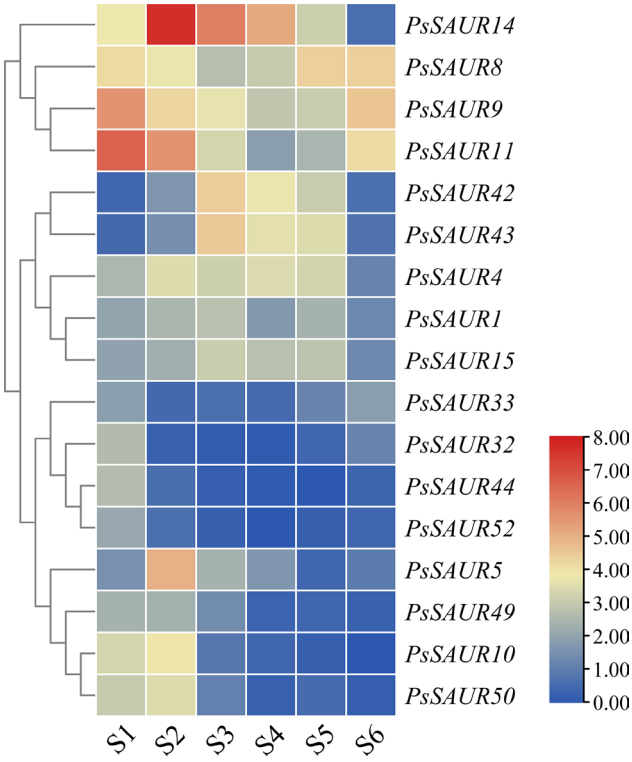

Fig. 7. Expression profile of *PsSAURs* in kernels at different developmental stages

stages. For example, *PsSAUR49* was downregulated during development. *PsSAUR9* and *PsSAUR11*, *PsSAUR5* and *PsSAUR14*, and *PsSAUR42* and *PsSAUR43* were highly expressed at S1, S2, and S3 respectively, which were the rapid growth stages, whereas the expression of *PsSAUR8* was higher at S5 and S6, which were the mature stage.

Discussion

Plant auxin is a hormone produced by cell regions with division and increased activity to regulate plant growth rate and direction. The primary role is to relax the plant cell wall so that cell growth and elongation can also increase RNA and protein synthesis in many plants (Zhao et al., 2018). *SAURs* are a significant class of early auxin-induced genes in plants, which can quickly respond to auxin induction and promote cell elongation (Stortenbeker et al., 2019). Along with the increasing number of plant genomes sequenced whole-genome identification and analysis of the *SAUR* family have been completed in many plants. However, the analysis of the *SAURs* in *P. sibirica* needed to be clarified; thus, we identified and analyzed the *PsSAURs* in the current study.

In this study, 57 *PsSAURs* were identified and characterized by bioinformatics methods, distributed on 8 chromosomes of the *P. sibirica* genome. Compared with the reported *SAUR* family members of

sweet cherry (86) (Qian-dong et al., 2023) and pear (116) (Wang et al., 2022) in Rosaceae, the number of *SAURs* in *P. sibirica* was less, which indicates that the *SAUR* family is conservative in the evolutionary process. Moreover, compared with the number of *SAURs* in apple (70), the number of *SAURs* in *P. sibirica* was less. This may be because of only one whole-genome duplication event in *P. sibirica*, whereas the apple genome has experienced two whole-genome duplication events (Velasco et al., 2010).

Understanding the structure of genes is essential for elucidating their function. Studies have shown that *SAURs* generally have no introns and are usually clustered (Zhu & Kong, 2014). Analysis of *PsSAURs* showed that 85% of these have no introns, which was associated with the previous reports, such as cucumber (67/73) (*Cucumis sativus* L.) (Wang & Shang, 2019), grape (59/64) (Li et al., 2018), tomato (90/99) (*Lycopersicon esculentum* Miller), potato (131/134) (*Solanum tuberosum* L.) (Wu et al., 2012), and apple (44/70) (Wang, Lu, et al., 2020). Chromosomal location analysis showed that *PsSAURs* were distributed in clusters on chromosomes 1, 4, 6, 7 and 8, indicating that they experienced significant fragment replication and tandem replication events during species evolution (Chen et al., 2014; Zhang et al., 2021; Wong et al., 2019). Phylogenetic and conserved structure analysis of 57 *PsSAURs* showed that the family was divided into five different groups. Some differences in conserved domains between different groups in *P. sibirica* indicate that *PsSAURs* have functional diversity during evolution.

Previous studies have identified tandem and segmental duplications as the primary driving force of plant evolution and gene family expansion, with segmental duplications being the most common in angiosperms (Cannon et al., 2004). Colinearity analysis showed that 35 and 59 gene pairs were found in *P. sibirica, P. mume* and *P. persica*, respectively. The Ka and Ks analysis of the replicated gene pairs showed that most were negative selection. Our results are consistent with the study of the pear (Wang et al., 2022) *SAUR* gene.

The *SAUR* gene family is one of the early auxin response gene families. Previous studies have shown that the *SAUR* gene family plays a vital role in plant growth and development. It can regulate hypocotyl elongation, cell expansion, auxin synthesis and transport (Gendreau et al., 1997; Chae et al., 2012). In the expression pattern analysis of *PsSAURs* in different floral organs, there were 4, 1, 13 and 2 highly expressed genes in pistil, stamen, petal and sepal, respectively, and two highly expressed genes in kernels of different developmental stages of *PsSAURs*, each of which was found in S1, S2 and S3 stages. It is shown that the expression pattern of *PsSAURs* is organ and developmental stage-specific. Our results are consistent with the study of sweet cherry (Qian-dong et al., 2023) and apple (Wang, Lu et al., 2020) *SAUR* gene.

Conclusions

In this study, genomics and bioinformatics analysis of 57 *PsSAURs* genes in *P. sibirica* were carried out. Specifically, we studied their gene structure, evolutionary relationships, collinearity, promoter *cis*-elements, and tissue-specific and growth-stage-specific expression patterns. It provides a theoretical basis and an important reference for further research on the function of *PsSAURs* in Rosaceae.

Author contributions

X. Y. and R. Y. designed the project. X. Y., R.Y. and W. B. wrote the paper. S. R., Y. B., J. C., Y. L., D. H. and H. H. helped with the data analysis. X. Y. and R.Y. lead the revision of the manuscript. All authors have reviewed and approved the manuscript for publication.

Funding

This research received no external funding.

Acknowledgements

This work is supported by the Young Scientific and Technological Talents in Universities Program of the Inner Mongolia Autonomous Region (NJYT22032) and the National Key R&D Program of China (2018YFD1000606). The Postgraduate Research and Innovation In Inner Mongolia Autonomous Region Program (S20210208Z).

Conflicts of interest

The authors declare no conflict of interest.

References

Bailey TL & Elkan C (1994) Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proceedings of the International Conference on Intelligent Systems Molecular Biology 2: 28–36.

Cannon SB, Mitra A, Baumgarten A, Young ND & May G (2004) The roles of segmental and tandem gene duplication in the evolution of large gene families in *Arabidopsis thaliana*. BMC Plant Biology 4: 10. doi:10.1186/1471-2229-4-10.

Chae K, Isaacs CG, Reeves PH, Maloney GS, Muday GK, Nagpal P & Reed JW (2012) Arabidopsis small auxin up RNA63 promotes hypocotyl and stamen filament elongation. The Plant Journal 71: 684–697. doi:10.1111/j.1365-313X.2012.05024.x.

Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y & Xia R (2020) TBtools: An integrative toolkit developed for interactive analyses of big biological data. Molecular Plant 13: 1194–1202. doi:10.1016/j.molp.2020.06.009.

Chen Y, Hao X & Cao J (2014) Small auxin upregulated RNA (*SAUR*) gene family in maize: identification, evolution, and its phylogenetic comparison with *Arabidopsis*, rice, and sorghum. Journal of Integrative Plant Biology 56: 133–150. doi:10.1111/jipb.12127.

Franco AR, Gee MA & Guilfoyle TJ (1990) Induction and superinduction of auxin-responsive mRNAs with auxin and protein synthesis inhibitors. Journal of Biological Chemistry 265: 15845–15849. doi:10.1016/S0021-9258(18)55475-2.

Gasteiger E, Hoogland C, Gattiker A, Duvaud, S, Wilkins, MR, Appel RD & Bairoch A (2005) Protein identification and analysis tools on the ExPASy server: The Proteomics Protocols Handbook (ed. by JM Walker) pp. 571–607. doi:10.1385/1-59259-890-0:571.

Gendreau E, Traas J, Desnos T, Grandjean O, Caboche M & Höfte H (1997) Cellular basis of hypocotyl growth in *Arabidopsis thaliana*. Plant Physiology 114: 295–305. doi:10.1104/pp.114.1.295.

Hagen G & Guilfoyle T (2002) Auxin-responsive gene expression: genes, promoters and regulatory factors. Plant Molecular Biology 49: 373–385. doi:10.1023/A:1015207114117.

Hou K, Wu W & Gan SS (2013) *SAUR36*, a SMALL AUXIN UP RNA gene, is involved in the promotion of leaf senescence in Arabidopsis. Plant Physiology 161: 1002–1009. doi:10.1104/pp.112.212787.

Hu B, Jin J, Guo AY, Zhang H, Luo J & Gao G (2015) GSDS 2.0: an upgraded gene feature visualization server. Bioinformatics 31: 1296–1297. doi:10.1093/bioinformatics/btu817.

Huang MZ, Zhu XC, Bai HK, Wang C, Gou NN, Zhang YJ, Chen C, Yin MY, Wang L & Wuyun T (2023) Comparative anatomical and transcriptomics reveal the larger cell size as a major contributor to larger fruit size in apricot. International Journal of Molecular Sciences–24: 8748. doi:10.3390/ijms24108748.

Kumar S, Stecher G, Li M, Knyaz C & Tamura K (2018) MEGA X: Molecular evolutionary genetics analysis across computing platforms. Molecular Biology and Evolution 35: 1547–1549. doi:10.1093/molbev/msy096.

Larkin MA, Blackshields G, Brown NP, Chenna R, Mcgettigan PA, Mcwilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ & Higgins DG (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23: 2947–2948. doi:10.1093/bioinformatics/btm404.

Lescot M, Dehais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouzé P & Rombauts S (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: 325–327. doi:10.1093/nar/30.1.325.

Li A, Cui M, Chen K, Xu Y, Jia H & Fang J (2018) Identification and bioinformatics analysis of the SAUR gene family in grape. Journal of Plant Genetic Resources 19: 326–337.

Liu H, Yu Y, Feng G, Xu K, Zhao S, Zhang S, Yang X & Zhao Y (2022) Identification and expression analysis of the SAUR gene family in *triticum aestivurn* L. Molecular Plant Breeding 20: 4525–4538. doi:10.13271/j.mpb.020.004525.

McClure BA & Guilfoyle T (1987) Characterization of a class of small auxin-inducible soybean polyadenylated RNAs. Plant Molecular Biology 9: 611–623. doi:10.1007/BF00020537.

Niu J, An J, Wang L, Fang C, Ha D, Fu C, Qiu L, Yu H, Zhao H, Hou X, Xiang Z, Zhou S, Zhang Z, Feng X & Lin S (2015) Transcriptomic analysis revealed the mechanism of oil dynamic accumulation during developing Siberian apricot (*Prunus sibirica* L.) seed kernels for the development of woody biodiesel. Biotechnology for Biofuels 8: 29. doi:10.1186/s13068-015-0213-3.

Qian-dong H, Yi H, Zhuang W, Wen Z, Shang CQ, Li ZC, Cai XW, Qiao G & Wen XP (2023) Molecular characterization of the *SAUR* gene family in sweet cherry and functional analysis of *PavSAUR55* in the process of abscission. Journal of Integrative Agriculture 22: 1720–1739. doi:10.1016/j.jia.2023.04.031.

Qiu T, Qi M, Ding X, Zheng Y, Zhou T, Chen Y, Han N, Zhu M, Bian H & Wang J (2020) The *SAUR41* subfamily of *SMALL AUXIN UP RNA* genes is abscisic acid inducible to modulate cell expansion and salt tolerance in *Arabidopsis thaliana* seedlings. Annals of Botany 125: 805–819. doi:10.1093/aob/mcz160.

Spartz AK, Lee SH, Wenger JP, Gonzalez N, Itoh H, Inzé D, Peer WA, Murphy AS, Overvoorde PJ & Gray WM (2012) The *SAUR19* subfamily of *SMALL AUXIN UP RNA* genes promote cell expansion. The Plant Journal 70: 978–990. doi:10.1111/j.1365-313X.2012.04946.x.

Stortenbeker N & Bemer M (2019) The *SAUR* gene family: the plant's toolbox for adaptation of growth and development. Journal of Experimental Botany 70: 17–27. doi:10.1093/jxb/ery332.

Subramanian B, Gao S, Lercher MJ, Hu S & Chen WH (2019) Evolview v3: a webserver for visualization, annotation, and management of phylogenetic trees. Nucleic Acids Research 47: W270–W275. doi:10.1093/nar/gkz357.

Velasco R, Zharkikh A, Affourtit J, Dhingra A, Cestaro A, Kalyanaraman A, Fontana P, Bhatnagar SK, Troggio M, Pruss D, Salvi S, Pindo M, Baldi P, Castelletti S, Cavaiuolo M, Coppola G, Costa F, Cova V, Dal Ri A, Goremykin V, Komjanc M, Longhi S, Magnago P, Malacarne G, Malnoy M, Micheletti D, Moretto M, Perazzolli M, Si-Ammour A, Vezzulli S, Zini E, Eldredge G, Fitzgerald LM, Gutin N, Lanchbury J, Macalma T, Mitchell JT, Reid J, Wardell B, Kodira C, Chen Z, Desany B, Niazi F, Palmer M, Koepke T, Jiwan D, Schaeffer S, Krishnan V, Wu C, Chu VT, King ST, Vick J, Tao Q, Mraz A, Stormo A, Stormo K, Bogden R, Ederle D, Stella A, Vecchietti A, Kater MM, Masiero S, Lasserre P, Lespinasse Y, Allan AC, Bus V, Chagné D, Crowhurst RN, Gleave AP, Lavezzo E, Fawcett JA, Proost S, Rouzé P, Sterck L, Toppo S, Lazzari B, Hellens RP, Durel C-E, Gutin A, Bumgarner RE, Gardiner SE, Skolnick M, Egholm M, Van de Peer Y, Salamini F & Viola R (2010) The genome of the domesticated apple (*Malus* x *domestica* Borkh.). Nature Genetics 42: 833–839. doi:10.1038/ng.654.

Wang F, Yu H, Hu Z, Guan D, Zhang P, Zhu S & Zhao X (2020) Genome-wide analysis of SAUR gene family in citrus. Acta Horticulturae Sinica 47: 23–40. doi:10.16420/j.issn.0513-353x.2019-0581.

Wang H & Shang Q (2019) Genome-wide identification and expression analysis of the SAUR gene family in *Cucumis sativus*. Acta Horticulturae Sinica 46: 1093–1111. doi:10.16420/j.issn.0513-353x.2018-0873.

Wang L, Bao W, Xu W, Huang M, Zhu G, Zhao H, Chen C, Bai H, Chen J, Liu H & Wu TN (2022) Transcriptome analysis reveals genes associated with kernel size in apricots cultivated for kernel consumption (*Prunus armeniaca* × *Prunus sibirica*). Scientia Horticulturae 302: 111141. doi:10.1016/j.scienta.2022.111141.

Wang M, Manzoor MA, Wang X, Feng X, Zhao Y, He J & Cai Y (2022) Comparative genomic analysis of SAUR gene family, cloning and functional characterization of two genes (*PbrSAUR13* and *PbrSAUR52*) in *Pyrus bretschneideri*. International Journal of Molecular Sciences 23: 7054. doi:10.3390/ijms23137054.

Wang P, Lu S, Xie M, Wu M, Ding S, Khaliq A, Ma Z, Mao J & Chen B (2020) Identification and expression analysis of the small auxin-up RNA (SAUR) gene family in apple by inducing of auxin. Gene 750: 144725. doi:10.1016/j.gene.2020.144725.

Wang Z, Kang M, Liu H, Gao J, Zhang Z, Li Y, Wu R & Pang X (2014) High-level genetic diversity and complex population structure of Siberian apricot (*Prunus sibirica* L.) in China as revealed by nuclear SSR markers. Plos One 9: e87381. doi:10.1371/journal.pone.0087381.

Wang Z, Song S, Sheng SB, Tian J, Wu RL & Pang XM (2019) Comparative transcriptome analysis identifies differentially expressed genes between normal and late-blooming Siberian apricot. Journal of Forestry Research 30: 2277–2288. doi:10.1007/s11676-018-0825-0.

Wong JH, Spartz AK, Park MY, Du M & Gray WM (2019) Mutation of a conserved motif of PP2C.D phosphatases confers SAUR immunity and constitutive activity. Plant Physiology 181: 353–366.

Wu J, Liu S, He Y, Guan X, Zhu X, Cheng L, Wang J & Lu G (2012) Genome-wide analysis of SAUR gene family in Solanaceae species. Gene 509: 38–50. doi:10.1016/j.gene.2012.08.002.

Zhang H, Yu Z, Yao X, Chen J, Chen X, Zhou H, Lou Y, Ming F & Jin Y (2021) Genome-wide identification and characterization of small auxin-up RNA (*SAUR*) gene family in plants: Evolution and expression profiles during normal growth and stress response. BMC Plant Biology 21: 4. doi:10.1186/s12870-020-02781-x.

Zhao Y (2018) Essential roles of local auxin biosynthesis in plant development and in adaptation to environmental changes. Annual Review of Plant Biology 69: 417–435. doi:10.1146/annurev-arplant-042817-040226.

Zhu YB, Kong YY & Wang JH (2014) Research advances in auxin-responsive SAUR genes. Chinese Bulletin of Life Sciences 26: 407–413.