Open Access

SCHOLARENA

Journal of Advancements in Plant Science

RESEARCH ARTICLE

ISSN: 2639-1368

Genome-wide Identification and Expression Analysis of Polygalacturonase Gene Family in *Euscaphis japonica*

Peilan Zhang[1,2], Lin Xiao[1,2], Yixun LE[1,2], Hang Zhang[1,2], Miaoli WU[3], Siren LAN[1] and Shuangquan Zou[1,2*]

[1]College of Forestry, Fujian Agriculture and Forestry University, Fuzhou, Fujian, 350002, China

[2]Fujian Colleges and Universities Engineering Research Institute of Conservation & Utilization of Natural Bioresources, Fuzhou, Fujian, 350002, China

[3]Fuzhou University of International Studies and Trade, Fuzhou, Fujian, 350002, China

[*]**Corresponding author:** Shuangquan Zou, College of Forestry, Fujian Agriculture and Forestry University, Fuzhou, Fujian, 350002, China, Tel: +8613860668887, E-mail: zou@fafu.edu.cn

Citation: Peilan Zhang, Lin Xiao, Yixun LE, Hang Zhang, Miaoli WU, et al. (2022) Genome-wide Identification and Expression Analysis of Polygalacturonase Gene Family in *Euscaphis japonica*. J Adv Plant Sci 5: 101

Abstract

Organ abscission takes part in the entire life cycle of plants, exerting great influence on plant survival and reproduction. The activation of polygalacturonase (PG) is one of the important signs of plant organ abscission. To investigate the role of PGs in fruit and flower abscission of *Euscaphis japonica*, we conducted genome wide identification and expression analysis for PG gene family in *E. japonica*. A total of 59 EkPG genes were identified in *E. japonica* genome which were further clustered into seven groups based on phylogenetic analysis. RAD-seq combined with qRT-PCR results showed that *EkPG6*, a homologous gene of *AtQRT3* that regulates pollen development, and *EkPG4, EkPG16, EkPG25, EkPG34* were highly expressed in abscission zone (AZ) where young fruits and floral organs are about to falling off, respectively, suggesting these homologous genes play an important role in the shedding of young fruits and flowers. This study could provide a theoretical foundation for *E. japonica* breeding and further promote the research on the abscission of plant organs.

Keywords: *Euscaphis japonica;* Polygalacturonase; Gene family; Expression analysis; Abscission

Introduction

Abscission of leaves, flowers, fruits, and other organs is common during the entire life cycle of plants. It is a highly coordinated regulation regime that plays essential roles in nutrition cycle and reproductive development, as well as becomes one of the key mechanisms assisting plants to adapt to the environment [1-4]. The specific area where plant organ abscission occurs is called the abscission zone (AZ), and abscission is a result of the degradation of the primary cell wall and intercellular layer in the AZ [5,6]. The primary cell wall is mainly composed of celluloses, hemicelluloses, pectins, and pectin is the major component of intercellular layer. Therefore, cell wall hydrolytic enzymes such as cellulase, pectinase, and xylanase are closely related to plant organs abscission [7]. The increasing enzymatic activity of the cell wall hydrolase is the signature of abscission, which results in decomposing celluloses, hemicelluloses, pectins, and other compounds in the cell wall [8] and reduction in the adhesion force between cells decreases, ultimately, the plant organs are detached from the plant body by gravity [9].

Polygalacturonase (PG) is a plant cell wall hydrolase that catalyzes the random hydrolysis and decomposition of pectins, leading to cell separation and promoting plant organs abscission [7]. In oil palm (*Elaeis guineensis*) fruit, 14 identified PG genes exhibit high expression in AZ at fruit ripening and mesocarp ripening [13]. A total of 38 PG genes were identified in citrus (*Citrus sinensis*), of which three *CitPG* genes may be related to the citrus fruitlet abscission [14]. In the study of tomato (*Lycopersicon esculentum*), *TAPG1*, *TAPG2*, and *TAPG4* are all expressed in the AZ of flowers and leaves, indicating they involved in the senescence and abscission of flowers and leaves [15]. In addition, members of PG family have a strong functional diversity in plants [10-12]. In many fruits, PGs are induced by ethylene and play an active role in fruit softening, such as *FaPG1* and *FaPG2* in strawberries (*Fragaria* × *ananassa*), *Ps-PG1* and *Ps-PG2* in pear (*Pyrus communis* L.), *MaPG3* and *MaPG4* in banana (*Musa accuminata*), *PpPG21* and *PpPG22* in peach (*Prunus persica* L.) [16-19]. In *Arabidopsis thaliana*, *ADPG1* and *ADPG2* are essential for silique and anther dehiscence [11]. *PGX1* and *PGX2* play a role in cell expansion and increase the elongation of hypocotyls. However, *PGX3* affects the shape of seedling cotyledons and the spacing and pore size of the developing stomata [20-22]. *OsPG1* in rice (*Oryza sativa* L.) regulates cell death and the plant immune system through cell wall remodeling, thereby enhancing rice resistance to bacterial blight [23].

Euscaphis japonica, a species in Staphyleaceae, is an evergreen or deciduous tree mainly distributed in southern China. The flowering time of evergreen *E. japonica* is in May, with fruit developing in July, and the pericarp color changes in August [24]. The fruits get mature in September, of which the exocarp turns red, cracks along the abdominal sutures, and turns over, revealing the red endocarp and mature black seeds, resembling a butterfly. In November, the fruits begin abscise, but there are still fruits hanging on the branches until March of the following year. Therefore, *E. japonica* is widely planted in southern China as a garden ornamental tree [25,26]. To reveal the molecular mechanism underlying *E. japonica* flower and fruit abscission, we combine whole-genome data and transcriptome data of five fruit development stages to identify and analyze the PGs gene family. The length of the fruit-bearing period has a greater impact on the ornamental value of fruit tree species. Therefore, this study sheds insights into the biosynthetic pathways related to the abscission of flowers and fruits of *E. japonica*, reveals the molecular mechanism of the ultra-long fruit-bearing period, as well as provides valuable data for the evolution and biological functions of PGs in plants.

Materials and Methods

Identification of PG gene family

PG protein sequences of *A. thaliana* were downloaded from the TAIR database (https://www.arabidopsis.org/) as the query. The BLAST in TBtools was used to preliminarily screen the PG gene family members from the *E. japonica* genome data [27]. The genome data was obtained from BioProject/GWH (https://bigd.big.ac.cn/gwh) under the accession codes PRJCA005268/GWH-BCHS00000000. Then, we used the NCBI Blastp database (https://blast.ncbi.nlm.nih.gov/) to perform sequence alignment on the identification of PG gene family members. Redundant sequences were discarded manually, including the alternative splicing

of genes. Finally, All EkPG candidate protein sequences were examined using the CDD (https://www.ncbi.nlm.nih.gov/) and the SMART database (http://smart.embl-heidelberg.de/) with an E value cutoff of 1.0 by the domain analysis programs [28].

Analysis of Protein Sequence Properties

We used the GFF3 file of *the E. japonica* to obtain the EkPG sequence name and position information through BLAST. The theoretical isoelectric point and molecular weight were predicted in the Expasy tool (https://web.expasy.org/protparam/) [29].

Phylogenetic analysis of EkPGs

We used Clustal W application in MEGA7.0 to compare and analyze the amino acid sequence of the protein [30]. MEGA was further used to construct the phylogenetic tree of the aligned sequences with the Neighbor-joining method. Poisson correction is used, and the Bootstrap value is 1000 repetitions [31].

Evolutionary analysis of EkPGs

The segmental and tandem duplications analysis of EkPG genes were conducted by MCScanX [32]. Parameters take default values. The *K*a and *K*s were computed between pairs of genes identified as homeologous by MCScanX. TBtools was used to visualize gene structure, gene locations, and replication events.

EkPG protein sequences were analyzed for novel domains with the MEME server [33]. The motif sites were set at 10 and other default. MEME results were visualized by TBtools [34].

EkPGs expression analysis

We use the transcriptome data of the five developing stages of fruits to analyze the expression of the EkPGs gene. The TBtools was used to draw heat maps. The five periods are 50d(Fr_I), 100d(Fr_II), 130d (Fr_III), 160d(Fr_IV), 180d(Fr_V) after blooming.

RNA extraction and real-time fluorescence quantitative PCR

Six genes (*EkPG4, EkPG6, EkPG16, EkPG21, EkPG25, EkPG34*) were selected from the EkPGs that may be related to the organ abscission of *E. japonica*. We analyzed their expression in AZ of flowers and young fruits by qRT-PCR. In this study, we collected the AZ in four periods: full-flowering (F1_1), final-flowering (F2_1), young fruit (F3_1), and young fruits about to abscise(F3_2). There are three biological replicates for each sample. Young fruit about to abscise: grow slowly but not abscise, the young fruit will fall off when the finger is lightly touched.

The total RNA was extracted with a polysaccharide polyphenol plant total RNA extraction kit (Hangzhou Bioer Technology, China). RNA integrity was verified by 1.5% agar gel electrophoresis. We used TransScript® One-Step gDNA Removal and cDNA Synthesis SuperMix kit to reverse transcription and synthesize cDNA. The Premier was used to design qRT-PCR primers for EkPG gene family members (Table 1). *EkLPP* was used as an internal reference gene [35]. *EkPLL* forward sequence: TTGGCCTCATCTATTGCTACTG; *EkLPP* reverse sequence: GTTCTCCTGTGCCCTCTAATG. The qRT-PCR used Genious 2X SYBR Green Fast qPCR Mix kit, and through ABIQuantStudio3 to detect the relative expression of the target genes in the sample. The PCR amplification cycle was as follows: 95 °C for 3min, 40 cycles at 95 °C for 5 s, and 60 °C for 30 s. We calculated the relative expression of the target gene according to the $2^{-\triangle\triangle Ct}$ method [36], and used Excel 2020 to analyze the relevant data.

Gene	Primer sequences (5'-3')	Primer sequences (5'-3')
*EkPG*4	F:TGGTGGATTGGTTGATTACAGAC	R:GGCATGAAATCTAAAGGGTTGT
*EkPG*6	F:CCGTCCAACTCCACATAACC	R:CCACCACCACCTGCTCTATA
*EkPG*16	F:AGATGTCAGGCGGTGTATCA	R:GTGTTCGTTGTAGTCTGTCTTCAT
*EkPG*21	F:CAGTGACAGCCTTCCATGTATAG	R:GTGCTCAACTCTCCGAATGT
*EkPG*25	F:GGATCATGTCGTTACCGTAGG	R:CAGTCAACCGTCACTTCCTT
*EkPG*34	F:GCTGTCGGAAGAGGAGGATA	R:CGCCGCTATTGACACATTCT

Table 1: Primers used to fluorescent quantitative PCR

Results

Identification of the PG gene of E. japonica

A total of 59 PG gene family members were obtained, and they were named *EkPG1~EkPG59* according to the chromosome location information. Physical and chemical analysis of EkPG protein was carried out through the ExPASy website. The results showed that the open reading frames (ORFs) lengths of EkPGs are between 519 and 1689 bp, the molecular weights (MWs) are between 19.60 and 61.23 kDa, and the *E. japonica* PG gene polypeptides showed long sequences with 172 and 562 aa (Table 2). The protein isoelectric points range from 4.55 to 9.51, among which 25 PG proteins are acidic (PI<6.5), 6 PG proteins are neutral (6.5<PI<7.5), and 28 PG proteins are alkaline (PI>7.5) (Table 2).

Group	Name	Gene ID	Proteins			ORF (bp)	Location
			pI	MW (kDa)	Length (aa)		
A	*EkPG7*	E.japonica.28428	4.92	51.83	477	1434	Chr02:1250231-1253430
	*EkPG*18	E.japonica.09948	5.27	46.22	428	1287	Chr03:28184651-28186764
	*EkPG*21	E.japonica.02733	8.58	56.47	519	1560	Chr03:54893678-54896616
	*EkPG*23	E.japonica.27973	9.16	51.97	473	1422	Chr04:96089485-96094733
	*EkPG*36	E.japonica.13905	5.01	51.17	464	1395	Chr04:96853756-96857721
	*EkPG*52	E.japonica.05491	4.55	52.38	485	1458	Chr08:68141115-68142969
	*EkPG*55	E.japonica.07549	6.35	50.14	461	1386	Chr09:42101952-42104589
B	*EkPG3*	E.japonica.02085	8.44	46.93	430	1293	Chr10:42175797-42177855
	*EkPG*11	E.japonica.10661	8.86	48.86	446	1341	Chr01:33039500-33042380
	*EkPG*15	E.japonica.17279	5.49	49.32	458	1377	Chr02:55608787-55612619
	*EkPG*22	E.japonica.02732	5.76	60.87	562	1689	Chr02:120604698-120607785
	*EkPG*31	E.japonica.02350	8.81	49.01	445	1338	Chr04:96214854-96223949
	*EkPG*32	E.japonica.02352	8.79	37.86	345	1038	Chr08:26263843-26296939
	*EkPG*33	E.japonica.29656	5.98	48.40	443	1332	Chr08:26396861-26404643
C	*EkPG*13	E.japonica.23553	9.17	37.05	349	1050	Chr08:54153996-54157128
	*EkPG*14	E.japonica.23556	9.51	41.39	389	1170	Chr02:85572488-85575763
	*EkPG*43	E.japonica.21044	9.45	44.24	416	1251	Chr02:85617998-85624180
	*EkPG*44	E.japonica.21046	6.91	41.26	393	1182	Chr09:5677648-5685727
	*EkPG*45	E.japonica.21048	8.19	40.70	391	1176	Chr09:5794508-5798944
	*EkPG*46	E.japonica.21049	8.85	40.62	391	1176	Chr09:5871075-5874223
	*EkPG*47	E.japonica.21051	8.77	40.95	391	1176	Chr09:5948234-5951574
	*EkPG*48	E.japonica.21052	9.2	34.75	329	990	Chr09:5994634-5998685

D	*EkPG*1	E.japonica.27188	9.28	44.12	410	1233	Chr09:6023895-6026465
	*EkPG*9	E.japonica.28320	9.03	53.61	491	1476	Chr01:12936379-12939225
	*EkPG*28	E.japonica.00829	9.2	43.75	400	1203	Chr02:3888916-3897455
	*EkPG*29	E.japonica.00522	7.88	42.92	396	1191	Chr06:20514345-20519766
	*EkPG*37	E.japonica.12423	8.43	42.68	403	1212	Chr06:28942810-28945128
	*EkPG*38	E.japonica.12422	9.24	43.11	404	1215	Chr08:82509171-82510862
	*EkPG*39	E.japonica.12419	8.85	42.14	396	1191	Chr08:82527922-82529977
	*EkPG*40	E.japonica.12418	8.18	36.30	333	1002	Chr08:82592983-82599814
	*EkPG*49	E.japonica.06798	9.41	44.98	415	1248	Chr08:82600830-82602817
	*EkPG*50	E.japonica.06797	5.92	44.59	420	1263	Chr09:36366684-36377462
	*EkPG*54	E.japonica.28555	4.87	43.40	406	1221	Chr09:36390878-36393232
	*EkPG*56	E.japonica.19208	8.75	42.81	400	1203	Chr10:15038017-15045819
	*EkPG*57	E.japonica.04422	7.44	43.48	407	1224	Chr11:29445715-29448297
E	*EkPG*4	E.japonica.02061	6.26	50.80	459	1380	Chr11:60761529-60769291
	*EkPG*5	E.japonica.02908	6.32	53.57	487	1464	Chr01:33775921-33784484
	*EkPG*8	E.japonica.28327	5.07	49.77	456	1371	Chr01:95435870-95439025
	*EkPG*10	E.japonica.28192	6.56	48.34	440	1232	Chr02:3810612-3814054
	*EkPG*16	E.japonica.13718	9.02	55.31	493	1482	Chr02:29886355-29890173
	*EkPG*19	E.japonica.17909	4.77	50.14	463	1392	Chr03:7837670-7844184
	*EkPG*20	E.japonica.15271	6.87	19.60	172	519	Chr03:125806837-125809113
	*EkPG*24	E.japonica.13048	8.75	54.46	489	1470	Chr04:67402262-67404103
	*EkPG*25	E.japonica.16818	8.64	61.23	562	1689	Chr05:56688252-56691708
	*EkPG*34	E.japonica.29802	7.52	52.66	481	1446	Chr05:72645578-72662482
	*EkPG*35	E.japonica.13879	5.42	51.48	465	1398	Chr08:58413027-58415556
	*EkPG*51	E.japonica.06761	5.42	34.04	314	945	Chr08:67558376-67562070
	*EkPG*53	E.japonica.16117	5.7	51.73	472	1419	Chr09:37577304-37580376
	*EkPG*58	E.japonica.23866	5.13	49.14	456	1371	Chr09:57864853-57867835
	*EkPG*59	E.japonica.04665	5.42	49.11	442	1329	Chr11:62161479-62164901
F	*EkPG*2	E.japonica.28966	6.66	42.69	397	1194	Chr11:70657191-70660061
	*EkPG*41	E.japonica.08984	8.44	42.70	407	1224	Chr01:26149450-26154075
	*EkPG*42	E.japonica.08983	7.1	43.11	402	1209	Chr09:595658-602735
	*EkPG*17	E.japonica.19146	5.49	42.79	402	1209	Chr09:654880-672199
G	*EkPG*6	E.japonica.22172	6.34	51.99	473	1422	Chr01:145359141-145360680
	*EkPG*12	E.japonica.10859	6.14	48.03	441	1326	Chr02:62827942-62831714
	*EkPG*26	E.japonica.06908	5.42	48.78	453	1362	Chr06:565512-568131
	*EkPG*27	E.japonica.06907	5.8	52.73	497	1494	Chr06:569337-571447
	*EkPG*30	E.japonica.09462	5.89	53.93	499	1500	Chr07:5139215-5143674

Table 2: Identify PG genes from *E. japonica* and their physicochemical and biochemical properties

Phylogenetic analysis of the EkPGs

We constructed a phylogenetic tree of PG genes from *E. japonica* and *A. thaliana*, and found that AtPGs and EkPGs are divided into seven groups: Group A, Group B, Group C, Group D, Group E, Group F, and Group G. EkPG gene family members exist in all branches (Figure 1). *ATQRT3*, belonging to Group G, has an important function in pollen development [37]. It was notably that the *ATQRT3*-like gene in *E. japonica* has expanded, leading to five genes, *EkPG6*, *EkPG12*, *EkPG26*, *EkPG27*, and *EkPG30*.

The conserved domains and gene structure of EkPGs

There are ten conserved motifs in EkPGs protein (Figure 2b), which are named Motif 1-10. PG protein has four characteristic conserved domains: I, SPNTDGIH; II, GDDC; III, CGPGHGIS; IV, RIK. The Motif 4, Motif 6, Motif 3, and Motif 1 correspond to 'SPNTDGIH', 'GDDC', 'CGPGHGIS', and 'RIK', respectively (Figure 2c). Most EkPGs in clades A, B, C, D, and F contain the above four conserved domains. However, Group E only contains three conserved domains (I, II, IV), except for *EkPG4*, *EkPG20*, and *EkPG51* which do not have IV domains. There are five *AtQRT3* homologous proteins in Group G, all of which contain the conserved domain Motif 5. Among them, IV domain presents in *EkPG6*, *EkPG27*, and *EkPG30*.

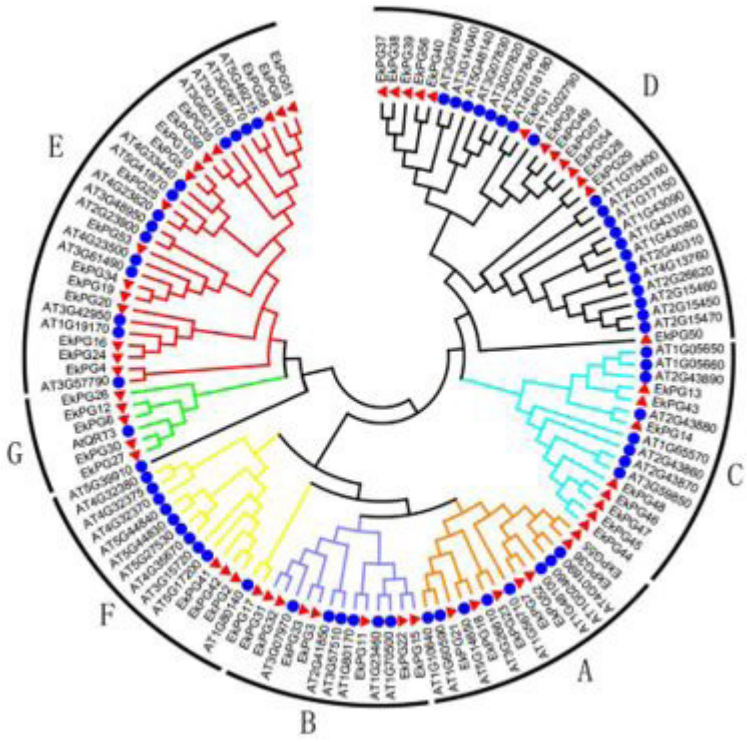

Figure 1: Phylogenetic analysis of Polygalacturonase (PG) proteins among *E. japonica* and *Arabidopsis*. These 127 sequences were used to construct a neighbor-joining (NJ) tree. The tree was divided into seven groups, represented by different colors. The triangle represents *E. japonica* and the circle represents *Arabidopsis*

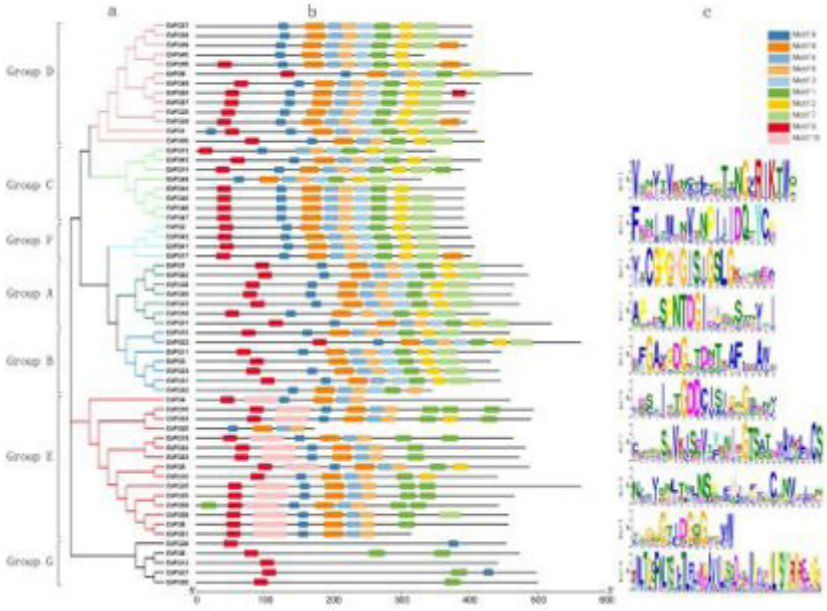

Figure 2: Analysis of the EkPG protein conserved domains. (a)Phylogenetic relationships, (b) motif compositions, and (c)site of amino acids of the 59 PG genes identified in the *E. japonica*. Phylogenetic relationships used the neighbor-joining(NJ) method, and different colors represent different groups. Colored boxes indicate conserved motifs and black lines represent nonconserved sequences. The lengths of motifs in each protein are shown proportionally

The gene structure of EkPGs was analyzed and we found that the number of exons ranges from two to nine (Figure 3b). *EkPG17*, *EkPG41*, and *EkPG42* from Group F contain eight introns and nine exons, while *EkPG6* from Group G has the least number of introns and exons. Based on the phylogenetic tree, Group F is distant from Group G. The closer the relationship of genes, the more similar the gene structure between them, such as Group A and B.

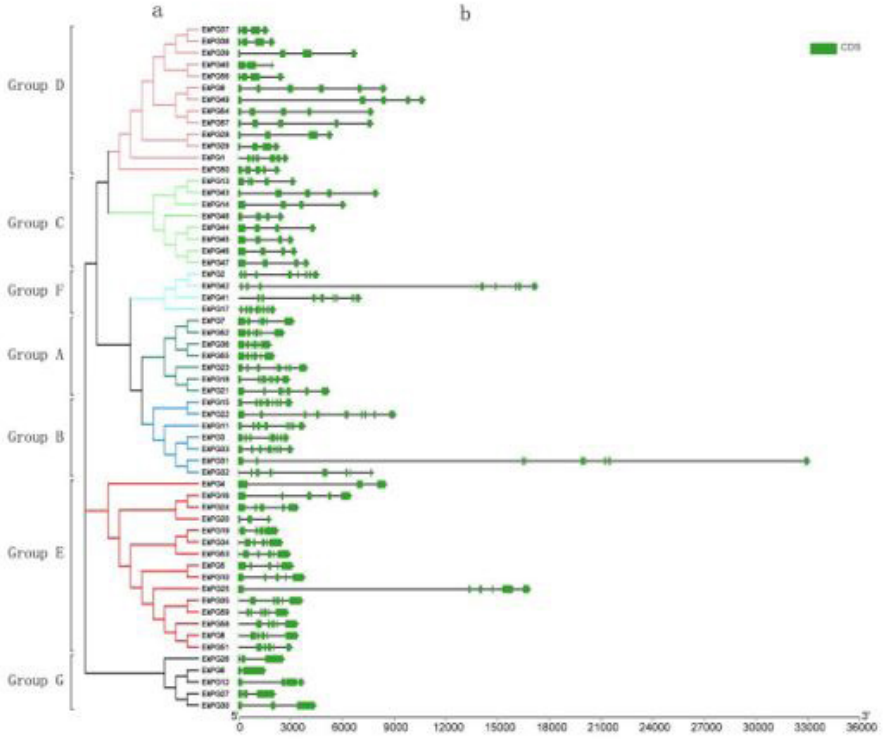

Figure 3: Analysis of the EkPG protein gene structure. (a)Phylogenetic relationships and (b) gene structures of the 59 PG genes identified in the *E. japonica*. Exons are represented by green boxes and introns by black lines

The chromosome localization of EkPGs

Fifty-nine EkPGs are unevenly distributed on each chromosome (Figure 4). Among them, chromosome 9 has the greatest number of genes, followed by chromosome 8, and chromosome 7 has only one gene. Chromosomes 2, 4, 6, 8, and 9 all contain gene clusters. There are three gene clusters on chromosome 9, two gene clusters are found on each of chromosomes 2, 8 and only one gene cluster is found on the remaining chromosomes. In addition, EkPGs contain a total of four pairs of fragment replication events, all of which are replication events between chromosomes.

The results of *K*a/*K*s of EkPGs repeat genes showed that the *K*a/*K*s of tandem repeats ranged from 0.1778 to 0.5770, with an average value of 0.282282973, and the *K*a/*K*s of fragment repeats ranged from 0.224868989 to 0.433590109, with an average value of 0.312837446 (Table 3). It is worth noting that *EkPG8-EkPG9*, *EkPG21-EkPG22*, *EkPG26-EkPG27* have no *K*s value.

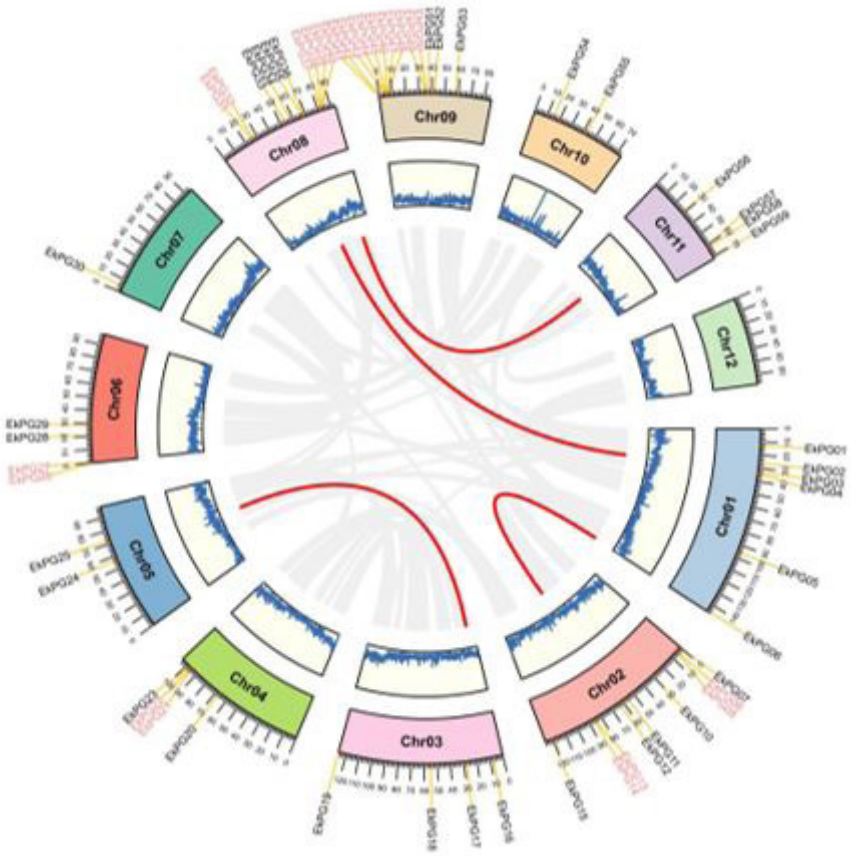

Figure 4: Genomic localization and gene duplication of PG genes in *E. japonica*. Tracks from outside to inside are as follows: (i) 12 pseudochromosomes and (ii) the gene density of *E. japonica* genome. Tandem duplicates are shown in a pink font, and segmental duplicated genes are linked by red lines

Gene1	Gene2	Ka	Ks	Ka/Ks	Duplication
EkPG3	*EkPG33*	0.275277145	0.634878747	0.433590109	S
EkPG6	*EkPG12*	0.201240026	0.894921199	0.224868989	S
EkPG8	*EkPG9*	1.0589238	NaN	NaN	T
EkPG13	*EkPG14*	0.252343725	1.339473246	0.188390269	T
EkPG16	*EkPG24*	0.117250873	0.415916599	0.281909579	S
EkPG21	*EkPG22*	0.662384297	NaN	NaN	T
EkPG26	*EkPG27*	0.384955736	NaN	NaN	T
EkPG31	*EkPG32*	0.190560772	0.330257406	0.57700681	T
EkPG37	*EkPG38*	0.080072148	0.255141282	0.313834543	T
EkPG38	*EkPG39*	0.20323028	0.590618295	0.344097501	T
EkPG39	*EkPG40*	0.4424913	1.725634958	0.256422309	T
EkPG40	*EkPG56*	0.377634387	1.214332245	0.310981108	S
EkPG41	*EkPG42*	0.214947911	0.647887612	0.331767281	T
EkPG43	*EkPG44*	0.496411074	2.669662311	0.185945268	T
EkPG44	*EkPG45*	0.127566741	0.373729111	0.341334773	T
EkPG45	*EkPG46*	0.074229839	0.417504009	0.177794314	T
EkPG46	*EkPG47*	0.07782996	0.384306397	0.202520595	T
EkPG47	*EkPG48*	0.136142843	0.529470621	0.257130117	T
EkPG49	*EkPG50*	0.53405773	2.529258564	0.211151891	T

S: segmental duplicated; T: tandem duplication

Table 3: Duplication models for EkPGs gene pairs

The expression patterns of EkPGs

Transcriptome sequencing was performed on fruits at five developmental stages, which were 50d (Fr_I), 100d (Fr_II), 130d (Fr_III), 160d (Fr_IV), 180d (Fr_V) after full bloom. Expression analysis showed that 30 out of 59 EkPGs were expressed in fruits at different developmental stages, their expression patterns were mainly divided into four categories and the gene expression patterns in the same category are the same (Figure 5).

The five member genes in group A showed the highest expression levels during the Fr_I period and almost no expression during the rest of the period. The expression patterns of the eleven member genes in group B were down-regulated with fruit development, while the expression levels of other members in group C were up-regulated except for *EkPG27*. In group D, *EkPGkPG5, EkPG12, EkPG25, EkPG53* have higher expression in Fr_IV and Fr_V.

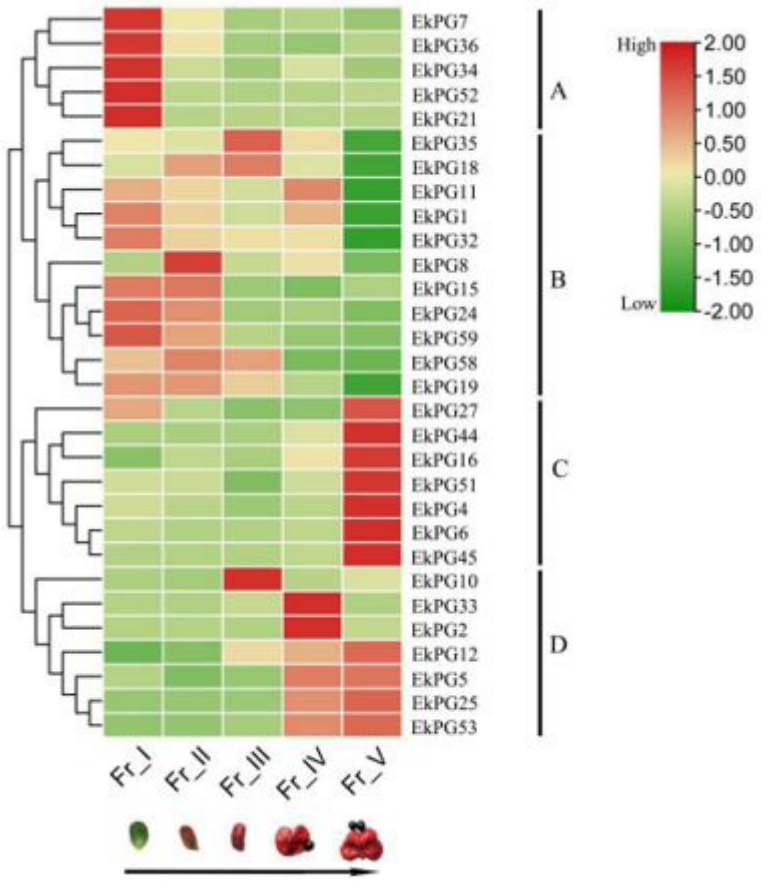

Figure 5: Expression analysis of the EkPGs genes. Five different developmental periods of the fruit: Fr_I Fr_II, Fr_III, Fr_IV, Fr_V

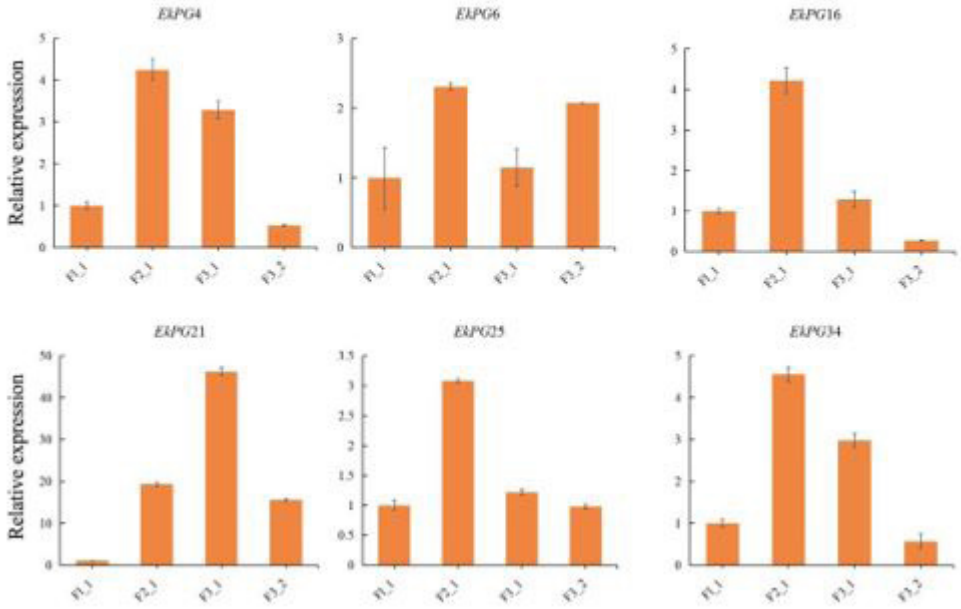

Figure 6: Expression analysis of six EkPGs in the AZ at different periods. Mean values and standard error were obtained from three replicates

qRT-PCR verification of abscission-related genes

We selected six EkPG genes based on gene structure, conserved domains, and functions of homologous genes in *A. thaliana* among the highly expressed genes in the early stage of floral organ abscission and the late stage of fruit ripening [38]. qRT-PCR results showed that (Figure 6), except for *EkPG6* and *EkPG21*, the other four EkPGs had the same expression patterns. The expression levels of *EkPG4*, *EkPG16*, *EkPG25*, and *EkPG34* were all the highest in F2_1, followed by F3_1. And the expression level is relatively low in F1_1 and F3_2. *EkPG6* has the highest expression in F2_1, followed by F3_2, and both were higher than F1_1 and F3_1. *EkPG21* is highly expressed in F3_1 period, while its expression in the F2_1 and F3_2 periods is relatively low, but both are significantly higher than that in F1_1 period.

Discussion

Evolution of EkPG members

PGs are a relatively large gene family and have been extensively studied. They play an important role in fruit softening, organ abscission, and improving plant resistance [39-41]. For example, inhibiting *PaPG1* expression is beneficial to reduce cell wall disintegration and delay fruit softening in apricot (*Prunus armeniaca* L.) during cold storage [42]; In rice, *OsPG1* functions as a suppressor of programmed cell death and affects cell wall integrity, which regulates plant resistance [23]. Based on phylogenetic and evolutionary analysis, many plants has seven groups of PGs (Group A-G), such as tomato (*Solanum lycopersicum*) [43], watermelon (*Cucumis sativus*) [44], and cucumber (*Citrullus lanatus*) [44]. Among them, Group G is generally composed of homologous genes of *AtQRT3* (*At4G20050*) in different species. *AtQRT3* is a special type of PG gene that can be expressed in yeast and mediate polygalacturonase activity, which can regulate the development of *A. thaliana* pollen [12,45]. Our study showed that all EkPGs clustered into seven clades through phylogenetic analysis, and group G contains five homologous genes of *AtQRT*3, compared to only one or two in other species [35].

Gene duplication event is prevalent during plant evolution [47,48]. We found that fragment duplication and tandem duplication mainly contribute to the expansion EkPG gene families. The tandem repetitive events mainly occurred in Group C, D, and F, indicating that the expansion of EkPGs in different clades varies. The substitution rate of non-synonymous (*K*a) and synonymous (*K*s) is the basis for evaluating the positive selection pressure of repeated events [49]. Selection pressure analysis showed that the *K*a/*K*s values of all repeated genes were less than one, suggesting that these genes were under purifying selection. However, *EkPG8-EkPG9*, *EkPG21-EkPG22*, *EkPG26-EkPG27* did not have *K*s values (NaN), which may be due to gene duplication leading to mutations at the nucleic acid level, but the amino acid sequence remains unchanged [50].

Further analysis revealed that the EkPGs of the same group or closely related group have similar gene structures. For example, the length and position of the introns of *EkPG34* and *EkPG53*, *EkPG37* and *EkPG38*, *EkPG46* and *EkPG47* are almost the same. The number of introns of EkPGs of Group A and Group F is between five and eight. Group G has a relatively distant relationship with them, and the number of introns is between one and three. Previous studies showed that evolutionarily conserved bases have longer intron lengths, and there is a positive correlation between evolutionary conservation levels and eukaryotic gene intron lengths [51]. In *E. japonica*, *EkPG4*, *EkPG25*, *EkPG42*, *EkPG45*, *EkPG49*, and *EkPG52* also have longer intron lengths than other genes, indicating that they may be more conserved in the evolutionary process, and involved in more important biological functions.

Conserved motif analysis showed that EkPGs members in the same group usually have similar motif composition. Among them, most EkPGs in Group A, B, C, D, and F clades contain four conserved domains commonly found in PGs. Domains I and II may be part of the catalytic site, domain III is considered to participate in the catalytic reaction, and domain IV may interact with the ionic group of the carboxylic acid group in the substrate [52-55]. Except for *EkPG4*, *EkPG20*, and *EkPG51* which do not contain IV domains, the rest members of Group E contain three conserved domains and lack III conserved domains. This is similar to the results of other species [43,55]. These EkPG members contain I, II and IV conserved domains, indicating that these genes may have catalytic

activity and at the same time have the ability to interact with substrates containing carboxylic acid groups. In Group G, there are three EkPGs containing IV domains, which are *EkPG6*, *EkPG27*, and *EkPG30*. The existence of this typical conserved domain may lead to new functions of EkPG genes in group G. Since protein domains are closely related to their functions, I, II, and IV domains may be more closely related to the identified EkPGs. In general, the similar motif composition and gene structure in the same subgroup strongly support the reliability of subgroup classification.

Validation of abscission-related genes

The analysis of the expression of the five different developmental stages of the *E. japonica* fruit shows that some genes show specific expression at different stages of the fruit. *EkPG7*, *EkPG21*, *EkPG34*, *EkPG36,* and *EkPG52* have the highest expression levels in the Fr_I stage and almost no expression in the rest of the period. During Fr_I period, the floral organs dropped and young fruits gradually formed. Therefore, it is speculated that these genes may be related to the abscission of floral organs. The expression levels of *EkPG4*, *EkPG6*, *EkPG16*, *EkPG44*, *EkPG45*, and *EkPG51* were the highest during Fr_V period, and the remaining genes were almost not expressed during the Fr_I to Fr_IV period. *EkPG5*, *EkPG12*, *EkPG25*, *EkPG53* have higher expressions in the later stages of fruit development. The expression level of these genes is positively associated with the fruit development stages, therefore it is suggested that these genes may be related to the fruit abscission of *E. japonica*.

To date, nine unique expression patterns of PGs have been found in the study of the process of floral organ abscission in *A. thaliana* [38]. In the process of flower abscission, high levels of gene expression may be related to cell separation or flower organ abscission. In our qRT-PCR quantitative analysis, *EkPG4*, *EkPG16*, *EkPG25*, and *EkPG34* all have high expression in F2_1 period, indicating that they may be involved in the shedding of floral organs. The expression pattern of *EkPG6* at different stages is different from the other EkPGs. It showed high expression level when young fruits are about to fall off, suggesting it may be related to the fruit abscission. The homologous gene *AtQRT*3 of *EkPG6* is a special type of PG gene, which does not contain the four characteristic conserved domains of the PG gene[34]. However, *EkPG6* contains the IV domain, which may lead to new functions of *EkPG6* in other tissues.

Conclusion

In summary, this study identified and analyzed the members of the PG family through the genome and transcriptome data of *E. japonica*. We identified 59 PG genes that are distributed on 11 chromosomes. Conserved domains and phylogenetic analysis confirmed the similarity and evolutionary relationship between the PG gene of *E. japonica* and *A. thaliana*. The deletion and addition of conserved domains may lead to the addition of new functions of EkPG on different clades. qRT-PCR results showed that *EkPG6* may be related to the abscission of flowers and fruits of *E. japonica*. And *EkPG4, EkPG16, EkPG25, EkPG34* may involve in the abscission of floral organs. This work provides a foundation for further research on the regulation mechanism of EkPGs in the process of flower and fruit abscission, a critical next step will be the functional verification of these related genes in *E. japonica*.

Data availability

The raw transcriptomes data of *E. japonica* have been deposited in BioProject/GSA (https://bigd.big.ac.cn/gsa.) under the accession codes PRJCA005298/CRA004272.

Author Contributions

S.Z. designed the research. P.Z. performed experiments and wrote the article. L.X., H.Z., Y.L. and M.W. performed the bioinformation analysis. S.Z. and S.L. revised and polished this manuscript. All authors have read and agreed to the published version of the manuscript.

Funding

This research was funded by the Fujian Provincial Department of Science *E. japonica* Evolution and Selection of Ornamental Medicinal Resources, China (No. 2020N5004), the Project of Forestry Peak Discipline at Fujian Agriculture and Forestry University, China (No. 712018007), and the Collection, Development and Utilization of *Eascaphis konlshli* Germplasm Resources, China (No. KSYLC004)

Conflicts of Interest

The authors declare no conflict of interest

References

1. XS Ma, CQ Li, Y Yuan, ML Zhao, JG Li (2021) Xyloglucan endotransglucosylase/hydrolase genes *LcXTH4/7/19* are involved in fruitlet abscission and are activated by *LcEIL2/3* in litchi, Physiol. Plant. 173: 1136-46.

2.ZH González-Carranza, E Lozoya-Gloria, JA Roberts (1998) Recent developments in abscission: shedding light on the shedding process, Trends Plant Sci. 3: 0-14.

3. Y Deng, Y Wu, YF Li, MD Yang, CB Shi, et al. (2007) Studies of postharvest berry abscission of 'Kyoho' table grapes during cold storage and high oxygen atmospheres, Postharvest Biol. Tec. 43: 95-101.

4. YQ Yu, EA Kellogg (2020) Inflorescence abscission zones in grasses: diversity and genetic regulation, Annual Plant Reviews 1: 497-532.

5. R Wang, RZ Li, LN Cheng, XY Wang, X Fu, et al. (2021) SlERF52 regulates *SlTIP1*;1 expression to accelerate tomato pedicel abscission, plant physiol. 185: 1829-46.

6. F Yan, Zh Gong, GJ Hu, XS Ma, RY Bai, et al. (2021) Wuriyanghan, Tomato *SlBL4* plays an important role in fruit pedicel organogenesis and abscission, Hortic Res 8: 78.

7. T Nakano, Y Ito (2013) Molecular mechanisms controlling plant organ abscission, Plant Biotechnol. 30: 209-16.

8. RJ Xie, L Deng, L Jing, SL He, YT Ma, et al. (2013) Recent advances in molecular events of fruit abscission, Biol Plantarum 57: 201-9.

9. PW Wu, FY Xin, HJL Xu, YY Chu, YL Du, et al. (2021) Chitosan inhibits postharvest berry abscission of 'Kyoho' table grapes by affecting the structure of abscission zone, cell wall degrading enzymes and SO2 permeation, Postharvest Biol Tec 176: 111507.

10. U Mahmood, YH Fan, SY Wei, Y Niu, YH Li, et al. (2020) Comprehensive analysis of polygalacturonase genes offers new insights into their origin and functional evolution in land plants, Genomics 33: 1096-108.

11. M Ogawa, P Kay, S Wilson, SM Swain (2009) ARABIDOPSIS DEHISCENCE ZONE POLYGALACTURONASE1 (ADPG1), ADPG2, and QUARTET2 Are Polygalacturonases Required for Cell Separation during Reproductive Development in *Arabidopsis*, Plant Cell 21: 216-33.

12. YJ Yu, Y Liang, ML Lv, J Wu, G Lu, et al. (2014) Genome-wide identification and characterization of polygalacturonase genes in *Cucumis sativus* and *Citrullus lanatus*, Plant Physiol. Bioch 74: 263-75.

13. P Roongsattham, Fe Morcillo, C Jantasuriyarat, Me Pizot, S Moussu, et al. (2012) Temporal and spatial expression of polygalacturonase gene family members reveals divergent regulation during fleshy fruit ripening and abscission in the monocot species oil palm, BMC 12: 150.

14. T Ge, X Huang, XT Pan, J Zhang, RJ Xie (2019) Genome-wide identification and expression analysis of citrus fruitlet abscission-related polygalacturonase genes. 3 Biotech 9: 250.

15. P Kalaitzis, T Solomos, ML Tucker (1997) Three Different Polygalacturonases Are Expressed in Tomato Leaf and Flower Abscission, Each with a Different Temporal Expression Pattern, Plant Physiol. 113: 1303-8.

16. C Paniagua, P Ric-Varas, JA García-Gago, G López-Casado, R Blanco-Portales, et al. (2020) Elucidating the role of polygalacturonase genes in strawberry fruit softening, J. Exp. Bot 7103-17.

17. K Hiwasa, Y Kinugasa, S Amano, A Hashimoto, R Nakano, et al. (2003) Ethylene is required for both the initiation and progression of softening in pear (*Pyrus communis* L.) fruit, J Exp Bot. 54: 771-9.

18. MH Asif, P Nath (2005) Expression of multiple forms of polygalacturonase gene during ripening in banana fruit, Plant Physiol Biochem, 43: 177-84.

19. M Qian, Z Xu, ZH Zhang, Q Li, XY Yan, et al. (2021) The downregulation of *PpPG21* and *PpPG22* influences peach fruit texture and softening, planta, 254: 22.

20. Y Rui, CW Xiao, HJ Yi, B Kandemir, JZ Wang, et al. (2017) POLYGALACTURONASE INVOLVED IN EXPANSION3 Functions in Seedling Development, Rosette Growth, and Stomatal Dynamics in *Arabidopsis thaliana*, The Plant cell, 29: 2413-32.

21. CW Xiao, WJ Barnes J, MS Zamil, H Yi, VM Puri, (2017) Activation tagging of *Arabidopsis POLYGALACTURONASE INVOLVED IN EXPANSION2* promotes hypocotyl elongation, leaf expansion, stem lignification, mechanical stiffening, and lodging, Plant J 89: 1159-73.

22. CW Xiao, C Somerville, CT Anderson (2014) POLYGALACTURONASE INVOLVED IN EXPANSION1 functions in cell elongation and flower development in *Arabidopsis*, The Plant cell 26: 1018-35.

23. YR Cao, Y Zhang, YY Chen, N Yu, S Liaqat, et al. (2021) *OsPG1* encodes a polygalacturonase that determines cell wall architecture and affects resistance to bacterial blight pathogen in rice, Rice, 14: 36.

24. WH Sun, XY Yuan, ZJ Liu, SR Lan, WC Tsai, et al. (2019) Multivariate analysis reveals phenotypic diversity of *Euscaphis japonica* population, PLoS One 14: e0219046.

25. CL Yang, XY Yuan, J Zhang, WH Sun, ZJ Liu, et al. (2020) Comprehensive transcriptome analysis of reference genes for fruit development of *Euscaphis konishii*, Peer J 8: e8474.

26. S Xiang, XD Liu, WH Sun, SR Lan, ZJ Liu, et al. (2019) The complete chloroplast genome sequence of *Euscaphis japonica* (Staphyleaceae), Mitochondrial DNA Part B 4: 3484-5.

27. WH Sun, Z Li, S Xiang, L Ni, DY Zhang, et al (2021) The *Euscaphis japonica* genome and the evolution of malvids, Plant J.

28. I Letunic, T Doerks, P Bork (2012) Smart 7: Recent updates to the protein domain annotation resource, Nucleic Acids Res. 40: 302-5.

29. KB Yang, Y Li, SN Wang, XR Xu, HY Sun, et al. (2019) Genome-wide identification and expression analysis of the MYB transcription factor in moso bamboo (*Phyllostachys edulis*), Peer J 6.

30. S Kumar, G Stecher, K Tamura (2016) MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets, Mol Biol Evol. 33.

31. WJ Huang, MY Chen, TT Zhao, F Han, Q Zhang, et al. (2020) Genome-Wide Identification and Expression Analysis of Polygalacturonase Gene Family in Kiwifruit (*Actinidia chinensis*) during Fruit Softening, Plants 9: 327.

32. YP Wang, HB Tang, JD DeBarry, X Tan, JP Li, et al. (2012) Mcscanx: A toolkit for detection and evolutionary analysis of gene synteny and collinearity, Nucleic Acids Res 40: e49.

33. TL Bailey, M Boden, FA Buske, M Frith, CE Grant, et al. (2009) MEME SUITE: tools for motif discovery and searching, Nucleic Acids Res 1-7.

34. CJ Chen, H Chen, Y Zhang, HR Thomas, MH Frank, et al. (2020) TB tools - an integrative toolkit developed for interactive analyses of big biological Data, Mol. Plant 13.

35.WX Liang, XX Zou, R Carballar-Lejaraz, LJ Wu, WH Sun, et al. (2018) Selection and evaluation of reference genes for qRT-PCR analysis in *Euscaphis konishii* Hayata based on transcriptome data, Plant Methods 14: 12.

36. JM Li, MH Zhang, J Sun, XR Mao, J Wang, et al. (2019) Genome-Wide Characterization and Identification of Trihelix Transcription Factor and Expression Profiling in Response to Abiotic Stresses in Rice (*Oryza sativa* L.), Int. J. Mol. Sci. 20: 251.

37. KE Francis, SY Lam, GP Copenhaver (2006) Separation of *Arabidopsis* Pollen Tetrads Is Regulated by Quartet1, a Pectin Methylesterase Gene, Plant Physiol. 142: 1004-13.

38. J Kim, SH Shiu, S Thoma, WH Li, SE Patterson (2006) Patterns of expansion and expression divergence in the plant polygalacturonase gene family, Genome Biol 7: R87.

39. KA Hadfield, AB Bennett (1998) Polygalacturonases: Many Genes in Search of a Function. Plant Physiol. 117: 337-43.

40. AJ Bowling, KC Vaugh (2011) Leaf Abscission in *Impatiens* (Balsaminaceae) is due to loss of highly de-esterified homogalacturonans in the middle lamellae, Am J Bot 98: 619-29.

41. G Peng, JY Wu, WJ Lu, JG Li (2013) A polygalacturonase gene clustered into clade E involved in lychee fruitlet abscission, Sci. Hortic. 150: 244-50.

42. YY Hou, F Wu, YT Zhao, L Shi, X Zhu (2019) Cloning and expression analysis of polygalacturonase and pectin methylesterase genes during softening in apricot (*Prunus armeniaca* L.) fruit. Sci. Hortic 256: 108607.

43. XB Ke, HS Wang, Y Li, B Zhu, YX Zang, et al. (2018) Genome-Wide Identification and Analysis of Polygalacturonase Genes in *Solanum lycopersicum*, Int. J. Mol. Sci. 19: 2290.

44. SY Rhee, E Osborne, PD Poindexter, CR Somerville (2003) Microspore Separation in the quartet 3 Mutants of *Arabidopsis* Is Impaired by a Defect in a Developmentally Regulated Polygalacturonase Required for Pollen Mother Cell Wall Degradation, Plant Physiol. 133: 1170-80.

45. ZH Gonzalez-Carranza, KA Elliott, JA Roberts (2007) Expression of polygalacturonases and evidence to support their role during cell separation processes in *Arabidopsis thaliana*, J. Exp. Bot. 58: 3719-30.

46. ML Lyu, J Iftikhar, RF Guo, BH Wu, JS Cao (2020) Patterns of Expansion and Expression Divergence of the Polygalacturonase Gene Family in *Brassica oleracea*, Int. J. Mol. Sci. 21: 5706.

47. SB Cannon, A Mitra, A Baumgarten, Ne D Young, G May (2004) The roles of segmental and tandem gene duplication in the evolution of large gene families in *Arabidopsis thaliana*, BMC Plant Biol 4: 10.

48. WJ Kent, R Baertsch, A Hinrichs, W Miller, D Haussler (2003) Evolution's cauldron: duplication, deletion, and rearrangement in the mouse and human genomes, PNAS 100: 11484-9.

49. ZH Yang (2007) PAML 4: phylogenetic analysis by maximum likelihood, Mol Biol Evol. 24: 1586-91.

50. J Wang, Y Liu, BQ Tang, XZ Dai, LL Xie, et al. (2020) Genome-Wide identification and capsaicinoid biosynthesis-related expression analysis of the R2R3-MYB gene family in *Capsicum annuum* L., Front. Genet. 11: 598183.

51. O Gorlova1, A Fedorov, C Logothetis, C Amos, I Gorlov (2014) Genes with a large intronic burden show greater evolutionary conservation on the protein level, BMC 14: 50.

52. M Torki, P Mandaron, RG Mache, D Falconet (2000) Characterization of a ubiquitous expressed gene family encoding polygalacturonase in *Arabidopsis thaliana*, Gene 242: 427-36.

53. WH Li, J Yang, X Gu (2005) Expression divergence between duplicate genes, Trends Genet. 21: 602-7.

54. MN Rao, AA Kembhavi, A Pant (1996) Implication of tryptophan and histidine in the active site of endo-polygalacturonase from *Aspergillus ustus*: elucidation of the reaction mechanism, BBA 1296: 167-73.

55. Y Liang, YJ Yu, XP Shen, H Dong, ML Lyu, et al. (2015) Dissecting the complex molecular evolution and expression of polygalacturonase gene family in *Brassica rapa* ssp. Chinensis, Plant Mol. Biol 89: 629-46.