Transcriptomic analysis reveals ethylene signal transduction genes involved in pistil development of pumpkin

Qingfei Li [Corresp., 1, 2], Li Zhang [1, 2], Feifei Pan [1, 2], Weili Guo [1, 2], Bihua Chen [1, 2], Helian Yang [1, 2], Guangyin Wang [1, 2], Xinzheng Li [Corresp. 1, 2]

[1] College of Horticulture and Landscape,Henan Institute of Science and Technology, Xin Xiang, China

[2] Henan Province Engineering Research Center of Horticultural Plant Resource Utilization and Germplasm Enhancement, Xin Xiang, China

Corresponding Authors: Qingfei Li, Xinzheng Li
Email address: lqf1988@hist.edu.cn, liuzhw@hist.edu.cn

Development of female flowers is an important process that directly affects the yield of *Cucubits*. Little information is available on the sex determination and development of female flowers in pumpkin, a typical monoecious plant. In the present study, we used aborted and normal pistils of pumpkin for RNA-Seq analysis and determined the differentially expressed genes (DEGs) to gain insights into the molecular mechanism underlying pistil development in pumpkin. A total of 3817 DEGs were identified, among which 1341 were upregulated and 2476 were downregulated. The results of transcriptome analysis were confirmed by real-time quantitative RT-PCR. KEGG enrichment analysis showed that the DEGs were significantly enriched in plant hormone signal transduction and phenylpropanoid biosynthesis pathway. Eighty-four DEGs were enriched in the plant hormone signal transduction pathway, which accounted for 12.54% of the significant DEGs, and most of them were annotated as predicted ethylene responsive or insensitive transcription factor genes. Furthermore, the expression levels of four ethylene signal transduction genes in different flower structures (female calyx, pistil, male calyx, stamen, leaf, and ovary) were investigated. The ethyleneresponsive DNA binding factor, *ERDBF3*, and ethylene responsive transcription factor, *ERTF10*, showed the highest expression in pistils and the lowest expression in stamens, and their expression levels were 78- and 162-times more than that in stamens, respectively. These results suggest that plant hormone signal transduction genes, especially ethylene signal transduction genes, play an important role in the development of pistils in pumpkin. Our study provides a theoretical basis for further understanding of the mechanism of regulation of ethylene signal transduction genes in pistil development and sex determination in pumpkin.

1   **Transcriptomic analysis reveals ethylene signal transduction genes**
2   **involved in pistil development of pumpkin**

3   **Qing-Fei Li[1,2*], Li Zhang[1,2], Fei-Fei Pan[1,2], Wei-Li Guo[1,2], Bi-Hua Chen[1,2], He-Lian Yang[1,2],**

4   **Guang-Yin Wang[1,2], Xin-Zheng Li[1,2*]**

5   [1] College of Horticulture and Landscape, Henan Institute of Science and Technology, Xinxiang, China;
6   [2] Henan Province Engineering Research Center of Horticultural Plant Resource Utilization and Germplasm
7   Enhancement, Xinxiang, China
8   [*] **Correspondence:**
9   Xin-Zheng Li, Qing-Fei Li
10
11   Email address: liuzhw@hist.edu.cn; lqf1988@hist.edu.cn

12   **Abstract**

13   Development of female flowers is an important process that directly affects the yield of *Cucubits*.
14   Little information is available on the sex determination and development of female flowers in
15   pumpkin, a typical monoecious plant. In the present study, we used aborted and normal pistils of
16   pumpkin for RNA-Seq analysis and determined the differentially expressed genes (DEGs) to
17   gain insights into the molecular mechanism underlying pistil development in pumpkin. A total of
18   3817 DEGs were identified, among which 1341 were upregulated and 2476 were downregulated.
19   The results of transcriptome analysis were confirmed by real-time quantitative RT-PCR. KEGG
20   enrichment analysis showed that the DEGs were significantly enriched in plant hormone signal
21   transduction and phenylpropanoid biosynthesis pathway. Eighty-four DEGs were enriched in the
22   plant hormone signal transduction pathway, which accounted for 12.54% of the significant DEGs,
23   and most of them were annotated as predicted ethylene responsive or insensitive transcription
24   factor genes. Furthermore, the expression levels of four ethylene signal transduction genes in
25   different flower structures (female calyx, pistil, male calyx, stamen, leaf, and ovary) were
26   investigated. The ethylene responsive DNA binding factor, *ERDBF3*, and ethylene responsive
27   transcription factor, *ERTF10*, showed the highest expression in pistils and the lowest expression
28   in stamens, and their expression levels were 78- and 162-times more than that in stamens,
29   respectively. These results suggest that plant hormone signal transduction genes, especially
30   ethylene signal transduction genes, play an important role in the development of pistils in
31   pumpkin. Our study provides a theoretical basis for further understanding of the mechanism of
32   regulation of ethylene signal transduction genes in pistil development and sex determination in
33   pumpkin.

34
35   **Key words: Pumpkin, Pistil development, Transcriptome, Ethylene, Plant hormone signal**
36   **transduction, Phenylpropanoid**
37

38   **Introduction**

39 Pumpkin (*Cucurbita moschata* Duch.) is a typical monoecious plant with distinct male and
40 female flowers. The number and proportion of female flowers can directly influence yields and
41 economic benefits of *Cucurbitaceae* crops. Pistil is the main characteristic structure of female
42 flowers. It is well established that ethylene promotes pistil and female flower development.
43 Application of ethylene, or inhibition of ethylene action, increases or decreases the number of
44 pistil-bearing buds (Papadopoulou et al., 2005). Ethylene promotes carpel development and
45 arrests stamen development in female flower (Chen et al., 2016). Treatment with exogenous
46 ethylene or ethylene releasing reagents can increase the numbers of female and bisexual flowers
47 in monoecious and andromonoecious lines, respectively (Iwahori et al., 1969;Malepszy and
48 Niemirowicz-Szczytt, 1991;Manzano et al., 2011). Auxin response factors (ARFs), an important
49 component in auxin signalling pathway, especially the *ARF13* and *ARF17* genes are essential for
50 pistil development in Japanese apricot (Song et al., 2015). The development of female flowers is
51 inseparable from sex differentiation in *Cucurbitaceae* crops.

52   A number of studies have found that the floral development and sex expression in
53 *Cucurbitaceae* crops can be affected by multiple phytohormones, including ethylene, auxin,
54 cytokinin, gibberellin, abscisic acid, brassinosteroid, jasmonic acid, and salicylic acid (Rudich
55 and Halevy, 1974;Trebitsh et al., 1987;Yamasaki et al., 2003;Menéndez et al., 2009;Pimenta
56 Lange and Lange, 2016;Mao et al., 2017;Zhang et al., 2017). Ethylene might be the major
57 hormone in sex determination. In cucumber, sex differentiation is mainly determined by the *F*
58 (*CsACS1G*), *M* (*CsACS2*), and *A* (*CsACS11*) genes. The *F*, *M*, and *A* genes encode 1-
59 aminocyclopropane-1-carboxylate (ACC) synthase (ACS), which is a key rate-limiting enzyme
60 in ethylene biosynthetic pathway (Pierce and Wehner, 1990;Pan et al., 2018). Among these, the
61 *F* gene promotes the development of female flowers (Mibus and Tatlioglu, 2004;Knopf and
62 Trebitsh, 2006), and the *M* gene inhibits the development of stamens (Yamasaki et al.,
63 2001;Saito et al., 2007). The *CsACS11*(*A*) is an androecious gene, and mutants with loss of
64 *CsACS11* function were found to be androecious, with no female flowers (Boualem et al., 2015).
65 The ACC oxidase gene (*ACO*), another key gene in ethylene biosynthesis, is also essential for
66 the development of female flowers. A transcription factor gene, *CsWIP1*, could directly bind the
67 promoter of *CsACO$_2$* to repress its expression (Chen et al., 2016). In melon, sex determination is
68 governed by the genes, andromonoecious (a) and gynoecious (g). *CmACS-7* is the
69 andromonoecious gene (Boualem et al., 2008). *CmACS11* controls the development of female
70 flowers in melon and its function is not exactly the same as of *CsACS11*.*CmWIP1*, the ortholog
71 of *CsWIP1*, can negatively regulate the femaleness and gynoecious plants are obtained when
72 there is a loss of function of *CmWIP1*. The expression patterns of *CmWIP1* and *CmACS11* are
73 opposite in melon. *CmACS11* can repress the expression of *CmWIP1* at an upstream step in the
74 sex determination pathway to control flower development. Ethylene could inhibit stamen
75 development through *CmACS-7* and *CmACS11* (Boualem et al., 2015). Thus, the sex
76 differentiation and floral development in melon could be a result of the interaction among
77 *CmACS11*, *CmWIP1*, and *CmACS-7*.

78   Besides the genes involved in ethylene synthesis, those involved in ethylene signal

79   transduction have also been implicated in the development of female flowers in cucumber.

80   Previous studies have revealed that there existed organ-specific DNA damage in primordial

81   anther of female flowers and that the DNA damage was induced via the ethylene signaling

82   pathway. The ethylene-receptor gene, *CsETR1*, located in the pistil primordia has also been

83   involved in the arrest of stamen development through induction of DNA damage in female

84   flowers (Hao et al., 2003;Yamasaki et al., 2003;Duan et al., 2008;Wang et al., 2010). It can be

85   bound and activated by *CsAP3 in vitro* and *in vivo* (Sun et al., 2016). Moreover, the expression

86   of C_S-*ETR2* and C_S-*ERS* was reported to be regulated by ethylene, because their mRNAs were

87   significantly elevated by the application of ethrel and their levels were lowered by the

88   application of an ethylene inhibitor, aminoethoxyvinyl glycine (Yamasaki et al., 2000). Thus, not

89   only the genes involved in ethylene synthesis, those involved in ethylene-mediated signal

90   transduction contribute to the development of female flowers. However, the evidence for how

91   the ethylene signal transduction genes regulate the development of female flowers is still weak.

92   Besides ethylene, other phytohormones are also involved in the development of female and

93   male flowers. Gibberellins (GA) can promote the male tendency. Its production in

94   andromonoecious cucumber is higher than that in monoecious and gynoecious cucumber

95   (Hemphill et al., 1972). A recent report indicated that GA could regulate sex expression via

96   ethylene-dependent and ethylene-independent pathways (Zhang et al., 2017). The application of

97   indole-3-acetic acid (IAA), except for the Beta -Alfa type, could enhance ethylene and ACC

98   production (Trebitsh et al., 1987). Cytokinins are possibly involved in determining the

99   morphological differences between sex types. The endogenous levels of the cytokinins were

100   found to be higher in female gametophytes than in male gametophytes (Menéndez et al., 2009).

101   Abscisic acid could promote the female tendency of gynoecious plants. It participates in the sex

102   regulation by inhibiting the GA activity of cucumber (Rudich and Halevy, 1974). Jasmonic acid

103   signaling also plays an important role in flower development in plants, especially in stamen

104   sterility, sex determination, female flower development, and seed maturation (Yuan and Zhang,

105   2015;Mao et al., 2017).

106   As evident from the above studies, floral development and sex differentiation is a result of the

107   interaction of various plant hormones and ethylene apparently plays a key role in these processes.

108   However, little is known about how the plant hormone signal transduction genes are involved in

109   female and male flower development. Herein, to explore the key genes involved in the

110   development of pistils, the aborted and normal pistils of pumpkin were used for RNA-Seq

111   analysis. The results showed that plant hormone signal transduction genes, especially ethylene

112   related genes, play important roles in the development of pistils in pumpkin. Fourteen DEGs,

113   which were annotated as plant hormone signal transduction genes, were chosen for qRT-PCR

114   verification. Furthermore, four ethylene signal transduction genes were appealing candidates for

115   investigation of their expression in different flower structures. Our results provide a foundation

116   for dissecting the molecular mechanism of regulation of pistil formation in pumpkin by the

117   candidate genes, identified herein.

Materials and Methods

Plant materials

120   The aborted plants were from the same inbred line as the normal plants and exhibited similar
121   plant characters except for the female development. The seeds of pumpkin were sterilized with
122   hot water (55 °C) for approximately 15 min, and then soaked in water for 4 h. The seeds were
123   germinated on a wet filter paper in a Petri dish at 28°C in the dark. When at least 80% of the
124   seeds had germinated, they were transferred into a substrate mixture of peat, vermiculite, and
125   perlite (3:1:0.5, v/v) and grown in an artificial climate room under conditions of 14-h light/10-h
126   dark and day/night temperatures of 25 °C/18 °C. When the seedlings had developed three true
127   leaves, they were transplanted to a field and grown under the natural environment. At flower
128   development period, the aborted female flowers exhibited different from those of the normal
129   plants. When the length of corolla of flower bud was less than 5 mm, there was no difference
130   between the aborted female flower and the normal one. As the buds grow bigger, in aborted
131   pistils the incomplete stigma can be seen when the corolla is peeled off, and ovary looks thinner
132   than in normal pistils, which is as shown in *Fig. 1A*. Thus, the aborted and normal pistils, which
133   were come from the female flower buds with 5 mm length of corolla and located in the shoot
134   apices, were used for RNA-Seq analysis and for the validation of gene expression. Moreover,
135   leaf, female calyx, pistil, ovary, male calyx, and stamen from the male and female flowers of the
136   same plants, were used to detect the expression of the candidate genes in different floral
137   structures. All the samples were immediately frozen in liquid nitrogen and stored at -80 °C for
138   further experiments.

RNA isolation

140   Total RNA was extracted using the Trizol™ reagent (Invitrogen, Carlsbad, CA, USA), as
141   described by the manufacturer. DNase I was then used to remove DNA contamination in the
142   RNA preparations. The purity and integrity of RNA samples were evaluated by electrophoresis
143   on 1% RNase-free agarose gels, spectrophotometry (A260/A280 and A260/A230) on NanoDrop
144   DU8000, and analysis using the Agilent 2100 Bioanalyzer system (Agilent, USA).

Construction and sequencing of cDNA libraries

146   The NEBNext Ultra Directional RNA Library Prep Kit for Illumina (NEB, Ispawich, USA) was
147   used for mRNA fragmentation. In brief, poly (A) mRNA was extracted from total RNA using
148   oligo-dT magnetic beads, and further fragmented. First strand cDNA was synthesized using

149 random hexamer primers and reverse transcriptase. Second strand cDNA was synthesized using
150 dNTPs, DNA polymerase I, RNase H, and the first strand. To select the fragments that were
151 preferentially 150–200 bp in length, the cDNA fragments were purified with an AMPure XP
152 system (Beckman Coulter, USA). The high-quality libraries were then sequenced on the Illumina
153 HiSeq 4000 platform by Sagene Co. (Guangzhou, China). The libraries of two biological
154 replicates of aborted and normal pistils were prepared independently. The raw transcriptome data
155 were deposited in the National Center for Biotechnology Information Sequence Read Archive
156 under BioProject number PRJNA554766.

157 **Bioinformatics analysis of differentially expressed genes (DEGs) data**

158 The raw reads were pre-processed by removing the adaptor sequences, low-quality reads (more
159 than 50% bases with SQ≤20 in one sequence), and reads with more than 5% N bases (bases
160 unknown). The clean reads were mapped to the *Cucurbita moschata* genome
161 (http://cucurbitgenomics.org/organism/9) using TopHat2 v2.1.1   (Trapnell et al., 2012;Kim et
162 al., 2013;Sun et al., 2017), allowing up to two mismatch. Functional annotation was performed
163 by searching against the non redundant (NR), Swiss–Prot and clusters of orthologous groups for
164 eukaryotic complete genomes (KOG) databases using BLAST with an E-value of $1e^{-5}$. The R
165 package edgeR was used to identify the DEGs using raw read counts as input data (Robinson et
166 al., 2010). P-values were adjusted using Benjamini and Hochberg's method to control the false
167 discovery rate (FDR) (Benjamini and Hochberg, 1995). The |log2 (fold change)| > 1 and adjusted
168 P-value < 0.05 for multiple tests using the Benjamini method were used as significance cut-offs
169 for the expression differences.

170 **Functional enrichment analysis for DEGs**

171 Gene functional enrichment analysis of DEGs was implemented using the GOSeq R package
172 (Young et al., 2010). GO terms included cellular component, molecular function, and biological
173 process. After the hypergeometric test, Bonferroni correction was employed for P-value
174 correction with a cut-off of 0.05. The GO terms satisfying the condition were considered
175 significantly enriched by DEGs. Furthermore, the Kyoto Encyclopedia of Genes and Genomes
176 (KEGG) enrichment analysis was performed. KEGG is the main public database of pathways.
177 FDR control method was used to identify the threshold of the P-value in multiple tests
178 (Benjamini and Hochberg, 1995). Pathways with their Benjamini and Hochberg adjusted P-
179 values≤0.05 were defined as significantly enriched by DEGs (Mao et al., 2005;Kanehisa et al.,
180 2014). Using significant enrichment of pathways, the major biochemical metabolic pathways and
181 signal transduction pathways involving DEGs can be determined.

182 **Quantitative real-time RT-PCR**

183 RNA extraction and detection were done as described above. cDNA synthesis was conducted
184 using the PrimeScript™ RT Master Mix (Perfect Real Time) Reagent Kit (Takara, Dalian, China),
185 according to the manufacturer's instructions. The gene primers used in the quantitative real-time

186   RT-PCR (qRT-PCR) experiment are listed in *Supplemental Table 1*. *ACTIN* was used as an
187   internal control. The qRT-PCR system was consisted of 10 µL SYBR Premix Ex Taq™ II, 0.8
188   µL each primer (10 µM), 2 µL diluted cDNA (150 ng), and 6.4 µL nuclease-free water. The qRT-
189   PCR reactions were performed in three biological replicates. The qRT-PCR was carried out
190   using a Bio-Rad IQ5 instrument (Foster City, CA, USA), with the following conditions: 1 cycle
191   of 95 °C for 40 s, 40 cycles of 95 °C for 5 s and 61 °C for 30 s. The relative expression levels
192   were calculated using the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001). To verify the RNA-Seq
193   data, the correlation analysis and the Pearson correlation coefficient between log2 (fold change)
194   of RNA-Seq and qRT-PCR was calculated using the IBM SPSS statistics 22 software.

Results

Summary of the sequencing data

197   After filtering, 5.62, 6.60, 5.86, and 5.72 million high quality clean reads for aborted (A2, A4)
198   and normal (N2, N4) pistils were obtained. Among these high quality clean reads, the
199   percentages of Q20 were 97.46%, 97.58%, 96.61%, and 97.45%, respectively, and the
200   percentage of GC were 45.55%, 45.49%, 45.77%, and 45.81%, respectively. The total pair reads,
201   mapped pair reads (ratio) of these libraries according to the *Cucurbita moschata* genome, known
202   gene number, and number of novel genes were as shown in *Table 1*. Finally, 3817 DEGs were
203   identified, including 1341 up-regulated and 2476 down-regulated genes, in the aborted pistils
204   compared with the normal pistils (*Fig. 1B, C, Supplemental Table 2*).

GO enrichment analysis

206   The DEGs were enriched in cellular component, molecular function, and biological process
207   groups. A total of 749 DEGs were categorized into biological process category, which mainly
208   included metabolic process (213 up-regulated, 345 down-regulated), cellular process (179 up-
209   regulated, 307 down-regulated), and single-organism process (151 up-regulated, 235 down-
210   regulated) (*Fig. 2*). Among these, the most significantly enriched GO terms included single-
211   organism metabolic process (256, *P* = 9.21E-12), hydrogen peroxide metabolic process (16, *P* =
212   7.80E-07), phenylpropanoid metabolic process (10, *P* = 4.34 E-04), and carbohydrate metabolic
213   process (49, *P* = 4.41E-04). 580 DEGs were divided into the cellular component category, and
214   they mainly enriched in membrane (140 up-regulated, 260 down-regulated), membrane part (111
215   up-regulated, 188 down-regulated), and cell (94 up-regulated, 146 down-regulated) terms (*Fig.*
216   *2*). The most significantly enriched GO terms in the cellular component category included
217   extracellular region (19, *P* = 5.01E-05), membrane (400, *P* = 3.94E-04), photosystem II (6, *P* =
218   5.37E-03), and external encapsulating structure (23, *P* = 9.34E-03). A total of 762 DEGs were
219   divided into molecular function, which mainly enriched in the catalytic activity (205 up-
220   regulated, 362 down-regulated), binding (178 up-regulated, 235 down-regulated), and transporter
221   activity (26 up-regulated, 35 down-regulated). The most significantly enriched GO terms in the
222   molecular function category included oxidoreductase activity (152, *P* = 1.47E-17), tetrapyrrole

223 binding (56, P = 1.69E-12), ion binding (186, P = 6.80E-05), and cation binding (167, P =1.96E-
224 04) (*Fig. 2*). On the whole, in cellular component, molecular function, and biological process
225 groups, the down-regulated genes were all more than up-regulated ones.

226 **Plant hormone signal transduction and phenylpropanoid biosynthesis involved in pistil**
227 **development in pumpkin**

228 The KEGG pathway enrichment analysis indicated that the DEGs were significantly enriched in
229 the following pathways: phenylpropanoid biosynthesis, plant hormone signal transduction,
230 phenylalanine metabolism, tryptophan metabolism, linoleic acid metabolism (*Fig. 3*). It showed
231 that 84 DEGs were enriched in the plant hormone signal transduction pathway (*Supplemental*
232 *Table 3*), which accounted for 12.54% of all the pathway annotated DEGs. The plant hormones
233 involved in this pathway included ethylene, auxin, cytokinin, gibberellin, abscisic acid,
234 brassinosteroid, jasmonic acid, and salicylic acid. Most of DEGs enriched in the plant hormone
235 signal transduction pathway were annotated as predicated ethylene-responsive transcription
236 factor, ethylene insensitive transcription factor, and ethylene receptor genes in NCBI blast and
237 *Cucurbita moschata* genome (*Supplemental Table 4*, *Supplemental Fig. 1*).

238   In addition, we found that 68 DEGs were enriched in phenylpropanoid biosynthesis
239 (*Supplemental Table 5*), which accounted for 10.15% of all the pathway annotated DEGs.
240 Among these, seven DEGs were defined as phenylalanine ammonia-lyase genes (*PAL*), which
241 coded for the key first rate-limiting enzyme in the phenylpropanoid pathway. It is worth noting
242 that six *PAL* DEGs were downregulated in the aborted pistils (*Table 2*). The results indicated that
243 the phenylpropanoid pathway, especially *PAL*, might be implicated in pistil development in
244 pumpkin.

245 **Verification of RNA-seq data by qRT-PCR analyses**

246 The RNA-Seq results were validated by qRT-PCR assays. Fourteen annotated DEGs were
247 chosen for qRT-PCR analyses. Twelve of these genes were downregulated and two were
248 upregulated. It included auxin related genes, *AUX-IAA* (CmoCh04G015610, auxin-responsive
249 protein IAA gene) and *ARF* (CmoCh16G005850, auxin response factor), gibberellins-related
250 gene, *TF* (CmoCh05G010630, PIF4: phytochrome-interacting factor 4), ethylene related genes,
251 *ACO* (CmoCh02G000640, 1-aminocyclopropane-1-carboxylate oxidation), *ETR*
252 (CmoCh08G004320, ethylene receptor), *ERDBF3* (CmoCh17G005350, ethylene response DNA
253 binding factor), *ERTF10* (CmoCh11G005450, ethylene responsive transcription factor gene), and
254 *AP2* (CmoCh01G000430, ethylene responsive transcription factor gene), brassinosteroid
255 biosynthesis genes, *TCH4* (CmoCh11G004020, xyloglucan:xyloglucosyl transferase) and
256 *CYCD3* (CmoCh01G018690, cyclin D3), jasmonic acid pathway gene, *JAR1*
257 (CmoCh06G010560, jasmonic acid-amino synthetase), salicylic acid pathway gene, *NPR1*
258 (CmoCh03G001860, regulatory protein), and transcription factor, *TGA* (CmoCh13G003300),
259 and phenylpropanoid biosynthesis gene, *PAL* (CmoCh07G009540, phenylalanine ammonia-

260 lyase). As shown in *Fig. 4*, all the detected genes showed similar expression trends in qRT-PCR
261 analyses as in RNA-Seq data, with a relative coefficient of $R^2 = 0.7887$ (*Fig. 4A, B*). The Pearson
262 correlation analysis indicated that the RNA-Seq and qRT-PCR were strongly correlated ($R = $
263 0.888, $P = 0.00002$).

264 Expression levels of ethylene candidate genes in different flower structures

265 Ethylene signal transduction genes (*ERDBF3*, *ERTF10*, *AP2*, and *ETR*) were appealing
266 candidates for investigating their gene expression in different flower structures, including pistil,
267 female calyx, ovary, stamen, male calyx, and leaf. The results showed that the expression level
268 of ethylene response DNA binding factor, *ERDBF3*, was highest in pistil, followed by that in
269 male calyx, leaf, female calyx, ovary, and stamen (*Fig. 5*), and the expression level in pistil was
270 78-times more than that in stamen. The expression level of *ERTF10* was also highest in pistils,
271 followed by that in leaf, male calyx, female calyx, ovary, and stamen (*Fig. 5*), and the expression
272 level in pistil was 162-times more than that in stamen. *ERTF10* and *ERDBF3* showed the highest
273 expression in pistil and the lowest in stamen. *AP2* has been reported in other crops to be
274 important for flower and seed development (Jofuku et al., 1994; Zhang et al., 2018). In this study,
275 the expression level of *AP2* was found to be higher in the reproductive organ stamen, male calyx,
276 female calyx, pistil, and ovary compared to that in leaf (*Fig. 5*). This indicates that, in pumpkin,
277 *AP2* may also play an important role in flower development. The expression level of the ethylene
278 receptor, *ETR*, was highest in stamen, followed by that in leaf, ovary, female calyx, pistil, and
279 male calyx (*Fig. 5*). From the results, we can see that the four ethylene signal transduction genes
280 are important for the development of female and male flowers, and *ERTF10* and *ERDBF3* may
281 especially be more important in the pistils of female flowers than that in the stamens of male
282 flowers.

Discussion

284 The number and proportion of female flowers are directly related to the yield of *Cucurbitaceae*
285 crops. In previous studies, researchers have focused on sex differentiation to explore the key
286 genes and mechanisms regulating the development of female and male flowers. Ethylene is
287 known to promote the development of female flowers, which was mainly manifested in sexual
288 expression, the earliness and the larger number of female flowers per plant. Moreover, the
289 development of female flowers requires much more ethylene than the development of male
290 flowers (Pierce and Wehner, 1990;Malepszy and Niemirowicz-Szczytt, 1991;Manzano et al.,
291 2011;Manzano et al., 2014;Pan et al., 2018).

292   In this study, the key genes involved in pistil development were studied directly using aborted
293 pistil materials. The results of KEGG pathway analysis showed that the DEGs were enriched in
294 plant hormone signal transduction pathway, and most of them were annotated as predicated
295 ethylene-responsive transcription factor, ethylene insensitive transcription factor, and ethylene

296 receptor genes (*Supplemental Table 4*). This result echo a previous study which have suggested
297 that ethylene synthesis and signal transduction play important roles in sex expression of pumpkin
298 (*Cucurbita maxima*) (Wang et al., 2019). Furthermore, we found that the ethylene response DNA
299 binding factor, *ERDBF3*, and the ethylene responsive transcription factor, *ERTF10*, showed the
300 highest expression in pistils and the lowest expression in stamens (*Fig. 5*). These results indicate
301 that ethylene signal transduction genes, *ERTF10* and *ERDBF3*, may be more important in the
302 pistils of female flowers than in other floral structures. In addition, previous studies have also
303 shown that the ethylene receptor, *CsETR1*, and the ethylene-responsive transcription factor,
304 *CsAP3*, are involved in the development of female flowers (Hao et al., 2003;Yamasaki et al.,
305 2003;Wang et al., 2010;Sun et al., 2016). The results presented in *Fig. 5* show that *AP2* and *ETR*
306 were expressed in all the structures of flowers that were assessed, and although their expression
307 level in the stamens was higher than that in the pistils, there is not enough evidence to show as to
308 how important they are for the development of female and male flowers. Except for the ethylene
309 signal transduction, which might play a major role in the development of female flowers, the
310 results of KEGG analysis also indicated that other plant hormone signal transduction pathways,
311 including auxin, cytokinin, gibberellin, abscisic acid, brassinosteroid, jasmonic acid, salicylic
312 acid signaling pathways, are involved in flower development. There is a possibility for the
313 existence of interactions between various plant hormones through their signal transduction
314 pathways. For example, ethylene synthesis can be induced by auxin, gibberellin, and jasmonic
315 acid (Rudich and Halevy, 1974;Trebitsh et al., 1987;Zhang et al., 2017;Schubert et al., 2019).

316 Phenylpropanoids comprise many aromatic metabolites, including the cell wall structural
317 component, lignin, and many small phenolic molecules, such as coumarins, stilbenes, flavonoids,
318 anthocyanins, and condensed tannins (Vogt, 2010;Fraser and Chapple, 2011;Zhang et al., 2013).
319 Phenylalanine ammonia-lyase catalyzes the first rate-limiting step in the phenylpropanoid
320 pathway, which controls the carbon flux to aromatic compounds, and to lignin. The results of this
321 study indicate that six *PAL* DEGs are downregulated in aborted pistils (*Table 2*). Previous study
322 have indicated that transgenic tobacco plants carrying antisense and sense *pal* cDNAs resulted in
323 partial male sterility, with the reduction of pollen fertility (Matsuda et al., 1996). At present,
324 there has been no systematic research on the mechanisms through which phenylpropanoids take
325 part in the development of female flowers.

326 In addition, the DEGs in the present study were obtained from RNA-Seq analysis of the
327 aborted and normal pistils, which were come from the flower buds with 5 mm length of corolla.
328 More genes related to pistil development were mined. But, when the aborted female flower buds
329 grew bigger, they exhibited incomplete stigma and thinner ovary for the absence of ovules, and
330 which is the primary cause of pistil abortion is unclear. Further gene function analysis is needed
331 to illuminate the mechanism of pistil abortion.

Conclusions

332

333   In the present study, DEGs in aborted and normal pistils of pumpkin were identified. KEGG
334 enrichment analysis showed that the DEGs were significantly enriched in plant hormone signal
335 transduction, which accounted for 12.54% of the significant DEGs. Most of them were annotated
336 as predicated ethylene signal transduction genes. The expression analysis of *ERDBF3* and
337 *ERTF10* in different flower structures showed the highest expression in pistils and the lowest
338 expression in stamens. These results suggest that plant hormone signal transduction genes,
339 especially ethylene signal transduction genes, play an important role in the development of pistils
340 in pumpkin.

341

References

343 Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful
344   approach to multiple testing. *Journal of the Royal statistical society*: *series B*
345   *(Methodological)* 57:289-300. DOI: 10.1111/j.2517-6161.1995.tb02031.x.
346 Boualem A, Fergany M, Fernandez R, Troadec C, Martin A, Morin H, Sari MA, Collin F,
347   Flowers JM, Pitrat M, Purugganan MD, Dogimont C, Bendahmane A. 2008. A conserved
348   mutation in an ethylene biosynthesis enzyme leads to andromonoecy in melons. *Science*
349   321:836-838. DOI: 10.1126/science.1159023.
350 Boualem A, Troadec C, Camps C, Lemhemdi A, Morin H, Sari MA, Fraenkel-Zagouri R,
351   Kovalski I, Dogimont C, Perl-Treves R, Bendahmane A. 2015. A cucurbit androecy gene
352   reveals how unisexual flowers develop and dioecy emerges. *Science* 350:688-691. DOI:
353   10.1126/science.aac8370.
354 Chen H, SunJ, Li S, Cui Q, Zhang H, Xin F, Wang H, Lin T, Gao D, Wang S, Li X, Wang D,
355   Zhang Z, Xu Z, Huang S. 2016. An ACC oxidase gene essential for cucumber carpel
356   development. *Molecular Plant* 9:1315-1327. DOI: 10.1016/j.molp.2016.06.018.
357 Duan QH, Wang DH, Xu ZH, Bai SN. 2008. Stamen development in Arabidopsis is arrested by
358   organ-specific overexpression of a cucumber ethylene synthesis gene *CsACO2*. *Planta*
359   228:537-543. DOI: 10.1007/s00425-008-0756-7.
360 Fraser CM, Chapple C. 2011. The phenylpropanoid pathway in arabidopsis. *The Arabidopsis*
361   *Book/American Society of Plant Biologists* 9:e0152. DOI: 10.1199/tab.0152.
362 Hao YJ, Wang DH, Peng YB, Bai SL, Xu LY, Li YQ, Xu ZH, Bai SN. 2003. DNA damage in
363   the early primordial anther is closely correlated with stamen arrest in the female flower of
364   cucumber (*Cucumis sativus* L.). *Planta* 217:888-895. DOI: 10.1007/s00425-003-1064-x.
365 Hemphill DD, Baker LR, Sell HM. 1972. Different sex phenotypes of *Cucumis sativus* L. and *C.*
366   *melo* L. and their endogenous gibberellin activity. *Euphytica* 21:285-291. DOI:
367   10.1007/BF00036769.
368 Iwahori S, Lyons JM, Sims WL. 1969. Induced femaleness in cucumber by 2-
369   chloroethanephosphonic acid. *Nature* 222:271-272. DOI: 10.1038/222271a0.

370 Jofuku KD, den Boer BG, Van Montagu M, Okamuro JK. 1994. Control of Arabidopsis flower
371   and seed development by the homeotic gene *APETALA2*. *The Plant Cell* 6:1211-1225. DOI:
372   10.1105/tpc.6.9.1211.

373 Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M. 2014. Data, information,
374   knowledge and principle: back to metabolism in KEGG. *Nucleic Acids Research* 42: 199-
375   205. DOI: 10.1093/nar/gkt1076.

376 Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. 2013. Tophat2: accurate
377   alignment of transcriptomes in the presence of insertions, deletions and gene fusions.
378   *Genome biology* 14:R36. DOI: 10.1186/gb-2013-14-4-r36.

379 Knopf RR, Trebitsh T. 2006. The female-specific *CS-ACS1g* gene of cucumber. A case of gene
380   duplication and recombination between the non-sex-specific 1-aminocyclopropane-1-
381   carboxylate synthase gene and a branched-chain amino acid transaminase gene. *Plant &Cell*
382   *Physiology* 47, 1217-1228. DOI: 10.1093/pcp/pcj092.

383 Livak KJ, SchmittgenTD. 2001. Analysis of relative gene expression data using real-time
384   quantitative PCR and the 2(-Delta Delta C(T)) method. *Methods* 25:402-408. DOI:
385   10.1006/meth.2001.1262.

386 Malepszy S, Niemirowicz-Szczytt K. 1991. Sex determination in cucumber (*Cucumis sativus*) as
387   a model system for molecular biology. *Plant Science* 80:39-47. DOI: 10.1016/0168-
388   9452(91)90271-9.

389 Manzano S, Martínez C, García JM, Megías Z, Jamilena M. 2014. Involvement of ethylene in
390   sex expression and female flower development in watermelon (*Citrullus lanatus*). *Plant*
391   *Physiology and Biochemistry* 85:96-104. DOI: 10.1016/j.plaphy.2014.11.004.

392 Manzano S, Martínez C, Megías Z, Gómez P, Garrido D, Jamilena M. 2011. The role of ethylene
393   and brassinosteroids in the control of sex expression and flower development in *Cucurbita*
394   *pepo*. *Journal of Plant Growth Regulation* 65: 213-221. DOI: 10.1007/s10725-011-9589-7.

395 Mao X, Cai T, Olyarchuk J, Wei L. 2005. Automated genome annotation and pathway
396   identification using the KEGG Orthology (KO) as a controlled vocabulary. *Bioinformatics*
397   21:3787-3793. DOI: 10.1093/bioinformatics/bti430.

398 Mao Y, Liu W, Chen X, Xu Y, Lu W, Hou J, Ni J, Wang Y, Wu L. 2017. Flower development
399   and sex determination between male and female flowers in *Vernicia fordii*. *Frontiers in*
400   *Plant Science* 8:1291. DOI: 10.3389/fpls.2017.01291.

401 Matsuda N, Tsuchiya T, Kishitani S, Tanaka Y, Toriyama K. 1996. Partial male sterility in
402   transgenic tobacco carrying antisense and sense *PAL* cDNA under the control of a tapetum-
403   specific promoter. *Plant and Cell Physiology* 37:215-222. DOI:
404   10.1093/oxfordjournals.pcp.a028934.

405 Menéndez V, Revilla MA, Fal MA, Fernández H. 2009. The effect of cytokinins on growth and
406   sexual organ development in the gametophyte of *Blechnum spicant* L. *Plant Cell, Tissue*
407   *and Organ Culture* 96:245-250. DOI: 10.1007/s11240-008-9481-y.

408 Mibus H, Tatlioglu T. 2004. Molecular characterization and isolation of the F/f gene for
409   femaleness in cucumber (*Cucumis sativus* L.). *Theoretical And Applied Genetics* 109:1669-
410   1676. DOI: 10.1007/s00122-004-1793-7.
411 Pan J, Wang G, Wen H, Du H, Lian H, He H, Pan J, Cai R. 2018. Differential gene expression
412   caused by the *F* and *M* loci provides insight into ethylene-mediated female flower
413   differentiation in cucumber. *Frontiers in Plant Science* 9:1091. DOI:
414   10.3389/fpls.2018.01091.
415 Papadopoulou E, Little HA, Hammar SA, Grumet R. 2005. Effect of modified endogenous
416   ethylene production on sex expression, bisexual flower development and fruit production in
417   melon (*Cucumis melo* L.). *Sexual Plant Reproduction* 18: 131-142. DOI: 10.1007/s00497-
418   005-0006-0.
419 Pierce L, Wehner T. 1990. Review of genes and linkage groups in cucumber. *Horticultural*
420   *Science* 25:605-615. DOI: 10.21273/HORTSCI.25.6.605.
421 Pimenta Lange MJ, Lange T. 2016. Ovary-derived precursor gibberellin A9 is essential for
422   female flower development in cucumber. *Development* 143:4425-4429. DOI:
423   10.1242/dev.135947.
424 Robinson MD, McCarthy DJ, Smyth GK. 2010. EdgeR: a bioconductor package for differential
425   expression analysis of digital gene expression data. *Bioinformatics* 26: 139-140. DOI:
426   10.1093/bioinformatics/btp616.
427 Rudich J, Halevy AH. 1974. Involvement of abscisic acid in the regulation of sex expression in
428   the cucumber. *Plant &Cell Physiology* 15:635-642. DOI:
429   10.1093/oxfordjournals.pcp.a075048.
430 Saito S, Fujii N, Miyazawa Y, Yamasaki S, Matsuura S, Mizusawa H, FujitaY, Takahashi H.
431   2007. Correlation between development of female flower buds and expression of the *CS-*
432   *ACS2* gene in cucumber plants. *Journal of Experimental Botany* 58:2897-2907. DOI:
433   10.1093/jxb/erm141.
434 Schubert R, Grunewald S, von Sivers L, Hause B. 2019. Effects of jasmonate on ethylene
435   function during the development of tomato stamens. *Plants (Basel)* 8. DOI:
436   10.3390/plants8080277.
437 Song J, Gao ZH, Huo XM, Sun HL, Xu YS, Shi T, Ni ZJ. 2015. Genome-wide identification of
438   the auxin response factor (*ARF*) gene family and expression analysis of its role associated
439   with pistil development in Japanese apricot (*Prunus mume* Sieb. et Zucc). *Acta*
440   *Physiologiae Plantarum* 37: 145. DOI: 10.1007/s11738-015-1882-z.
441 Sun H, Wu S, Zhang G, Jiao C, Guo S, Ren Y, Zhang J, Zhang H, Gong G, Jia Z, Zhang F, Tian
442   J, Lucas WJ, Doyle JJ, Li H, Fei Z, Xu Y. 2017. Karyotype stability and unbiased
443   fractionation in the paleo-allotetraploid Cucurbita genomes. *Molecular Plant* 10:1293-1306.
444   DOI: 10.1016/j.molp.2017.09.003.
445 Sun JJ, Li F, Wang DH, Liu XF, Li X, Liu N, Gu HT, Zou C, Luo JC, He CX, Huang SW,
446   Zhang XL, Xu ZH, Bai SN. 2016. CsAP3: a cucumber homolog to Arabidopsis

447   APETALA3 with novel characteristics. *Frontiers in Plant Science* 7:1181. DOI:
448   10.3389/fpls.2016.01181.
449 Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL,
450   Pachter L. 2012. Differential gene and transcript expression analysis of RNA-seq
451   experiments with TopHat and Cufflinks. *Nature Protocols* 7:562-578. DOI:
452   10.1038/nprot.2012.016.
453 Trebitsh T, Rudich J, Riov J. 1987. Auxin, biosynthesis of ethylene and sex expression in
454   cucumber (*Cucumis sativus*). *Plant Growth Regulation* 5:105-113. DOI:
455   10.1007/BF00024738.
456 Vogt T. 2010. Phenylpropanoid biosynthesis. *Molecular Plant* 3:2-20. DOI: 10.1093/mp/ssp106.
457 Wang DH, Li F, Duan QH, Han T, Xu ZH, Bai SN. 2010. Ethylene perception is involved in
458   female cucumber flower development. *Plant Journal* 61:862-872. DOI: 10.1111/j.1365-
459   313X.2009.04114.x.
460 Wang YL, Yan CD, Zou BX, Wang CJ, Xu WL, Cui CS, Qu SP. 2019 Morphological,
461   transcriptomic and hormonal characterization of trimonoecious and subandroecious
462   pumpkin (*Cucurbita maxima*) suggests important roles of ethylene in sex expression.
463   *International Journal of Molecular Science* 20:3185. DOI: 10.3390/ijms20133185.
464 Yamasaki S, Fujii N, Takahashi H. 2000. The ethylene-regulated expression of *CS-ETR2* and
465   *CS-ERS* genes in cucumber plants and their possible involvement with sex expression in
466   flowers. *Plant and Cell Physiology* 41:608-616. DOI: 10.1093/pcp/41.5.608.
467 Yamasaki S, Fujii N, Takahashi H. 2003. Characterization of ethylene effects on sex
468   determination in cucumber plants. *Sexual Plant Reproduction* 16:103-111. DOI:
469   10.1007/s00497-003-0183-7.
470 Yamasaki S, Fujii N, Matsuura S, Mizusawa H, Takahashi H. 2001. The M locus and ethylene-
471   controlled sex determination in andromonoecious cucumber plants. *Plant and Cell*
472   *Physiology* 42:608-619. DOI: 10.1093/pcp/pce076.
473 Young MD, Wakefield MJ, Smyth GK, Oshlack A. 2010. Gene ontology analysis for RNA-seq:
474   accounting for selection bias. *Genome Biology* 11, R14. DOI: 10.1186/gb-2010-11-2-r14.
475 Yuan Z, Zhang D. 2015. Roles of jasmonate signalling in plant inflorescence and flower
476   development. *Current Opinion in Plant Biology* 27:44-51. DOI: 10.1016/j.pbi.2015.05.024.
477 Zhang X, Gou M, Liu CJ. 2013. Arabidopsis Kelch repeat F-box proteins regulate
478   phenylpropanoid biosynthesis via controlling the turnover of phenylalanine ammonia-lyase.
479   *Plant Cell* 25:4994-5010. DOI: 10.1105/tpc.113.119644.
480 Zhang Y, Huang S, Wang X, Liu J, Guo X, Mu J, Tian J, Wang X. 2018. Defective APETALA2
481   genes lead to sepal modification in brassica crops. *Frontiers in Plant Science* 9:367. DOI:
482   10.3389/fpls.2018.00367.
483 Zhang Y, Zhao G, Li Y, Mo N, Zhang J, Liang Y. 2017. Transcriptomic analysis implies that GA
484   regulates sex expression via ethylene-dependent and ethylene-independent pathways in
485   cucumber (*Cucumis sativus* L.). *Frontiers in Plant Science* 8:10. DOI:
486   10.3389/fpls.2017.00010.

Statistical results of reads mapped to reference genome and expressed genes in different libraries.

Statistical results of reads mapped to reference genome after rRNA filtering and expressed genes in different libraries.

Sample	Total Pair Reads	Mapped Pair Reads (Ratio)	Known Gene Num	Novel Gene Num
A2	55729080	49972699(89.67%)	26128	386
A4	65402194	58874465(90.02%)	26463	386
N2	58025908	51001759(87.89%)	26559	381
N4	56371344	50670138(89.89%)	25914	372

List of PAL in the significant DEGs.

List of PAL in the significant DEGs.

Gene ID	Gene annotation	Fold change	FDR
CmoCh03G012270	PREDICTED: phenylalanine ammonia-lyase-like [*Cucumis melo*]	0.439562	3.20E-05
CmoCh03G012290	PREDICTED: phenylalanine ammonia-lyase-like [*Cucumis melo*]	0.343162	1.13E-28
CmoCh07G009470	phenylalanine ammonia-lyase 3 [*Luffa aegyptiaca*]	0.458279	4.76E-08
CmoCh07G009540	phenylalanine ammonia-lyase 3 [*Luffa aegyptiaca*]	0.077247	1.86E-26
CmoCh07G009550	phenylalanine ammonia-lyase 3 [*Luffa aegyptiaca*]	0.209537	1.47E-13
CmoCh07G009560	PREDICTED: phenylalanine ammonia-lyase-like [*Cucumis melo*]	0.368197	4.93E-05
CmoCh20G005320	phenylalanine ammonia-lyase 4 [*Luffa aegyptiaca*]	2.000495	0

Figure 1

The aborted and normal pistils and differentially expressed genes (DEGs) that were significantly upregulated and downregulated between the aborted and normal pistils of pumpkin.

The aborted and normal pistils and differentially expressed genes (DEGs) that were significantly upregulated and downregulated between the aborted and normal pistils of pumpkin. (A) The aborted and normal pistils (bars=1 cm). **(B)** The number of upregulated and downregulated DEGs. **(C)** Volcano map for the gene expression. The |log2 (fold change)| > 1 and FDR <0.05 were used as the cut-offs for significance.
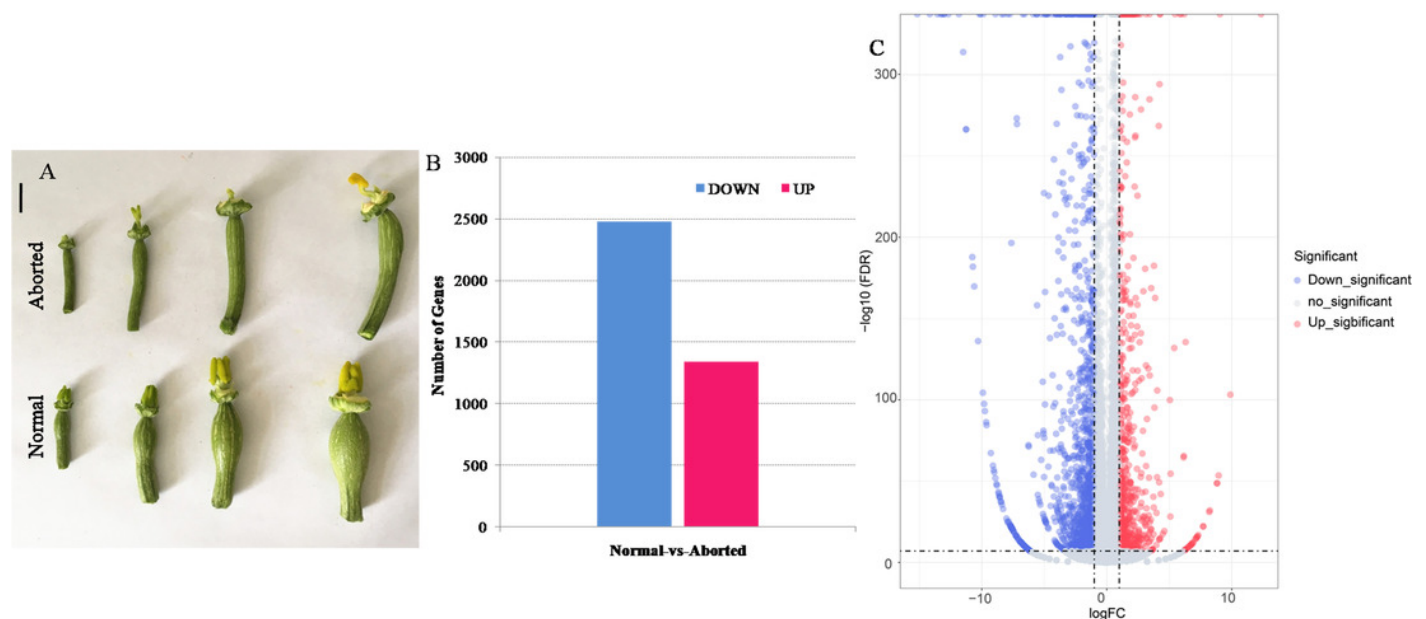
Figure 2

Gene ontology (GO) categories of differentially expressed genes (DEGs) between the aborted and normal pistils.

Gene ontology (GO) categories of differentially expressed genes (DEGs) between the aborted and normal pistils.
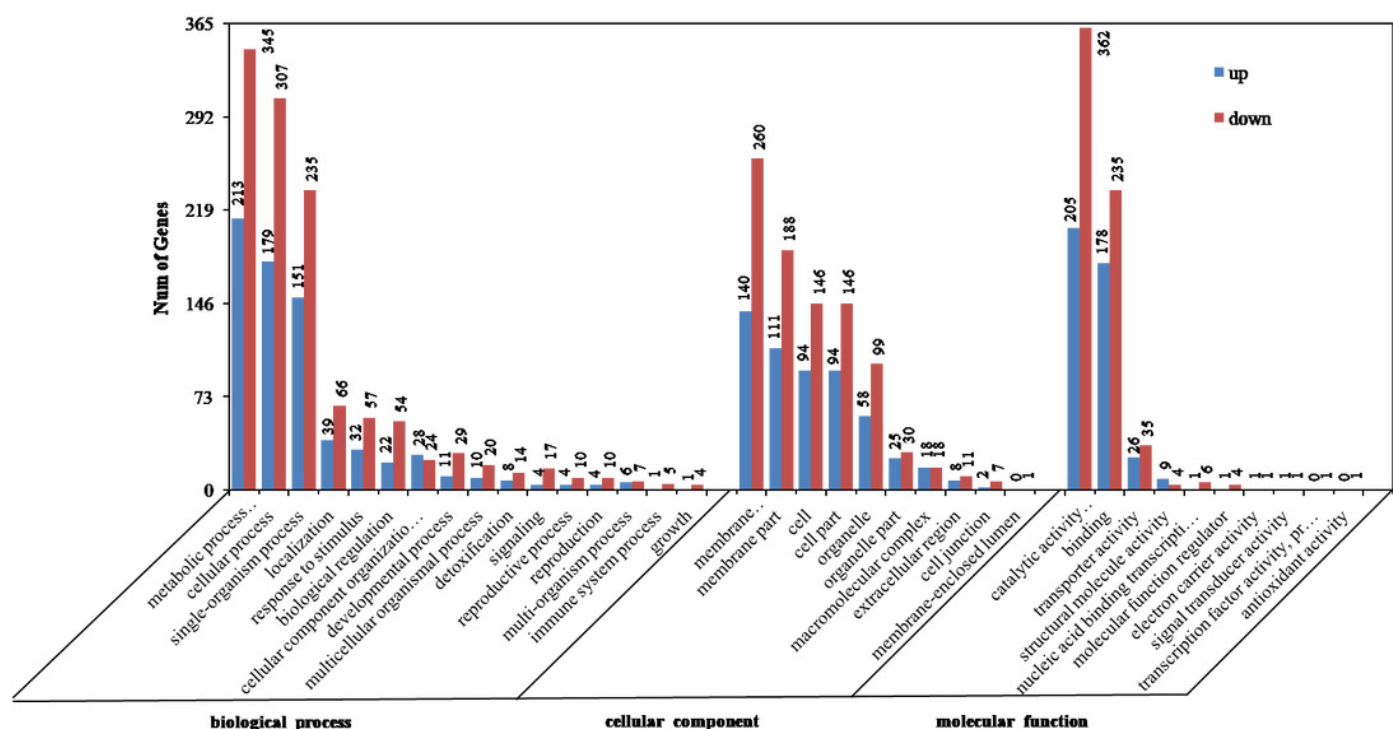
Figure 3

Kyoto encyclopedia of genes and genomes (KEGG) pathway enrichment of differentially expressed genes (DEGs) between the aborted and normal pistils.

Kyoto encyclopedia of genes and genomes (KEGG) pathway enrichment of differentially expressed genes (DEGs) between the aborted and normal pistils. Ggplot2 package for R language was used to generate this figure. The bigger the bubble, the more the DEGs. The smaller the pvalue, the more significant the KEGG enrichment. Rich Factor represents the ratio of DEGs number in the pathway to the total number of genes in the pathway. The larger the RichFactor, the higher the enrichment.

Top20 of pathway enrichment
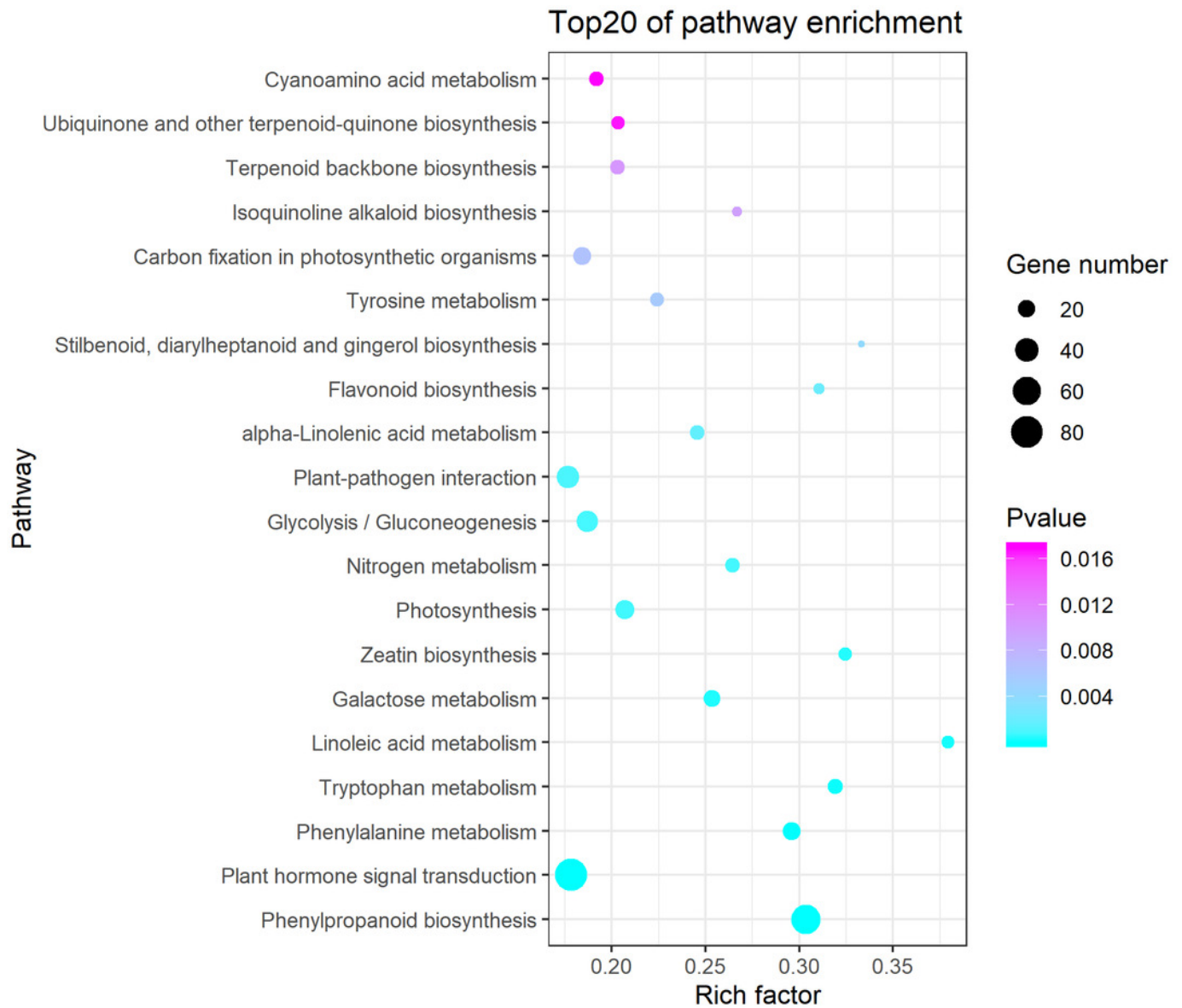
Figure 4

Verification of RNA-Seq Data by quantitative real-time PCR.

Verification of RNA-Seq Data by quantitative real-time PCR. (A) Relative normalized expression of 14 selected genes in the normal and aborted pistils using qRT-PCR. Error bars indicate the standard errors. *AUX-IAA* (auxin-responsive protein IAA), *ARF* (auxin response factor), *TF* (*PIF4*, phytochrome-interacting factor 4), *ACO* (1-aminocyclopropane-1-carboxylate oxidation), *ETR* (ethylene receptor gene), *ERDBF3* (ethyleneresponse DNA binding factor 3), *ERTF10* (ethylene responsive transcription factor 10), *AP2* (like ethylene-responsive transcription factor), *TCH4* (xyloglucan endotransglucosylase/hydrolase), *CYCD3* (cyclin D3), *JAR1* (jasmonic acid-amino synthetase), *NPR1* (BTB/POZ domain and ankyrin repeat-containing protein NOOT2), *TGA* (TGACG-sequence-specific DNA-binding protein), *PAL* (phenylalanine ammonia-lyase). **(B)** Comparison of the expression ratios of 14 selected genes in using RNA-seq and qRT-PCR.

OFF
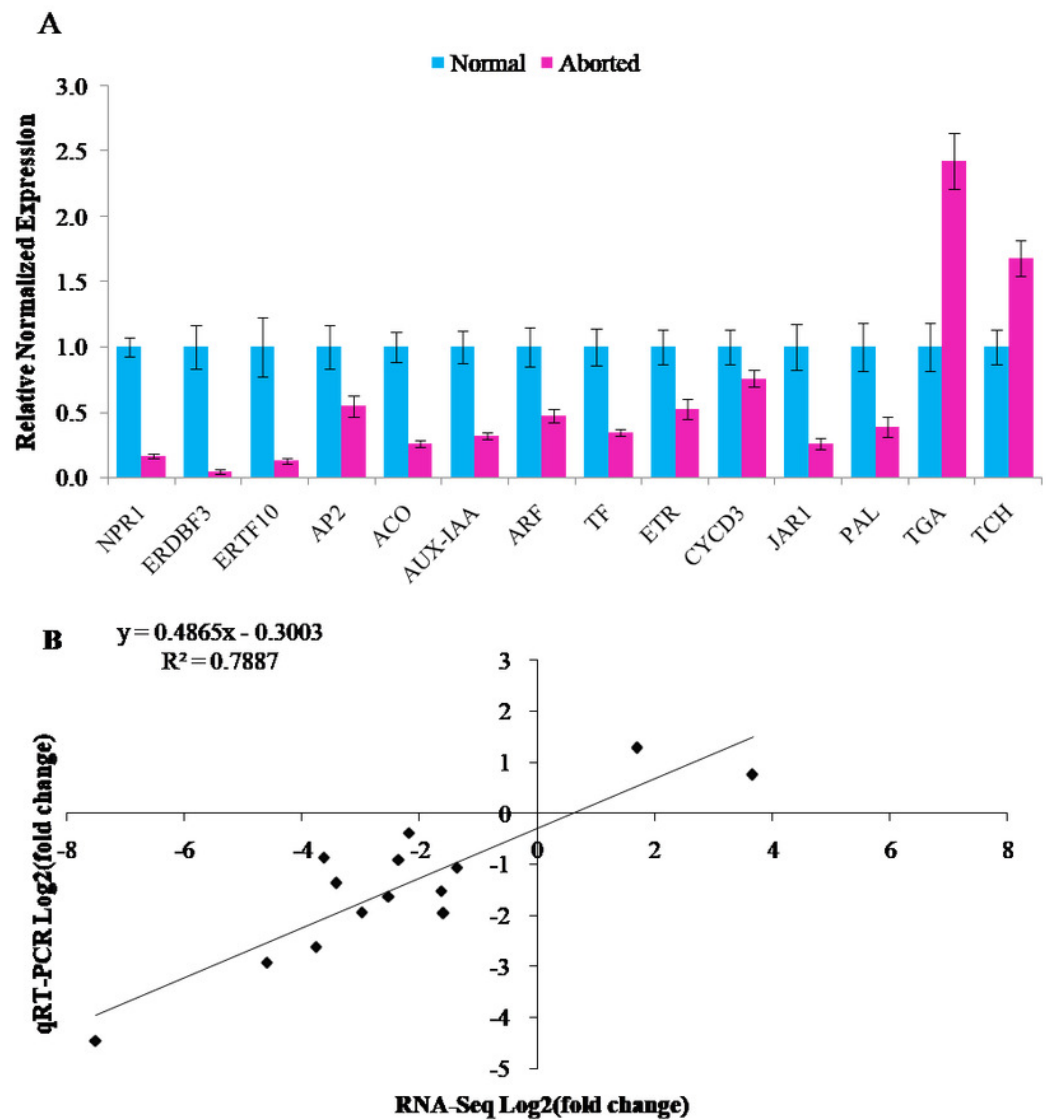
Figure 5

Expression levels of ethylene signal transduction candidate genes (*ERDBF3*, *ERTF10*, *AP2*, and *ETR*) in different flower structures.

Expression levels of ethylene signal transduction candidate genes (*ERDBF3*, *ERTF10*, *AP2*, and *ETR*) in different flower structures. Error bars indicate the standard errors. *ERDBF3* (ethyleneresponse DNA binding factor 3), *ERTF10* (ethylene responsive transcription factor 10), *AP2* (like ethylene-responsive transcription factor), *ETR* (ethylene receptor gene).