Genome-wide analysis suggests high level of microsynteny and purifying selection affect the evolution of *EIN3/EIL* family in Rosaceae (#15747)

First revision

Please read the **Important notes** below, the **Review guidance** on page 2 and our **Standout reviewing tips** on page 3. When ready **submit online**. The manuscript starts on page 4.

Important notes

Editor

Marion Röder

Files	1 Tracked changes manuscript(s) 1 Rebuttal letter(s) 7 Figure file(s) 2 Table file(s) 2 Other file(s) Please visit the overview page to **download and review** the files not included in this review PDF.
Declarations	No notable declarations are present

Review guidelines

Please read in full before you begin

How to review

When ready **submit your review online**. The review form is divided into 5 sections. Please consider these when composing your review:

1. **BASIC REPORTING**
2. **EXPERIMENTAL DESIGN**
3. **VALIDITY OF THE FINDINGS**
4. General comments
5. Confidential notes to the editor

You can also annotate this PDF and upload it as part of your review

To finish, enter your editorial recommendation (accept, revise or reject) and submit.

BASIC REPORTING

- Clear, unambiguous, professional English language used throughout.
- Intro & background to show context. Literature well referenced & relevant.
- Structure conforms to **PeerJ standards**, discipline norm, or improved for clarity.
- Figures are relevant, high quality, well labelled & described.
- Raw data supplied (see **PeerJ policy**).

EXPERIMENTAL DESIGN

- Original primary research within **Scope of the journal**.
- Research question well defined, relevant & meaningful. It is stated how the research fills an identified knowledge gap.
- Rigorous investigation performed to a high technical & ethical standard.
- Methods described with sufficient detail & information to replicate.

VALIDITY OF THE FINDINGS

- Impact and novelty not assessed. Negative/inconclusive results accepted. *Meaningful* replication encouraged where rationale & benefit to literature is clearly stated.
- Data is robust, statistically sound, & controlled.
- Conclusions are well stated, linked to original research question & limited to supporting results.
- Speculation is welcome, but should be identified as such.

The above is the editorial criteria summary. To view in full visit **https://peerj.com/about/editorial-criteria/**

7 Standout reviewing tips

The best reviewers use these techniques

Tip	*Example*
Support criticisms with evidence from the text or from other sources	*Smith et al (J of Methodology, 2005, V3, pp 123) have shown that the analysis you use in Lines 241-250 is not the most appropriate for this situation. Please explain why you used this method.*
Give specific suggestions on how to improve the manuscript	*Your introduction needs more detail. I suggest that you improve the description at lines 57- 86 to provide more justification for your study (specifically, you should expand upon the knowledge gap being filled).*
Comment on language and grammar issues	*The English language should be improved to ensure that your international audience can clearly understand your text. I suggest that you have a native English speaking colleague review your manuscript. Some examples where the language could be improved include lines 23, 77, 121, 128 – the current phrasing makes comprehension difficult.*
Organize by importance of the issues, and number your points	*1. Your most important issue* *2. The next most important item* *3. …* *4. The least important points*
Give specific suggestions on how to improve the manuscript	*Line 56: Note that experimental data on sprawling animals needs to be updated. Line 66: Please consider exchanging "modern" with "cursorial".*
Please provide constructive criticism, and avoid personal opinions	*I thank you for providing the raw data, however your supplemental files need more descriptive metadata identifiers to be useful to future readers. Although your results are compelling, the data analysis should be improved in the following ways: AA, BB, CC*
Comment on strengths (as well as weaknesses) of the manuscript	*I commend the authors for their extensive data set, compiled over many years of detailed fieldwork. In addition, the manuscript is clearly written in professional, unambiguous language. If there is a weakness, it is in the statistical analysis (as I have noted above) which should be improved upon before Acceptance.*

Genome-wide analysis suggests high level of microsynteny and purifying selection affect the evolution of *EIN3/EIL* family in Rosaceae

Yunpeng Cao Corresp., [1] , **Yahui Han** [2] , **Dandan Meng** [1] , **Dahui Li** [1] , **Qing Jin** [1] , **Yi Lin** [1] , **Yongping Cai** [1]

[1] college of Life Sciences, Anhui Agricultural University, Hefei, China

[2] State Key Laboratory of Tea Plant Biology and Utilization, Anhui Agricultural University, Hefei, China

Corresponding Author: Yunpeng Cao
Email address: 499981806@qq.com

The ethylene-insensitive3/ethylene-insensitive3-like (EIN3/EIL) protein, are a type of nuclear-localized proteins with DNA-binding activity in plants. Although the *EIN3/EIL* gene family has been studied in several plant species, little is known about comprehensive study of the *EIN3/EIL* gene family in Rosaceae. In this study, 10, 5, 4, and 5 *EIN3/EIL* genes were identified in the genomes of pear (*Pyrus bretschneideri*), mei (*Prunus mume*), peach (*Prunus persica*) and strawberry (*Fragaria vesca*), respectively. Twenty-eight chromosomal segments of *EIL/EIN3* gene family were found in four Rosaceae species, and these segments could form seven orthologous or paralogous groups based on interspecies or intraspecies gene colinearity (microsynteny) analysis. Moreover, the highly conserved regions of microsynteny were found in four Rosaceae species. Subsequent it was found that both whole genome duplication and tandem duplication events significantly contributed to the *EIL/EIN3* gene family expansion. Gene expression analysis of the *EIL/EIN3* genes in pear revealed subfunctionalization for several *PbEIL* genes derived from whole genome duplication. Noteworthy, according to environmental selection pressure analysis, the strong purifying selection should dominate the maintenance of the *EIL/EIN3* gene family in four Rosaceae species. These results provided useful information of Rosaceae *EIL/EIN3* genes, as well as insights into the evolution of this gene family in four Rosaceae species. Furthermore, high level of microsynteny in the four Rosaceae plants suggested that a large-scale genome duplication event in *EIL/EIN3* gene family was predated to speciation.

Genome-wide analysis suggests high level of microsynteny and purifying selection affect the evolution of *EIN3/EIL* family in Rosaceae

Yunpeng Cao [1], Yahui Han [2], Dandan Meng[1], Dahui Li [1], Qing Jin [1], Yi Lin [1]and Yongping Cai [1]*

Affiliation

1 School of Life Sciences, Anhui Agricultural University, Hefei 230036, China

2 State Key Laboratory of Tea Plant Biology and Utilization, Anhui Agricultural University, Hefei 230036, China

* Correspondence: ypcaiah@163.com; Tel.: +86 551 65786137; Fax: +86 551 65786340

Abstract

The ethylene-insensitive3/ethylene-insensitive3-like (EIN3/EIL) protein, are a type of nuclear-localized proteins with DNA-binding activity in plants. Although the *EIN3/EIL* gene family has been studied in several plant species, little is known about comprehensive study of the *EIN3/EIL* gene family in Rosaceae. In this study, 10, 5, 4, and 5 *EIN3/EIL* genes were identified in the genomes of pear (*Pyrus bretschneideri*), mei (*Prunus mume*), peach (*Prunus persica*) and strawberry (*Fragaria vesca*), respectively. Twenty-eight chromosomal segments of *EIL/EIN3* gene family were found in four Rosaceae species, and these segments could form seven orthologous or paralogous groups based on interspecies or intraspecies gene colinearity (microsynteny) analysis. Moreover, the highly conserved regions of microsynteny were found in four Rosaceae species. Subsequent it was found that both whole genome duplication and tandem duplication events significantly contributed to the *EIL/EIN3* gene family expansion. Gene expression analysis of the *EIL/EIN3* genes in pear revealed subfunctionalization for several *PbEIL* genes derived from whole genome duplication. Noteworthy, according to environmental selection pressure analysis, the strong purifying selection should dominate the maintenance of the *EIL/EIN3* gene family in four Rosaceae species. These results provided useful information of Rosaceae *EIL/EIN3* genes, as well as insights into the evolution of this gene family in four Rosaceae species. Furthermore, high level of microsynteny in the four Rosaceae plants suggested that a large-scale genome duplication event in *EIL/EIN3* gene family was predated to speciation.

Introduction

Rosaceae species such as pear (*Pyrus bretschneideri*), mei (*Prunus mume*), peach (*Prunus persica*) and strawberry (*Fragaria vesca*), ~~which are~~ important perennial tree, cultivated for commercial production of fruits

44 available worldwide. According to previous studies, the genomes of strawberry (X = 7), mei (X = 8), peach (X
45 = 8), and pear (X = 17) shared an ancestor, which had 9 pairs of chromosomes (Shulaev et al. 2011; Verde et al.
46 2013; Wu et al. 2013; Zhang et al. 2012). Recently, the researchers confirmed that chromosome inversions,
47 fusions, and translocations played an important role in the evolution of the Rosaceae genome (Illa et al. 2011).
48 Some extant "diploid" species of Rosaceae family are originated from their polyploid ancestors, others are
49 actually thought to be true polyploids (Wendel 2000). These studies indicate that the diploid species in
50 Rosaceae have evolved with a complex history. There are several gene families which share highly conserved
51 genome sequences with each other among the related species of family Rosaceae, as well as other taxonomic
52 families. In this study, the *EIL/EIN3* gene family was selected to investigate the specific evolutionary
53 relationships among the related species of family Rosaceae.

54 The *EIN3/EIL* gene family is a relatively small one in higher plants. Some *EIN3/EIL* genes have been
55 isolated from *Arabidopsis thaliana* (Chao et al. 1997a), tobacco (Kosugi & Ohashi 2000; Rieu et al. 2003),
56 banana (Jourda et al. 2014), tomato (Tieman et al. 2001; Yokotani et al. 2003) and rice (Hiraga et al. 2009;
57 Mao et al. 2006). These plant-specific EIN3/EIL proteins are located in the nuclei, with the highly conserved
58 amino acid sequences at the N-termini, including several important structural features, such as acidic amino
59 acid regions, proline-rich regions and 5-basic amino acid clusters (BD I-V) (Chao et al. 1997a). Compared to
60 the N-terminal sequences, the conservation of their C-termini is lower. For example, it was found that although
61 asparagine-enriched regions or glutamine-enriched regions were commonly distributed within the C-terminal
62 sequences of EIN3/EILs from Arabidopsis, carnation and mung beans (Chao et al. 1997b; Lee & Kim 2003;
63 Waki et al. 2001), they were not widely existed in other EIN3/EIL members, such as tobacco NtEILs (Rieu et
64 al. 2003).

Functions of the *EIN3/EIL* gene family have been studied in several plants, such as *Hevea brasiliensis* (Yang et al. 2014) and tomato (Tieman et al. 2001; Yokotani et al. 2003). Recently, a number of researches on application of comparative genome in the analysis of evolution and function of gene family have been reported. For example, based on the comparative genomic analysis, Wang Yiyi et al (2015) explored the evolution and functional differences of WRKY type-III transcription factor family of poplar, grape, Arabidopsis and rice (Wang et al. 2015). Qing jin et al (2016) explored the evolution of WRKY I subfamily in Gramineae (Jing et al. 2016). However, there is still lack of specific evolutionary relationships of the *EIN3/EIL* gene family in Rosaceae. To address this question, the evolutionary relationships and gene duplication events of *EIN3/EIL* genes from Rosaceae species, including pear (*Pyrus bretschneideri*), mei (*Prunus mume*), peach (*Prunus persica*) and strawberry (*Fragaria vesca*), were analyzed, based on their phylogenetic relationships, microsynteny and environmental selection pressures analysis. In addition, the expression patterns of pear *EIN3 / EIL* genes were investigated on a variety of organs/tissues including fruits at several developmental stages. The results obtained from this study provided valuable information about *EIN3/EIL* genes that will aid future functional research involved in many important biological processes of this important gene family in flowering plants, especially in pear.

Materials and methods

Sequence identification and collection

The genome data of four Rosaceae species were obtained from their respective genome sequence websites: *Pyrus bretschneideri* from the GigaDB database (http://gigadb.org/site/index); *Prunus mume* from the Genome Database for Rosaceae (http://www.rosaceae.org/); *Prunus persica* from the Phytozome database

(https://phytozome.jgi.doe.gov/pz/portal.html) and *Fragaria vesca* from the Joint Genome Institute (http://www.jgi.doe.gov/). The Hidden Markov Model (HMM) profiles of EIN3/EIL domain (PF04873) (Chang & Shockey 1999; Chao et al. 1997a) were obtained from the Pfam database (http://pfam.xfam.org) (Finn et al. 2006). The EIN3/EIL domain was used as query sequences to identify *EIN3/EIL* genes in four Rosaceae species by using DNAtools software (E-value <0.001). To verify the *EIN3/EIL* genes in four Rosaceae genomes, all putative proteins were validated by searching for the EIN3/EIL domain using InterPro online tool (http://www.ebi.ac.uk/interpro/) (Zdobnov & Apweiler 2001) and SMART database (http://smart.embl-heidelberg.de/) (Letunic et al. 2012). In our study, only the EIN3/EIL domain-containing sequences were retained.

Chromosomal location of *EIN3/EIL* genes

The genome annotation information was collected from GigaDB database (http://gigadb.org/site/index), Genome Database for Rosaceae (http://www.rosaceae.org/), Phytozome database (https://phytozome.jgi.doe.gov/pz/portal.html) and Joint Genome Institude (http://www.jgi.doe.gov/), respectively. Subsequently, the MapInspect software (http://mapinspect.software.informer.com/) was used for data visualization.

Gene structure and motif analysis

The exon-intron structure of each *EIN3/EIL* gene was determined by alignment of its CDS and genomic DNA sequence. Then a diagram was constructed using Structure Display Server website (Hu et al. 2014). Subsequently, the Online MEME server was used to screen the conserved motifs encoded by *EIN3/EIL* genes.

Additionally, Pfam website (Punta et al. 2011) and SMART tools (Letunic et al. 2012) were used to annotate these structural motifs.

Phylogenetic analysis of *EIN3/EIL* genes

EIN3/EIL sequences were aligned using ClustalX version 1.83 (Thompson et al. 1997) and evolutionary relationships were inferred by analyzing an unrooted phylogenetic tree using MEGA 5 and neighbor-joining (NJ) method (Tamura et al. 2011) with the following parameters: poisson correction, pairwise deletion and 1000 bootstrap replicate.

Microsynteny analysis

In order to reveal the sequence features of the *EIN3/EIL* gene-containing regions , microsynteny analysis was performed across the four Rosaceae species (*Pyrus bretschneideri*, *Prunus mume*, *Prunus persica* and *Fragaria vesca*), using MCScanx (Multiple Collinearity Scan toolkit) (Wang et al. 2012) with the gene identifier file, the gene list file and the coding sequence file. Subsequently, a syntenic block was defined as a region containing three or more conserved homologs which were located within 100-kb downstream and upstream of protein-coding sequences.

Environmental selection pressure analysis

The nucleotide coding sequences from segmentally duplicated pairs were aligned by Clustal X (Thompson et al. 1997). Then DnaSP (version 5.10) was used to calculate the nonsynonymous (Ka) and synonymous (Ks) substitution rates of the homologues (Librado & Rozas 2009). For each pair of duplicated regions, we

128 estimated the mean Ks values of the flanking conserved genes for individual homologs. To further understand
129 the selective pressure experienced by *EIN3/EIL* genes, Ka, Ks and Ka/Ks ratios were estimated using sliding
130 window (with parameters: window size, 150 bp; step size, 9 bp) over the entire aligned length (Cao et al.
131 2016a; Han et al. 2016).
132
133 *EIN3/EIL* gene expression analysis in pear different tissues
134 To verify the expression patterns of *EIN3/EIL* genes, the qRT-PCR analysis was carried out. The first-strand
135 cDNA was synthesized with Oligo18dT primer (Table S1) by using M-MLV reverse transcriptase (TakaRa,
136 Japan) following the manufacture introduction. The TransStart Tip Green qPCR SuperMix (TransGen Biotech,
137 China) with SYBR Green I as the fluorescent dye was used for the qPCR employing a Bio-rad CFX96
138 Touch™Deep Well Real-Time PCR Detection system (BioRad, USA). The relative transcript level relative to
139 the Pyrus tubulin gene (Wu et al. 2012) were estimated according to a previous workflow (Cao et al. 2016a;
140 Cao et al. 2016b). For each sample, three replicates were set up in parallel experiments.
141 **Results and discussion**
142 **Identification of *EIN3/EIL* genes in Rosaceae**
143 The genome data of pear (*P. bretschneideri*), mei (*P. mume*), peach (*P. persica*) and strawberry (*F. vesca*)
144 were recently published, respectively (Shulaev et al. 2011; Verde et al. 2013; Wu et al. 2013; Zhang et al.
145 2012). To identify the members of the *EIN3/EIL* gene families in these species, EIN3/EIL specific domain
146 (PF04873) was used to perform Blastp searches of the local protein databases. Sequences identified were
147 verified for EIN3/EIL domains through SMART database and InterPro online tool. Totally, 24 of *EIN3/EIL*
148 genes were identified, including 10 in pear, 4 in peach, 5 in mei and 5 in strawberry, and named as *PbEIL1*-

149 *PbEIL10*, *PpEIL1-PpEIL4*, *PmEIL1-PmEIL5* and *FvEIL1-FvEIL5*, according to their locations in chromosome,
150 respectively (Table 1 and Figure 1). This result suggested that *EIN3/EIL* gene family was relatively small
151 compared to other gene families in the studied species. Similar indication was also reported by the previous
152 studies in which 6, 5, 4, 6 and 17 *EIN3/EILs* genes were found in *Arabidopsis thaliana* (Chao et al. 1997a; Guo
153 & Ecker 2004), tobacco (Kosugi & Ohashi 2000; Rieu et al. 2003), tomato (Tieman et al. 2001; Yokotani et al.
154 2003), rice (Hiraga et al. 2009; Mao et al. 2006) and banana (Jourda et al. 2014), respectively. Furthermore, it
155 was found that the genome sizes and number of *EIN3/EIL* gene family members appeared not to have a direct
156 relevance. For example, although there was no significant variety in genome size of pear (271.9 Mb) and
157 strawberry (240 Mb), the number of *EIN3/EIL* genes obviously changed. Contrarily, the number of *EIN3/EIL*
158 genes of the peach (224.6 Mb) and strawberry (240 Mb) had a corresponding relationship with their genome
159 size. Remarkably, compared with those in peach, mei and strawberry, the numbers of *EIN3/EIL* genes in pear
160 were found to be almost doubled. Moreover, the chromosome numbers of peach, mei and strawberry are 16, 16
161 and 14, respectively (Shulaev et al. 2011; Verde et al. 2013; Wu et al. 2013; Zhang et al. 2012), whereas the
162 chromosome number of pear is 34, indicating that the *EIN3/EIL* gene family has undergone an expansion
163 corresponding to the variation in chromosome number. However, a recent whole genome duplication event
164 (30–45 million years ago) (Wu et al. 2013) occurred in pear but not in peach, mei and strawberry, probably
165 contributed to the expansion of *EIN3/EIL* gene family in the pear.

166 To determine the distribution of *EIN3/EI*L genes on chromosomes among pear, peach, mei and strawberry,
167 respectively, a chromosome map (Figure 1) was drawn based on genome annotation (Wu et al. 2013). In pear,
168 two *EIN3/EIL* genes were located on chromosome 2, and one gene on chromosome 3, 8, 11 and 15,
169 respectively, with the remaining genes localized on different scaffold regions (Figure 1C). In peach, two

170 *EIN3/EIL* genes were found on chromosome 2 and 6, respectively (Figure 1D). In mei, two *EIN3/EIL* genes
171 were distributed on chromosome 1 and 5, respectively, with the remaining one localized on a scaffold region
172 (Figure 1B). And in strawberry, two *EIN3/EIL* genes were distributed on chromosome 1 and 7, respectively,
173 with the remaining one localized on chromosome 3 (Table 1 and Figure 1A).
174
175 **Phylogenetic analysis of *EIN3/EIL* genes**
176 The phylogenetic tree containing *EIN3/EIL* gene homologs from a variety of species, including *Arabidopsis*
177 *thaliana*, rice, banana, sorghum, maize, *Brachypodium distachyon*, *Thellungiella parvula*, and four Rosaceae
178 species, was also constructed (Figure S1). As shown in the phylogenetic tree, most *EIN3/EIL* genes from four
179 Rosaceae species were clustered together. To further understand the evolutionary history of *EIN3/EIL* genes in
180 Rosaceae, phylogenetic analysis was carried out using neighbor joining (NJ) method. As shown in Figure 2, 24
181 EIN3/EIL sequences were divided into two subfamilies, designated as A and B, which contained four classes
182 (Classes A1, A2, B1 and B2). Classes A1, B1 and B2 were composed of the *EIN3/EIL* genes from the four
183 species (pear, peach, mei and strawberry), while the Class A2 contained the members only from pear.
184 According to our study, a whole genome duplication event happened 30–45 million years ago in pear, but not
185 in peach, yang mei and strawberry (Wu et al. 2013). This result suggested the probable reason for occurrence
186 of Class A2 genes in pear. Remarkably, *EIN3/EIL* genes from peach and mei showed higher similarity with
187 each other according to genetic distance, which was consistent with a previous study reporting that the closer
188 relationship between peach and mei versus peach and pear/ strawberry (Cao et al. 2016a; Cao et al. 2016b; Cao
189 et al. 2016d).
190 Each of the four Rosaceae species contributed at least one members of *EIN3/EIL* gene to each class, with

191 the exception for Class A2 (Figure 2). Therefore, we deduced that *EIN3/EIL* genes had rapidly been duplicated
192 before these dicotyledon species diverged. However, only the *EIN3/EIL* genes in class A2 revealed an internal
193 duplication. In addition, we identified three pairs of orthologous genes among the *EIN3/EIL* genes: *PmEIL3*
194 and *PpEIL1*, *PpEIL2* and *PmEIL4*, *PpEIL3* and *PmEIL5* based on the phylogenetic analysis.
195
196 **Structural analysis of *EIN3/EIL* genes**
197 Previous studies have suggested that gene structural diversity is the primary resource for the evolution of
198 multigene families (Cao et al. 2016c; Leitch & Leitch 2013; Mercereau-Puijalon et al. 2002). To characterize
199 the structural diversity of the *EIN3/EIL* gene family, exon-intron organization of each *EIN3/EIL* genes was
200 analyzed. As shown in Figure 3A, most genes did not contain introns, such as *FvEIL1*, *PbEIL1*, *PmEIL3* and
201 *PpEIL4* et al. Furthermore, *PbEIL7* contained 8 introns, followed by *FvEIL3* (5), whereas *FvEIL5* had three
202 introns, *PmEIL2* had two introns and 8 *EIN3/EIL* genes contained one intron (Figure 3A). These results
203 implied that the intron/exon loss and acquire has occurred in the evolution of the *EIN3/EIL* gene family, which
204 may be able to explain the functional divergence of closely related *EIN3/EIL* genes. In present study, the gene
205 structures of the *EIN3/EIL* homologous gene pairs were investigated. We found the exon number of two gene
206 pairs (*PbEIL2*/*PbEIL8* and *PmEIL1*/*PmEIL2*) had changed. Further analysis indicated that *PbEIL8* and
207 *PmEIL2* obtained one exon during evolution, while *PbEIL2* and *PmEIL1* lost one exon. These diversities
208 might be due to single intron loss or obtain events during evolution.
209 Because 24 *EIN3/EIL* genes did not have high similarity, MEME web server was used to find conserved
210 motifs. Subsequently, we identified twenty conserved motifs, which were shown in Table S2 and Figure 3B.
211 The Pfam and SMART database were used to annotate the individual of the putative motifs. Motif 1 and motif

2 were identified to encode a conserved EIN3/EIL domain (Chang & Shockey 1999; Chao et al. 1997a), whereas the remaining motifs did not get function annotation. Most EIN3/EIL proteins have motifs 1, 2, 4, 7, 10 and 19. In addition, several proteins from clade B2 contained unique motif 12, which might imply its specific functions. Remarkably, most of the closely related EIN3/EIL proteins in the same clade exhibited similar motif compositions (e.g. PmEIL5/PpEIL3 and PbEIL3/PbEIL5), indicating their functional similarity among these EIN3/EIL proteins. Summarily, the similarity in motif distribution and exon-intron structure of most EIN3/EIL proteins supported the results from phylogenetic analysis of the *EIN3/EIL* genes, whereas the differences of the related characteristics in the different classes indicated that their functions were diversified.

Conserved microsynteny of *EIN3/EIL* genes in the four Rosaceae species

Based on the whole-genome data, species microsynteny can be used to identify the location of orthologous genes and/or paralogous genes (Cao et al. 2016d; Lin et al. 2014). To identify the homologous genes (orthology or paralogy) within the *EIN3/EIL* genes from four Rosaceae species (pear, peach, mei and strawberry), as well as their evolutionary history, microsynteny analysis was performed. By pairwise comparisons of flanking sequences in the chromosomal regions containing *EIN3/EIL* genes, three or more pairs were present in this region, which were considered as either conserved microsynteny or high levels of microsynteny.

In this study, a total of 55 flanking sequences containing *EIN3/EIL* genes could be assembled into 28 regions and divided into seven microsynteny groups. It was supposed that *EIN3/EIL* genes from the same group should evolve from the most recent common ancestor. Based on this criterion, orthology and/or paralogy relationships, as well as their evolutionary origins, were detected among *EIN3/EIL* genes of the four Rosaceae

species. Nine, five, four and four out of the seven microsynteny groups are from pear, peach, mei and strawberry, respectively (Figure 4). In class B2, two gene pairs (*PbEIL1* and *PbEIL9*, *PbEIL9* and *PbEIL10*) with both a higher level of microsynteny and a noticeable inverted duplication, were identified. Interestingly, it was also found that some duplication rules in several regions were in disorder, such as *FvEIL1* and *FvEIL2*, *FvEIL1* and *PpEIL4* (Figure 4). Similar microsynteny was identified in other classes with a concordant inverted microsynteny (Figure 4). In addition, according to the constructed phylogenetic tree, conservation of microsynteny between different families appeared gradually. However, some flanking genes in each microsyntenic group were not conserved, indicating they arose later than this duplication event (Figure 4). Furthermore, we only identified four pairs of intraspecies microsynteny groups from pear (*PbEIL1* and *PbEIL9*, *PbEIL3* and *PbEIL5*, *PbEIL4* and *PbEIL6*, *PbEIL9* and *PbEIL10*), but peach, mei and strawberry were excluded (Figure 5). This difference might be resulted from the expansion of pear *EIN3/EIL* genes. However, no similar gene expansion was identified in peach, mei and strawberry. Some previous studies has hypothesized that transcription factors should be generally and preferentially retained after genome duplications (Blanc & Wolfe 2004), with a lower frequency of tandem duplication events in a number of transcription factor (Freeling 2009). Additionally, genes from whole-genome duplication events are more easily retained into genomes. With the stoichiometric relationships, these genes were strongly retained by stabilizing selection (Lynch & Conery 2000). Our results were not only consistent with this hypothesis, but also strong evidences for it.

Subsequently, the quality of the synteny was estimated in four Rosaceae plants based on previous research methods (Cannon et al. 2003; Cannon et al. 2006). As shown in Table 2, the relative synteny quality of the *EIN3/EIL* genes from these Rosaceae four species genomes was 24.43 % for orthologous regions. The

highest value of synteny quality found between pear and peach was 33.29%. And the lower value of synteny quality was obtained between strawberry and peach (10.26%) and mei (13.26%) The relative synteny quality in the pear/mei and pear/strawberry syntenic regions was 31.35% and 26.17%, which was substantially lower than 32.13% found in the pear/peach synteny blocks. Our results were essentially consistent with their evolutionary relationship (Xiang et al. 2016; Zhong et al. 2015).

Strong purifying selection for *EIN3/EIL* genes in four Rosaceae species

In general, Ks values can be used to estimate evolutionary data of the whole genome duplication events or segmental duplication events. Previous studies showed that pear had experienced two whole genome duplication events, including an ancient whole genome duplication (Ks ~1.5–1.8) estimated at ~140 MYA (Fawcett et al. 2009) and a recent whole genome duplication (Ks ~0.15–0.3) estimated at 30–45 MYA (Wu et al. 2013)), while peach, mei and strawberry only experienced an ancient whole genome duplication event. Therefore, Ks values were applied to analyze the whole genome duplication or segmental duplication events in *EIN3/EILs* of four Rosaceae species. As shown in Table S3, the mean Ks values of each duplication pairs in the syntenic region were lists. In pear, we found the mean Ks values of *EIN3/EIL* gene pairs were 0.0363, 0.1717 and 0.2836, respectively. It was obvious that these duplications might be resulting from the latest whole genome duplication (30–45 MYA; Ks ~0.15–0.3), but an ancient whole genome duplication (~140MYA; Ks ~1.5–1.8) in pear.

In addition, the Ka/Ks values are widely used to represent the gene selection pressure and evolution rate [40]: Ka/Ks value with > 1 indicates positive selection with accelerated evolution, Ka/Ks < 1 indicates negative/purifying selection with the functional constraint, and Ka/Ks = 1 suggests that the genes are drifting

neutrally. In this study, all paralogs was found with Ka/Ks ratios < 1 (Figure 6), indicating their purifying selection. Furthermore, to better understand the delineate regions of diversifying and purifying selection in the *EIN3/EIL* gene family, a sliding window analysis of the Ka/Ks values between paralogs was performed (Figure 4), the EIN3/EIL domains in the N-termini exhibited stronger purifying selection compared with the whole gene regions (C-termini). These results suggested that the *EIN3/EIL* genes had undergone strongly purifying selection, especially for EIN3/EIL domains in the N-termini (Figure 4Q). Overall, strong evolutionary constraints were involved in *EIN3/EIL* gene evolution, which may contribute to their functional stability. On the other hand, some parts of protein-coding genes had undergone positive selection, implying the generation of innovative gene functions.

Expression profiles analysis of *PbEIL* genes in different tissues

To increase our understanding the potential functions of pear *EIN3/EIL* genes during development, the qRT-PCR analysis was carried out to determine the expression profiles of 10 *PbEIL* genes in different tissues. As shown in Figure 7 and Table S4, 10 pear *EIN3/EIL* genes showed significantly different tissue-specific expression patterns in eight samples from root, stem, leaves and fruits in several development stages. Among the 10 pear *EIN3/EIL* genes, three (*PbEIL5*, *PbEIL6* and *PbEIL10*) showed the highest transcript accumulation in the leaves, three (*PbEIL2*, *PbEIL3* and *PbEIL9*) in 145 DAF (days after flowering), two (*PbEIL1* and *PbEIL4*) in 79 DAF, one (*PbEIL7*) in roots. Additionally, the duplication gene pairs showed different expression patterns; for example, *PbEIL4* was highly expressed in 79 DAF, while its duplication gene, *PbEIL6*, was expressed at a high level in the leaves. Thus, the pear *EIN3/EIL* duplicates resulting from recent whole genome duplication have different expression patterns in several different tissues, indicating

296 subfunctionalization after duplication. At the same time, this phenomenon was also observed among other
297 *EIN3/EIL* duplication genes (Jourda et al. 2014).

298 **Conclusions**

299 In this study, we identified 24 *EIN3/EIL* genes from four Rosaceae species (pear, peach, mei and strawberry).
300 Subsequently, a systematic analysis, including their chromosomal location, evolutionary relationship,
301 conserved microsynteny, gene structure and sliding window, was carried out. According to phylogenetic
302 analysis, the *EIN3/EIL* genes divided into four classes. Remarkably, high level of microsynteny of *EIL/EIN3*
303 family in Rosaceae was found, indicating that the genome duplication plays a key role in the expansion of the
304 *EIL/EIN3* genes in the Rosaceae. In these *EIL/EIN3* genes, all paralogs have experienced purifying selection,
305 especially the EIL/EIN3 domains in the Rosaceae. Furthermore, the expression profiles of the *PbEIL* genes
306 suggested that the recent whole genome duplication derived genes show indications of subfunctionalization.
307 These results may help promote the extrapolation of *EIL/EIN3* gene functions in future.

308

309 **ACKNOWLEDGEMENTS**

310 We would like to thank Muhammad Abdullah for his careful reading and helpful comments on this manuscript.
311 We extend our thanks to the reviewers and editors for their careful reading and helpful comments on this
312 manuscript.

313

314 **References**

315 Blanc G, and Wolfe KH. 2004. Widespread paleopolyploidy in model plant species inferred from age distributions of
316 duplicate genes. *Plant Cell* 16:1667-1678.
317 Cannon SB, Mccombie WR, Sato S, Tabata S, Denny R, Palmer L, Katari M, Young ND, and Stacey G. 2003. Evolution
318 and microsynteny of the apyrase gene family in three legume genomes. *Molecular Genetics and Genomics*

270:347-361.

Cannon SB, Sterck L, Rombauts S, Sato S, Cheung F, Gouzy J, Wang X, Mudge J, Vasdewani J, and Schiex T. 2006. Legume genome evolution viewed through the Medicago truncatula and Lotus japonicus genomes. *Proceedings of the National Academy of Sciences of the United States of America* 103:14959-14964.

Cao Y, Han Y, Li D, Lin Y, and Cai Y. 2016a. MYB Transcription Factors in Chinese Pear (Pyrus bretschneideri Rehd.): Genome-Wide Identification, Classification, and Expression Profiling during Fruit Development. *Frontiers in Plant Science* 7.

Cao Y, Han Y, Li D, Lin Y, and Cai Y. 2016b. Systematic Analysis of the 4-Coumarate:Coenzyme A Ligase (4CL) Related Genes and Expression Profiling during Fruit Development in the Chinese Pear. *Genes* 7:89.

Cao Y, Han Y, Meng D, Li D, Jin Q, Lin Y, and Cai Y. 2016c. Structural, Evolutionary, and Functional Analysis of the Class III Peroxidase Gene Family in Chinese Pear (Pyrus bretschneideri). 7.

Cao YP, Han Y, Jin Q, Lin Y, and Cai Y. 2016d. Comparative genomic analysis of the GRF genes in Chinese pear (Pyrus bretschneideri Rehd), poplar (Populous), grape (Vitis vinifera), Arabidopsis and rice (Oryza sativa). *Frontiers in Plant Science* 7. 10.3389/fpls.2016.01750

Chang C, and Shockey JA. 1999. The ethylene-response pathway: signal perception to gene regulation. *Current opinion in plant biology* 2:352-358.

Chao Q, Rothenberg M, Solano R, Roman G, Terzaghi W, and Ecker JR. 1997a. Activation of the ethylene gas response pathway in Arabidopsis by the nuclear protein ETHYLENE-INSENSITIVE3 and related proteins. *Cell* 89:1133-1144.

Chao Q, Rothenberg M, Solano R, Roman G, Terzaghi W, and Ecker JR. 1997b. Activation of the ethylene gas response pathway in Arabidopsis by the nuclear protein ETHYLENE-INSENSITIVE3 and related proteins. *Cell* 89:1133-1144.

Fawcett JA, Maere S, and Van dPY. 2009. Plants with double genomes might have had a better chance to survive the Cretaceous-Tertiary extinction event. *Proceedings of the National Academy of Sciences* 106:5737-5742.

Finn RD, Mistry J, Schuster-Böckler B, Griffiths-Jones S, Hollich V, Lassmann T, Moxon S, Marshall M, Khanna A, and Durbin R. 2006. Pfam: clans, web tools and services. *Nucleic acids research* 34:D247-D251.

Freeling M. 2009. Bias in plant gene content following different sorts of duplication: tandem, whole-genome, segmental, or by transposition. *Annual Review of Plant Biology* 60:433-453.

Guo H, and Ecker JR. 2004. The ethylene signaling pathway: new insights. Curr Opin Plant Biol 7:40-49. *Current opinion in plant biology* 7:40-49.

Han Y, Ding T, Su B, and Jiang H. 2016. Genome-Wide Identification, Characterization and Expression Analysis of the Chalcone Synthase Family in Maize. *International journal of molecular sciences* 17:161.

Hiraga S, Sasaki K, Hibi T, Yoshida H, Uchida E, Kosugi S, Kato T, Mie T, Ito H, and Katou S. 2009. Involvement of two rice ETHYLENE INSENSITIVE3-LIKE genes in wound signaling. *Molecular Genetics and Genomics* 282:517-529.

Hu B, Jin J, Guo YA, Zhang H, Luo J, and Gao G. 2014. GSDS 2.0: an upgraded gene feature visualization server. *Bioinformatics* 31:1296.

Illa E, Sargent DJ, Girona EL, Bushakra J, Cestaro A, Crowhurst R, Pindo M, Cabrera A, van der Knaap E, and Iezzoni A. 2011. Comparative analysis of rosaceous genomes and the reconstruction of a putative ancestral genome for the family. *BMC Evolutionary biology* 11:1.

360 Jing J, Kong J, Qiu J, Zhu H, Peng Y, and Jiang H. 2016. High level of microsynteny and purifying selection affect the
361 evolution of WRKY family in Gramineae. *Development Genes & Evolution* 226:15-25.
362 Jourda C, Cardi C, Mbéguiéambéguié D, Bocs S, Garsmeur O, D'Hont A, and Yahiaoui N. 2014. Expansion of banana
363 (Musa acuminata) gene families involved in ethylene biosynthesis and signalling after lineage-specific
364 whole-genome duplications. *New Phytologist* 202:986-1000.
365 Kosugi S, and Ohashi Y. 2000. Cloning and DNA-binding properties of a tobacco Ethylene-Insensitive3 (EIN3)
366 homolog. *Nucleic acids research* 28:960-967.
367 Lee JH, and Kim WT. 2003. Molecular and biochemical characterization of VR-EILs encoding mung bean ETHYLENE
368 INSENSITIVE3-LIKE proteins. *Plant Physiology* 132:1475-1488.
369 Leitch IJ, and Leitch AR. 2013. *Genome Size Diversity and Evolution in Land Plants*.
370 Letunic I, Doerks T, and Bork P. 2012. SMART 7: recent updates to the protein domain annotation resource. *Nucleic*
371 *acids research* 40:D302-D305.
372 Librado P, and Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data.
373 *Bioinformatics* 25:1451-1452.
374 Lin Y, Cheng Y, Jin J, Jin X, Jiang H, Yan H, and Cheng B. 2014. Genome Duplication and Gene Loss Affect the
375 Evolution of Heat Shock Transcription Factor Genes in Legumes. *PLoS One* 9:e102825.
376 Lynch M, and Conery JS. 2000. The evolutionary fate and consequences of duplicate genes. *Science* 290:1151-1155.
377 Mao C, Wang S, Jia Q, and Wu P. 2006. OsEIL1, a rice homolog of the Arabidopsis EIN3 regulates the ethylene
378 response as a positive component. *Plant molecular biology* 61:141-152.
379 Mercereau-Puijalon O, Barale JC, and Bischoff E. 2002. Three multigene families in Plasmodium parasites: facts and
380 questions. *International Journal for Parasitology* 32:1323-1344.
381 Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, Pang N, Forslund K, Ceric G, and Clements J. 2011.
382 The Pfam protein families database. *Nucleic acids research*:gkr1065.
383 Rieu I, Mariani C, and Weterings K. 2003. Expression analysis of five tobacco EIN3 family members in relation to
384 tissue‐specific ethylene responses. *Journal of Experimental Botany* 54:2239-2244.
385 Shulaev V, Sargent DJ, Crowhurst RN, Mockler TC, Folkerts O, Delcher AL, Jaiswal P, Mockaitis K, Liston A, and
386 Mane SP. 2011. The genome of woodland strawberry (Fragaria vesca). *Nature genetics* 43:: 109–116.
387 Tamura K, Peterson D, Peterson N, Stecher G, Nei M, and Kumar S. 2011. MEGA5: molecular evolutionary genetics
388 analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular*
389 *biology and evolution* 28:2731-2739.
390 Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, and Higgins DG. 1997. The CLUSTAL_X windows interface:
391 flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic acids research*
392 25:4876-4882.
393 Tieman DM, Ciardi JA, Taylor MG, and Klee HJ. 2001. Members of the tomato LeEIL (EIN3‐like) gene family are
394 functionally redundant and regulate ethylene responses throughout plant development. *The Plant Journal*
395 26:47-58.
396 Verde I, Abbott AG, Scalabrin S, Jung S, Shu S, Marroni F, Zhebentyayeva T, Dettori MT, Grimwood J, and Cattonaro
397 F. 2013. The high-quality draft genome of peach (Prunus persica) identifies unique patterns of genetic
398 diversity, domestication and genome evolution. *Nature genetics* 45:487-494.
399 Waki K, Shibuya K, Yoshioka T, Hashiba T, and Satoh S. 2001. Cloning of a cDNA encoding EIN3‐like protein
400 (DC‐EIL1) and decrease in its mRNA level during senescence in carnation flower tissues. *Journal of*

401 *Experimental Botany* 52:377-379.
402 Wang Y, Feng L, Zhu Y, Li Y, Yan H, and Xiang Y. 2015. Comparative genomic analysis of the WRKY III gene family in
403 populus, grape, arabidopsis and rice. *Biology direct* 10:1.
404 Wang Y, Tang H, DeBarry JD, Tan X, Li J, Wang X, Lee T-h, Jin H, Marler B, and Guo H. 2012. MCScanX: a toolkit for
405 detection and evolutionary analysis of gene synteny and collinearity. *Nucleic acids research* 40:e49-e49.
406 Wendel JF. 2000. Genome evolution in polyploids. *Plant molecular evolution*: Springer, 225-249.
407 Wu J, Wang Z, Shi Z, Zhang S, Ming R, Zhu S, Khan MA, Tao S, Korban SS, and Wang H. 2013. The genome of the
408 pear (Pyrus bretschneideri Rehd.). *Genome research* 23:396-408.
409 Wu T, Zhang R, Gu C, Wu J, Wan H, Zhang S, and Zhang S. 2012. Evaluation of candidate reference genes for real
410 time quantitative PCR normalization in pear fruit. *Afr J Agric Res* 7:3701-3704.
411 Xiang Y, Huang CH, Yi H, Wen J, Li S, Yi T, Chen H, Xiang J, and Hong M. 2016. Evolution of Rosaceae fruit types
412 based on nuclear phylogeny in the context of geological times and genome duplication. *Molecular Biology*
413 *& Evolution*.
414 Yang ZP, Li HL, Guo D, and Peng SQ. 2014. Identification and characterization of the EIN3/EIL gene family in Hevea
415 brasiliensis. *Plant Physiology & Biochemistry* 11:1-11.
416 Yokotani N, Tamura S, Nakano R, Inaba A, and Kubo Y. 2003. Characterization of a novel tomato EIN3 - like gene
417 (LeEIL4). *Journal of Experimental Botany* 54:2775-2776.
418 Zdobnov EM, and Apweiler R. 2001. InterProScan--an integration platform for the signature-recognition methods
419 in InterPro. *Bioinformatics* 17:847-848.
420 Zhang Q, Chen W, Sun L, Zhao F, Huang B, Yang W, Tao Y, Wang J, Yuan Z, and Fan G. 2012. The genome of Prunus
421 mume. *Nature Communications* 3:1318.
422 Zhong Y, Yin H, Sargent DJ, Malnoy M, and Cheng ZM. 2015. Species-specific duplications driving the recent
423 expansion of NBS-LRR genes in five Rosaceae species. *BMC genomics* 16:1-16.

424

425

426

427

428

Figure 1(on next page)

Figure 1. Chromosomal location of *EIN3/EIL* genes in the genomes of strawberry (A), mei (B), pear (C) and peach (D).

The distribution of *EIN3/EIL* genes among the chromosomes in each species was diverse. The chromosome number was represented at the top of each chromosome. The left scale indicates the megabases (Mb).

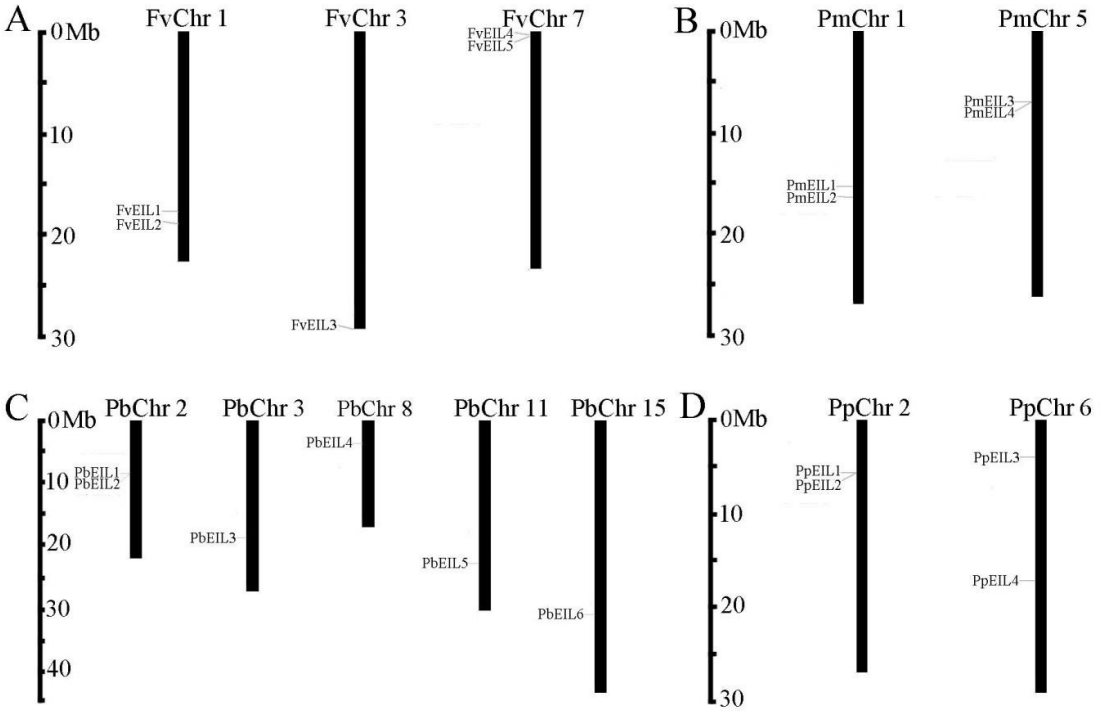
A
0Mb
10
20
30
FvChr 1
FvEIL1
FvEIL2
FvChr 3
FvEIL3
FvChr 7
FvEIL4
FvEIL5
B
0Mb
10
20
30
PmChr 1
PmEIL1
PmEIL2
PmChr 5
PmEIL3
PmEIL4
C
0Mb
10
20
30
40
PbChr 2
PbEIL1
PbEIL2
PbChr 3
PbEIL3
PbChr 8
PbEIL4
PbChr 11
PbEIL5
PbChr 15
PbEIL6
D
0Mb
10
20
30
PpChr 2
PpEIL1
PpEIL2
PpChr 6
PpEIL3
PpEIL4

Figure 2(on next page)

Figure 2. Phylogenetic tree of EIN3/EIL proteins from pear, peach, yangmei and strawberry. The neighbor-joining (NJ) tree was constructed by using MEGA 5 software.

The neighbor-joining (NJ) tree was constructed by using MEGA 5 software (1000 bootstrap replicates). The different colors suggest the different species background for each EIN3/EIL protein. Gene names are listed in Table 1. The scale bar represents 0.1 amino acid changes per site.

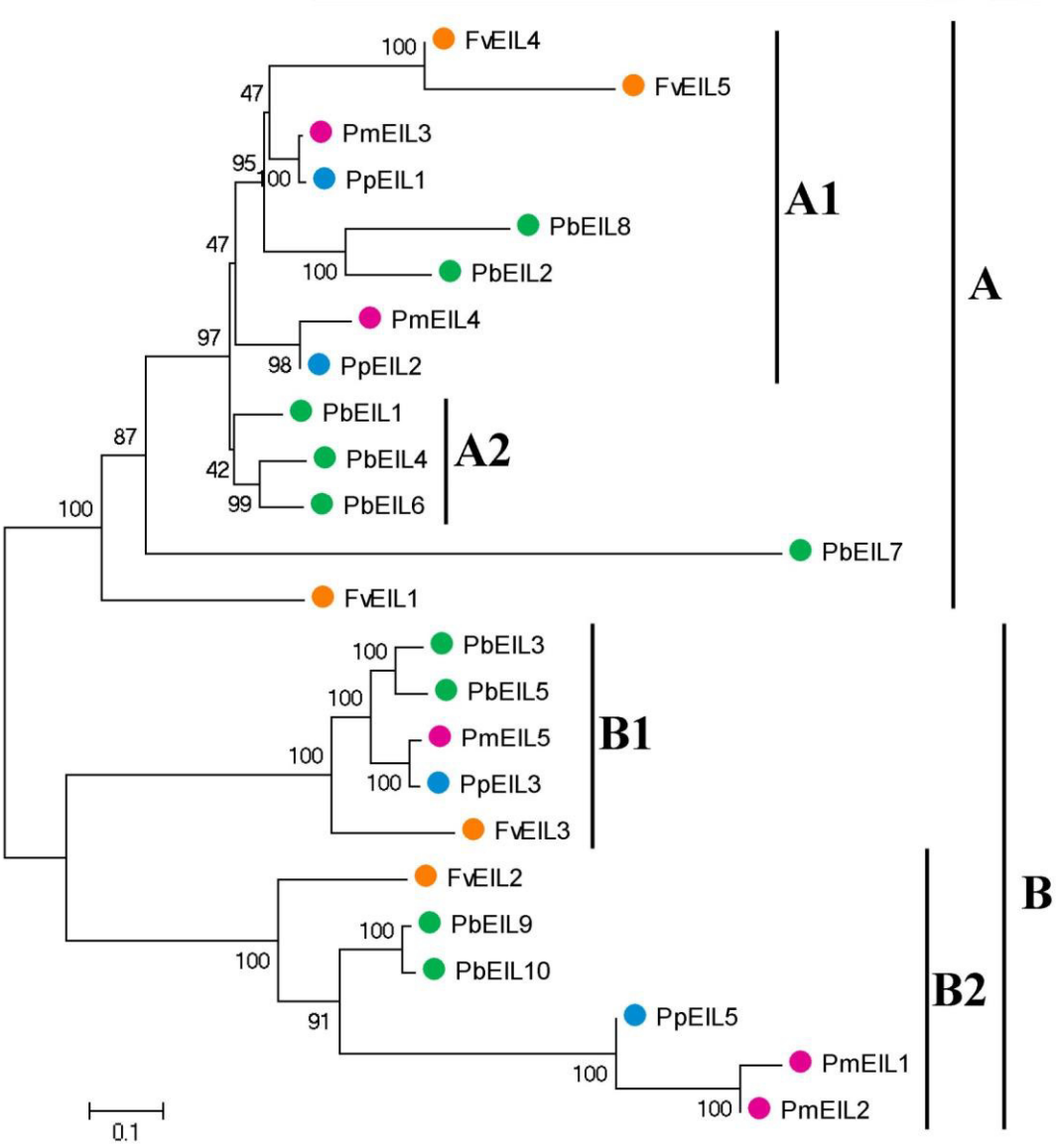
FvEIL4
FvEIL5
PmEIL3
PpEIL1
PbEIL8
PbEIL2
PmEIL4
PpEIL2
PbEIL1
PbEIL4
PbEIL6
PbEIL7
FvEIL1
PbEIL3
PbEIL5
PmEIL5
PpEIL3
FvEIL3
FvEIL2
PbEIL9
PbEIL10
PpEIL5
PmEIL1
PmEIL2
A1
A2
A
B1
B2
B
0.1

Figure 3(on next page)

Figure 3. Gene structure (A) and conserved motif compositions (B) of *EIN3/EIL* genes in Rosaceae species.

Untranslated regions (UTRs), introns and exons are represented by blue boxes, thin lines and green rectangles, respectively. Note that the gene or protein lengths can be estimated by using the scale at the bottom.

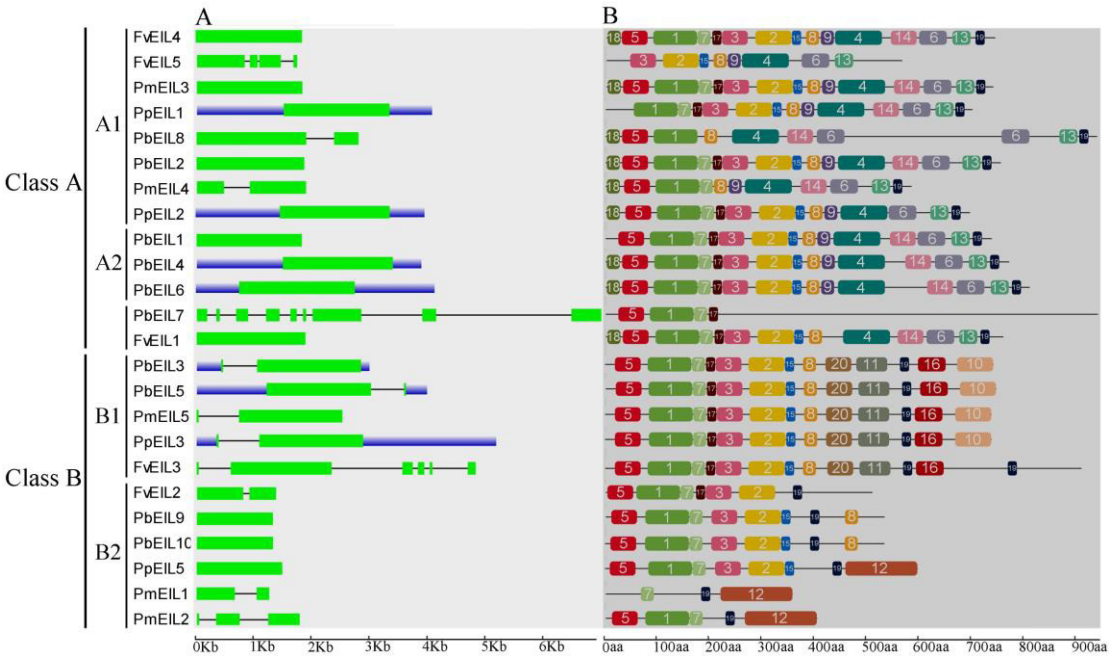
A
B
Class A
A1
A2
Class B
B1
B2
FvEIL4
FvEIL5
PmEIL3
PpEIL1
PbEIL8
PbEIL2
PmEIL4
PpEIL2
PbEIL1
PbEIL4
PbEIL6
PbEIL7
FvEIL1
PbEIL3
PbEIL5
PmEIL5
PpEIL3
FvEIL3
FvEIL2
PbEIL9
PbEIL10
PpEIL5
PmEIL1
PmEIL2
0Kb
1Kb
2Kb
3Kb
4Kb
5Kb
6Kb
0aa
100aa
200aa
300aa
400aa
500aa
600aa
700aa
800aa
900aa

Figure 4(on next page)

Figure 4. Interspecies microsynteny related to EIN3/EIL families in four Rosaceae.

The relative positions of all flanking protein-coding genes were defined by anchored *EIN3/EIL* genes, highlighted in red. The gene's orientations are shown as triangle, with gray lines corresponding to chromosomal segments.

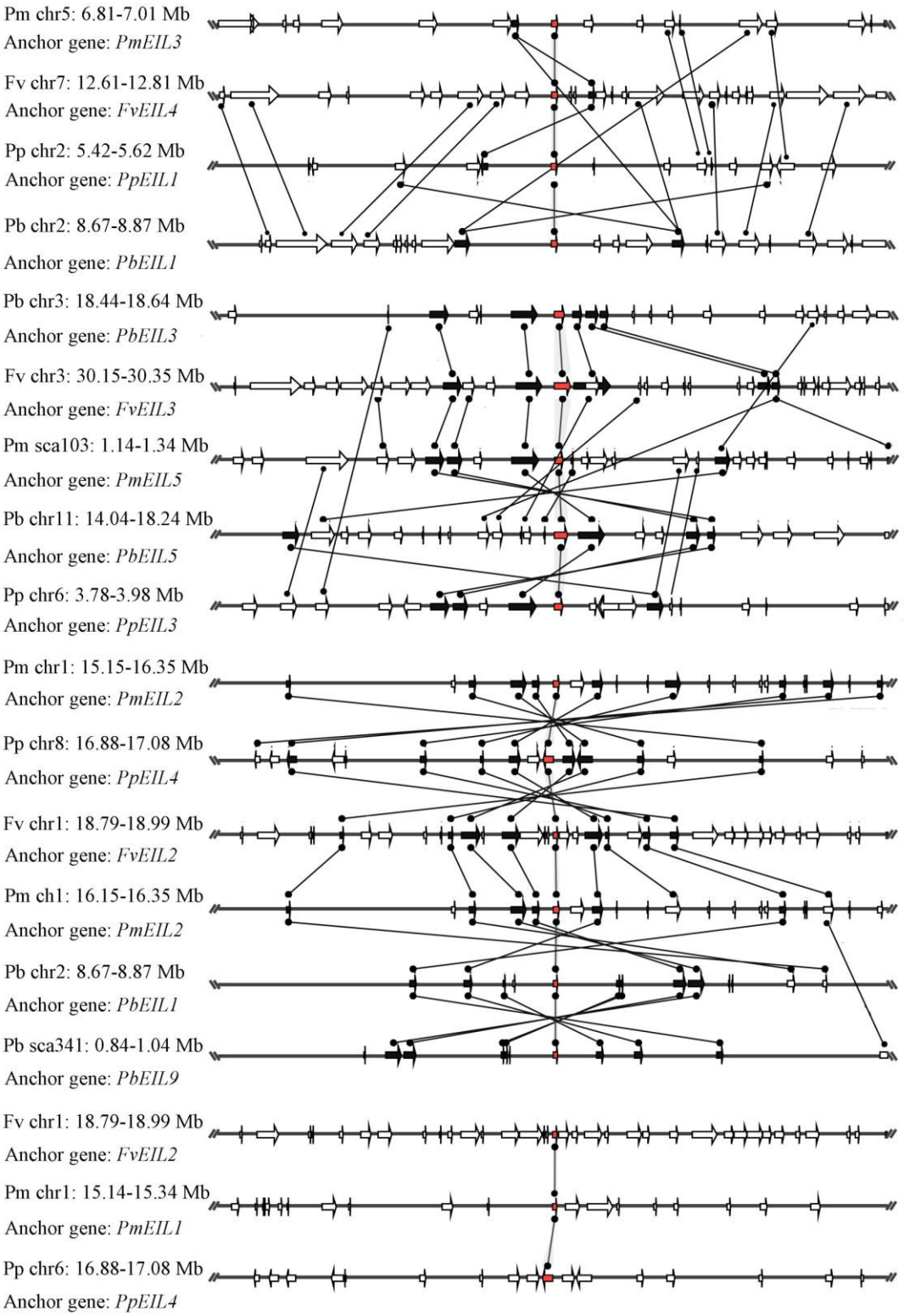
Pm chr5: 6.81-7.01 Mb
Anchor gene: PmEIL3
Fv chr7: 12.61-12.81 Mb
Anchor gene: FvEIL4
Pp chr2: 5.42-5.62 Mb
Anchor gene: PpEIL1
Pb chr2: 8.67-8.87 Mb
Anchor gene: PbEIL1
Pb chr3: 18.44-18.64 Mb
Anchor gene: PbEIL3
Fv chr3: 30.15-30.35 Mb
Anchor gene: FvEIL3
Pm sca103: 1.14-1.34 Mb
Anchor gene: PmEIL5
Pb chr11: 14.04-18.24 Mb
Anchor gene: PbEIL5
Pp chr6: 3.78-3.98 Mb
Anchor gene: PpEIL3
Pm chr1: 15.15-16.35 Mb
Anchor gene: PmEIL2
Pp chr8: 16.88-17.08 Mb
Anchor gene: PpEIL4
Fv chr1: 18.79-18.99 Mb
Anchor gene: FvEIL2
Pm ch1: 16.15-16.35 Mb
Anchor gene: PmEIL2
Pb chr2: 8.67-8.87 Mb
Anchor gene: PbEIL1
Pb sca341: 0.84-1.04 Mb
Anchor gene: PbEIL9
Fv chr1: 18.79-18.99 Mb
Anchor gene: FvEIL2
Pm chr1: 15.14-15.34 Mb
Anchor gene: PmEIL1
Pp chr6: 16.88-17.08 Mb
Anchor gene: PpEIL4

Figure 5(on next page)

Figure 5. Intraspecific microsynteny related to *EIN3/EIL* families with the same species.

The relative positions of all flanking protein-coding genes were defined by anchored *EIN3/EIL* genes, highlighted in red. The gene's orientations are shown as triangle, with gray lines corresponding to chromosomal segments.

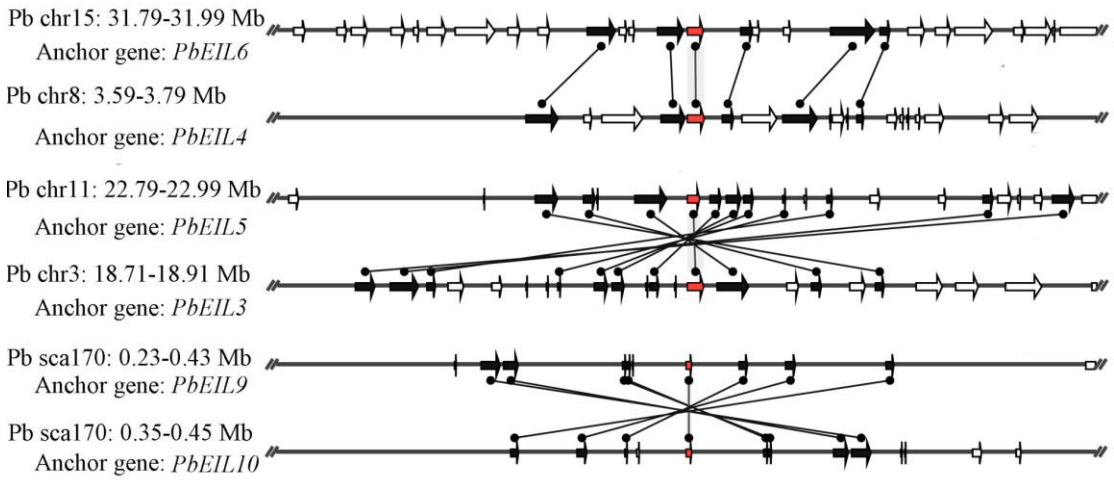
Pb chr15: 31.79-31.99 Mb
Anchor gene: PbEIL6
Pb chr8: 3.59-3.79 Mb
Anchor gene: PbEIL4
Pb chr11: 22.79-22.99 Mb
Anchor gene: PbEIL5
Pb chr3: 18.71-18.91 Mb
Anchor gene: PbEIL3
Pb sca170: 0.23-0.43 Mb
Anchor gene: PbEIL9
Pb sca170: 0.35-0.45 Mb
Anchor gene: PbEIL10

Figure 6(on next page)

Figure 6. Sliding window plots of duplicated EIN3/EIL genes in Rosaceae species.

The gray blocks indicate the positions of the EIN3/EIL domains. The window size is 150 bp, and the step size is 9 bp. The x-axis denotes the synonymous distances within each gene.

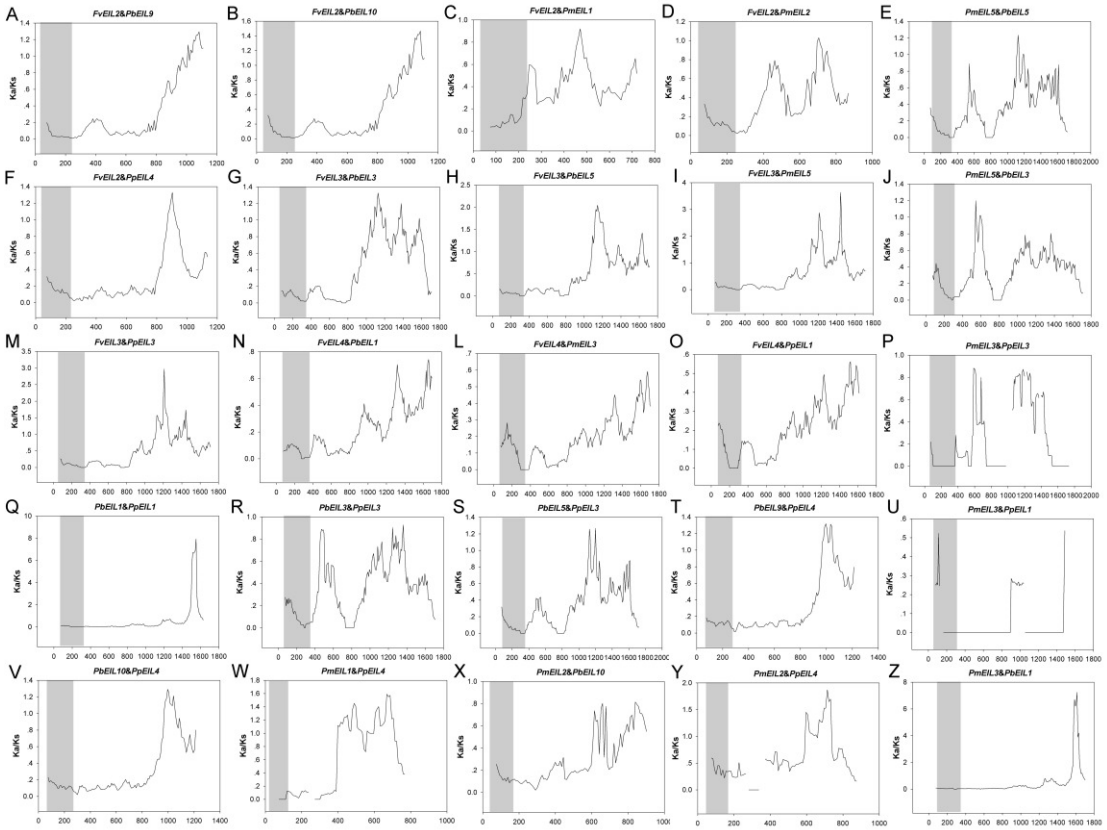
A
FvEIL2&PbEIL9
B
FvEIL2&PbEIL10
C
FvEIL2&PmEIL1
D
FvEIL2&PmEIL2
E
PmEIL5&PbEIL5
F
FvEIL2&PpEIL4
G
FvEIL3&PbEIL3
H
FvEIL3&PbEIL5
I
FvEIL3&PmEIL5
J
PmEIL5&PbEIL3
M
FvEIL3&PpEIL3
N
FvEIL4&PbEIL1
L
FvEIL4&PmEIL3
O
FvEIL4&PpEIL1
P
PmEIL3&PpEIL3
Q
PbEIL1&PpEIL1
R
PbEIL3&PpEIL3
S
PbEIL5&PpEIL3
T
PbEIL9&PpEIL4
U
PmEIL3&PpEIL1
V
PbEIL10&PpEIL4
W
PmEIL1&PpEIL4
X
PmEIL2&PbEIL10
Y
PmEIL2&PpEIL4
Z
PmEIL3&PbEIL1
Ka/Ks

Figure 7(on next page)

Figure 7. Expression profiling of pear *PbEIL* genes in eight samples from root, stem, leaves and fruits in several development stage .

The expression profile data was obtained qRT-PCR experiment. Blue and red colors indicate low-expression and high-expression, respectively.

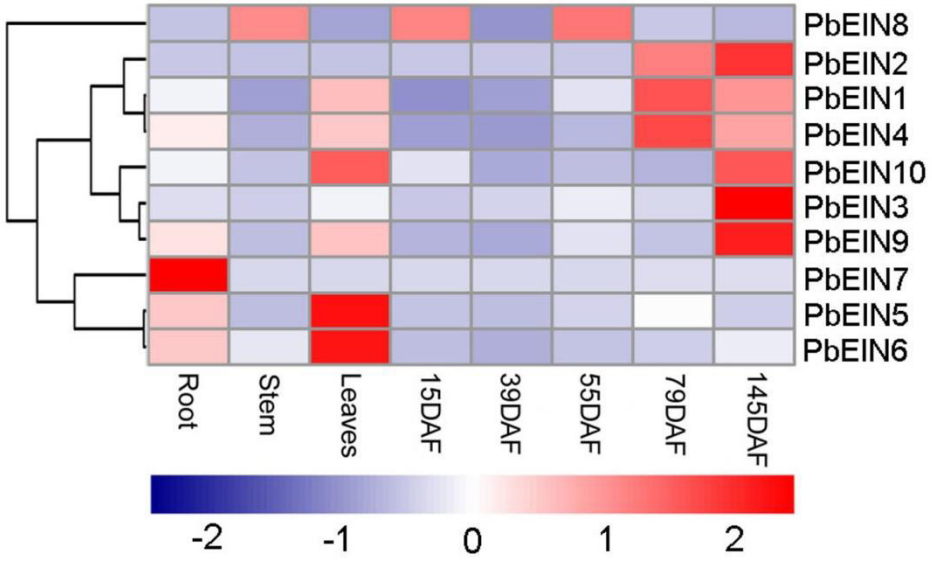
PbEIN8
PbEIN2
PbEIN1
PbEIN4
PbEIN10
PbEIN3
PbEIN9
PbEIN7
PbEIN5
PbEIN6
Root
Stem
Leaves
15DAF
39DAF
55DAF
79DAF
145DAF
-2
-1
0
1
2

Table 1(on next page)

Table 1. List of *EIN3/EIL* genes identified in pear, peach, mei and strawberry.

Note: Pear gene models are found in the GigaDB Genome database; mei and peach gene models are found in the Rosaceae Genome Database; strawberry gene models are found in the Phytozome database.

1 **Table 1**. List of EIN3/EIL genes identified in pear, peach, mei and strawberry.

Name	Gene Model	Chromosme	5' End	3' End
FvEIL1	mrna25474.1	Chr1	17653671	17655527
FvEIL2	mrna16361.1	Chr1	18891616	18892965
FvEIL3	mrna20650.1	Chr3	29248944	29253704
FvEIL4	mrna00379.1	Chr7	290967	292781
FvEIL5	mrna00392.1	Chr7	349495	351202
PmEIL1	Pm001950	Chr1	15239829	15241073
PmEIL2	Pm002057	Chr1	16248534	16250294
PmEIL3	Pm017009	Chr5	6907428	6909233
PmEIL4	Pm017011	Chr5	6933006	6934874
PmEIL5	Pm028171	scaffold103	1235430	1246520
PpEIL1	ppa003493m	Chr2	5516222	5518429
PpEIL2	ppa003550m	Chr2	5549949	5552334
PpEIL3	ppa003113m	Chr6	3882268	3885188
PpEIL4	ppa016118m	Chr6	16979366	16982360
PbEIL1	Pbr024739.1	Chr2	8493409	8495211
PbEIL2	Pbr024740.1	Chr2	8506285	8508129
PbEIL3	Pbr000646.1	Chr3	18718500	18721454
PbEIL4	Pbr026603.1	Chr8	3598382	3602224

PbEIL5	Pbr004535.1	Chr11	22794386	22798326
PbEIL6	Pbr033210.1	Chr15	31009957	31014042
PbEIL7	Pbr010447.1	scaffold170.2.1	239188	246113
PbEIL8	Pbr010448.1	scaffold170.2.1	259361	262132
PbEIL9	Pbr022557.1	scaffold341.0	56672	57973
PbEIL10	Pbr039294.1	scaffold837.0	82840	84144

Note: Pear gene models are found in the GigaDB Genome database; mei and peach gene models are found in the Rosaceae Genome Database; strawberry gene models are found in the Phytozome database.

Table 2(on next page)

Table 2. The relative syntenic quality of *EIN3/EIL* genes in four Rosaceae plants.

Note: the relative syntenic quality was estimated as twice the number of matches divided by the sum of the total number of genes in both conserved gene regions, based on the previous methods (Cannon et al. 2003 ; Cannon et al. 2006) .

Table 2. The relative syntenic quality of *EIN3/EIL* genes in four Rosaceae plants.

	Clade A1	Clade A2	Clade B1	Clade B2	Average
Pb-Pp			22.50%	44.07%	33.29%
Pb-Pm			21.51%	41.18%	31.35%
Pb-Fv			24.07%	28.26%	26.17%
Pp-Pm	26.67%		40.00%	29.73%	32.13%
Pp-Fv	10.26%				10.26%
Pm-Fv	10.00%		26.32%	3.77%	13.36%
					24.43%

Note: the relative syntenic quality was estimated as twice the number of matches divided by the sum of the total number of genes in both conserved gene regions, based on the previous methods (Cannon et al. 2003; Cannon et al. 2006).

References

Cannon SB, Mccombie WR, Sato S, Tabata S, Denny R, Palmer L, Katari M, Young ND, and Stacey G. 2003. Evolution and microsynteny of the apyrase gene family in three legume genomes. *Molecular Genetics and Genomics* 270:347-361.

Cannon SB, Sterck L, Rombauts S, Sato S, Cheung F, Gouzy J, Wang X, Mudge J, Vasdewani J, and Schiex T. 2006. Legume genome evolution viewed through the Medicago truncatula and Lotus japonicus genomes. *Proceedings of the National Academy of Sciences of the United States of America* 103:14959-14964.