

## 梨III型聚酮合成酶家族的比较基因组学研究及表达模式分析

李国辉<sup>#</sup>, 李亿红<sup>#</sup>, 钱冉, 苏雪强, 程曦, 赵宇, 孙燕铭, 金青, 林毅, 蔡永萍<sup>\*</sup>

安徽农业大学生命科学学院, 合肥230036

**摘要:** 植物III型聚酮合成酶(PKS III)在植物次生代谢产物生物合成中起着非常重要的作用。本研究从7种蔷薇科果树中共鉴定出57个PKS家族成员, 随后进行种间比较基因组学分析。基因结构和保守基序分析表明, 57个PKS多数由两个外显子和一个内含子组成, 都具有PKS特有的保守结构域Chal-sti-synt-N和Chal-sti-synt-C。进化分析显示, 57个PKS可明显分为4个亚家族, 其中I亚家族成员数目最多。染色体定位及基因复制事件表明, PKS不均匀地分布在2~5条染色体上, 且PKS家族进化过程主要受纯化选择作用。荧光定量分析发现, *PbPKS4*和*PbPKS5*表达模式与‘砀山酥梨’(*Pyrus bretschneideri* cv. ‘Dangshan Su’)不同发育时期果实中木质素及石细胞含量变化趋势类似, 因此推测这2个基因在梨果实木质素合成及石细胞发育过程中发挥一定作用。本研究为今后研究蔷薇科果树PKS家族建立了基础, 也为从分子水平改善梨果实品质提供了理论依据。

**关键词:** 梨; 植物III型聚酮合成酶; 生物信息学; 木质素; 表达分析

植物聚酮类化合物(polyketide, PK)是一类由碳和氧原子构成的具有大环结构的次生代谢产物, 主要包括酚类、芪类及类黄酮化合物等(Abe和Morita 2010)。根据蛋白的结构可将聚酮合成酶(polyketide synthase, PKS)分为I、II和III型。其中, I型(模块型)、II型(迭代型) PKS仅存在于微生物界; III型(查尔酮型) PKS广泛分布于植物界, 主要由查尔酮合酶(chalcone synthase, CHS)和类CHS超家族共同组成(Eom和Hyun 2016)。CHS是类黄酮代谢途径中的第一个关键酶, 可催化一分子的4-香豆酰辅酶A(4-香豆酰CoA)与三分子的丙二酰辅酶A(丙二酰CoA)缩合形成查尔酮(Martinez等2014)。类CHS(CHS-like, CHSL)是通过基因复制和功能分化衍生出的一个新类群, 与CHS在进化上起源于共同的祖型(Eom和Hyun 2016)。因此, CHSL与CHS成员之间的序列同源性较高, 两者在结构和催化机制上也具有极大的相似性, 例如, 它们构成同源二聚体的亚基大小相近, 并且在催化活性位点处均具有由3个保守氨基酸Cys-His-Asn构成的催化三联体等(Su等2017)。CHSL与CHS的功能分歧仅在于催化反应时对不同底物的偏好性及产物的环化方式不同(Schröder 2000)。

自首次从欧芹中获得CHS后, 多个物种的CHS均被克隆和功能分析, 这些研究均聚焦在CHS与类黄酮代谢途径的关系上, 例如陆地棉(*Gossypium hirsutum*) (Su等2017)、苹果(*Malus domestica*) (Dare等2013)、大豆(*Glycine max*) (Tuteja等2004)等。目

前研究发现, 陆地棉中CHS主要在纤维中表达, 与棕色棉纤维色素的生物合成相关(Su等2017); 红茶藨子(*Ribes rubrum*)中的CHS在果实成熟的后期表达量达到峰值, 进而影响果实的发育(Zorenc等2017)。最近的研究发现, CHS除了在类黄酮生物合成中发挥重要作用以外, 在植物木质素合成以及细胞壁发育中也充当重要角色。例如在亚麻(*Linum usitatissimum*)中, CHS表达受到抑制会导致木质素含量降低, 且细胞壁组分发生改变(Zuk等2016)。由此可见, 通过调节CHS表达可以改变木质素含量和细胞壁发育。CHS在苯丙烷代谢的两大分支途径——类黄酮和木质素生物合成中均发挥重要作用, 可作为一个潜在的协同调控果实类黄酮和木质素含量的关键位点。

‘砀山酥梨’(*Pyrus bretschneideri* cv. ‘Dangshan Su’)是我国特色名优品种, 也是栽培面积最大的梨品种(Cai等2010)。但由于近年来人们生活水平的提高以及梨种类的丰富, 果肉细腻的梨品种越发受到消费者的青睐, 从而使得‘砀山酥梨’石细胞团较大的固有缺陷日益凸显, 阻碍了其产业的发展。石细胞是梨果实的一个特征性状, 其含量及大小是决定梨果实品质的关键因素之一(黄成和李

收稿 2018-03-20 修定 2018-06-11

资助 国家自然科学基金(31640068)、安徽省大学生创新创业训练项目(201710364098)和安徽农业大学研究生创新基金(2018yjs-41)。

<sup>#</sup> 并列第一作者。

<sup>\*</sup> 通讯作者(y pcaiah@163.com)。

来庚2016; 徐超等2015)。因此, 深入探究梨石细胞形成机理, 对于人为调控石细胞发育、改善我国梨品质有重要意义。

石细胞的发育与木质素合成及沉积有密切关系(Cheng等2017)。*‘砀山酥梨’*果实中含有多种类黄酮物质, 具有良好的保健作用, 因此研究*‘砀山酥梨’*中的CHS家族对于调控木质素和类黄酮代谢具有重要意义, 可为改善*‘砀山酥梨’*果实中石细胞团大小和类黄酮含量奠定基础。迄今为止, 有关梨III型PKS家族的研究鲜有报道, 为此, 本研究对*‘砀山酥梨’*和其他6种蔷薇科果树的III型PKS家族成员进行了比较基因组学分析, 并分析了*‘砀山酥梨’*中III型PKS的表达模式, 筛选出与梨果实木质素和类黄酮代谢相关候选基因, 不仅为明确蔷薇科植物PKS家族特征奠定基础, 而且可为今后调控梨品质提供目标基因。

## 1 材料与方法

### 1.1 植物材料

供试材料*‘砀山酥梨’* (*Pyrus bretschneideri* Rehd. *‘Dangshan Su’*) 种植于安徽省砀山县园艺场。于花后15 d (15 DAF) 开始定期取样, 取15、39、47、55、63、79和145 DAF (成熟期) 的梨果实各20个, 立即用液氮速冻后保存在-80℃超低温冰箱中备用。

### 1.2 七种蔷薇科果树PKS家族成员的筛选与鉴定

桃(*Prunus persica*)和草莓(*Fragaria vesca*)基因序列信息从Phytozomes数据库(<https://phytozome.jgi.doe.gov/pz/portal.html>)中获得, 梅(*Prunus mume*)基因序列信息来源于<http://prunusmumegenome.bjfu.edu.cn/index.jsp>, 樱桃(*Cerasus pseudocerasus*)、西洋梨(*Pyrus communis*)和黑树莓(*Rubus occidentalis*)基因序列信息从蔷薇科基因组数据库(GDR) (<http://www.rosaceae.org/>)中获得, *‘砀山酥梨’*基因组来源于<http://gigadb.org/dataset/100083>。以PKS家族特征结构域(Chal-sti-synt-N和Chal-sti-synt-C)作为查询序列, 通过DNATOOLS软件在每个基因组数据库中进行比对筛选( $E$ 值=0.001), 初步筛选出PKS候选序列。将候选序列利用SMART (<http://smart.embl-heidelberg.de/>)软件与Pfam (<http://pfam.xfam.org/>)软件在线分析保守结构域, 鉴别是否具有PKS家族特征结构域, 去除不含有保守结构域的序列及重

复冗余序列, 最终获得这7种蔷薇科果树PKS家族成员序列。

## 1.3 比较基因组学分析

### 1.3.1 序列属性及分子进化分析

为分析蔷薇科果树PKS家族成员序列属性和特征, 使用在线分析工具ExPASy ([http://web.expasy.org/compute\\_pi/](http://web.expasy.org/compute_pi/))对从7种蔷薇科果树中获得的PKS所编码氨基酸序列的等电点、蛋白分子量等属性进行预测。使用在线工具MBC (<http://cello.life.nctu.edu.tw/>)预测亚细胞定位。系统进化树利用MEGA 6.0软件, 采用邻接法(neighbor-joining method)构建, Bootstrap值设为1 000。

### 1.3.2 基因结构及保守基序分析

利用GSDS软件(<http://gsds.cbi.pku.edu.cn/>)绘制基因结构图, 通过在线软件MEME ([http://meme.sdsc.edu/meme4\\_3\\_0/intro.html](http://meme.sdsc.edu/meme4_3_0/intro.html))对保守基序(motif)进行分析, 基序的最大数目设为20, 基序长度设置在6~200之间。

### 1.3.3 染色体定位及复制事件分析

染色体信息从基因组数据库中获得, 再使用MapInspect (<http://mapinspect.software.informer.com/>)软件将所有PKS在染色体上的位置绘制成图。

基因复制事件的确定: 两个基因匹配部分长度大于较长基因的80%, 且匹配区域相似性大于70%时, 认为发生基因复制事件; 再根据两个基因在染色体上位置情况, 判断是片段复制事件还是串联复制事件。针对复制基因对, 利用MEGA 6.0和DnaSP v5.0软件计算其非同义替换率(non-synonymous substitution rate,  $Ka$ )、同义替换率(synonymous substitution rate,  $Ks$ )和 $Ka/Ks$ 比值, 并绘制滑动窗口图。

## 1.4 *‘砀山酥梨’* PKS (*PbPKSs*)多序列比对

本研究以水稻(*Oryza sativa*) PKS家族中OsCHS的氨基酸序列为模板, 使用MEGA 6.0中的ClustalW进行多序列比对, 对*PbPKSs*编码的氨基酸序列进行分析, 以此为基础探明PKS家族功能多样性的原因。

## 1.5 *PbPKSs*家族成员启动子元件分析

*PbPKSs*家族成员上游1 500 bp的启动子序列从*‘砀山酥梨’*的基因组数据库获得, 然后使用在线软件PlantCARE (<http://bioinformatics.psb.ugent.be/web-tools/plantcare/html/>)分析启动子区顺式作用元件。

### 1.6 *PbPKSs*家族成员表达模式分析

‘砀山酥梨’总RNA的提取采用多糖多酚植物RNA提取试剂盒(天根生物技术有限公司)。利用PrimeScript<sup>RT</sup> reagent Kit with gDNA Eraser (Perfect Real Time)试剂盒操作说明书将RNA反转录为cDNA。使用荧光定量PCR仪(Bio-Rad)进行反应,反应体系含10  $\mu$ L的SYBR<sup>®</sup> Premix Ex Taq<sup>TM</sup> II (2) (TaKaRa)、2  $\mu$ L的cDNA、*PbPKSs*上下游引物各0.8  $\mu$ L和6.4  $\mu$ L水。PCR反应程序为: 95 $^{\circ}$ C预变性30 s; 95 $^{\circ}$ C变性5 s、60 $^{\circ}$ C退火20 s、72 $^{\circ}$ C延伸10 min, 40个循环。最后按照 $2^{-\Delta\Delta C_t}$ 法计算基因的相对表达量。荧光定量PCR引物见表1。

## 2 实验结果

### 2.1 七种蔷薇科果树*PKS*鉴定及序列属性

根据已报道的*PKS*家族保守结构域(Chal-sti-synt-N和Chal-sti-synt-C) (Han等2014), 利用Pfam网站(<http://pfam.xfam.org/>)获得对应的隐马尔可夫模型, 以此为目标序列, 用DNATOOLS软件在基因组数据库中进行比对, 去除重复冗余序列, 最终在这7种蔷薇科果树中获得57个*PKS*, 并进行命名。由表2可见, 鉴定出的蔷薇科果树*PKS*家族成员分子量为42~49 kDa, 长度在390个氨基酸左右, 其理论等电点大于7的成员有6个, 小于7的成员有51个, 说明以酸性蛋白为主; 亚细胞定位预测表明, *PKS*

主要分布在细胞质(41个)中, 其次分布在线粒体(10个)中, 少数定位在细胞核(5个)和内质网(1个)中。

### 2.2 七种蔷薇科果树*PKS*家族成员基因结构与保守基序

由图1-A可知, 57个*PKS*可分为4个亚家族(I~IV), 共形成17个基因对。其中, I、III、IV亚家族成员来源于7种被鉴定的蔷薇科果树, II来源于西洋梨、桃、梅、樱桃和‘砀山酥梨’。基于其系统进化树分析, 进一步分析*PKS*家族成员基因结构和保守基序。结果显示, ‘砀山酥梨’中所有*PKS*中均由两个外显子和一个内含子组成。结合基因结构和进化树分析, 发现亲缘关系较近的成员往往具有类似的基因结构, 特别是同一小亚类的基因之间, 其基因结构几乎相同(图1-B)。这一现象进一步揭示这些成员之间的进化关系较近, 同时也说明进化树构建的可靠性。

在鉴定得到的57个*PKS*中, 利用MEME软件识别了20个保守Motif (6~200氨基酸残基宽度) (图1-C)。其中Motif 1、2、8编码Chal-sti-synt-N和Chal-sti-synt-C结构域(图1-D), 这3个基序中至少有一个全部存在于所有的*PKS*家族成员中, 表明家族成员筛选鉴定的可信度较高。值得注意的是, 一些亚家族特有的结构域也被鉴定, 如Motif 5仅存在II亚家族中, 而Motif 10和Motif 12分别仅存在I和III亚

表1 本研究所用的引物序列

Table 1 Primer sequences used in this study

基因名称	上游引物(5'→3')	下游引物(5'→3')
<i>PbPKS1</i>	TTGGACCTGGAATAACCTT	GGAGAACAAAACATCACCATA
<i>PbPKS2</i>	CCTCACTGTCGAGACTGT	GCAGCAGATAGATAAATCAACAC
<i>PbPKS3</i>	AGAAAGGACTCAAAACAAC	TTAAACCCACGCTATGAA
<i>PbPKS4</i>	ATTTGGGCCTGGCCTCAC	TCAACGCTGATTTTCATGTCT
<i>PbPKS5</i>	ATTTGGGCCTGGCCTCAC	ATAAATCGACATTGCTACTGGTGT
<i>PbPKS6</i>	GAGACAGTCGTGCTTCAC	AGGAACACAGTACAAATGATC
<i>PbPKS7</i>	GGACTTACAGTCGAGACA	CCAATGTATGAATCAITTTTATGC
<i>PbPKS8</i>	ATCGGACCAGGACTCACT	ATACACCGTTCCATGCAAGTA
<i>PbPKS9</i>	GGACTCACCGTGGAGACAG	TGTTTGAATCGTTTCATGCAGGTA
<i>PbPKS10</i>	GGACCAGGACTCACTGTT	ACCGTTTCATGCAGGTTT
<i>PbPKS11</i>	ATCGGACCAGGACTTACAG	ACCAACGTATGAATCAITTTTATGC
<i>PbPKS12</i>	GTTGAGACAGTCGTGCTA	TGCATAACCAATGTAATCAT
<i>PbPKS13</i>	AGGACTCACTGTCGAGAC	CCAATGTATGAATCAITTTTATGC
<i>PbPKS14</i>	GGACTTACAGTCGAGACA	CCAATGTATGAATCAITTTTATGC
<i>Tublin</i>	AGAACAAGAACTCGTCTCTAC	GAACGTCTCGTCACTCTCC

表2 七种蔷薇科果树PKS成员基本信息

Table 2 Basic information of PKS genes in seven fruit trees of Rosaceae

基因名称	基因编号	氨基酸序列长度/aa	相对分子量/kDa	理论等电点	染色体位置	亚细胞定位
<i>PbPKS1</i>	Pbr008591.1	391	45.57	6.48	未定位	细胞质
<i>PbPKS2</i>	Pbr019531.1	389	42.46	5.77	Chr 4	细胞质
<i>PbPKS3</i>	Pbr019532.1	391	42.66	6.05	Chr 4	细胞质
<i>PbPKS4</i>	Pbr019533.1	396	43.21	6.05	Chr 4	细胞质
<i>PbPKS5</i>	Pbr020913.1	389	42.41	5.90	Chr 15	细胞质
<i>PbPKS6</i>	Pbr020914.1	390	42.49	5.90	Chr 15	细胞质
<i>PbPKS7</i>	Pbr021494.1	390	43.07	5.50	Chr 10	细胞质
<i>PbPKS8</i>	Pbr021495.1	390	42.93	5.97	Chr 10	细胞质
<i>PbPKS9</i>	Pbr027907.1	388	43.00	6.28	Chr 15	细胞质
<i>PbPKS10</i>	Pbr027908.1	390	43.14	6.10	Chr 15	细胞质
<i>PbPKS11</i>	Pbr027910.1	390	43.03	5.59	Chr 15	细胞质
<i>PbPKS12</i>	Pbr027912.1	301	32.89	5.11	Chr 15	细胞质
<i>PbPKS13</i>	Pbr027916.1	390	42.98	5.97	Chr 15	细胞质
<i>PbPKS14</i>	Pbr027917.1	418	46.11	5.28	Chr 15	细胞质
<i>PmPKS1</i>	Pm009565.1	391	42.91	6.12	Chr 2	细胞质
<i>PmPKS2</i>	Pm009568.1	391	42.87	6.28	Chr 2	细胞质
<i>PmPKS3</i>	Pm012349.1	390	43.29	5.51	Chr 3	细胞质
<i>PmPKS4</i>	Pm012347.1	297	32.53	5.31	Chr 3	线粒体
<i>PmPKS5</i>	Pm017616.1	395	43.68	7.55	Chr 5	细胞核
<i>PmPKS6</i>	Pm021033.1	393	43.23	5.73	Chr 6	细胞质
<i>PpPKS1</i>	Pp006852.1	393	43.03	5.62	Chr 8	细胞质
<i>PpPKS2</i>	Pp006888.1	391	42.65	5.84	Chr 1	细胞质
<i>PpPKS3</i>	Pp006899.1	391	42.74	5.84	Chr 1	细胞质
<i>PpPKS4</i>	Pp006930.1	390	43.06	5.51	Chr 4	细胞质
<i>PpPKS5</i>	Pp008402.1	333	36.17	6.02	未定位	细胞质
<i>PpPKS6</i>	Pp019432.1	362	39.72	6.56	Chr 2	线粒体
<i>PpPKS7</i>	Pp025745.1	390	42.96	5.97	Chr 4	细胞质
<i>FvPKS1</i>	Fv04837.1	387	42.67	6.92	Chr 7	线粒体
<i>FvPKS2</i>	Fv07378.1	393	43.43	5.92	未定位	细胞质
<i>FvPKS3</i>	Fv10965.1	388	41.97	5.47	Chr 2	细胞质
<i>FvPKS4</i>	Fv10966.1	388	42.23	6.03	Chr 2	细胞质
<i>FvPKS5</i>	Fv26825.1	389	42.63	6.04	Chr 7	细胞质
<i>FvPKS6</i>	Fv26826.1	389	42.63	6.18	Chr 7	细胞质
<i>RoPKS1</i>	Ro01512.1	387	42.44	5.63	Scaffold 0004	细胞质
<i>RoPKS2</i>	Ro08007.1	411	44.99	5.73	Scaffold 0045	细胞质
<i>RoPKS3</i>	Ro16736.1	387	42.74	6.92	Scaffold 0045	线粒体
<i>RoPKS4</i>	Ro24959.1	391	42.85	6.28	Scaffold 0583	细胞质
<i>RoPKS5</i>	Ro26206.1	406	44.92	6.33	Scaffold 0742	细胞质
<i>PaPKS1</i>	Pa_g280.1	294	32.48	8.72	Chr 1	细胞核
<i>PaPKS2</i>	Pa_g350.1	262	28.78	8.80	Chr 1	细胞核
<i>PaPKS3</i>	Pa_g150.1	672	73.81	5.29	Chr 4	线粒体
<i>PaPKS4</i>	Pa_g020.1	392	43.01	5.91	Chr 8	细胞质
<i>PaPKS5</i>	Pa_g060.1	441	48.69	7.56	Chr 2	内质网
<i>PaPKS6</i>	P_ga010.1	182	20.19	8.55	未定位	细胞质
<i>PcPKS1</i>	Pc011959.1	394	43.03	5.67	Scaffold 00468	细胞质
<i>PcPKS2</i>	Pc012534.1	320	35.14	6.17	Scaffold 02268	线粒体
<i>PcPKS3</i>	Pc012650.1	360	39.07	5.23	Scaffold 04168	细胞质
<i>PcPKS4</i>	Pc014287.1	257	27.63	5.02	Scaffold 09589	细胞质
<i>PcPKS5</i>	Pc015190.1	1024	114.79	8.19	Scaffold 0005	细胞核
<i>PcPKS6</i>	Pc019697.1	388	42.98	6.15	Scaffold 00233	细胞质
<i>PcPKS7</i>	Pc019698.1	390	43.01	6.28	Scaffold 00233	细胞核
<i>PcPKS8</i>	Pc019699.1	257	27.70	4.94	Scaffold 00233	细胞质
<i>PcPKS9</i>	Pc019701.1	390	43.02	5.59	Scaffold 00233	细胞质
<i>PcPKS10</i>	Pc022135.1	390	42.57	5.90	Scaffold 02234	线粒体
<i>PcPKS11</i>	Pc023048.1	391	42.63	5.90	Scaffold 00375	线粒体
<i>PcPKS12</i>	Pc023049.1	571	62.89	6.18	Scaffold 00375	线粒体
<i>PcPKS13</i>	Pc026641.1	389	42.71	5.89	Scaffold 00917	线粒体

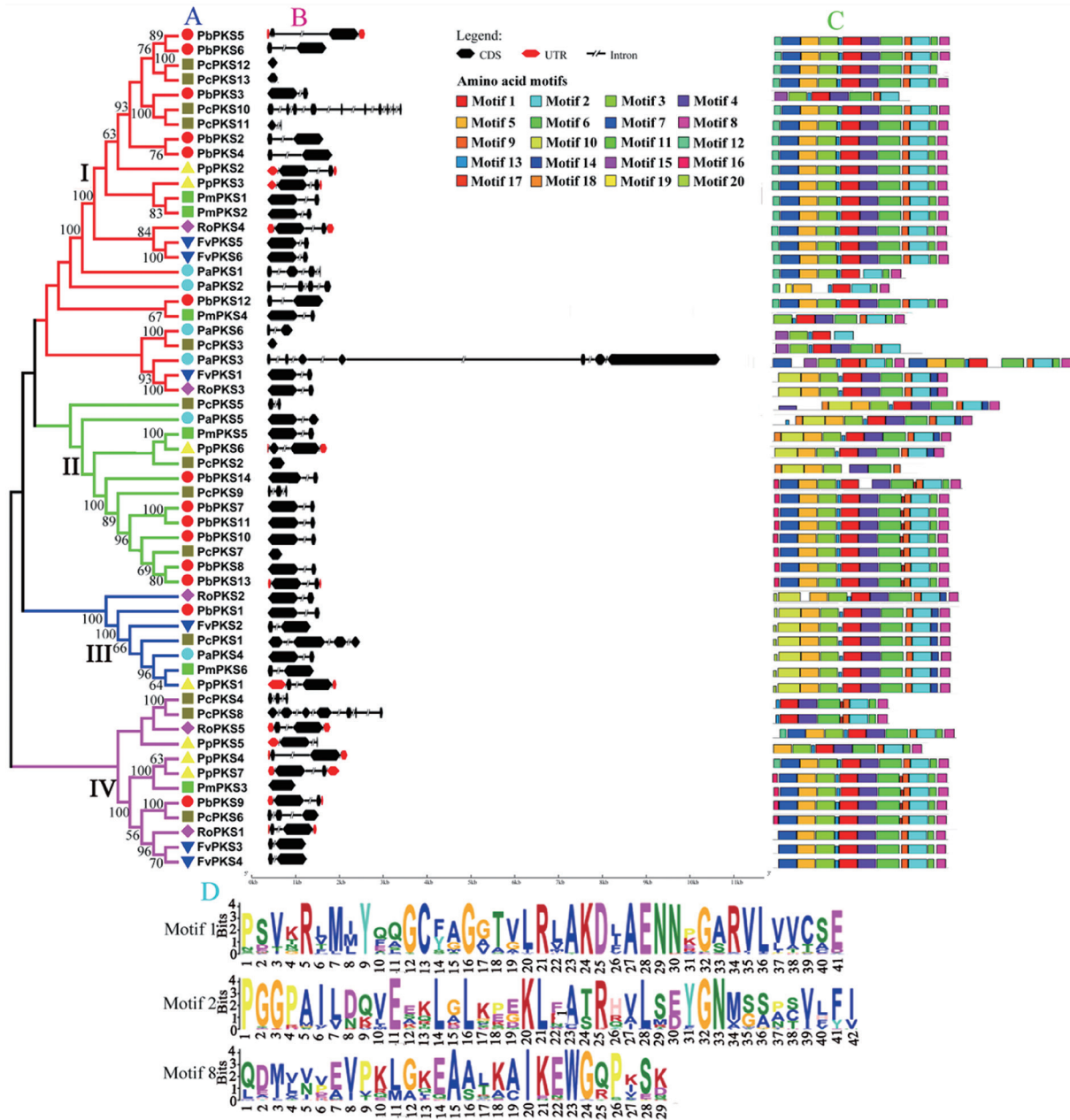


图1 七种蔷薇科果树PKS系统进化树、基因结构及保守基序

Fig.1 Phylogenetic relationship, the exon/intron structure and conserved motifs of *PKS* genes in seven fruit trees of Rosaceae

A: 系统进化树; B: 基因结构; C: 保守基序; D: 编码Chal-sti-synt-N和Chal-sti-synt-C结构域。

家族中。表明这些基序可能赋予这些亚族成员特有的功能。

### 2.3 七种蔷薇科果树PKS家族成员染色体定位和基因复制事件

根据各物种基因组信息,得到了57个PKS在染色体上的分布情况。由图2可知,草莓Chr 2上分布2个PKS, Chr 7上分布3个PKS,而FvPKS2没有在染

色体上定位;梅Chr 5和Chr 6上各分布1个PKS, Chr 2和Chr 4各分布2个PKS;桃的5个PKS分布在Chr 1、Chr 2、Chr 4染色体上, PpPKS1未定位在染色体上;樱桃6个PKS分布比较分散, Chr 2、Chr 4、Chr 8上各分布1个PKS, Chr 1上分布2个PKS, PaPKS6未定位在染色体上;‘砧山酥梨’的14个PKS (PbPKS1~14)主要分布在Chr 4和Chr 15上,以基因

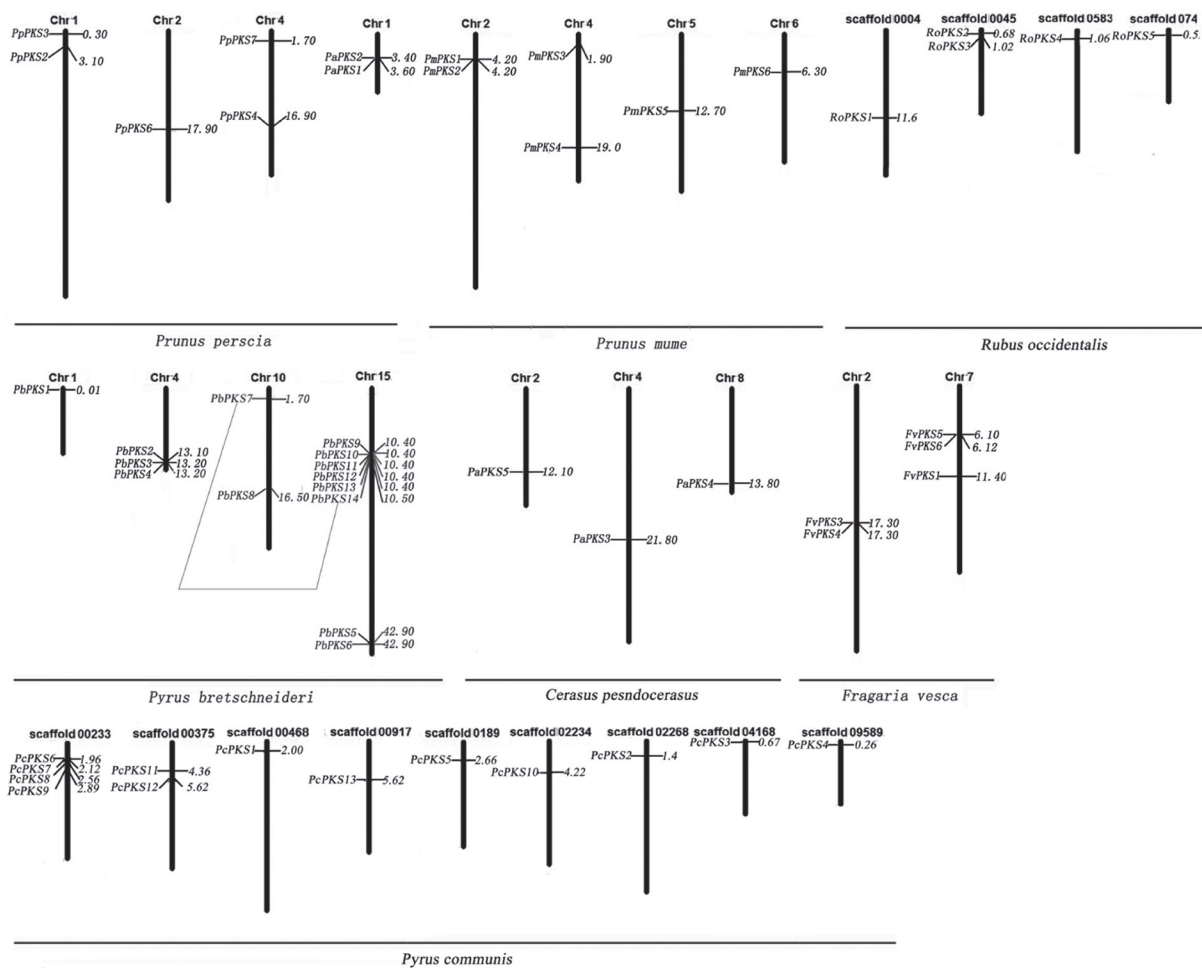


图2 七种蔷薇科果树PKS的染色体定位

Fig.2 Chromosomal location of *PKS* genes in seven fruit trees of Rosaceae

簇的形式存在, Chr 1和Chr 10只分布1~2个PKS;西洋梨和黑树莓中PKS均定位在scaffold。

一般在基因进化过程中会有基因复制事件的发生,包括片段复制和串联复制,这在一定程度上可反映基因进化机制。在7种蔷薇科果树中鉴定出的57个PKS中,发现6个基因复制事件,除了PbPKS2/PbPKS4、PbPKS10/PcPKS7属于串联复制外,其余均为片段复制(图3),说明蔷薇科果树PKS家族成员扩张的主要原因是片段复制。

为了探究基因复制事件的选择压力,计算复制事件的 $Ka/Ks$ 比值,发现所有基因对 $Ka/Ks$ 均小于1(表3),这说明蔷薇科PKS在进化过程中主要受纯化选择作用。

#### 2.4 PbPKSs多序列比对

将PbPKS1~14序列与先前已经报道的水稻

OsCHS的氨基酸序列进行多序列比对,结果显示,在PbPKS1~14蛋白中,从酮酸酰基合酶III (ketoacyl carrier protein synthases III, KASIII)处继承而来的Cys-His-Asn催化三联体(Austin和Noel 2002)在每个序列中均高度保守,没有发生替换现象。但是,在4个活性氨基酸位置处出现了较多的氨基酸替换现象,如Gly处PbPKSs出现多处由Ala代替Gly, Ser处PbPKS1由Met代替Ser, Thr处PbPKSs未发生改变。活性氨基酸Phe在催化活性中心腔内有两个位点存在,与丙二酰CoA的脱羧反应密切相关(Jez等2002)(图4中黑色三角形处)。第一个Phe活性位点在PbPKSs蛋白中高度保守,但是第二个Phe活性位点处出现Tyr代替Phe的现象(图4)。在PbPKSs中,活性氨基酸Gly和Ser在一个或多个蛋白中发生替换,而Thr均没有发生替换,这可能与PKS家族功能多样性密切相关。

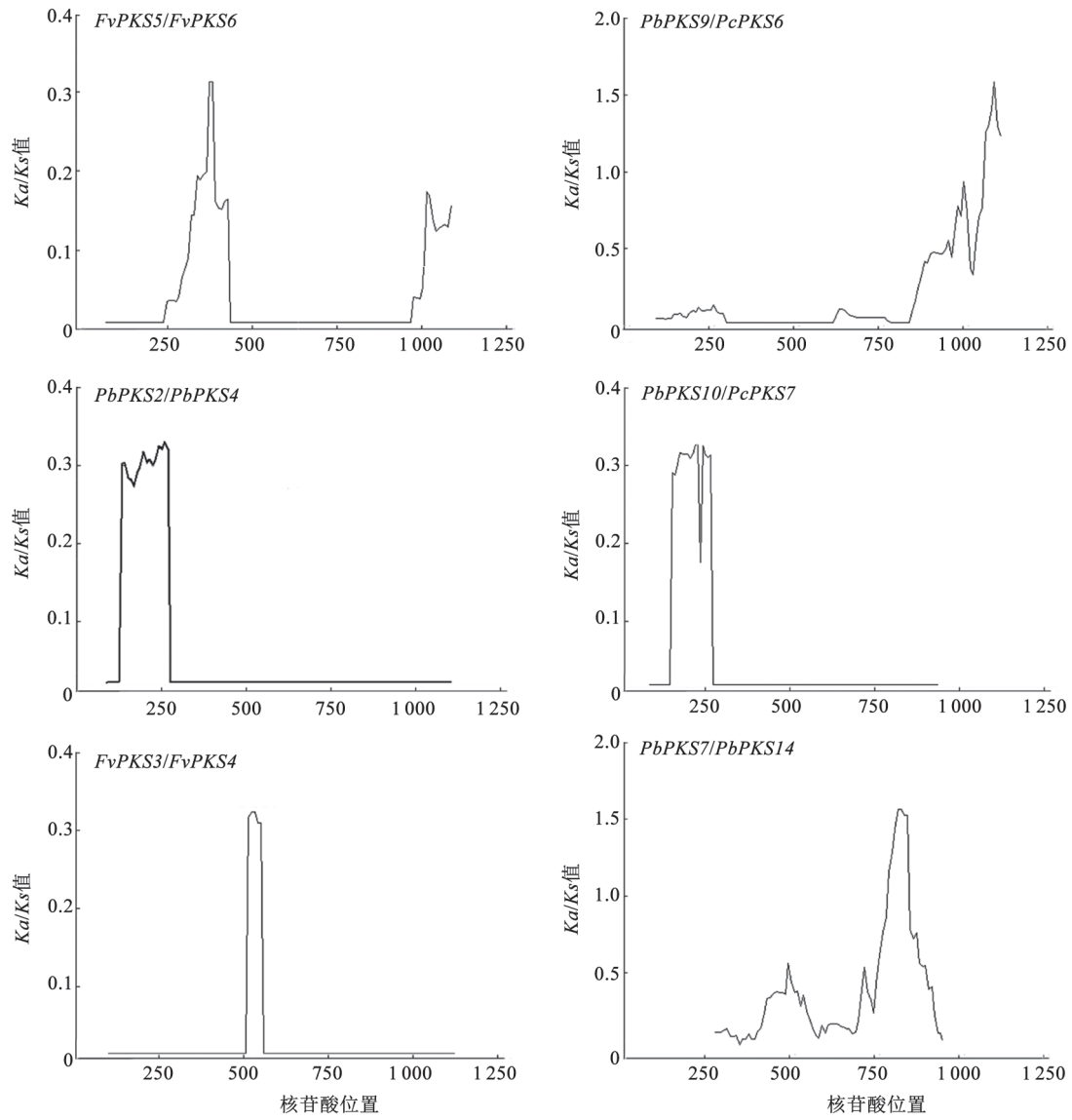


图3 *PKS*复制基因对的滑动窗口分析  
Fig.3 The sliding window analysis of *PKS* duplicated genes

表3 七种蔷薇科果树*PKS*基因复制事件分析

Table 3 *Ka/Ks* analysis for the *PKS* duplicated genes in seven fruit trees of Rosaceae

复制基因对	<i>Ka</i>	<i>Ks</i>	<i>Ka/Ks</i>	纯化选择	复制方式
<i>FvPKS5/FvPKS6</i>	0.0068	0.2039	0.0333	是	片段复制
<i>PbPKS2/PbPKS4</i>	0.1813	0.5888	0.5889	是	串联复制
<i>FvPKS3/FvPKS4</i>	0.0393	0.1518	0.2589	是	片联复制
<i>PbPKS9/PcPKS6</i>	0.0011	0.0150	0.0733	是	片段复制
<i>PbPKS10/PcPKS7</i>	0.0056	0.0072	0.7778	是	串联复制
<i>PbPKS7/PcPKS14</i>	0.0012	0.0280	0.0429	是	片段复制

2.5 *PbPKSs*家族启动子序列

为明确*PbPKSs*启动子特征, 我们分析了其顺

式作用元件, 发现这14个*PbPKSs*除存在大量基本启动子元件TATA-box外, 还存在参与光响应顺式

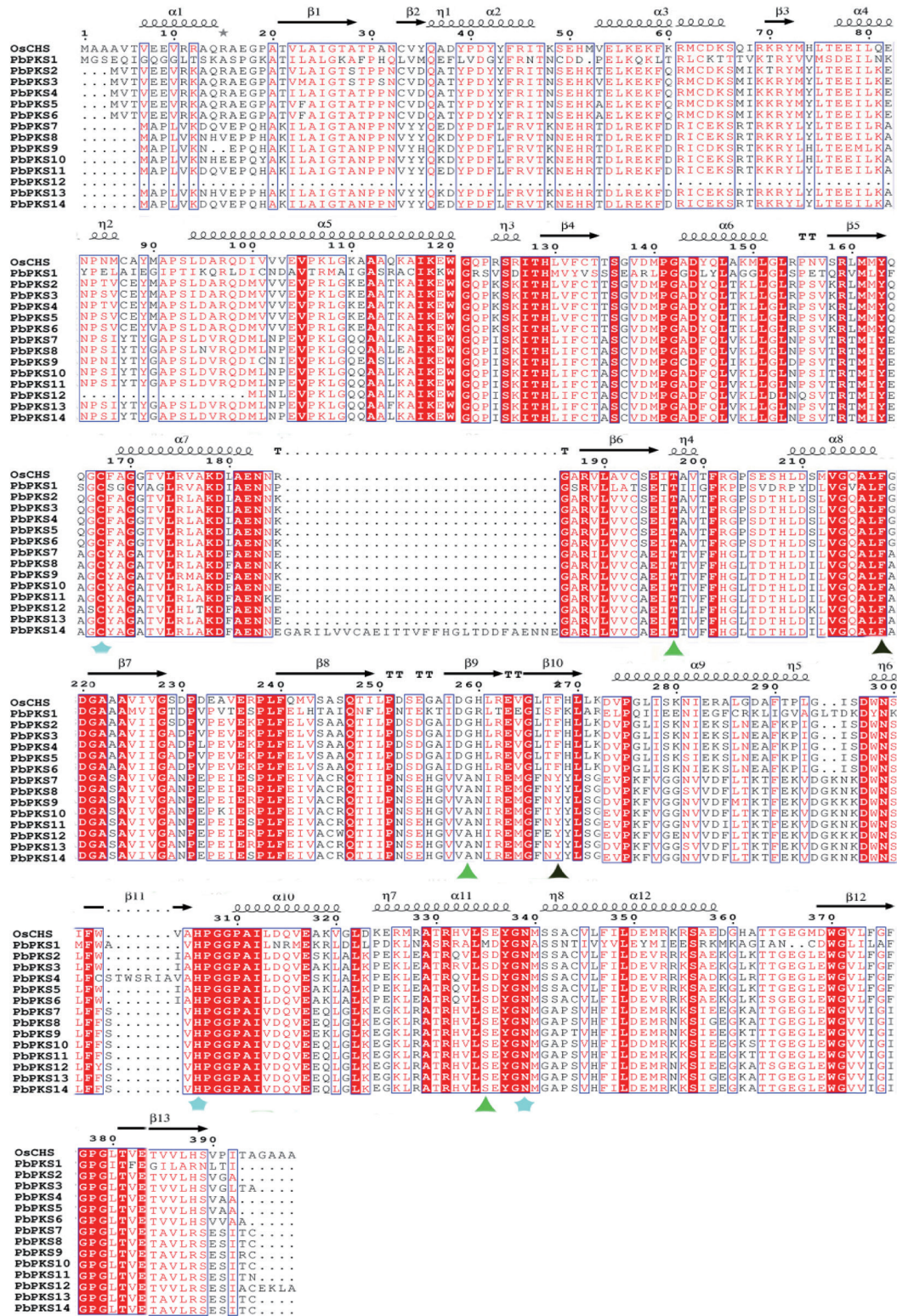


图4 PbpPKs与OsCHS序列比对

Fig.4 Sequence alignment of PbpPKs against the OsCHS

黑色的波浪线和箭头分别表示 $\alpha$ -螺旋和 $\beta$ -折叠; 蓝色框和红色的字均代表保守氨基酸残基, 红色区域的序列表示氨基酸残基保守性极高; 蓝色五角星表示催化三联体(Cys-His-Asn), 位于催化活性中心腔内的活性氨基酸(Thr、Phe、Gly、Ser)用绿色或者黑色的三角形表示。



表4 *PbPKSs*启动子顺式作用元件的数目

Table 4 The number of *cis*-acting elements in the promoter of *PbPKSs*

基因	AC-I	AC-II	GAG-motif	MBS	SP1	Box I	HSE	TATA-box
<i>PbPKS1</i>	0	0	2	2	5	1	0	3
<i>PbPKS2</i>	0	0	1	3	2	1	1	3
<i>PbPKS3</i>	0	0	0	2	0	1	1	4
<i>PbPKS4</i>	1	1	0	1	1	0	2	3
<i>PbPKS5</i>	0	0	1	0	4	2	1	4
<i>PbPKS6</i>	0	0	0	3	11	1	1	0
<i>PbPKS7</i>	0	0	2	1	0	2	2	0
<i>PbPKS8</i>	0	0	1	1	0	0	1	5
<i>PbPKS9</i>	1	1	2	0	3	4	0	2
<i>PbPKS10</i>	1	1	7	0	2	2	2	4
<i>PbPKS11</i>	0	0	0	0	0	0	3	3
<i>PbPKS12</i>	0	0	1	1	2	0	2	5
<i>PbPKS13</i>	0	0	4	1	0	0	1	4
<i>PbPKS14</i>	0	0	2	2	0	0	1	3

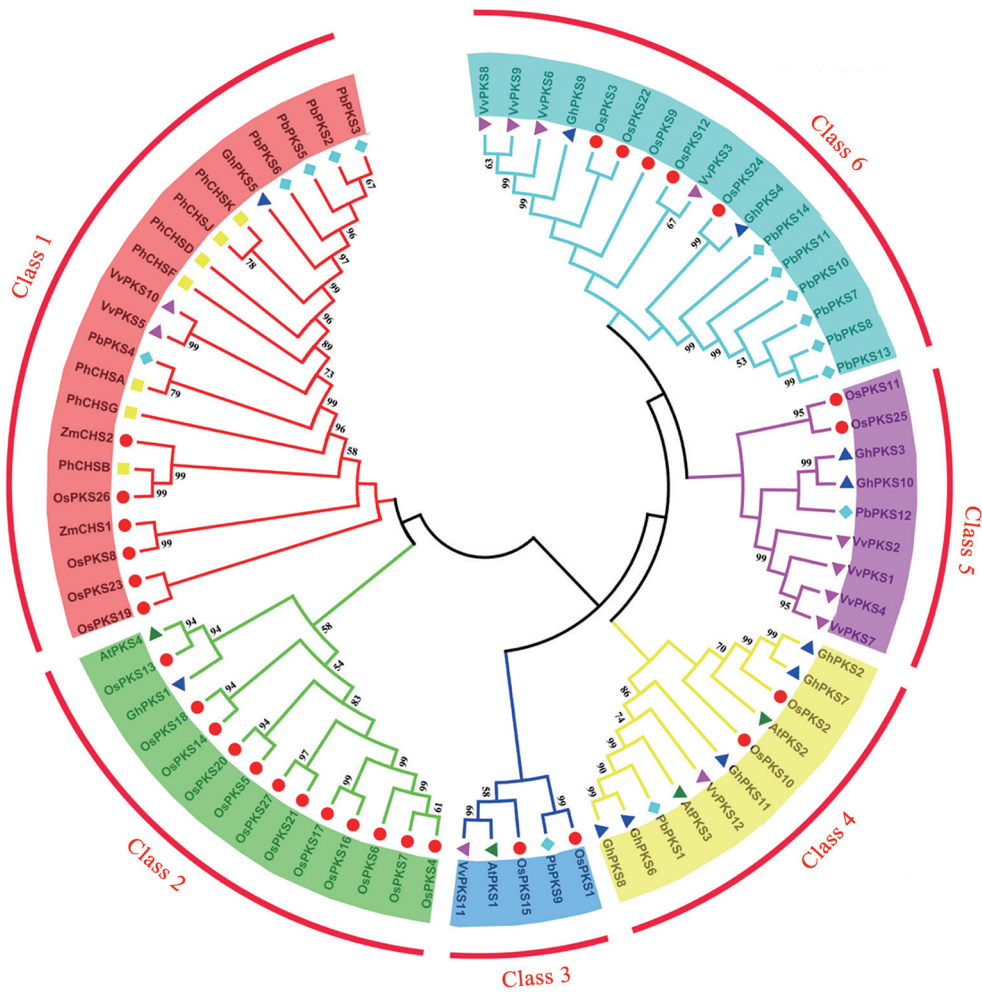


图5 不同植物来源PKS家族成员的系统进化树分析

Fig.5 Phylogenetic tree of PKS family members from different plant species

调控元件SP1、GAG-motif、Box I等,多个与胁迫响应相关的顺式作用元件(包括干旱胁迫响应元件MBS、热激反应响应元件HSE),与MYB结合的元件AC-I和AC-II(表4)。这些结果表明PKS不仅在木质素合成中发挥作用,还可能参与光响应过程及胁迫反应。

## 2.6 *PbPKSs*家族成员的种间进化树及功能预测

将*PbPKSs*家族成员和其他物种中已知功能的PKS一起构建进化树,对*PbPKSs*的功能进行预测。种间系统进化树分为Class 1~6六大亚家族(图5)。由图5可知,‘砀山酥梨’的14个PKS基因主要分布在Class 1和Class 6这两个亚家族中。Class 2亚家族中不含有*PbPKSs*,水稻中的PKS基因主要存在于该亚家族中,且与花青素代谢相关的成员*AtPKS4*(Owens等2008)也在该亚家族中。Class 3~5中均只存在1个*PbPKS*。值得注意的是,矮牵牛和玉米中的*CHS*都包括在Class 1亚家族中。

## 2.7 *PbPKSs*表达模式

由图6可知,*PbPKSs*具有多种表达模式,*PbPKS1*在‘砀山酥梨’果实发育各个时期中几乎均不表达,*PbPKS2*、7~10表达量较低,且无明显有规律的变化趋势。*PbPKS3*、6、13、14在果实发育中期表达量较高,在果实发育初期和成熟期表达量均较低,且未呈现明显先升高后下降的趋势,与相应时期果实中木质素和石细胞含量变化趋势不符。*PbPKS11*、12在果实发育初期表达量较低,在果实成熟期表达量较高,与相应时期果实中木质素和石细胞含量变化趋势相反。*PbPKS4*和*PbPKS5*的表达趋势呈先上升后下降,与果实中木质素和石细胞含量变化趋势基本相符,说明它们可能在果实木质素合成和石细胞发育中扮演重要角色。

## 3 讨论

植物III型聚酮合成酶(PKS III)与植物生命活动息息相关。目前,PKS III基因家族已经在多个物种被鉴定和克隆。例如,在水稻中已经报道了27个PKS(Hu等2017),这是目前报道的PKS数量最多的物种。本研究在7种蔷薇科果树中共鉴定到57个PKS,其中,在梨基因组中鉴定出的PKS成员(27个)多于从桃(7个)、草莓(6个)、梅(6个)、黑树莓(5个)、樱桃(6个),这可能是由于桃、草莓、

梅、黑树莓和樱桃只存在古老的全基因组复制方式,而梨可能存在古老的全基因组复制和近代全基因组复制两种方式(乔鑫等2014)。本文在从7种蔷薇科果树中鉴定出的57个PKS中,有6对基因发生复制事件,而4对基因属于片段复制,说明蔷薇科果树PKS家族成员扩张的主要原因是片段复制,这与Su等(2017)报道的植物中PKS家族扩张动力主要以片段复制为主的结果相一致。

57个PKS中有41个成员由2个外显子和1个内含子组成(图1-B),这与已报道的植物中大部分的PKS含有2个外显子和1个内含子的结果相一致(Durbin等2000)。7种蔷薇科果树PKS在染色体上的分布是无规律的(图2),有部分基因是以基因簇的形式存在,如15号染色体,基因簇的形成一般是基因复制的结果,它们大多数聚类在一个亚家族中,同一性比较高。但是,大部分PKS在染色体上分布比较分散,这与Han等(2016)的研究结果一致。在7种蔷薇科果树中发现有6对基因发生基因复制,在这6对重复基因中,只有‘砀山酥梨’和西洋梨2对重复基因是串联复制,其余均是片段复制。串联复制有助于新基因的产生,片段复制则会导致基因家族的进化变得较为缓慢(Cao等2016)。在‘砀山酥梨’中PKS的复制主要以片段复制的方式进行,这表明*PbPKSs*家族的进化较为缓慢。对7种蔷薇科果树的6对重复基因Ka/Ks值的分析结果显示,6个重复基因对的Ka/Ks值均小于1,所以这些重复基因在完成复制后经历了较为强烈的纯化选择,这对维持PKS家族的功能具有积极的意义。

在14个*PbPKSs*启动子中,*PbPKS4*、*PbPKS9*、*PbPKS10*均存在与MYB结合的AC-I和AC-II元件,现已发现与木质素合成相关的启动子元件中均存在AC-I和AC-II元件(郭光艳等2015)。在亚麻中*CHS*下调导致木质素合成减少(Zuk等2016),可通过调节*CHS*活性来改变木质素含量。有报道称,拟南芥中*CHS*受到MYB转录因子的调节(Chezem和Clay 2016),同时与木质素合成相关的启动子元件AC-I和AC-II也能够与MYB结合,这就意味着*PbPKSs*的表达可能受到MYB转录因子的调节,且MYB可能调控‘砀山酥梨’木质素的合成。

前人研究表明,石细胞和木质素含量在梨果实发育中呈先上升后下降趋势(Cai等2010; Jin等2013),

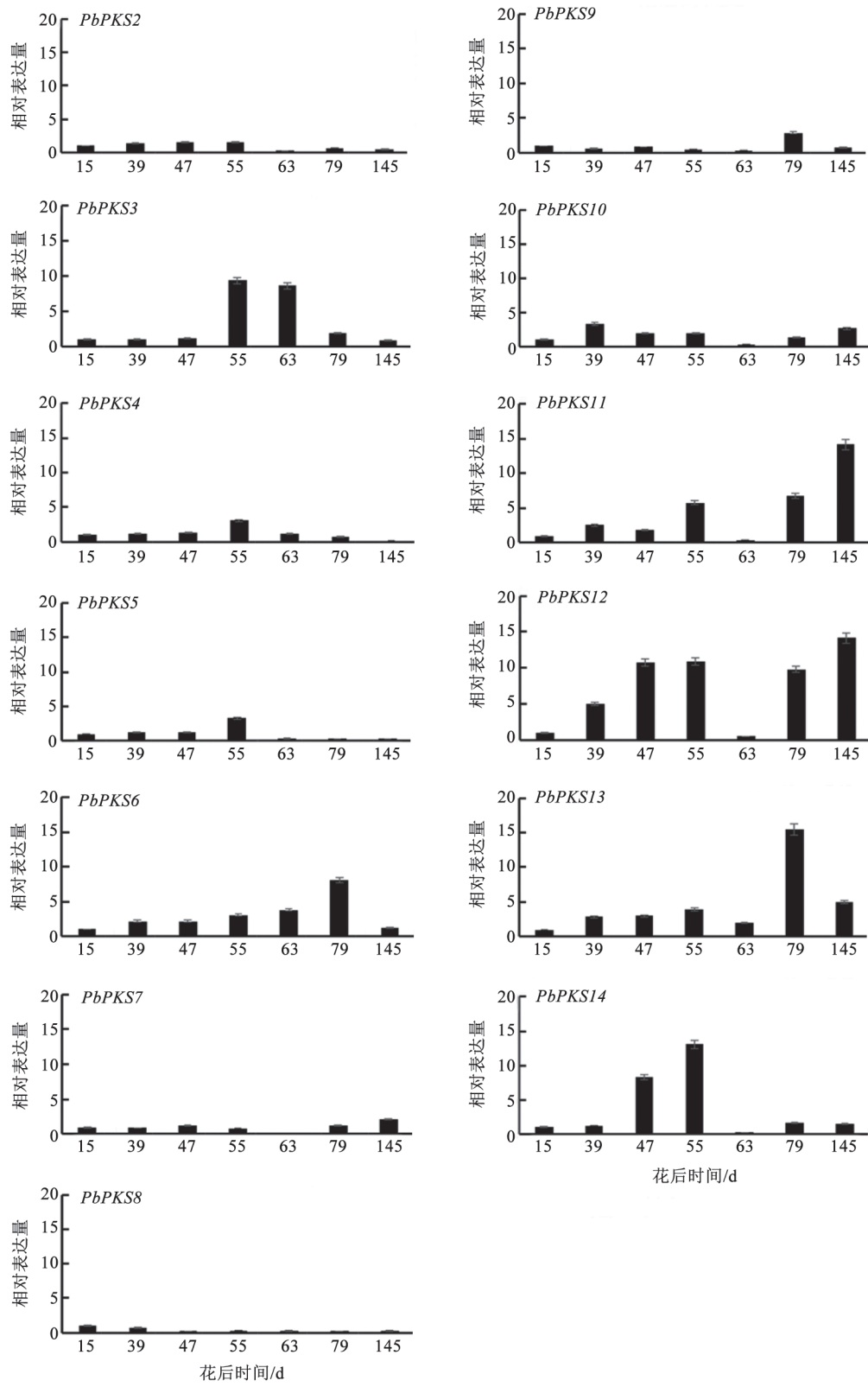


图6 *PbPKSs*表达分析

Fig.6 Expression patterns of *PbPKSs*

在‘砀山酥梨’果实发育过程中, Class 1中*PbPKS4*和*PbPKS5*这两个基因表达量变化与梨果实石细胞和木质素含量变化规律相符, 推测它们可能参与梨果实木质素生物合成。在Class 1中其他功能已知的成员主要是负责木质素和类黄酮的代谢反应, 例如*PhCHSK*表达和木质素合成相关(Zuk等2012, 2016), 且在玉米中通过沉默*ZmCHS2*导致类黄酮产物合成受阻, 而木质素含量增加(Eloy等2017)。推测在亚家族Class 1中包含的5个*PbPKSs* (*PbPKS2~6*)可能参与木质素和类黄酮的合成。

由此可见, CHS在苯丙烷代谢的两大分支途径——类黄酮和木质素生物合成中均发挥重要作用, 且类黄酮和木质素合成之间可能存在拮抗的关系。推测*PbPKS4*可能不仅在木质素生物合成中起重要作用, 而且在类黄酮生物合成中也发挥作用。

#### 参考文献(References)

- Abe I, Morita H (2010). Structure and function of the chalcone synthase superfamily of plant type III polyketide synthases. *Nat Prod Rep*, 41 (37): 809–838
- Austin MB, Noel JP (2002). The chalcone synthase superfamily of type III polyketide synthases. *Nat Prod Rep*, 20 (1): 79–110
- Cai YP, Li GQ, Nie JQ, et al (2010). Study of the structure and biosynthetic pathway of lignin in stone cells of pear. *Sci Hortic*, 125 (3): 374–379
- Cao YP, Han YH, Meng DD, et al (2016). Structural evolutionary, and functional analysis of the class III peroxidase gene family in Chinese pear (*Pyrus bretschneideri*). *Front Plant Sci*, 7: 1874
- Cheng X, Yan C, Zhang JY, et al (2017). The effect of different pollination on the expression of Dangshan Su pear microRNA. *BioMed Res Int*, 2017: 2794040
- Chezem WR, Clay NK (2016). Regulation of plant secondary metabolism and associated specialized cell development by MYBs and bHLHs. *Phytochemistry*, 131: 26–43
- Dare AP, Tomes S, Jones M, et al (2013). Phenotypic changes associated with RNA interference silencing of chalcone synthase in apple (*Malus domestica*). *Plant J*, 74 (3): 398–410
- Durbin ML, Mccaig B, Clegg MT (2000). Molecular evolution of the chalcone synthase multigene family in the morning glory genome. *Plant Mol Biol*, 42 (1): 79–92
- Eloy NB, Voorend W, Lan W, et al (2017). Silencing chalcone synthase impedes the incorporation of tricetin in lignin and increases lignin content. *Plant Physiol*, 173 (2): 998–1016
- Eom SH, Hyun TK (2016). Genome-wide identification and transcriptional expression analysis of chalcone synthase in flax (*Linum usitatissimum* L.). *Gene Rep*, 5: 51–56
- Guo GY, Bai F, Liu W, et al (2015). Research progress of transcription factors regulating lignin biosynthesis. *Sci Agric Sin*, 48 (7): 1277–1287 (in Chinese with English abstract) [郭光艳, 柏峰, 刘伟等(2015). 转录因子对木质素生物合成调控的研究进展. *中国农业科学*, 48 (7): 1277–1287]
- Han Y, Zhao W, Wang Z, et al (2014). Molecular evolution and sequence divergence of plant *chalcone synthase* and *chalcone synthase-like* genes. *Genetica*, 142 (3): 215–225
- Han YH, Ding T, Su B, et al (2016). Genome-wide identification, characterization and expression analysis of the chalcone synthase family in maize. *Int J Mol Sci*, 17 (2): 161–176
- Hu LF, He HH, Zhu CL, et al (2017). Genome-wide identification and phylogenetic analysis of the chalcone synthase gene family in rice. *J Plant Res*, 130 (1): 95–105
- Huang C, Li LG (2016). Advances in research on thickening and regulation of secondary cell wall in plants. *Plant Physiol J*, 52 (1): 8–18 (in Chinese with English abstract) [黄成, 李来庚(2016). 植物次生细胞壁加厚调控研究进展. *植物生理学报*, 52 (1): 8–18]
- Jez JM, Bowman ME, Noel JP (2002). Expanding the biosynthetic repertoire of plant type III polyketide synthases by altering starter molecule specificity. *Proc Natl Acad Sci USA*, 99 (8): 5319–5324
- Jin Q, Yan CC, Qiu JX, et al (2013). Structural characterization and deposition of stone cell lignin in Dangshan Su pear. *Sci Hortic*, 155: 123–130
- Owens DK, Alerding AB, Crosby KC, et al (2008). Functional analysis of a predicted flavonol synthase gene family in *Arabidopsis*. *Plant Physiol*, 147 (3): 1046–1061
- Qiao X, Li M, Yin H, et al (2014). Advances on whole genome sequencing in fruit trees. *Acta Hortic Sin*, 41 (1): 165–177 (in Chinese with English abstract) [乔鑫, 李梦, 殷豪等(2014). 果树全基因组测序研究进展. *园艺学报*, 41 (1): 165–177]
- Schröder J (2000). The family of chalcone synthase-related proteins: functional diversity and evolution. *Rec Adv Phytochem*, 34: 55–89
- Su XQ, Sun X, Cheng X, et al (2017). Comparative genomic analysis of the PKS genes in five species and expression analysis in upland cotton. *PeerJ*, 5 (6): e3974
- Tuteja JH, Clough SJ, Chan WC, et al (2004). Tissue-specific gene silencing mediated by a naturally occurring chalcone synthase gene cluster in glycine max. *Plant Cell*, 16 (4): 819–835
- Xu C, Fang Z, Yang FM, et al (2015). Analysis of *F5H* expression and stone cell development in *Pyrus bretschneideri* cv. Dangshan Su fruit. *Plant Physiol J*, 51 (5): 778–784 (in Chinese with English abstract) [徐超, 方志, 杨芳梅等

- (2015). 砀山酥梨果实 $F5H$ 表达与石细胞发育的分析. 植物生理学报, 51 (5): 778–784]
- Zorenc Z, Veberic R, Koron D, et al (2017). Polyphenol metabolism in differently colored cultivars of red currant (*Ribes rubrum* L.) through fruit ripening. *Planta*, 246: 217–226
- Zuk M, Działo M, Richter D, et al (2016). Chalcone synthase (CHS) gene suppression in flax leads to changes in wall synthesis and sensing genes, cell wall chemistry and stem morphology parameters. *Front Plant Sci*, 7: 894
- Zuk M, Prescha A, Stryczewska M, et al (2012). Engineering flax plants to increase their antioxidant capacity and improve oil composition and stability. *J Agric Food Chem*, 60 (19): 5003–5012

## Comparative genomics study of type III polyketide synthase family of pear and its expression pattern analysis

LI Guo-Hui<sup>#</sup>, LI Yi-Hong<sup>#</sup>, QIAN Ran, SU Xue-Qiang, CHENG Xi, ZHAO Yu, SUN Yan-Ming, JIN Qing, LIN Yi, CAI Yong-Ping<sup>\*</sup>

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

**Abstract:** Type III polyketide synthase (PKS) plays an important role in the biosynthesis of secondary metabolites in plants. In this study, a total of 57 *PKS* family members were identified from 7 species of fruit trees in family Rosaceae, followed by a comparative genome analysis. Analysis of gene structure and conserved motifs showed that most of these *PKSs* were composed of two exons and one intron, containing the *PKS*-specific conserved domains Chal-sti-synt-N and Chal-sti-synt-C. Phylogenetic analysis showed that 57 *PKSs* were divided into four subfamilies with the maximal number in the subfamily I. And chromosomal location analysis showed that *PKSs* were unevenly distributed across chromosomes 2 to 5. Furthermore, it was found that the evolution of *PKS* family was mainly affected by the purification and selection under gene-duplication analysis. Fluorescence quantitative analysis showed that the expression patterns of *PbPKS4* and *PbPKS5* were in a similar trend to the change in both lignin and stone cell contents at the different developmental stages of *Pyrus bretschneideri* cv. ‘Dangshan Su’. Therefore, both of *PbPKS4* and *PbPKS5* might be associated with the lignin synthesis and stone cell development in fruits of *P. bretschneideri* cv. ‘Dangshan Su’. These results primarily reveal the important function of *PKSs* in lignin metabolism in some Rosaceae species, and provide a theoretical foundation for improving the quality of pear fruits at the molecular level.

**Key words:** pear; plant type III polyketide synthase; bioinformatics; lignin; expression analysis

Received 2018-03-20 Accepted 2018-06-11

This work was supported by the National Natural Science Foundation of China (31640068), Anhui Provincial College Students Innovation and Entrepreneurship Training Project (201710364098) and Anhui Agricultural University Postgraduate Innovation Fund (2018yjs-41).

<sup>#</sup>Co-first authors.

<sup>\*</sup>Corresponding author (ypcaiah@163.com).