

# 桃 WRKY 基因家族全基因组鉴定和表达分析

谷彦冰, 冀志蕊, 迟福梅, 乔壮, 徐成楠, 张俊祥, 周宗山, 董庆龙

中国农业科学院果树研究所, 兴城 125100

**摘要:** WRKY 转录因子是植物中最大的转录调控家族之一, 在生物和非生物胁迫以及植物生长和发育过程中起着重要调控作用。本文利用 HMMER 3.0 软件, 使用 WRKY 保守域全蛋白序列(Pfam 数据库编号: PF03106)鉴定桃(*Prunus persica* L.)基因组中的 WRKY 基因; 利用 DNAMAN 5.0、WebLogo 3、MEGA5.1、MapInspect 和 MEME 等软件对其蛋白序列进行生物信息学分析。本文共鉴定得到 61 个桃 WRKY 基因。进化树分析结果显示, 桃 WRKY 蛋白分为 I、II 和 III 类型, 类型 I 分为 I-C 亚组和 I-N 亚组, 类型 II 分为 II-a、II-b、II-c、II-d 和 II-e 亚组。WRKY 结构域分析显示, WRKY 结构域高度保守, 绝大多数都含有 WRKYGQK 七肽和锌指结构。染色体定位分析显示, 桃 WRKY 基因分布于 8 条染色体中, 呈不均匀分布。内含子和外显子结构分析表明, WRKY 基因结构进化高度保守。保守元件分析表明, 桃 WRKY 基因家族包含 5 个保守元件, 元件 1、2 和 3 为 WRKY 盒, 元件 4、5 为未知盒。桃 WRKY 基因家族都包含有 WRKY 盒, 类型 I 中含有 2 个 WRKY 盒; II-d 亚组中含有未知元件 5。半定量和荧光定量 PCR 结果显示, 16 个 WRKY 基因均在桃的根、茎、叶、花和果中表达, 但其相对表达水平不同。

**关键词:** 桃; WRKY; 转录因子; 基因家族; 表达分析

## Genome-wide identification and expression analysis of the WRKY gene family in peach

Yanbing Gu, Zhirui Ji, Fumei Chi, Zhuang Qiao, Chengnan Xu, Junxiang Zhang,  
Zongshan Zhou, Qinglong Dong

Institute of Pomology, Chinese Academy of Agricultural Sciences, Xingcheng 125100, China

**Abstract:** The WRKY transcription factors are one of the largest families of transcriptional regulators and play diverse regulatory roles in biotic and abiotic stresses, plant growth and development processes. In this study, the WRKY DNA-binding domain (Pfam Database number: PF03106) downloaded from Pfam protein families database was exploited to identify WRKY genes from the peach (*Prunus persica* 'Lovell') genome using HMMER 3.0. The

收稿日期: 2015-05-27; 修回日期: 2015-12-30

基金项目: 国家自然科学基金项目(编号: 31401852)和农业部公益性行业项目(编号: 201203034)资助[Supported by the National Natural Science Foundation of China (No. 31401852) and Special Fund for Agro-scientific Research in the Public Interest (No. 201203034)]

作者简介: 谷彦冰, 硕士研究生, 专业方向: 果树分子病理学。E-mail: guyanbing2010@126.com

冀志蕊, 助理研究员, 研究方向: 果树分子病理学。E-mail: xinyu\_jzr@163.com

谷彦冰和冀志蕊为并列第一作者。

通讯作者: 董庆龙, 助理研究员, 研究方向: 果树分子生物学。E-mail: dong19850412@163.com

周宗山, 研究员, 研究方向: 果树病理学。E-mail: zszhouqrj@163.com

DOI: 10.16288/j.yczs.15-235

网络出版时间: 2015-12-30 17:12:51

URL: <http://www.cnki.net/kcms/detail/11.1913.R.20151230.1712.006.html>

obtained amino acid sequences were analyzed with DNAMAN 5.0, WebLogo 3, MEGA 5.1, MapInspect and MEME bioinformatics softwares. Totally 61 peach *WRKY* genes were found in the peach genome. Our phylogenetic analysis revealed that peach *WRKY* genes were classified into three Groups: I, II and III. The WRKY N-terminal and C-terminal domains of Group I (group I-N and group I-C) were monophyletic. The Group II was sub-divided into five distinct clades (group II-a, II-b, II-c, II-d and II-e). Our domain analysis indicated that the WRKY regions contained a highly conserved heptapeptide stretch WRKYGQK at its N-terminus followed by a zinc-finger motif. The chromosome mapping analysis showed that peach *WRKY* genes were distributed with different densities over 8 chromosomes. The intron-exon structure analysis revealed that structures of the *WRKY* gene were highly conserved in the peach. The conserved motif analysis showed that the conserved motifs 1, 2 and 3, which specify the WRKY domain, were observed in all peach WRKY proteins, motif 5 as the unknown domain was observed in group II-d, two WRKY domains were assigned to Group I. SqRT-PCR and qRT-PCR results indicated that 16 *PpWRKY* genes were expressed in roots, stems, leaves, flowers and fruits at various expression levels. Our analysis thus identified the *PpWRKY* gene families, and future functional studies are needed to reveal its specific roles.

**Keywords:** peach; WRKY; transcription factor; expression analysis; gene family

桃(*Prunus persica* L.)是我国水果产业的重要支柱产业之一,其种植面积和总产量仅次于苹果和梨。但是在桃产业高速发展的同时,各种病害也造成了巨大的危害和损失,已成为制约我国桃产业可持续发展的一道门槛。大量研究表明,植物 WRKY 转录因子作为在寄主-病原物互作中一个重要的转录调控因子,是植物防卫反应信号传递网络的重要组成部分,通过其正调控和负调控的协同作用能够使植物的抗病能力广谱持久。WRKY 是植物中最大的转录因子家族,参与植物诸多生理生化过程,其名称源于一段靠近于 C 端由约 60 个氨基酸残基组成的高度保守的 WRKY 保守域,这段区域包括 N 端的 WRKYGQK 序列和 C 端的锌指结构(Zinc finger motif),能够与靶基因启动子区域中的 W 盒(T)TGACC(A/T)<sup>[1,2]</sup>或者 SURE(糖响应顺式元件)<sup>[3]</sup>核苷酸序列特异性结合进而调控下游目标基因的表达。

根据 WRKY 保守域的数量和锌指结构的类型,WRKY 蛋白最初被分为 3 种类型(Type I、II 和 III)<sup>[1,2]</sup>。典型的类型 I 蛋白包含 2 个 WRKYGQK 序列和 C<sub>2</sub>H<sub>2</sub> 锌指结构,类型 II 蛋白包含 1 个 WRKYGQK 七肽和 C<sub>2</sub>H<sub>2</sub> 型锌指结构,依据其结构特征可进一步分为 5 个亚组(II a、II b、II c、II d 和 II e);类型 III 蛋白含有一个 WRKY 保守域,但其锌指结构为 C<sub>2</sub>HC<sup>[1,2]</sup>。进一步研究表明,依据进化关系可将高等植物 WRKY 家族分为 I、II a+II b、II c、II d+II e 和 III 5 组,并且 II a+II b、II c 和 II d+II e 三组是由 I

组 C 端 WRKY 结构域进化而来<sup>[4]</sup>。此外除了 WRKY 结构域和锌指结构,WRKY 蛋白还含有核定位信号、亮氨酸拉链结构、激酶域、TIR-NBS-LRR 以及富含谷氨酸酰胺域、丝氨酸/苏氨酸域和脯氨酸域等结构域<sup>[5]</sup>。

自 1994 年 Ishiguro 等<sup>[6]</sup>首次发现甘薯(*Ipomoea batatas* L.)WRKY 转录因子 SPF1 以来,目前已在野燕麦(*Avena fatua* L.)<sup>[7]</sup>、欧芹(*Petroselinum crispum* L.)<sup>[8]</sup>和拟南芥(*Arabidopsis thaliana*)<sup>[9]</sup>等多种植物中克隆获得该类基因。大量研究表明,WRKY 蛋白在植物的多重生理过程中发挥重要调控作用,例如在生长发育过程中,调控表皮毛的起始<sup>[10]</sup>、胚胎形态的发生<sup>[11]</sup>、叶片的衰老<sup>[12]</sup>;同时在糊粉细胞中,还能调控 GAMYB(GA 诱导转录激活子)介导的 GA 信号途径<sup>[13]</sup>;另外,在非生物胁迫包括寒冷、干旱、盐等<sup>[5,14~16]</sup>和生物胁迫包括细菌、真菌、病毒等<sup>[1,2]</sup>逆境响应及信号转导过程中也起着重要调控作用。

近年来随着高通量测序技术的发展,已经完成了多种植物的基因组测序。截止目前为止,WRKY 基因家族已经在拟南芥(*A. thaliana*)、水稻(*Oryza sativa* L.)、番茄(*Solanum lycopersicum* L.)、黄瓜(*Cucumis sativus* L.)、杨树(*Populus tremula* × *Populus alba*)、苹果(*Malus domestica* L.)、棉花(*Gossypium arboreum* L.)、橡胶树(*Hevea brasiliensis* L.)和蒺藜苜蓿(*Medicago truncatula* L.)等物种中被鉴定出来<sup>[15,17~24]</sup>,但有关桃(*Prunus persica* ‘Lovell’)WRKY 基因家族

相关鉴定研究尚未进行。因此,本研究首次对桃 *WRKY* 基因(*PpWRKY*)进行全基因组挖掘和分析,全面、系统地解析桃 *WRKY* 基因家族的基本信息、保守域结构、进化关系、亚组分类、内含子与外显子结构、染色体定位和组织表达分析等,以期对桃 *WRKY* 基因家族功能的解析、调控网络的构建以及为植物和病原物的协同进化研究提供一定的理论基础。

## 1 材料和方法

### 1.1 材料

供试材料为 2013 年山东省果树研究所天平湖试验示范基地(山东泰安)3 年生‘超红’桃树(*Prunus persica* ‘Chaohong’, 是山东省果树研究所从美国引进“White Lady”桃的自然杂交实生种中选出的早熟桃)取新生根系(4 月 5 日)、一年生的枝条(茎)(4 月 5 日)、新叶(4 月 5 日)、花(盛花期)(4 月 5 日)和果实(花后 80 d)(6 月 24 日,以盛花期 4 月 5 日开始计算)于液氮中速冻,−80℃ 保存备用。

### 1.2 方法

#### 1.2.1 *PpWRKY* 基因鉴定

下载 *WRKY* 保守域蛋白序列 PF03106(来自 Pfam 数据库, <http://pfam.sanger.ac.uk/>)HMM 文件,利用该结构域 HMM 文件,使用 HMMER 3.0 软件(<http://hmmer.janelia.org/>)对桃功能基因组数据库([http://www.rosaceae.org/gb/gbrowse/prunus\\_persica/](http://www.rosaceae.org/gb/gbrowse/prunus_persica/))中注释基因组序列进行 *WRKY* 基因的比对搜索,参数设置为默认参数;利用保守域预测软件 Pfam 和 SMART(<http://smart.embl-heidelberg.de/>)确保上述候选基因含有 *WRKY* 保守域<sup>[25]</sup>。

#### 1.2.2 *PpWRKY* 蛋白 *WRKY* 保守域序列分析

利用序列分析软件 DNAMAN 5.0 和在线软件 WebLogo 3(<http://weblogo.threeplusone.com/>)<sup>[26]</sup>分析 *PpWRKY* 蛋白的 *WRKY* 保守序列;利用在线软件 SWISS-MODEL 对 *PpWRKY* 蛋白的 *WRKY* 保守域的 3D 结构进行模型匹配( $\beta$ 1-4 拓扑结构;PDB ID: 2AYD)(<http://swissmodel.expasy.org/>);3D 结构可视化采用 RasTop 软件(<http://www.geneinfinity.org/ras-top/>)<sup>[26]</sup>。

#### 1.2.3 *PpWRKY* 的进化分析

利用进化树分析软件 MEGA 5.1(<http://www.megasoftware.net>)<sup>[27]</sup>,根据 NJ 方法(执行参数: Poisson correction、pairwise deletion 和 bootstrap (1000 次重复)),对 *PpWRKY* 的 *WRKY* 保守域蛋白序列进行进化树分析。引入 62 条拟南芥 *WRKY* 蛋白对 *PpWRKY* 进行亚组分类。拟南芥 *WRKY* 蛋白序列从拟南芥网站(<http://www.arabidopsis.org/>)下载。

#### 1.2.4 *PpWRKY* 内含子、外显子和基因组定位

*PpWRKY* 内含子、外显子和基因组定位信息均从桃功能基因组数据库下载。*PpWRKY* 基因内含子和外显子结构示意图根据桃功能基因组数据库中 Exon-Tron 结构图整理而成;*PpWRKY* 基因的基因组定位利用软件 MapInspect 制作而成(<http://mapinspect.software.informer.com/>)<sup>[28]</sup>。

#### 1.2.5 *PpWRKY* 蛋白的保守域分析

利用在线软件 MEME(<http://meme.nbcr.net/meme/tools/meme>)对 *PpWRKY* 蛋白保守域结构进行分析<sup>[25]</sup>。

#### 1.2.6 *PpWRKY* 基因的组织表达

使用 RNeasy Plant Mini Kit(QIAGEN 公司)从桃的根、茎、新叶、花和果实组织提取 RNA,用反转录试剂盒 PrimeScript™ 1st Strand cDNA Synthesis Kit(购自宝生物工程(大连)有限公司)合成 cDNA。采用 *PPERACTIN-2*<sup>[29,30]</sup>和 *TEF2*(Translation elongation factor 2; EST 登录号: TC3544)<sup>[31]</sup>分别作为半定量 PCR(SqRT-PCR)和实时荧光定量 PCR(qRT-PCR)内参基因。参照 Tian 等<sup>[25]</sup>方法, SqRT-PCR 扩增体系为 25  $\mu$ L: 1  $\mu$ L cDNA; 2.5  $\mu$ L 10 $\times$  *Taq* 缓冲液; 200  $\mu$ mol dNTP; 各 10 pmol 上、下游引物; 0.5 U *Taq* DNA 聚合酶(购自全式金公司); 剩余用 ddH<sub>2</sub>O 补齐。循环参数为: 94℃ 5 min; 94℃ 25 s; 58℃ 25 s; 72℃ 30 s; 目的基因为 36 个循环,内参基因为 28 个循环。每一基因的 SqRT-PCR 反应均为 3 次重复,得到类似结果。qRT-PCR 反应体系为: SYBR Green I 10  $\mu$ L(购自康为试剂公司), 上、下游引物(5  $\mu$ mol/L)各 1  $\mu$ L, cDNA 1  $\mu$ L, ddH<sub>2</sub>O 7  $\mu$ L。PCR 扩增程序为: 95℃ 预变性 3 min; 95℃ 10 s, 58℃ 30 s, 72℃ 15 s, 40 个

循环,最后 55℃进行复性,每隔 7 s 上升 0.5℃至 95℃  
共 81 个循环。所得数据采用  $2^{-\Delta\Delta CT}$  进行计算分析。

引物设计均避开 WRKY 保守域,所用引物具体序列  
详见表 1。

表 1 实时荧光定量 PCR 和半定量 PCR 引物序列

Table 1 Primer sequences of qRT-PCR and SqRT-PCR

用途	基因名称	上游引物(5'→3')	下游引物(5'→3')
sqRT-PCR	<i>PpWRKY2</i>	ATGGCTAAGGGAGGTGGACT	TGCAGAGTTGGAGGAGGAAG
	<i>PpWRKY3</i>	CTTCTACTGGTGGTGGTGGT	ACCGTACACCATTTGAGGCT
	<i>PpWRKY5</i>	TCTGTCTTTGTTGCTCTTCTGC	GCCAAGTTTGACGGGGAATT
	<i>PpWRKY6</i>	AGTCATAACACTGCCACCGA	GGTTTTCGTTAAATTGGATTGGC
	<i>PpWRKY8</i>	ATGAAACCCTCTCGCCAGAA	CTGCACCTGTAGTAGCCTCT
	<i>PpWRKY12</i>	GAGGCAGAGCGGGTAGTTAT	GACATCTCTGCCTGGGAAGA
	<i>PpWRKY13</i>	TTCTTTTGCATCATCGGGCG	AGTGGTGGTTGTTGTTGCTG
	<i>PpWRKY14</i>	ACCCTTCTGCATCCCAAGAA	CACCTGGGCATTGTTTTCCT
	<i>PpWRKY17</i>	GCTGGAATTCAAAGCGTCGA	CAGGAGCTCTTCTGAACCGA
	<i>PpWRKY19</i>	GATCCAGCTAGAACGACCGA	ACAAGCTGGCTACTAGTTGC
	<i>PpWRKY20</i>	CAAGCTCCTGTGAAGCAACA	GCGGTCATCAACTCCTTCAC
	<i>PpWRKY21</i>	AGGCTAAGTCTAAGTGCAGCA	AGGTCCTGATGATCCTTGGG
	<i>PpWRKY23</i>	CCTCATCACCATTTCAGGCA	AAGAGTTTGGTGTTCAGGC
	<i>PpWRKY25</i>	GCAACAACAGCAACAGCAGA	TGCTGTTGTGCGCTTTGATG
	<i>PpWRKY26</i>	TGATCCGCGTCTATCGAAT	CGAAGTTGTCTGAGGCTTGG
	<i>PpWRKY28</i>	AGGACGAAGAAGGAAAGCGA	ACCCCAAGTTGCTCCTGTAA
	<i>PpWRKY30</i>	CGTTTTCACTTTGACAGCGC	TTTTGCCCGTATTTTCGCCA
	<i>PpWRKY32</i>	GCCCAACTTGCTCTCAAAC	CAGCTACAAGTTGGGTTTGA
	<i>PpWRKY33</i>	ATCCAAACCCTAGCTCTGCC	CTCCTCCCTCATTTCTGCA
	<i>PpWRKY35</i>	GAGTCAATCACGATGCATGAGA	TCTTGTGACACCAGAAGTCT
	<i>PpWRKY36</i>	ACCACCAGCTTCTTCTCCT	GTCCCTAACCTCACAGCAGT
	<i>PpWRKY37</i>	GGAGAGAGAAGGAGATGGGC	TCTGTGCTCTTGGTTTCATGC
	<i>PpWRKY39</i>	ACCAGTCTCTTCAGATGCATT	TCTTGATCGTACCTTTCCCA
	<i>PpWRKY40</i>	AAACTCTTCCACCACTCCGT	GTCTTGTGGGGTTGACGAC
	<i>PpWRKY41</i>	TCTTTGTGCTACCAGAGGCG	CCTTGCTCAAACCCCTTCC
	<i>PpWRKY43</i>	GGGGAAGGGTGAAGGAAGA	AGCCACCACCACTACTGATC
	<i>PpWRKY44</i>	AGCTGAAATGGGTGAGGTGA	GGGCTCCCTGGACTTCTTC
	<i>PpWRKY46</i>	CTCAGACCATCCTACCCAG	CAACTTGGCTCCTCGTTTGA
	<i>PpWRKY47</i>	TTATCCATCCAAACTCCTCACT	GTTGCTGCTGTTGTTGCTGA
	<i>PpWRKY49</i>	AGCAGCCTCTACTTTCTCCG	TCCATGGATGATTTGTCGCT
	<i>PpWRKY50</i>	ACATGCCCATCACCTGTTTG	TGGGATGAAGGGTTAGGAGAG
	<i>PpWRKY52</i>	GTCAGCCTTCTCTCTGGTCA	GCCCATATTTCTCCAGCGG
	<i>PpWRKY53</i>	GGATTTGCAAGCTGTGGTGA	AACAGACATGGAGGAGGCAA
	<i>PpWRKY55</i>	AGCCCATGACGTCAGAGAAT	TTTGCCGTACTTCTCCTCAA
	<i>PpWRKY58</i>	CAGACGTTGCAAGGACTCAA	TGACCATACTTTTCTTCACTGCT
	<i>PpWRKY60</i>	CAACAGGCCAAGCTTTCCAT	GTGATCTCTACAGCAGCCT
	<i>PpWRKY61</i>	ACCAACAGACGAAGAGAGCA	AGGTGGAATAGTGAGGCAGG
	<i>PPERACTIN</i>	GTGGGGATGGGACAGAAAGATG	GAGGTCAAGCCGGAGGATGG



续表 1

用途	基因名称	上游引物(5'→3')	下游引物(5'→3')
qRT-PCR	<i>PpWRKY5</i>	GCTTCTTCCTCAACAAGCCC	GCCAAGTTTGACGGGGAATT
	<i>PpWRKY6</i>	AGTCATAACACTGCCACCGA	GGGTATGTGTCGGGGAAGAA
	<i>PpWRKY17</i>	GCCGCTACAGAGTACAAAGC	CAGGAGCTCTTCTGAACCGA
	<i>PpWRKY19</i>	GATCCAGCTAGAACGACCGA	TTCGGTGAGGAGTGTGTTGT
	<i>PpWRKY26</i>	CTCGTCGGGTCTCAAGGAT	TTGATTGCATAGGACGCGGA
	<i>PpWRKY28</i>	CCAGAAGGTCAGGCATCCAA	GACCAAGAATCAGAAGGCGG
	<i>PpWRKY35</i>	GAGTCAATCACGATGCATGAGA	TTTCTCTGCAGCTCCTCCT
	<i>PpWRKY36</i>	ACCACCAGCTTCTTCTTCCT	ATCAGCCCCAAGAACCCATT
	<i>PpWRKY44</i>	ACCAGTCTCTCCAATGTCGT	GGGCTCCCTTGGACTTCTTC
	<i>PpWRKY46</i>	CTCAGACCATCCTCACCCAG	GAAGTTGGTGGCGAGGTTG
	<i>PpWRKY53</i>	ACTGCCAATATTGATGAAGCCA	TCCTCCAATTCATCTAAAGCTGC
	<i>PpWRKY55</i>	AGCCCATGACGTCAGAGAAT	ACTGTCAACCACCTCTTCTG
	<i>TEF2</i>	GGTGTGACGATGAAGAG TGATG	TGAAGGAGAGGGAAGGTGAAAG

## 2 结果与分析

### 2.1 *PpWRKY* 基因序列鉴定

利用 WRKY 结构域 HMM 文件,使用 HMMER 3.0 软件对桃注释基因组序列进行 *WRKY* 基因的全基因组序列比对搜索,得到 61 个 *PpWRKY* 基因(表 2)。利用保守域预测软件 Pfam 和 SMART 确定桃 *WRKY* 基因存在 WRKY 保守域,用以验证序列比对搜索的正确性。

### 2.2 *PpWRKY* 基因进化树分析和分组鉴定

通过分析已知类型的拟南芥 WRKY 蛋白 WRKY 保守域特点,利用 MEGA 5.1 软件对 *PpWRKY* 蛋白的 WRKY 结构域序列进行聚类分析。利用已知拟南芥 WRKY 蛋白的类型确定 *PpWRKY* 蛋白的分组情况。进化树结果(图 1)显示,相比于拟南芥类型 I、II、III, *PpWRKY* 蛋白也分为类型 I、II 和 III,类型 I 分为 I-C 亚组和 I-N 亚组;类型 II 分为 II-a、II-b、II-c(除 AtWRKY58N)、II-d 和 II-e 亚组。每个 *PpWRKY* 基因所属类型详见表 2。

### 2.3 *PpWRKY* 基因 WRKY 保守域分析

通过 DNAMAN5.0 序列分析软件和在线软件 WebLogo 3 对 *PpWRKY* 基因的 WRKY 结构域进行分析。结果表明, *PpWRKY* 基因的 WRKY 结构域是高度保守的(图 2,图 3)。其中,类型 I 中 C 端 WRKY

七肽域和锌指结构是 WRKYGQK 和 CX<sub>4</sub>C<sub>23</sub>HXH 形式,N 端 WRKY 七肽域和锌指结构则是 WRKYGQK 和 CX<sub>4</sub>C<sub>22</sub>HXH 形式(除 *PpWRKY61N* 外);类型 II 中 a、b、d 和 e 亚组 C 端 WRKY 七肽域和锌指结构为 WRKYGQK 和 CX<sub>5</sub>C<sub>23</sub>HXH 形式,而 c 亚组 C 端 WRKY 七肽域和锌指结构为 WRKYGQK(除 *PpWRKY33* 和 *PpWRKY35* 外)和 CX<sub>4</sub>C<sub>23</sub>HXH 形式(除 *PpWRKY31* 和 *PpWRKY35* 外);类型 III 中 C 端 WRKY 七肽域和锌指结构为 WRKYGQK 和 CX<sub>7</sub>C<sub>23</sub>HXC 形式(除 *PpWRKY46* 和 *PpWRKY58* 外)(表 3)。

利用 SWISS-MODEL 在线软件分析发现, *PpWRKY* 蛋白的 WRKY 结构域与拟南芥 WRKY1 蛋白(AtWRKY1)WRKY 结构域的晶体结构相似<sup>[32]</sup>。因此,利用 AtWRKY1 蛋白 WRKY 结构域的晶体结构为模板,使用 3D 作图软件 RasTop 对 *PpWRKY* 蛋白的 WRKY 结构域的结构进行作图,结果表明桃 WRKY 结构域包括 β 折叠 1、2、3 和 4(图 3)。以上数据表明, *PpWRKY* 基因家族在结构上是高度保守的。

### 2.4 *PpWRKY* 基因家族内含子与外显子分析

利用进化树分析软件 MEGA5 对 *PpWRKY* 全长蛋白序列进行进化分析。从图 4 中可以看出, *PpWRKY* 基因家族分为 I、II 和 III 组,其中 II 组又分为 II-a、II-b、II-c、II-d 和 II-e 共 5 个亚组,这与图 1 进化树结果相一致(除 *PpWRKY35*, 41, 46 和 58)。同时,

通过桃功能基因组数据库下载内含子和外显子结构分布图，整理数据得到图 4。PpWRKY 基因家族成员由 2~6 个外显子构成，大多数翻译成 C 端部分 WRKY 结构域的外显子是该基因组结构倒数第 2 个外显子(图 4 右侧灰色区域)，并且不同小组编码成 WRKY 的外显子长度是不同的，统计结果见表 4。

表 2 PpWRKY 基因家族成员信息  
Table 2 Information of the PpWRKY gene family in peach

基因名称	定位名称	类型	染色体定位	氨基酸大小(aa)	拟南芥同源基因	GenBank 登录号
<i>PpWRKY1</i>	ppa001924m	I	scaffold_6:2468104..2471544 +	740	<i>AtWRKY2</i>	XM_007206365
<i>PpWRKY2</i>	ppa002619m	II -b	scaffold_3:16460557..16463047 +	651	<i>AtWRKY6</i>	XM_007216926
<i>PpWRKY3</i>	ppa003305m	I	scaffold_4:14951801..14956955 +	586	<i>AtWRKY2</i>	XM_007214607
<i>PpWRKY4</i>	ppa003333m	I	scaffold_6:1828828..1831501 -	584	<i>AtWRKY3</i>	XM_007208228
<i>PpWRKY5</i>	ppa003574m	II -b	scaffold_4:820970..825131 +	564	<i>AtWRKY6</i>	XM_007213885
<i>PpWRKY6</i>	ppa003809m	I	scaffold_3:19419766..19423127 +	547	<i>AtWRKY4</i>	XM_007215479
<i>PpWRKY7</i>	ppa004009m	I	scaffold_6:24639609..24641627 -	535	<i>AtWRKY3</i>	XM_007204965
<i>PpWRKY8</i>	ppa004042m	I	scaffold_7:22197495..22201494 +	533	<i>AtWRKY3</i>	XM_007204037
<i>PpWRKY9</i>	ppa004312m	I	scaffold_1:27443164..27446022 -	517	<i>AtWRKY3</i>	XM_007222810
<i>PpWRKY10</i>	ppa004535m	II -e	scaffold_4:5152653..5155480 -	504	<i>AtWRKY14</i>	XM_007211770
<i>PpWRKY11</i>	ppa004905m	I	scaffold_3:15682525..15685261 -	486	<i>AtWRKY1</i>	XM_007215222
<i>PpWRKY12</i>	ppa005152m	I	scaffold_6:22336216..22339034 +	475	<i>AtWRKY44</i>	XM_007205075
<i>PpWRKY13</i>	ppa006909m	II -c	scaffold_1:8640398..8642407 -	390	<i>AtWRKY4</i>	XM_007222977
<i>PpWRKY14</i>	ppa007708m	III	scaffold_2:19059670..19062101 -	358	<i>AtWRKY5</i>	XM_007218131
<i>PpWRKY15</i>	ppa007745m	II -e	scaffold_2:18665651..18667118 +	357	<i>AtWRKY2</i>	XM_007218136
<i>PpWRKY16</i>	ppa007838m	II -d	scaffold_6:27763305..27764742 +	354	<i>AtWRKY4</i>	XM_007205381
<i>PpWRKY17</i>	ppa007930m	II -d	scaffold_1:35966149..35967624 +	351	<i>AtWRKY4</i>	XM_007222195
<i>PpWRKY18</i>	ppa007986m	III	scaffold_2:25613734..25614985 -	349	<i>AtWRKY4</i>	XM_007218173
<i>PpWRKY19</i>	ppa008480m	II -c	scaffold_1:27581955..27585622 +	330	<i>AtWRKY5</i>	XM_007222551
<i>PpWRKY20</i>	ppa008566m	II -a	scaffold_1:34009489..34011202 +	326	<i>AtWRKY4</i>	XM_007222589
<i>PpWRKY21</i>	ppa008573m	III	scaffold_5:11878403..11880061 -	326	<i>AtWRKY5</i>	XM_007209289
<i>PpWRKY22</i>	ppa008613m	II -d	scaffold_6:21500388..21502030 +	325	<i>AtWRKY7</i>	XM_007205490
<i>PpWRKY23</i>	ppa008731m	II -c	scaffold_10:526239..528128 -	321	<i>AtWRKY2</i>	XM_007221428
<i>PpWRKY24</i>	ppa008846m	II -d	scaffold_1:37472989..37474731 +	317	<i>AtWRKY1</i>	XM_007222883
<i>PpWRKY25</i>	ppa009647m	II -e	scaffold_4:3206039..3207150 +	283	<i>AtWRKY6</i>	XM_007211708
<i>PpWRKY26</i>	ppa009700m	II -d	scaffold_1:4817218..4818580 +	281	<i>AtWRKY1</i>	XM_007222139
<i>PpWRKY27</i>	ppa009939m	II -a	scaffold_1:34029320..34030841 +	271	<i>AtWRKY4</i>	XM_007222815
<i>PpWRKY28</i>	ppa010217m	II -e	scaffold_8:21517865..21519013 -	258	<i>AtWRKY3</i>	XM_007201303
<i>PpWRKY29</i>	ppa010796m	II -a	scaffold_3:6662580..6663950 +	236	<i>AtWRKY4</i>	XM_007215847
<i>PpWRKY30</i>	ppa012414m	II -c	scaffold_5:13581887..13583455 +	170	<i>AtWRKY7</i>	XM_007209633
<i>PpWRKY31</i>	ppa012514m	II -c	scaffold_5:13581887..13583455 +	166	<i>AtWRKY7</i>	XM_007209632
<i>PpWRKY32</i>	ppa014826m	III	scaffold_2:23638750..23643347 +	348	<i>AtWRKY5</i>	XM_007220984
<i>PpWRKY33</i>	ppa015480m	II -c	scaffold_6:15635236..15637392 +	196	<i>AtWRKY5</i>	XM_007206175
<i>PpWRKY34</i>	ppa015755m	II -c	scaffold_2:18683476..18685191 +	221	<i>AtWRKY5</i>	XM_007219053

续表 2

基因名称	定位名称	类型	染色体定位	氨基酸大小(aa)	拟南芥同源基因	GenBank 登录号
<i>PpWRKY35</i>	ppa015765m	II -c	scaffold_5:12459036..12459687 +	134	<i>AtWRKY3</i>	XM_007210154
<i>PpWRKY36</i>	ppa016153m	II -c	scaffold_1:22733682..22735582 +	185	<i>AtWRKY7</i>	XM_007226135
<i>PpWRKY37</i>	ppa016401m	II -b	scaffold_1:43717207..43719518 -	561	<i>AtWRKY9</i>	XM_007224089
<i>PpWRKY38</i>	ppa016459m	II -c	scaffold_3:21511993..21513894 -	223	<i>AtWRKY7</i>	XM_007216875
<i>PpWRKY39</i>	ppa017919m	II -b	scaffold_1:26674923..26676963 +	533	<i>AtWRKY7</i>	XM_007226734
<i>PpWRKY40</i>	ppa018075m	II -b	scaffold_10:162785..165574 +	567	<i>AtWRKY4</i>	XM_007221507
<i>PpWRKY41</i>	ppa019550m	II -c	scaffold_6:23123896..23125761 -	299	<i>AtWRKY4</i>	XM_007206465
<i>PpWRKY42</i>	ppa020629m	II -c	scaffold_6:15369732..15373318 -	231	<i>AtWRKY1</i>	XM_007208140
<i>PpWRKY43</i>	ppa020736m	III	scaffold_6:25162460..25164336 -	335	<i>AtWRKY7</i>	XM_007207576
<i>PpWRKY44</i>	ppa020780m	II -b	scaffold_3:19839824..19842047 -	481	<i>AtWRKY7</i>	XM_007217278
<i>PpWRKY45</i>	ppa021195m	II -b	scaffold_5:15492412..15495700 +	646	<i>AtWRKY6</i>	XM_007210689
<i>PpWRKY46</i>	ppa021559m	III	scaffold_16:32444..34188 +	119	<i>AtWRKY5</i>	XM_007221211
<i>PpWRKY47</i>	ppa021657m	II -d	scaffold_8:19829267..19831289 -	299	<i>AtWRKY2</i>	XM_007199137
<i>PpWRKY48</i>	ppa022287m	II -c	scaffold_2:15597322..15600214 +	244	<i>AtWRKY1</i>	XM_007219434
<i>PpWRKY49</i>	ppa022758m	II -b	scaffold_4:13666776..13669306 -	513	<i>AtWRKY4</i>	XM_007212931
<i>PpWRKY50</i>	ppa023256m	II -d	scaffold_5:8848352..8851174 -	340	<i>AtWRKY2</i>	XM_007210612
<i>PpWRKY51</i>	ppa024027m	II -c	scaffold_4:3695838..3697364 +	337	<i>AtWRKY2</i>	XM_007214189
<i>PpWRKY52</i>	ppa024540m	II -c	scaffold_2:21972625..21973268 -	174	<i>AtWRKY4</i>	XM_007219721
<i>PpWRKY53</i>	ppa025013m	II -e	scaffold_5:11272595..11273677 +	283	<i>AtWRKY2</i>	XM_007211052
<i>PpWRKY54</i>	ppa025042m	II -e	scaffold_2:25381962..25384144 +	402	<i>AtWRKY2</i>	XM_007219579
<i>PpWRKY55</i>	ppa025586m	II -e	scaffold_3:8609429..8613217 -	277	<i>AtWRKY6</i>	XM_007216667
<i>PpWRKY56</i>	ppa025912m	III	scaffold_6:25157893..25159753 -	350	<i>AtWRKY7</i>	XM_007206750
<i>PpWRKY57</i>	ppa026125m	III	scaffold_2:23642489..23644464 -	323	<i>AtWRKY7</i>	XM_007219677
<i>PpWRKY58</i>	ppa026260m	III	scaffold_5:11029791..11030494 +	76	<i>AtWRKY7</i>	XM_007211071
<i>PpWRKY59</i>	ppa026704m	III	scaffold_6:25167296..25170022 -	286	<i>AtWRKY7</i>	XM_007208435
<i>PpWRKY60</i>	ppa026831m	II -c	scaffold_3:13803400..13805785 +	360	<i>AtWRKY2</i>	XM_007217759
<i>PpWRKY61</i>	ppb022557m	I	scaffold_6:28587954..28591793 -	751	<i>AtWRKY2</i>	XM_007206947

注：*PpWRKY23*、*PpWRKY40* 和 *PpWRKY46* 没有定位到 8 条染色体上。+表示正向链，-表示反义链。

2.5 PpWRKY 基因家族保守元件分析

利用在线软件 MEME 鉴定 PpWRKY 基因家族的保守结构域。结果显示，PpWRKY 基因家族包含 5 个保守元件：元件 1、2 和 3 为 WRKY 盒；元件 4、5 为未知盒(图 5)。PpWRKY 基因家族都包含 WRKY 盒，Group I 中含有 2 个 WRKY 盒；Group II -d 中含有未知元件 5。

2.6 PpWRKY 基因家族染色体定位分析

通过桃功能基因组数据库下载 PpWRKY 基因家族相关染色体定位信息，利用软件 MapInspect 制

作染色体定位分布图。从图 6 中可以看出，PpWRKY 基因分布于桃 8 条染色体中，呈不均匀分布。染色体 6 上分布最多，为 13 个；其次是染色体 1，分布 11 个；染色体 2、3、5、4 和 8，分别分布 9、8、8、6 和 2 个；染色体 7 分布最少，为 1 个。

2.7 PpWRKY 基因表达分析

随机选择 37 个 PpWRKY 基因进行组织(根、茎、叶、花和果实)表达分析(由于 PpWRKY30/31 和 PpWRKY20/29 基因序列相似性较高，本文选择 PpWRKY30 和 PpWRKY20 进行组织表达分析)。半定

量 PCR(SqRT-PCR)结果显示, 15 个 *PpWRKY* 基因在桃根、茎、叶、花和果中均有表达(图 7), 但相对表达水平有所差异。其中, *PpWRKY50*、53、55 和 60 在根、茎、叶和花中表达; *PpWRKY20*、33 和 41 在叶和花中有表达; 另外, *PpWRKY32*、39、44、46 和 52 只在叶中表达; *PpWRKY6* 和 35 分别只在果和花中表达(图 7)。此外, *PpWRKY23*、37 和 58 在上

述组织中均未能检测到表达(图 7), 这可能是由于存在假基因或者这些基因在特定的发育阶段和特定的环境下进行表达<sup>[25, 33]</sup>。为了验证上述基因表达的准确性, 随机选择了 12 个 *PpWRKY* 基因进行实时荧光定量 PCR(qRT-PCR)进行组织表达分析。qRT-PCR 结果与 SqRT-PCR 结果相对应, *PpWRKY* 基因表达结果是一致的(图 8)。

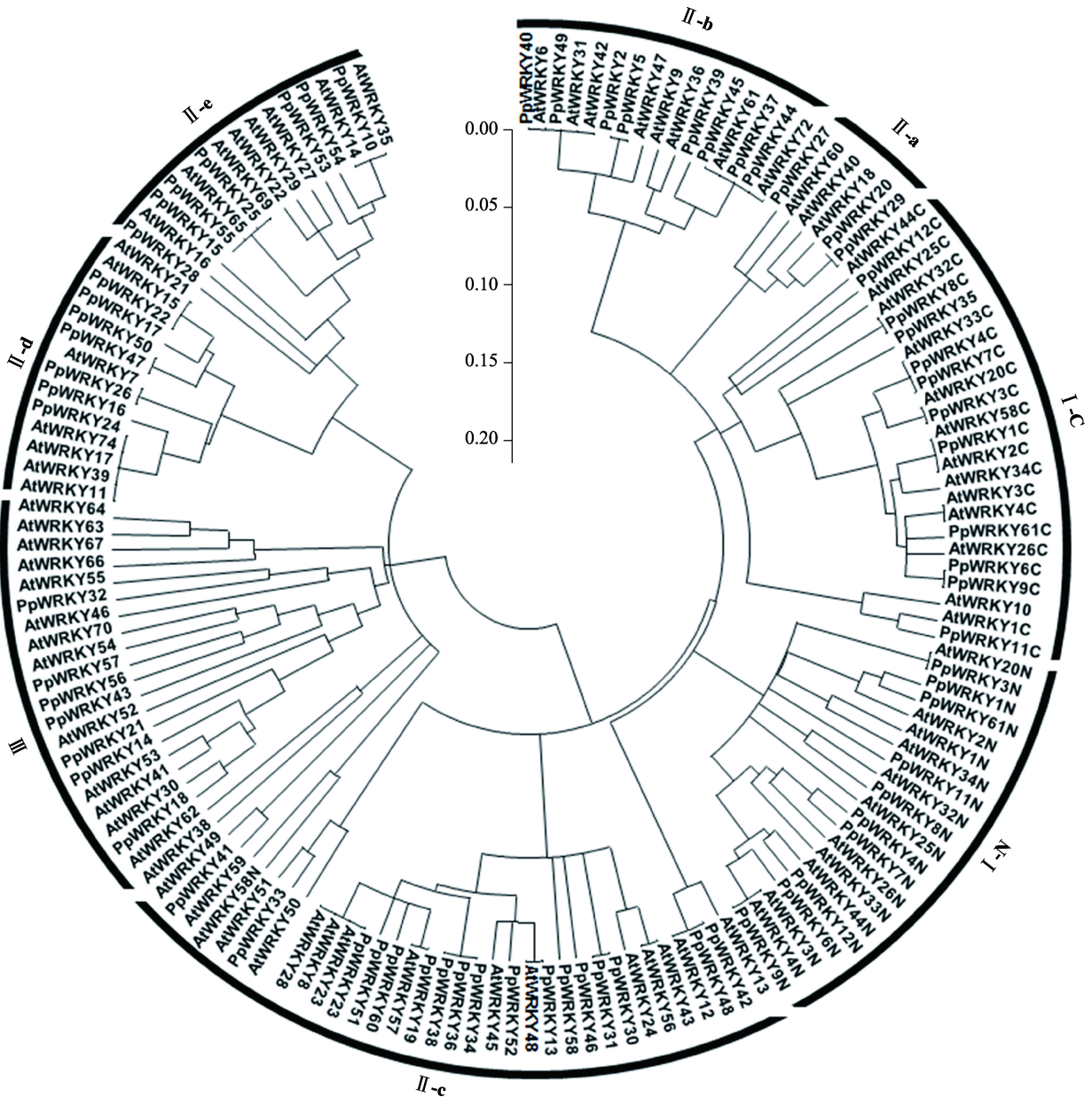


图 1 桃和拟南芥 WRKY 蛋白的 WRKY 保守域进化关系和亚组分类

Fig. 1 Phylogenetic relationship and subgroup classification of WRKY conserved domains in *Prunus persica* L. and *Arabidopsis* WRKY proteins

系统进化树采用邻接法由 MEGA 5.0 程序构建, 自举值(Bootstrap)设置为 1000。



序列比对采用 DNAMAN 5.0 软件生成。序列标签采用在线软件 WebLogo 3.0 构建,每个氨基酸高度表示此位点此氨基酸的相对频率,红色三角号表示内含子插入位点。

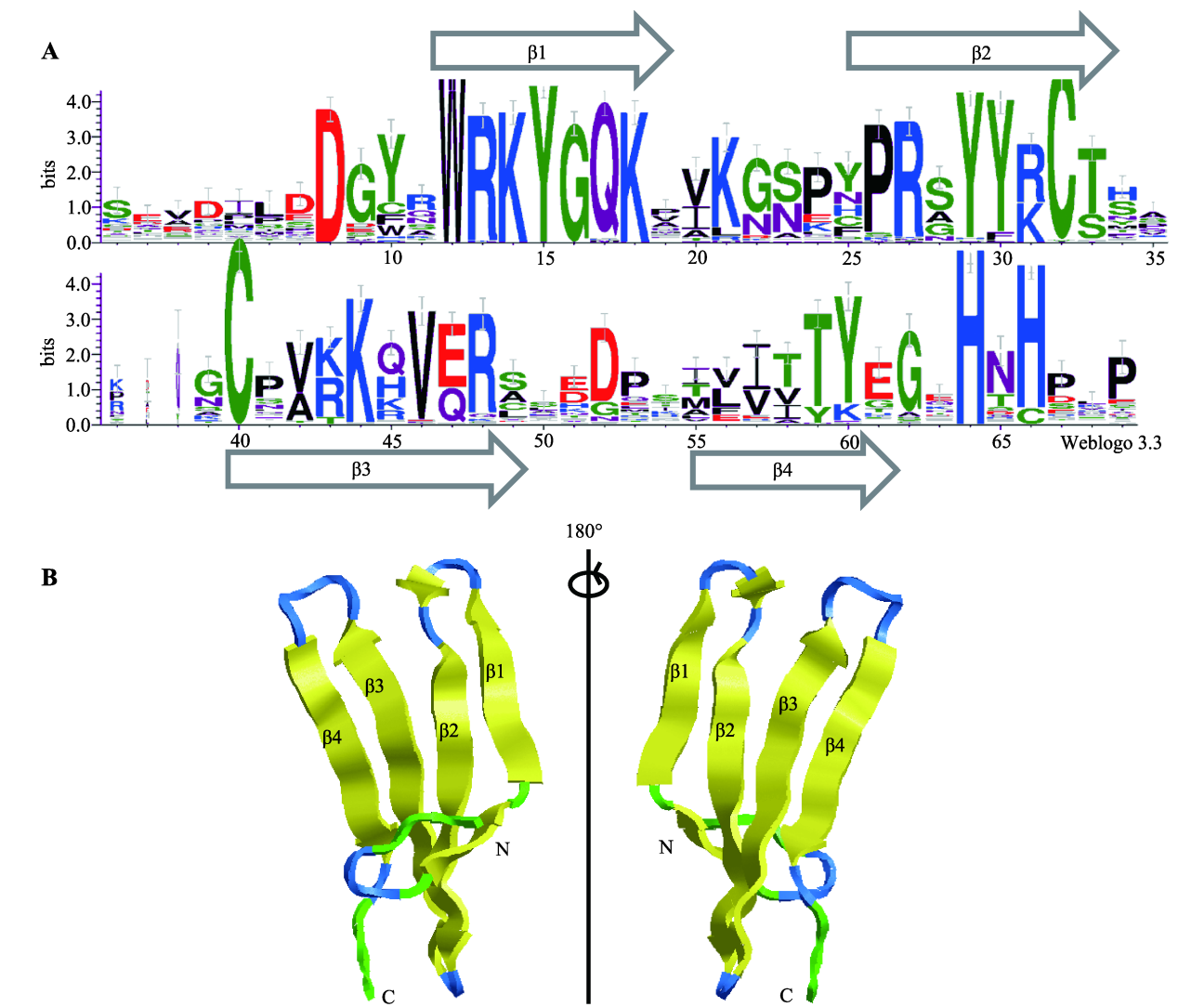


图 3 PpWRKY 蛋白的 WRKY 保守结构域

Fig. 3 The conserved domain logo and 3D structural model of the WRKY conserved domains in peach WRKY proteins

A: 序列标签。采用在线软件 WebLogo 3.0 构建, 每个氨基酸高度表示此位点此氨基酸的相对频率,  $\beta$  折叠标记在氨基酸相应位置, 桃 WRKY 保守结构氨基酸序列是从全长蛋白序列比对而来; B: 3D 结构模型。采用在线软件 SWISS-MODEL 进行模型匹配( $\beta$ 1-4 拓扑结构; PDB ID: 2AYD), 3D 结构可视化采用 RasTop 软件,  $\beta$  折叠标记黄色箭头区域, 肽链标记蓝色链条。

表 3 PpWRKY 蛋白的 WRKY 结构域中 WRKY 七肽域和锌指结构的变异

变异	<i>PpWRKY</i> 基因	分组
WRKYGQK 变异		
WRKYGKK	<i>PpWRKY33</i>	II -c
WRKYGMK	<i>PpWRKY35</i>	II -c
锌指结构变异		
CX4CX22	<i>PpWRKY35</i>	II -c
None	<i>PpWRKY46</i>	III
None	<i>PpWRKY58</i>	III



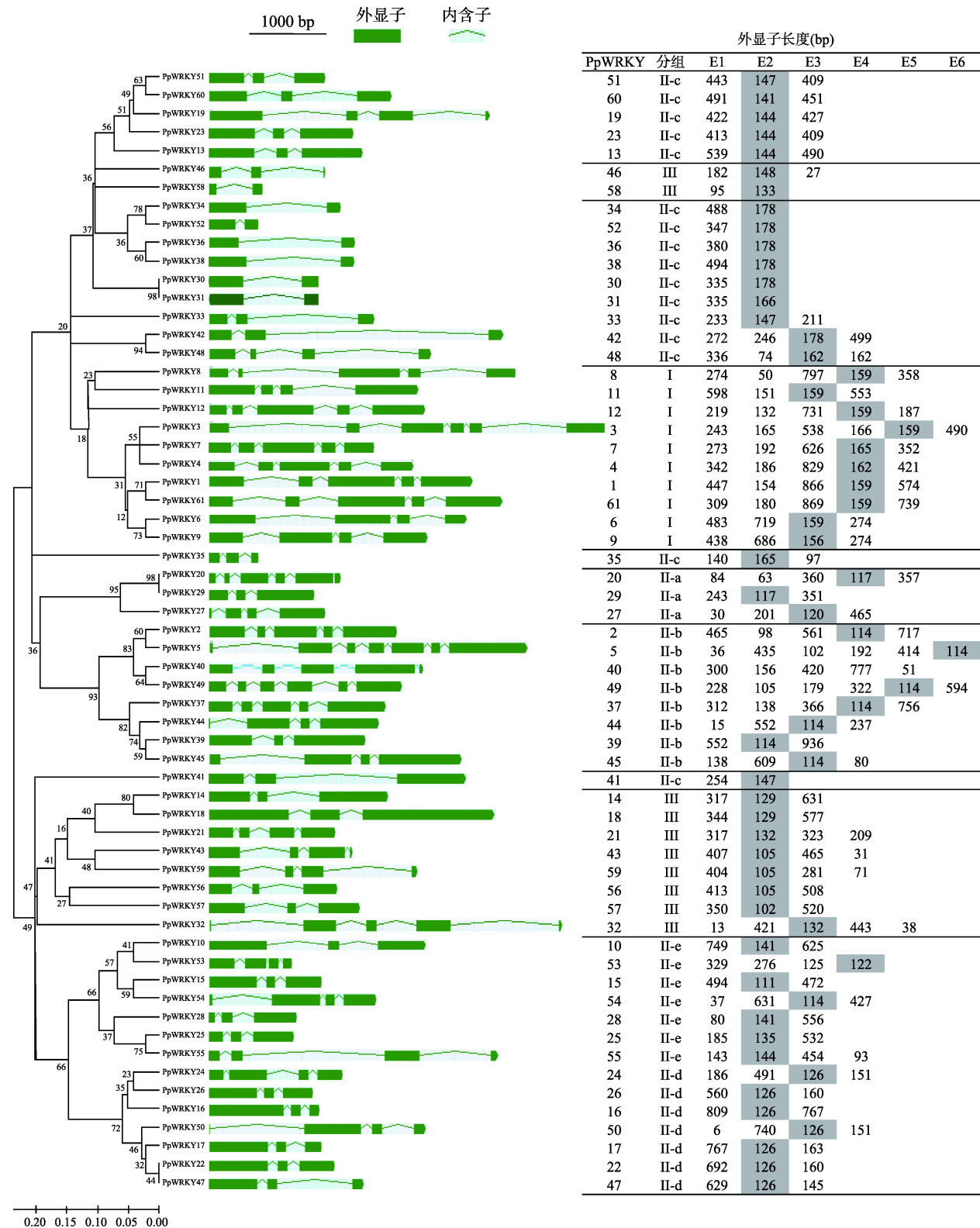


图 4 PpWRKY 基因家族内含子和外显子结构图

Fig. 4 Intron-exon structures of PpWRKY gene family

系统进化树采用邻接法由 MEGA5.1 序构建; 基因结构由桃基因组网站下载; 右侧表代表每个基因的外显子数目和长度, 阴影部分的数字表示表 4 统计该外显子翻译成部分 WRKY 保守域的长度。

表 4 不同组核苷酸序列翻译成部分 WRKY 保守结构域的外显子长度统计  
Table 4 Statistics of exon length of the deduced WRKY fragments in different groups

	I	II -a	II -b	II -c		II -d	II -e	III		
外显子长度(bp)	159	117	114	144	147	178	126	141	105	129
百分比(%)	70	66.7	100	18.8	18.8	37.5	100	28.6	30	20

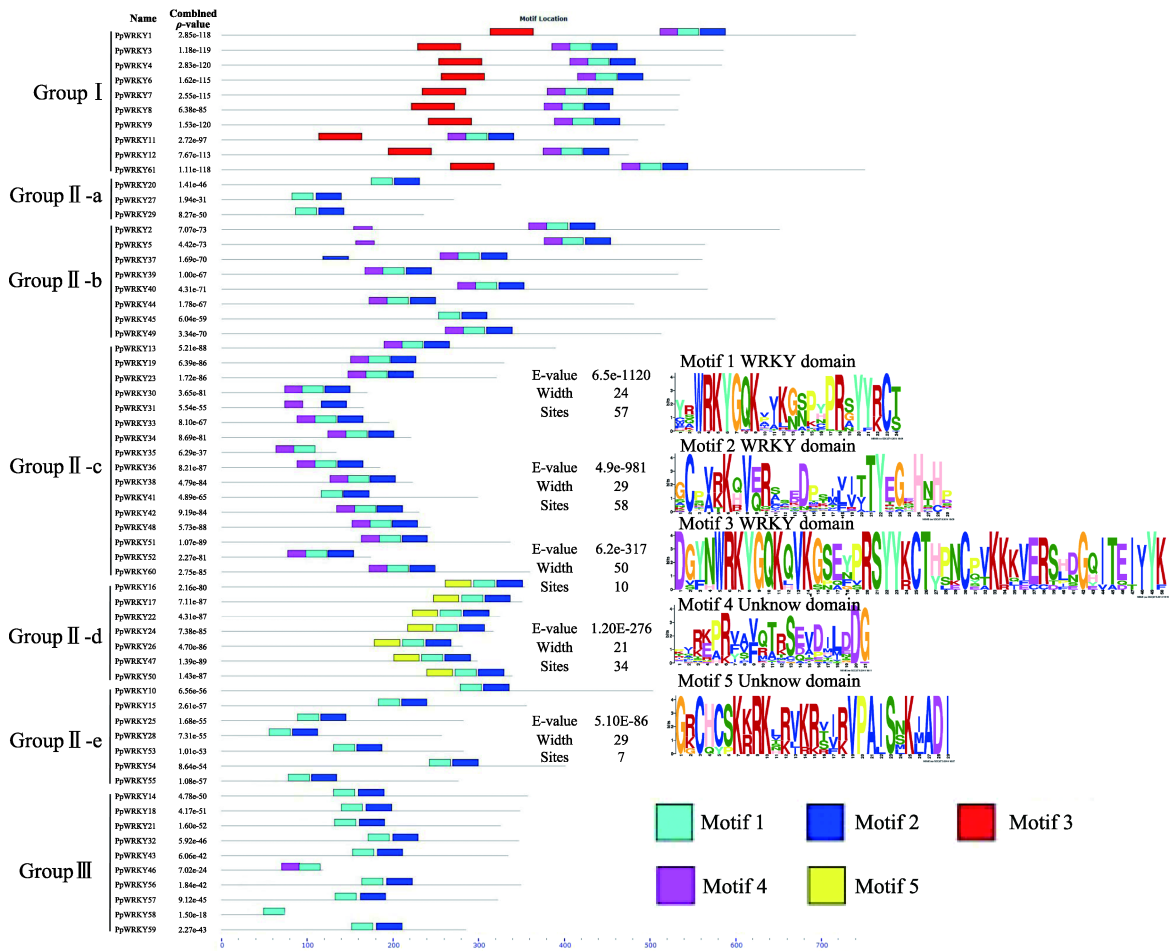


图 5 PpWRKY 基因家族保守元件分布图  
Fig. 5 Distribution of the conserved motifs in the PpWRKY gene family  
使用在线软件 MEME 预测 PpWRKY 蛋白中的保守元件。Motif 1 和 Motif 2 组成 C 端 WRKY 盒；Motif 3 为 N 端 WRKY 盒，Motif 4 和 Motif 5 是未知盒。

### 3 讨 论

作为转录因子家族，WRKY 基因广泛参与植物多种生理过程的调控。大量研究表明，模式植物拟南芥和水稻 WRKY 基因家族的复杂相关特征和生物学功能已经得到了广泛的挖掘及功能鉴定和分析<sup>[1,2]</sup>。并且，随着不同物种基因组测序的不断完成，已经在多个物种中挖掘和鉴定出 WRKY 基因家族。截止目前为止，在拟南芥<sup>[15]</sup>、籼稻<sup>[17]</sup>、粳稻<sup>[17]</sup>、番茄<sup>[18]</sup>、

黄瓜<sup>[19]</sup>、毛果杨<sup>[20]</sup>、苹果<sup>[21]</sup>、棉花<sup>[22]</sup>、橡胶树<sup>[23]</sup>和蒺藜苜蓿<sup>[24]</sup>中分别鉴定出 72、102、98、81、55、104、132、116、81 和 81 个 WRKY 基因(详细信息见表 5)。通过物种间比较发现，不同物种的 WRKY 基因数目与其基因组大小并不成正比例关系。例如，拟南芥基因组大小为 125 Mb，含有 72 个 WRKY 基因；黄瓜和蒺藜苜蓿基因组大小相近，前者含有 55 个 WRKY 基因，后者含有 81 个 WRKY 基因；桃基

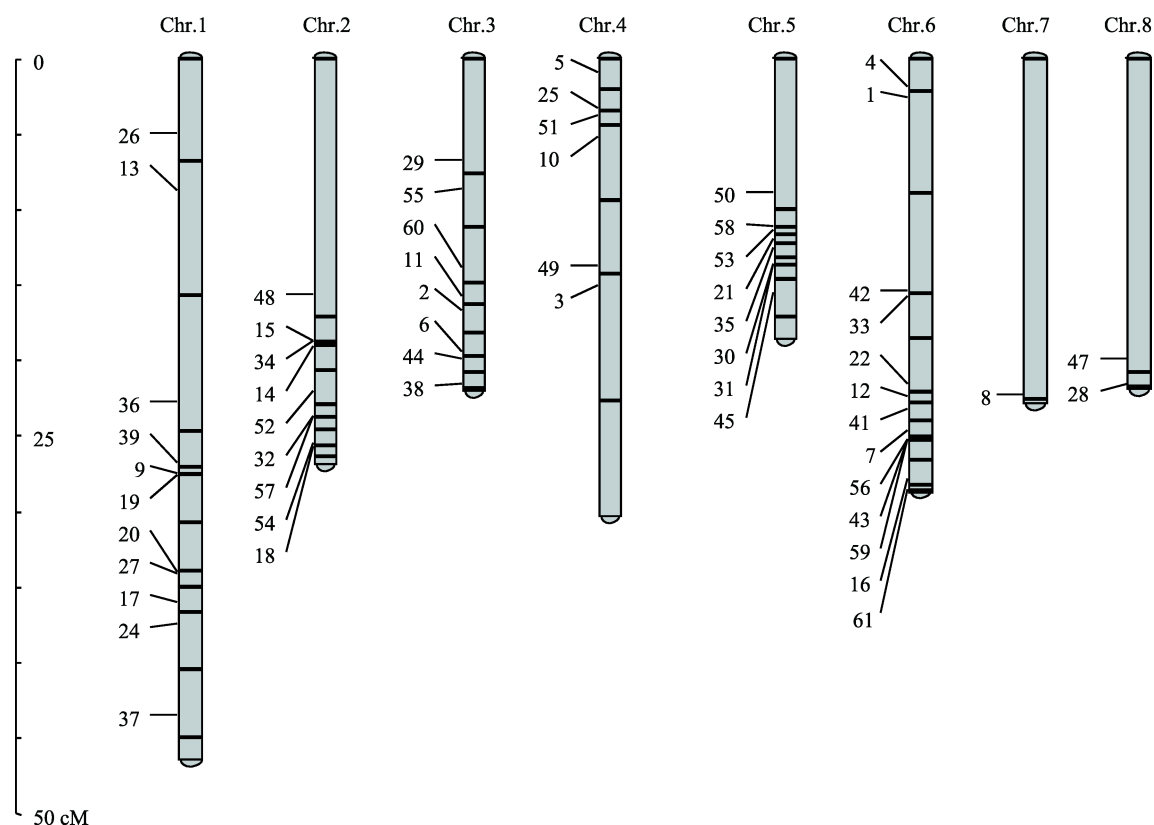


图 6 *PpWRKY* 基因家族染色体定位

Fig. 6 Chromosomal locations of the *PpWRKY* gene family

数字代表该基因的名称(如 1 代表 *PpWRKY1*)标注在染色体的左侧。

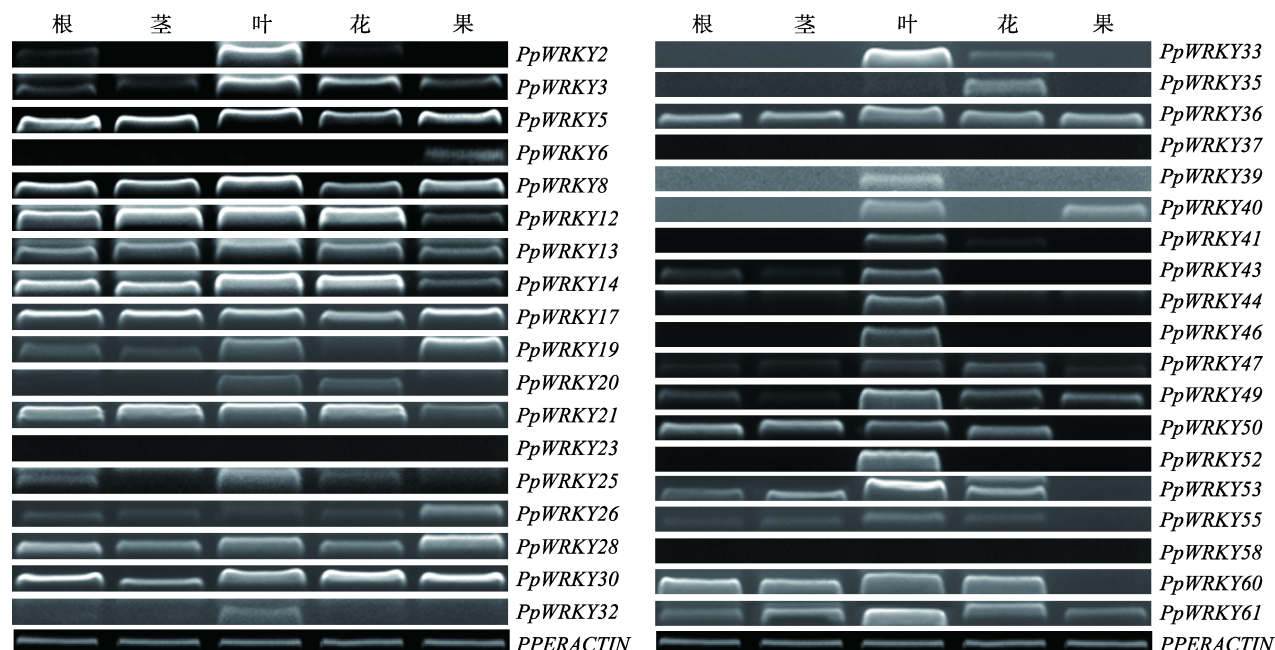
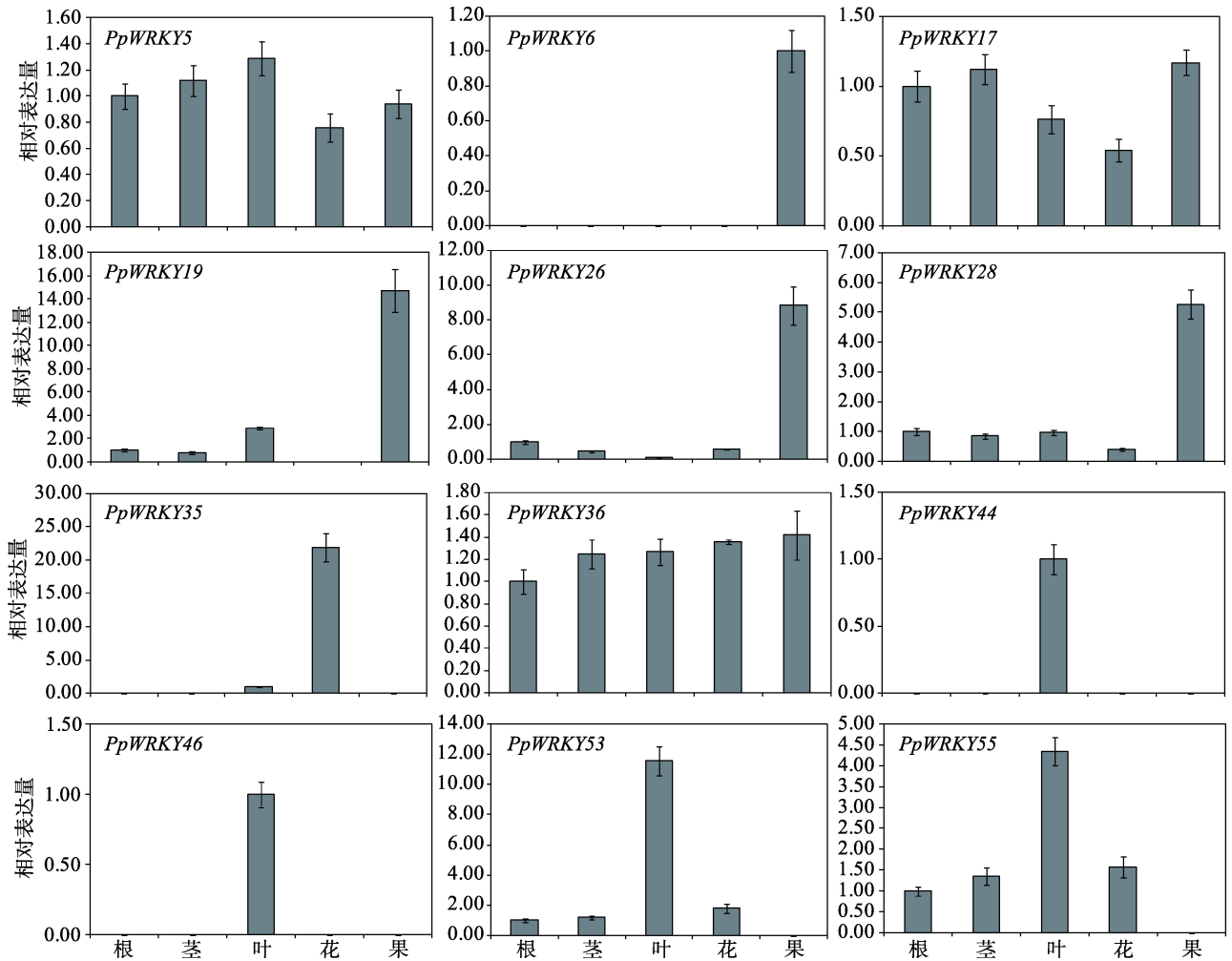


图 7 *PpWRKY* 基因在桃不同组织中的半定量表达分析

Fig. 7 Sqrt-PCR expression analysis of *PpWRKY* genes in different peach tissues

图 8 *PpWRKY* 在桃不同组织中的荧光定量表达分析Fig. 8 qRT-PCR expression analysis of *PpWRKY* genes in different peach tissues

基因组大小是 265 Mb 却仅包含 61 个 *WRKY* 基因。最近研究提出基因复制(Gene duplication)(包括串联复制(Tandem duplication)、片段复制(Segmental duplication)和全基因组复制(Whole-genome duplication))不仅在基因组重排和扩张过程中而且在基因功能的多样化和基因家族的大量产生等过程中都起到重要的作用<sup>[34]</sup>。在水稻<sup>[17]</sup>、番茄<sup>[18]</sup>、毛果杨<sup>[20]</sup>、苹果<sup>[21]</sup>、棉花<sup>[22]</sup>和蒺藜苜蓿<sup>[24]</sup> *WRKY* 基因家族均发现了基因复制事件的发生,但在黄瓜基因组中并没有发现该事件的发生<sup>[19]</sup>。Verde 等<sup>[35]</sup>经过与其他物种对比分析发现,桃基因组大小为 265 Mb,在进化过程中没有像其他物种经历过近期全基因组复制事件,这可能部分解释了桃基因组中 *WRKY* 基因少于其他物种的原因。此外,随着基因组测序精度与组装的提高、搜索和分析软件的升级与更新以及基

因组中存在的可变性剪切,各物种中都有可能发现新的 *WRKY* 成员,例如拟南芥和黄瓜中最初分别鉴定存在 72 个和 55 个 *WRKY* 成员<sup>[15,19]</sup>,而到目前为止已经分别发现了 90 个和 88 个 *WRKY* 成员(<http://plantfdb.cbi.pku.edu.cn/>, 植物转录因子数据库统计数目)。

本研究在桃中鉴定得到 61 个 *WRKY* 成员,这与植物转录因子数据库中关于桃 *WRKY* 成员数目的统计是一致的,但随着桃基因组测序精度的加深和组装的提高,搜索和分析软件的升级与更新以及其基因组中可变性剪切的发现,我们有理由相信桃基因组中还存在新的 *WRKY* 成员。

进化分析研究表明,在植物 *WRKY* 基因家族进化过程中 I 组被认为是 II 和 III 组的原始祖先, *WRKY* 基因家族通过其 I 组 C 末端的 *WRKY* 结构

域的变化以及 N 末端 WRKY 结构域的保持与缺失演

表 5 不同物种 WRKY 基因个数比较

Table 5 Number of WRKY genes in different species

物种	基因组大小	I	II					III	总数
			II -a	II -b	II -c	II -d	II -e		
拟南芥( <i>Arabidopsis thaliana</i> ) <sup>[15]</sup>	125 Mb	13	3	8	19	7	8	14	72
粳稻( <i>Oryza sativa</i> L. ssp. <i>japonica</i> ) <sup>[17]</sup>	466 Mb	13	4	8	17	7	11	38	98
籼稻( <i>Oryza sativa</i> L. ssp. <i>indica</i> ) <sup>[17]</sup>	466 Mb	13	4	8	14	6	10	47	102
番茄( <i>Solanum lycopersicum</i> ) <sup>[18]</sup>	900 Mb	15	5	8	16	6	17	11	81
黄瓜( <i>Cucumis sativus</i> ) <sup>[19]</sup>	350 Mb	10	4	4	16	8	7	6	55
毛果杨( <i>Populus trichocarpa</i> ) <sup>[20]</sup>	480 Mb	50	5	9	13	13	4	10	104
苹果( <i>Malus domestica</i> ) <sup>[21]</sup>	742 Mb	24	8	12	31	14	14	29	132
棉花( <i>Gossypium raimondii</i> ) <sup>[22]</sup>	761.4 Mb	22	10	18	29	15	14	8	116
橡胶树( <i>Hevea brasiliensis</i> ) <sup>[23]</sup>	2.15 Gb	16	6	11	17	9	8	14	81
蒺藜苜蓿( <i>Medicago truncatula</i> ) <sup>[24]</sup>	375 Mb	19	6	8	19	8	8	13	81
桃( <i>Prunus persica</i> )	265 Mb	10	3	8	16	7	7	10	61

变为 II 组和 III 组<sup>[4]</sup>。同时 WRKY 结构域中 WRKYGQK 七肽的变异和锌指结构的变异也表明植物 WRKY 基因家族进化过程的多样性。本研究利用已知物种 WRKY 基因家族的 WRKY 保守域的结构特征对桃基因组中的 WRKY 基因进行分组和鉴定(表 1、图 1 和图 2),同时分析了 WRKY 结构域的结构(图 3),发现桃 WRKY 基因家族结构是高度保守的,WRKY 结构域发生了 WRKYGQK 变异(*PpWRKY33* 和 *PpWRKY35*)和锌指结构变异(*PpWRKY35*、*PpWRKY46* 和 *PpWRKY58*)(表 2)。相似结果也在拟南芥、番茄和橡胶树等多个物种基因组中<sup>[15,17-23]</sup>被发现。以上结果表明:植物 WRKY 基因家族结构上是高度保守的,尽管 WRKY 结构域发生少量变异,但也表明植物 WRKY 基因家族进化过程出现了多样性。

内含子和外显子在基因组的定位模式能为其进化关系提供重要的证据。本研究系统、全面地分析了桃 WRKY 基因家族成员的外显子和内含子的结构分布和长度大小。通过分析其在染色体上的基因结构发现 *PpWRKY* 基因家族成员由 2~6 个外显子构成,大多数翻译成 WRKY 区域的外显子是倒数第 2 个外显子,并且不同亚组编码成保守 WRKY 的外显子长度也不同(图 4,表 3),这是首次发现桃 WRKY 基因家族成员编码 WRKY 结构域的外显子呈现此

规律,对于研究桃和其他物种关于 WRKY 基因家族进化过程提供了有价值的信息。

基因表达模式能够为功能研究提供线索,拟南芥<sup>[15]</sup>、番茄<sup>[18]</sup>、黄瓜<sup>[19]</sup>、棉花<sup>[22]</sup>和橡胶树<sup>[23]</sup>中 WRKY 基因家族成员的组织表达检测都相继完成。本研究检测到 15 个 *PpWRKY* 基因均在根、茎、叶、花和果中表达,但相对表达水平有差异(图 7)。在拟南芥<sup>[15]</sup>、番茄<sup>[18]</sup>、黄瓜<sup>[19]</sup>、棉花<sup>[22]</sup>和橡胶树<sup>[23]</sup>WRKY 基因家族中也发现了相似的结果。因此,植物 WRKY 基因家族参与了植物营养和生殖生长发育等多个方面的调控<sup>[1,2]</sup>。有趣的是,*PpWRKY32*、39、44、46 和 52 只在叶中表达;*PpWRKY6* 只在果实中表达;*PpWRKY35* 只在花中表达(图 7),推测这些基因可能在这些组织中起到重要的作用。

#### 参考文献(References):

- [1] Eulgem T, Rushton PJ, Robatzek S, Somssich IE. The WRKY superfamily of plant transcription factors. *Trends Plant Sci*, 2000, 5(5): 199-206. [DOI]
- [2] Rushton PJ, Somssich IE, Ringler P, Shen QJ. WRKY transcription factors. *Trends Plant Sci*, 2010, 15(5): 247-258. [DOI]
- [3] Sun CX, Palmqvist S, Olsson H, Borén M, Ahlandsberg S, Jansson C. A novel WRKY transcription factor, SUSIBA2, participates in sugar signaling in barley by binding to the



- sugar-responsive elements of the *iso1* promoter. *Plant Cell*, 2003, 15(9): 2076–2092. [DOI]
- [4] Zhang Y, Wang L. The WRKY transcription factor superfamily: its origin in eukaryotes and expansion in plants. *BMC Evol Biol*, 2005, 5(1): 1. [DOI]
- [5] Chen LG, Song Y, Li SJ, Zhang LP, Zou CS, Yu DQ. The role of WRKY transcription factors in plant abiotic stresses. *BBA-Gene Regul Mech*, 2012, 1819(2): 120–128. [DOI]
- [6] Ishiguro S, Nakamura K. Characterization of a cDNA encoding a novel DNA-binding protein, SPF1, that recognizes SP8 sequences in the 5' upstream regions of genes coding for sporamin and  $\beta$ -amylase from sweet potato. *Mol Gen Genet*, 1994, 244(6): 563–571. [DOI]
- [7] Rushton PJ, Macdonald H, Huttly AK, Lazarus CM, Hookey R. Members of a new family of DNA-binding proteins bind to a conserved cis-element in the promoters of  $\alpha$ -Amy2 genes. *Plant Mol Biol*, 1995, 29(4): 691–702. [DOI]
- [8] Rushton PJ, Torres JT, Parniske M, Wernert P, Hahlbrock K, Somssich IE. Interaction of elicitor-induced DNA-binding proteins with elicitor response elements in the promoters of parsley *PR1* genes. *EMBO J*, 1996, 15(20): 5690–5700. [DOI]
- [9] de Pater S, Greco V, Pham K, Memelink J, Kijine J. Characterization of a zinc-dependent transcriptional activator from *Arabidopsis*. *Nucleic Acids Res*, 1996, 24(23): 4624–4631. [DOI]
- [10] Johnson SC, Kolevski B, Smyth DR. *Transparent testa glabra2*, a trichome and seed coat development gene of *Arabidopsis*, encodes a WRKY transcription factor. *Plant Cell*, 2002, 14(6): 1359–1375. [DOI]
- [11] Lagacé M, Matton DP. Characterization of a WRKY transcription factor expressed in late torpedo-stage embryos of *Solanum chacoense*. *Planta*, 2004, 219(1): 185–189. [DOI]
- [12] Robatzek S, Somssich IE. Targets of *AtWRKY6* regulation during plant senescence and pathogen defense. *Genes Dev*, 2002, 16(9): 1139–1149. [DOI]
- [13] Xie Z, Zhang ZL, Zou XL, Yang GX, Komatsu S, Shen QJ. Interactions of two abscisic-acid induced *WRKY* genes in repressing gibberellin signaling in aleurone cells. *Plant J*, 2006, 46(2): 231–242. [DOI]
- [14] Dong JX, Chen CH, Chen ZX. Expression profiles of the *Arabidopsis WRKY* gene superfamily during plant defense response. *Plant Mol Biol*, 2003, 51(1): 21–37. [DOI]
- [15] Ülker B, Somssich IE. WRKY transcription factors: from DNA binding towards biological function. *Curr Opin Plant Biol*, 2004, 7(5): 491–498. [DOI]
- [16] Fu QT, Yu DQ. Expression profiles of *AtWRKY25*, *AtWRKY26* and *AtWRKY33* under abiotic stresses. *Hereditas (Beijing)*, 2010, 32(8): 848–856.
- 付乾堂, 余迪求. 拟南芥 *AtWRKY25*、*AtWRKY26* 和 *AtWRKY33* 在非生物胁迫条件下的表达分析. *遗传*, 2010, 32(8): 848–856. [DOI]
- [17] Ross CA, Liu Y, Shen QJ. The *WRKY* gene family in rice (*Oryza sativa*). *J Integr Plant Biol*, 2007, 49(6): 827–842. [DOI]
- [18] Huang SX, Gao YF, Liu JK, Peng XL, Niu XL, Fei ZJ, Cao SQ, Liu YS. Genome-wide analysis of WRKY transcription factors in *Solanum lycopersicum*. *Mol Genet Genomics*, 2012, 287(6): 495–513. [DOI]
- [19] Ling J, Jiang WJ, Zhang Y, Yu HJ, Mao ZC, Gu XF, Huang SW, Xie BY. Genome-wide analysis of WRKY gene family in *Cucumis sativus*. *BMC Genomics*, 2011, 12(1): 471. [DOI]
- [20] He HS, Dong Q, Shao YH, Jiang HY, Zhu SW, Cheng BJ, Xiang Y. Genome-wide survey and characterization of the *WRKY* gene family in *Populus trichocarpa*. *Plant Cell Rep*, 2012, 31(7): 1199–1217. [DOI]
- [21] Gu YB, Ji ZR, Chi FM, Qiao Z, Xu CN, Zhang JX, Dong QL, Zhou ZS. Bioinformatics and expression analysis of the WRKY gene family in apple. *Sci Agric Sin*, 2015, 48(16): 3221–3238.
- 谷彦冰, 冀志蕊, 迟福梅, 乔壮, 徐成楠, 张俊祥, 董庆龙, 周宗山. 苹果 WRKY 基因家族生物信息学及表达分析. *中国农业科学*, 2015, 48(16): 3221–3238. [DOI]
- [22] Dou LL, Zhang XH, Pang CY, Song MZ, Wei HL, Fan SL, Yu SX. Genome-wide analysis of the WRKY gene family in cotton. *Mol Genet Genomics*, 2014, 289(6): 1103–1121. [DOI]
- [23] Li HL, Guo D, Yang ZP, Tang X, Peng SQ. Genome-wide identification and characterization of *WRKY* gene family in *Hevea brasiliensis*. *Genomics*, 2014, 104(1): 14–23. [DOI]
- [24] Song H, Nan ZB. Genome-wide identification and analysis of WRKY transcription factors in *Medicago truncatula*. *Hereditas (Beijing)*, 2014, 36(2): 152–168.
- 宋辉, 南志标. 蒺藜苜蓿全基因组中 WRKY 转录因子的鉴定和分析. *遗传*, 2014, 36(2): 152–168. [DOI]
- [25] Tian Y, Dong QL, Ji ZR, Chi FM, Cong P, Zhou ZS. Genome-wide identification and analysis of the MADS-box gene family in apple. *Gene*, 2015, 555(2): 277–290. [DOI]
- [26] Dong QL, Wang HR, An M, Yu XM, Wang CJ. Cloning, sequence and expression analysis of NADP-malic enzyme genes in apple. *Sci Agric Sin*, 2013, 46(9): 1857–1866.



- 董庆龙, 王海荣, 安森, 余贤美, 王长君. 苹果 NADP 依赖的苹果酸酶基因克隆、序列和表达分析. 中国农业科学, 2013, 46(9): 1857–1866. [DOI]
- [27] Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol*, 2011, 28(10): 2731–2739. [DOI]
- [28] Dong QL, Ji ZR, Chi FM, Tian Y, An XH, Xu CN, Zhou ZS. Bioinformatics of the MADS-box transcription factor and their expression in different apple tissues. *Sci Agric Sin*, 2014, 47(6): 1151–1161.
- 董庆龙, 冀志蕊, 迟福梅, 田义, 安秀红, 徐成楠, 周宗山. 苹果 MADS-box 转录因子的生物信息学及其在不同组织中的表达. 中国农业科学, 2014, 47(6): 1151–1161. [DOI]
- [29] Tani E, Polidoros AN, Tsaftaris AS. Characterization and expression analysis of *FRUITFULL*- and *SHATTERPROOF*-like genes from peach (*Prunus persica*) and their role in split-pit formation. *Tree Physiol*, 2007, 27(5): 649–659. [DOI]
- [30] Tani E, Polidoros AN, Flemetakis E, Stedel C, Kalloniati C, Demetriou K, Katinakis P, Tsaftaris AS. Characterization and expression analysis of *AGAMOUS*-like, *SEEDSTICK*-like, and *SEPALLATA*-like MADS-box genes in peach (*Prunus persica*) fruit. *Plant Physiol Bioch*, 2009, 47(8): 690–700. [DOI]
- [31] Tong ZG, Gao ZH, Wang F, Zhou J, Zhang Z. Selection of reliable reference genes for gene expression studies in peach using real-time PCR. *BMC Mol Biol*, 2009, 10(1): 71. [DOI]
- [32] Duan MR, Nan J, Liang YH, Mao P, Lu L, Li LF, Wei CH, Lai LH, Li Y, Su XD. DNA binding mechanism revealed by high resolution crystal structure of *Arabidopsis thaliana* WRKY1 protein. *Nucleic Acids Res*, 2007, 35(4): 1145–1154. [DOI]
- [33] Su HY, Zhang SZ, Yuan XW, Chen CT, Wang XF, Hao YJ. Genome-wide analysis and identification of stress-responsive genes of the NAM–ATF1,2–CUC2 transcription factor family in apple. *Plant Physiol Bioch*, 2013, 71: 11–21. [DOI]
- [34] Cannon SB, Mitra A, Baumgarten A, Young ND, May G. The roles of segmental and tandem gene duplication in the evolution of large gene families in *Arabidopsis thaliana*. *BMC Plant Biol*, 2004, 4(1): 10. [DOI]
- [35] Verde I, Abbott AG, Scalabrin S, Jung S, Shu SQ, Marroni F, Zhebentyayeva T, Dettori MT, Grimwood J, Cattonaro F, Zuccolo A, Rossini L, Jenkins J, Vendramin E, Meisel LA, Decroocq V, Sosinski B, Prochnik S, Mitros T, Policriti A, Cipriani G, Dondini L, Ficklin S, Goodstein DM, Xuan PF, del Fabbro C, Aramini V, Copetti D, Gonzalez S, Horner DS, Falchi R, Lucas S, Mica E, Maldonado J, Lazzari B, Bielenberg D, Pirone R, Miculan M, Barakat A, Testolin R, Stella A, Tartarini S, Tonutti P, Arús P, Orellana A, Wells C, Main D, Vizzotto G, Silva H, Salamini F, Schmutz J, Morgante M, Rokhsar DS. The high-quality draft genome of peach (*Prunus persica*) identifies unique patterns of genetic diversity, domestication and genome evolution. *Nat Genet*, 2013, 45(5): 487–494. [DOI]

(责任编辑: 赵方庆)