

周恩强, 周 瑶, 姚梦楠, 等. 基于全长转录组的蚕豆 *WRKY* 基因家族分析及耐盐胁迫相关候选基因挖掘[J]. 江苏农业学报, 2024, 40(1): 14-30.

doi: 10.3969/j.issn.1000-4440.2024.01.002

基于全长转录组的蚕豆 *WRKY* 基因家族分析及耐盐胁迫相关候选基因挖掘

周恩强, 周 瑶, 姚梦楠, 王学军, 赵 娜, 缪亚梅, 王永强, 薛 冬, 李 波, 汪凯华, 顾春燕, 魏利斌

(江苏沿江地区农业科学研究所, 江苏 南通 226012)

摘要: *WRKY* 基因是植物特有的转录因子基因, 能够调控植物的生长发育和胁迫响应。为了鉴定蚕豆 *WRKY* 基因家族成员, 揭示其进化关系并挖掘与盐胁迫相关的候选 *WRKY* 基因, 本研究在完成蚕豆全长转录组测序(9 个样品)和二代转录组测序(27 个样品)的基础上, 利用生物信息学方法对 *WRKY* 转录因子基因进行鉴定与分析, 并通过拟南芥同源基因比对挖掘盐胁迫相关的候选 *VjWRKY* 基因。结果表明, 蚕豆全长转录组测序共获得 53.84 Gb 数据量, 通过比对和校正最终获得 58 885 条转录本序列信息; 基于蚕豆全长转录组共鉴定出 113 个 *WRKY* 家族成员, 氨基酸数目为 153~737 aa, 等电点为 4.84~9.87, 113 个 *WRKY* 家族蛋白质全部定位于细胞核中; 根据拟南芥 *WRKY* 家族系统发育特征, *VjWRKY* 基因家族可分为 3 组, 分别为 group 1(38 个 *VjWRKY*)、group 2(61 个 *VjWRKY*)、group 3(14 个 *VjWRKY*); Motif 1 和 Motif 3 是 *VjWRKY* 基因家族的特征基序, 并对应 *WRKY* 保守结构域, 在进化过程中较为保守; *VjWRKY* 基因家族主要富集在植物 MAPK 信号通路、植物与病原菌相互作用和剪接体 3 个通路中; 通过同源比对, 在蚕豆中发现 14 个盐胁迫相关候选 *WRKY* 基因, 且主要在根中高表达。该研究结果为蚕豆的遗传学研究提供了丰富的参考数据, 也为创制耐盐蚕豆新品种提供了基因信息。

关键词: *WRKY* 基因; 蚕豆; 全长转录组; 盐胁迫; 进化分析

中图分类号: S551⁺.4

文献标识码: A

文章编号: 1000-4440(2024)01-0014-17

Analysis of *WRKY* gene family based on full-length transcriptome and mining of salt stress candidate genes in *Vicia faba*

ZHOU En-qiang, ZHOU Yao, YAO Meng-nan, WANG Xue-jun, ZHAO Na, MIAO Ya-mei, WANG Yong-qiang, XUE Dong, LI Bo, WANG Kai-hua, GU Chun-yan, WEI Li-bin

(Jiangsu Yanjiang Institute of Agricultural Sciences, Nantong 226012, China)

收稿日期: 2023-05-05

基金项目: 江苏省种业振兴揭榜挂帅项目[JBSG(2021)056]; 江苏省农业科技自主创新基金项目[CX(22)2011]; 国家食用豆产业技术体系项目(CARS-08-Z10); 江苏沿江地区农业科学研究所青年科技基金项目[YJ(2022)002]; 江苏沿江地区农业科学研究所博士基金项目[YJBS(2021)001]

作者简介: 周恩强(1993-), 男, 安徽蚌埠人, 硕士, 研究实习员, 主要从事豆类遗传育种研究。(E-mail) zhouenqiang0526@163.com

通讯作者: 魏利斌, (Tel) 15936288927; (E-mail) libinwei2013@aliyun.com

Abstract: The *WRKY* gene is a plant-specific transcription factor gene that regulates plant growth and development and stress response. In order to identify the *WRKY* gene family members of *Vicia faba*, reveal their evolutionary relationship and mine candidate *WRKY* genes related to salt stress, in this study, the *WRKY* transcription factor genes were identified and analyzed by bioinformatics methods based on the full-length transcriptome (nine samples) and second-generation transcriptome sequencing (27 samples) of *Vicia faba*, and the candidate *VjWRKY* genes

related to salt stress were excavated by homologous gene alignment of *Arabidopsis thaliana*. The results showed that 53.84 Gb data were obtained from the full-length transcriptome sequencing of *Vicia faba*, and 58 885 transcript sequence information was finally obtained through alignment and correction. A total of 113 *WRKY* family members were identified based on the full-length transcriptome of *Vicia faba*, with amino acid numbers of 153–737 aa and isoelectric points of 4.84–9.87, all 113 *WRKY* family proteins were localized in the nucleus. According to the phylogenetic characteristics of *Arabidopsis WRKY* family, the *VjWRKY* gene family was divided into three groups: group 1 (38 *VjWRKY*), group 2 (61 *VjWRKY*), and group 3 (14 *VjWRKY*). Motif 1 and Motif 3 were characteristic motifs of the *VjWRKY* gene family and corresponded to the *WRKY* domain, which were relatively conserved during evolution. The *VjWRKY* gene family was mainly enriched in three pathways: plant-MAPK signaling pathway, plant-pathogen interaction pathway and spliceosome pathway. Fourteen candidate *WRKY* genes related to salt stress were found in *Vicia faba* by homology alignment, and they were mainly highly expressed in the roots. The results of this study can provide rich reference data for the genetics research of *Vicia faba*, and also provide gene information for creating new varieties of salt-tolerant *Vicia faba*.

Key words: *WRKY* gene; *Vicia faba*; full-length transcriptome; salt stress; evolutionary analysis

非生物胁迫,包括盐胁迫、干旱胁迫和低温胁迫等,限制了作物的生产力,并影响了植物中许多基因的表达^[1],其中盐胁迫是制约全球农业发展和作物生产的最重要因素之一^[2]。盐胁迫使植物遭受离子毒性、渗透胁迫和氧化胁迫,从而破坏各种细胞和生理过程。为了适应或抵抗盐胁迫,植物进化出几种策略,包括渗透调节(如渗透保护剂积累)、离子平衡(Na^+/K^+ 平衡)和抗氧化(抗氧化酶的积累和活性的增强)^[3]。植物对盐胁迫的适应性也可以通过复杂的信号通路改变许多基因的表达来实现,研究发现 bZIP、WRKY、MYB 等家族的转录因子通过结合特定的顺式作用元件形成一个复杂的调控网络,对植物的盐胁迫反应至关重要^[4-6],这些转录因子过表达通常会增加植物对盐胁迫的适应性。

WRKY 转录因子是植物中最大的转录因子家族之一,通过与保守的 DNA 结合位点 *W-box* 结合来调控基因的转录,在植物的生长发育、生物与非生物胁迫响应中发挥着多种作用^[7]。所有的 WRKY 蛋白由 1 个或 2 个约 60 个氨基酸的 WRKY 保守结构域组成,其 N 端具有高度保守的 WRKYGQK 七肽序列, C 端具有 C2H2 基序或 C2HC 基序的锌指结构^[8]。根据 WRKY 结构域的数量和类锌指基序的特征可将 WRKY 基因家族分为 3 组(I、II 和 III),第 I 组具有 2 个 WRKY 结构域,第 II 组具有 1 个 WRKY 结构域和 C2H2 基序,第 III 组具有 1 个 WRKY 结构域和 C2HC 基序^[9]。目前 WRKY 基因家族已在多种植物中被鉴定出,其中拟南芥鉴定出 72 个 WRKY 基因^[8],水稻中包含 102 个 WRKY 基因^[10],大豆基因组中鉴定出 197 个 WRKY 基因^[11],

油茶中鉴定出 89 个 WRKY 基因^[12],绿豆中鉴定出 79 个 WRKY 基因^[13]。同时研究结果表明,WRKY 转录因子基因在多种植物的盐胁迫响应中发挥着重要作用,例如,在拟南芥中过表达陆地棉 *GhWRKY34* 增强了拟南芥转基因植株的耐盐性^[14];在烟草中过表达 *NbWRKY79* 增强了植株盐胁迫耐受性;在玉米中抑制 *ZmZIP111* 的表达提高了玉米幼苗对盐胁迫的敏感性^[15]。

WRKY 转录因子基因参与植物生长发育的多个方面,在盐胁迫的多种不同响应途径中发挥重要作用,在模式植物拟南芥中,已发现多个受盐胁迫调控的 WRKY 基因。Babitha 等^[16]研究发现共表达 *At-bHLH17* 和 *AtWRKY28* 能够增强拟南芥对氯化钠、甘露醇胁迫和氧化胁迫的耐受性。在盐胁迫和脱落酸(ABA)处理下,相比于野生型, *Atwrky66* 突变体表现出对 ABA 和盐胁迫更高的敏感性^[17]。*Atwrky25* 突变体与 *Atwrky33* 突变体的盐胁迫敏感性与野生型没有太大差异,但过表达 *AtWRKY25* 或 *AtWRKY33* 增加了拟南芥对氯化钠胁迫的耐受性^[18]。研究发现 *JMJ15* 通过调控 *AtWRKY46* 和 *AtWRKY70* 的表达水平来增强其耐盐性,并证明 *AtWRKY46* 和 *AtWRKY70* 在盐胁迫中起负调控作用^[19]。在盐胁迫下, *AtWRKY1* 的表达被诱导,且 *AtWRKY1* 功能缺失导致了突变体拟南芥的盐敏感性增加^[20]。Li 等^[21]利用基因编辑技术(CRISPR/Cas9)对 *AtWRKY3* 和 *AtWRKY4* 进行了编辑,发现突变体对盐和茉莉酸甲酯(MeJA)胁迫的耐受性降低。与野生型相比,过表达 *AtWRKY30* 的拟南芥植株对氧化胁迫和盐胁迫表现出更强的耐受性^[22]。*AtWRKY9* 通过调控 *AtC-*

YP94B3 和 AtCYP86B1 来控制软木脂的沉积,从而增加拟南芥的耐盐性^[23]。Chen 等^[24]通过对 AtWRKY18、AtWRKY40 和 AtWRKY60 进行单突、双突、三突和过表达分析发现,AtWRKY18 和 AtWRKY60 对抑制种子萌发和根系生长有积极作用,同时 2 个 WRKY 基因也提高了拟南芥对盐胁迫和渗透胁迫的敏感性。拟南芥 AtWRKY8 主要在根中表达,并在盐处理后显著上调表达,AtWRKY8 的功能缺失使拟南芥在盐胁迫环境下受到抑制作用,表现出萌发延迟,并且抑制萌发后的生长发育^[25]。在盐胁迫条件下,与野生型相比,Atwrky11 和 Atwrky17 突变体的萌发速度较慢,根系生长受损,表现出盐胁迫敏感性^[26]。

蚕豆 (*Vicia faba* L.) 属于豆科野豌豆属一年生或越年生草本植物^[27],因其较高的营养价值和有效的生物固氮作用在作物生产中占据着重要地位。随着人们生活质量的不断提高,市场对蚕豆的需求呈增加和多样化趋势^[28],由于中国耕地资源有限,粮、菜、油争地矛盾突出,现有的蚕豆种植面积已无法满足蚕豆生产发展的需求,可以通过挖掘基因资源、培育适宜盐碱地种植的蚕豆新品种,从而扩大蚕豆种植面积来满足市场需求。

目前在多种植物中已相继有关于 WRKY 基因家族的报告,而蚕豆作为重要的食用豆类和生物固氮作物,关于其 WRKY 转录因子基因家族的内容却未见报道。本研究拟利用 PacBio 平台的单分子实时测序 (SMRT) 技术对蚕豆的根、茎叶、花、种子 (开花后 20 d、30 d、40 d)、果皮 (开花后 20 d、30 d、40 d) 进行全长转录组测序,基于蚕豆全长转录组,利用生物信息学方法对蚕豆 WRKY 基因家族进行鉴定,并对该家族进行理化性质、亚细胞定位、染色体分布、保守结构域、保守基序、京都基因与基因组百科全书 (KEGG) 通路等分析,以期全面解析遗传信息特征。同时利用拟南芥已知功能基因在蚕豆中进行同源比对,挖掘与盐胁迫相关的候选 VjWRKY 基因,并分析其表达模式,为后期深入分析蚕豆 WRKY 基因响应盐胁迫的分子机制提供参考,同时也为创制耐盐蚕豆新品种提供基因信息。

1 材料与方法

1.1 试验材料

本研究使用的蚕豆品种启豆 2 号种植于江苏沿江地区农业科学研究所种植基地。采集根 (盛花

期)、茎叶 (盛花期)、花 (盛花期)、种子 (开花后 20 d、30 d、40 d)、果皮 (开花后 20 d、30 d、40 d) 等样本保存于液氮中备用。委托北京百迈客生物科技有限公司进行二代转录组 (9 个样本,3 次重复,27 个样品) 和全长转录组 (9 个样本 RNA 等量混合) 测序、检测、分析等。

1.2 全长转录组数据处理

根据条件 *full passes* (全票) ≥ 3 且序列准确性大于 0.9 从原始序列提取 CCS (Circular consensus sequencing read) 序列,并对 CCS 序列进行校正。检测 CCS 序列中是否包含 5' 引物、3' 引物及 poly A 尾,3 个都包含的为全长非嵌合序列 (Full-length no chimera, FLNC)。使用 SMRTLink 软件中的 IsoSeq 模块将全长非嵌合序列中相似的序列聚类到一簇 (Cluster),每个 Cluster 得到一个一致序列 (Consensus isoform),通过 minimap2 将得到的校正后的一致序列与蚕豆 Tiffany 参考基因组 (<https://projects.audk/fabagenome/genomics-data>) 进行序列比对 (设置参数 *-ax splice -uf --secondary=no -C5*),使用 cDNA Cupcake 软件对比对结果去冗余,过滤 *Identity* (一致性) 小于 0.9、*Coverage* (优势度) 小于 0.85 的序列,合并仅 5' 端外显子有差异的比对,最终得到非冗余转录本。

1.3 VjWRKY 基因家族成员的鉴定

本研究以蚕豆全长转录组为研究对象。在 pfam 蛋白质家族数据库 (<http://pfam.xfam.org/>) 中查找获取 WRKY 基因家族保守结构域的序列号 (PF03106),并下载其对应的隐马尔可夫模型文件^[29],利用 TBtools 软件在蚕豆全长转录组蛋白质序列中初步检索具有 WRKY 保守结构域的 WRKY 转录因子。分别从拟南芥 (<https://www.arabidopsis.org/index.jsp>) 和 PlantTFDB v5.0 (<http://planttfdb.gao-lab.org/index.php>) 网站下载拟南芥和蒺藜苜蓿的 WRKY 家族基因序列,并将其比对到蚕豆全长转录组中获取序列相似性最高的基因。将获取的 WRKY 蛋白质序列提交到 NCBI CDD (<https://www.ncbi.nlm.nih.gov/cdd>) 数据库,明确是否含 WRKY 保守结构域,并剔除重复、冗余和注释不完整的序列,保留下来的转录因子即为蚕豆 WRKY 家族成员。

1.4 VjWRKY 基因家族成员理化性质及亚细胞定位分析

利用 TBtools 软件计算蚕豆 WRKY 家族成员编

码蛋白质的氨基酸大小、相对分子质量、等电点、不稳定指数、脂肪系数和亲疏水性。利用在线网站 Cell-PLoc 2.0 (<http://www.csbio.sjtu.edu.cn/bioinf/Cell-PLoc-2/>) 对 *WRKY* 家族成员进行亚细胞定位预测。

1.5 *VfWRKY* 基因家族成员保守结构域及系统进化分析

提交拟南芥和蚕豆 *WRKY* 家族基因蛋白质序列到 MEGA 7.0 软件中进行多重比对,将比对结果使用最大似然法 (Maximum likelihood, ML) 构建进化树。提交进化树参数到网站 Evolview (<https://evolgenius.info/evolview-v2/#login>) 上美化进化树^[30],并根据拟南芥 *WRKY* 基因家族分类方法进行分组^[8]。单独提交蚕豆 *WRKY* 家族基因蛋白质序列到 MEGA 7.0 软件中进行多重比对,使用最大似然法构建进化树,将进化树参数和 *WRKY* 家族基因保守结构域信息输入到 TBtools 软件中进行数据可视化分析^[31]。

1.6 *VfWRKY* 基因家族成员染色体分布及 Motif 鉴定

将 113 个 *WRKY* 基因家族成员的基因编号提交到 TBtools 软件中,即可将每个基因定位到相应的染色体上。使用在线软件 MEME (<https://meme-suite.org/meme/tools/meme>) 预测蚕豆 *WRKY* 基因的保守基序^[32],Motif 最大值设置为 10。下载 Motif 分析文件与蚕豆进化树参数共同提交到 TBtools 软件中进行数据可视化分析。

1.7 *VfWRKY* 基因家族成员 GO 富集和 KEGG 通路分析

利用 BLAST 比对工具,将蚕豆 *WRKY* 家族转录本与基因本体 (Gene Ontology, GO) (<http://www.geneontology.org/>) 和 KEGG (<https://www.kegg.jp/>) 公共数据库进行比对,从而得到 *WRKY* 蛋白功能注释信息,并将该注释信息提交到百迈客云在线分析平台 (<https://international.biocloud.net/zh/software/tools/>) 进行可视化分析。

1.8 蚕豆盐胁迫候选 *WRKY* 基因的挖掘

通过文献查找模式植物拟南芥中与盐胁迫相关的 *WRKY* 基因,下载其蛋白质序列与蚕豆 *WRKY* 蛋白序列进行同源比对,从而获得蚕豆盐胁迫相关候选 *WRKY* 基因。利用全长转录组表达量数据对盐胁迫相关候选 *WRKY* 基因进行表达模式分析。

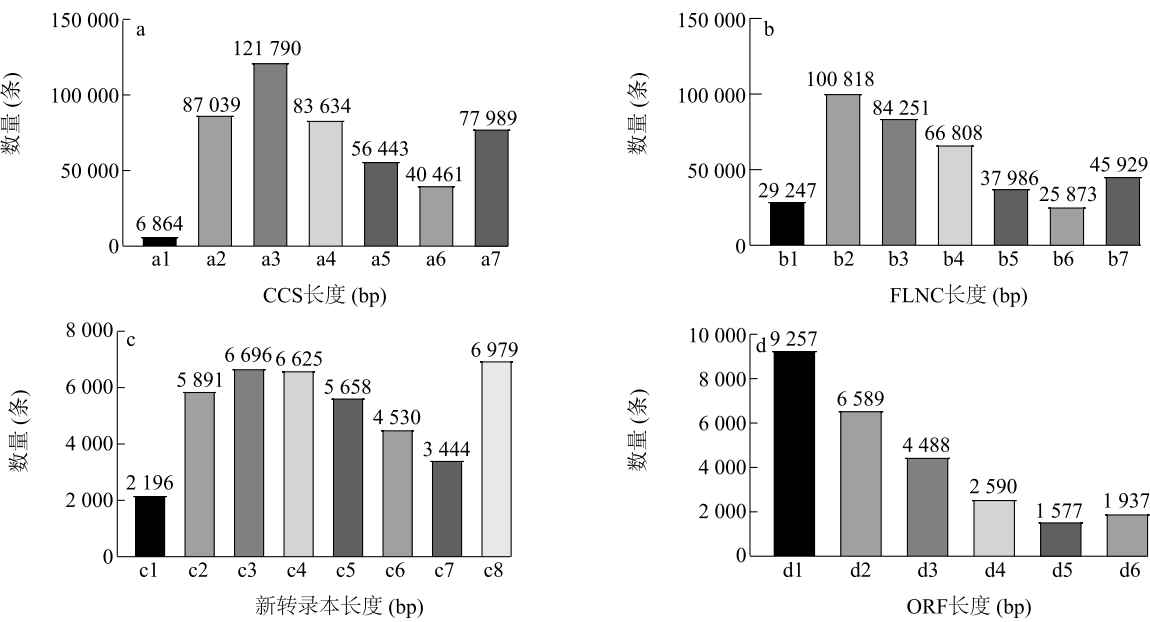
2 结果与分析

2.1 全长转录组数据统计

通过 PacBio 高通量测序技术对蚕豆 2 号的根、茎叶、花、种子、果皮等组织进行全长转录组测序,共获得 53.84 Gb 原始数据。全长转录组测序获得 474 220 条环形一致性序列 (CCS),平均长度 1 936 bp,分布于 1 000~1 499 bp 的 CCS 数量最多,为 121 790 条 (图 1a);全长非嵌合序列 (FLNC) 有 390 912 条,平均长度 1 646 bp,分布于 500~999 bp 的 FLNC 数量最多,为 100 818 条 (图 1b)。对全长非嵌合序列进行聚类分析和去冗余分析,最终得到 58 885 条转录本序列,其中新转录本数量为 42 019 个,平均长度为 2 209 bp,分布于 81~9 800 bp,主要集中在 500~2 499 bp 区间内,1 000~1 499 bp 数量最多,为 6 696 个转录本 (图 1c);我们对新发现的 42 019 个转录本进行序列结构分析,预测出 26 438 条完整开放阅读框 (Open reading frame, ORF) 序列,平均长度为 1 038 bp,分布于 500 bp 以下的 ORF 数量最多,为 9 257 条 (图 1d)。

2.2 *VfWRKY* 转录因子家族的鉴定及理化性质分析

基于蚕豆全长转录组数据,在 TBtools 软件中使用 pfam 程序搜索出 121 个 *WRKY* 转录本,经数据库比对以及去除冗余和缺失序列的转录因子,最终鉴定筛选出 113 个 *WRKY* 转录本,并重新命名为 *VfWRKY1*~*VfWRKY113*,其中相对蚕豆 Tiffany 参考基因组新发现 *WRKY* 转录本 31 个 (以 PB 命名)。*VfWRKY* 转录因子基因家族编码的蛋白质的理化性质见表 1,蚕豆 113 个 *WRKY* 转录因子编码的氨基酸数目最小为 *VfWRKY61* (153 aa),最大为 *VfWRKY36* (737 aa);相对分子质量为 17 729.49~81 727.72;等电点为 4.84~9.87,76 个蛋白质等电点小于 7.00,为酸性蛋白质,37 个蛋白质等电点大于 7.00,为碱性蛋白质;不稳定指数为 29.87~67.85,7 个蛋白质不稳定指数小于 40.00,为稳定蛋白质,106 个蛋白质不稳定指数大于 40.00,为不稳定蛋白质;脂肪系数为 45.60~81.99;平均亲疏水性为 -1.42~-0.40,均小于 0,表明蚕豆 113 个 *WRKY* 家族蛋白质均属于疏水性蛋白质。亚细胞定位预测发现 113 个 *WRKY* 家族蛋白质均定位于细胞核中,间接证明 113 个 *WRKY* 基因作为转录因子调控下游基因的表达。



a: 环形一致性序列 (CCS) 长度分布; b: 全长非嵌合序列 (FLNC) 长度分布; c: 新转录本长度分布; d: 开放阅读框 (ORF) 长度分布。a1: <500; a2: 500~999; a3: 1 000~1 499; a4: 1 500~1 999; a5: 2 000~2 499; a6: 2 500~2 999; a7: ≥3 000; b1: <500; b2: 500~999; b3: 1 000~1 499; b4: 1 500~1 999; b5: 2 000~2 499; b6: 2 500~2 999; b7: ≥3 000; c1: <500; c2: 500~999; c3: 1 000~1 499; c4: 1 500~1 999; c5: 2 000~2 499; c6: 2 500~2 999; c7: 3 000~3 499; c8: ≥3 500; d1: <500; d2: 500~999; d3: 1 000~1 499; d4: 1 500~1 999; d5: 2 000~2 499; d6: ≥2 500。

图 1 全长转录组序列长度分布
Fig.1 Length distribution of full-length transcriptome sequence

表 1 蚕豆 WRKY 转录因子家族成员信息

Table 1 Information of the *Vicia faba* WRKY transcription factor family members

转录本名称	WRKY 家族	氨基酸数目 (aa)	相对分子质量	等电点	不稳定指数	脂肪系数	亲疏水性
PB.1060.1	VfWRKY1	317	34 931.35	6.60	57.85	50.35	-1.04
PB.1060.2	VfWRKY2	317	34 977.44	6.60	58.91	50.35	-1.03
PB.1060.3	VfWRKY3	317	34 931.35	6.60	57.85	50.35	-1.04
PB.1061.1	VfWRKY4	317	35 025.40	6.27	58.82	51.58	-1.04
PB.1061.2	VfWRKY5	317	35 025.40	6.27	58.82	51.58	-1.04
PB.1621.2	VfWRKY6	390	42 829.02	6.47	48.29	54.23	-0.94
PB.17567.1	VfWRKY7	631	70 396.91	5.54	50.53	58.24	-0.79
PB.18329.1	VfWRKY8	444	49 604.33	8.54	45.73	57.77	-0.91
PB.18329.2	VfWRKY9	361	40 747.20	8.45	47.02	52.94	-1.09
PB.360.1	VfWRKY10	404	45 386.87	8.13	53.56	48.49	-1.17
PB.360.3	VfWRKY11	403	45 335.83	8.12	54.15	48.61	-1.18
PB.4114.2	VfWRKY12	516	56 989.22	6.80	65.21	57.27	-0.83
PB.4114.3	VfWRKY13	516	56 989.22	6.80	65.21	57.27	-0.83
PB.6637.1	VfWRKY14	686	74 934.92	5.99	51.26	53.79	-0.92
PB.6637.2	VfWRKY15	686	74 934.92	5.99	51.26	53.79	-0.92
PB.7187.1	VfWRKY16	325	36 060.09	5.94	40.32	63.51	-0.72
PB.7187.3	VfWRKY17	441	48 405.86	5.97	42.61	61.66	-0.71
PB.7187.4	VfWRKY18	441	48 405.86	5.97	42.61	61.66	-0.71
PB.7187.5	VfWRKY19	338	37 552.68	5.88	38.40	61.66	-0.74
PB.7187.7	VfWRKY20	406	44 687.17	6.40	42.29	53.30	-0.95
PB.7187.8	VfWRKY21	479	52 738.18	6.47	40.94	52.90	-0.86

续表 1 Continued 1

转录本名称	<i>WRKY</i> 家族	氨基酸数目 (aa)	相对分子质量	等电点	不稳定指数	脂肪系数	亲疏水性
<i>Vfaba. Tiffany. R1. Ig006440.1</i>	VfWRKY22	489	54 068.12	7.74	66.50	57.44	-0.83
<i>Vfaba. Tiffany. R1. Ig183400.1</i>	VfWRKY23	598	66 667.84	6.41	56.79	51.52	-1.05
<i>Vfaba. Tiffany. R1. Ig231040.1</i>	VfWRKY24	317	34 931.35	6.60	57.85	50.35	-1.04
<i>Vfaba. Tiffany. R1. Ig231080.1</i>	VfWRKY25	317	35 025.40	6.27	58.82	51.58	-1.04
<i>Vfaba. Tiffany. R1. Ig264640.1</i>	VfWRKY26	427	46 931.60	6.70	50.26	55.25	-0.95
<i>Vfaba. Tiffany. R1. Ig002400.1</i>	VfWRKY27	686	74 952.95	5.99	51.44	53.22	-0.92
<i>Vfaba. Tiffany. R1. Ig002400.2</i>	VfWRKY28	678	74 056.94	5.99	51.66	52.70	-0.92
<i>Vfaba. Tiffany. R1. Ig038280.1</i>	VfWRKY29	582	63 629.36	6.27	44.32	54.95	-0.81
<i>Vfaba. Tiffany. R1. Ig038280.2</i>	VfWRKY30	583	63 716.44	6.27	44.37	54.85	-0.81
<i>Vfaba. Tiffany. R1. Ig038280.3</i>	VfWRKY31	569	62 178.82	6.24	44.78	56.03	-0.80
<i>Vfaba. Tiffany. R1. Ig048240.1</i>	VfWRKY32	514	58 020.84	6.77	50.61	51.50	-1.00
<i>Vfaba. Tiffany. R1. Ig202000.1</i>	VfWRKY33	459	52 598.98	6.30	54.44	53.68	-0.99
<i>Vfaba. Tiffany. R1. Ig173640.1</i>	VfWRKY34	507	54 791.53	9.07	54.82	64.22	-0.79
<i>Vfaba. Tiffany. R1. Ig046440.1</i>	VfWRKY35	530	59 277.82	7.05	56.55	49.09	-0.94
<i>Vfaba. Tiffany. R1. Ig061160.1</i>	VfWRKY36	737	81 727.72	5.61	52.15	60.46	-0.75
<i>Vfaba. Tiffany. R1. Ig108040.1</i>	VfWRKY37	444	49 618.36	8.54	46.91	58.00	-0.91
<i>Vfaba. Tiffany. R1. Ig067040.1</i>	VfWRKY38	378	42 098.80	6.97	56.30	60.58	-0.88
<i>PB.1060.4</i>	VfWRKY39	197	21 517.52	5.37	65.67	47.46	-0.98
<i>PB.1621.3</i>	VfWRKY40	188	20 992.09	7.07	40.73	58.46	-1.02
<i>PB.1645.2</i>	VfWRKY41	292	32 283.55	6.51	60.64	45.72	-0.94
<i>PB.17182.1</i>	VfWRKY42	280	31 761.40	6.07	47.47	67.18	-0.84
<i>PB.17567.2</i>	VfWRKY43	184	21 381.46	9.05	67.58	45.60	-1.42
<i>PB.17567.3</i>	VfWRKY44	225	25 297.93	5.02	38.77	54.53	-0.79
<i>PB.3367.1</i>	VfWRKY45	557	59 900.40	7.02	44.37	65.57	-0.59
<i>PB.4114.4</i>	VfWRKY46	241	27 121.66	5.63	67.02	53.36	-1.15
<i>Vfaba. Tiffany. R1. Ig038080.1</i>	VfWRKY47	426	47 364.16	5.23	48.41	58.43	-0.86
<i>Vfaba. Tiffany. R1. Ig039960.1</i>	VfWRKY48	510	56 211.90	5.20	45.60	61.06	-0.75
<i>Vfaba. Tiffany. R1. Ig048400.1</i>	VfWRKY49	273	29 889.25	5.62	67.85	59.23	-0.68
<i>Vfaba. Tiffany. R1. Ig137720.1</i>	VfWRKY50	235	26 882.37	9.25	43.76	64.64	-0.70
<i>Vfaba. Tiffany. R1. Ig217520.1</i>	VfWRKY51	321	37 119.50	6.80	58.81	46.23	-1.24
<i>Vfaba. Tiffany. R1. Ig243320.1</i>	VfWRKY52	430	48 438.38	6.84	51.24	72.35	-0.62
<i>Vfaba. Tiffany. R1. Ig265920.1</i>	VfWRKY53	292	32 283.55	6.51	60.64	45.72	-0.94
<i>Vfaba. Tiffany. R1. Ig300800.1</i>	VfWRKY54	281	31 221.67	9.39	49.34	74.95	-0.47
<i>Vfaba. Tiffany. R1. Ig354360.1</i>	VfWRKY55	389	42 711.65	6.50	52.19	51.13	-0.87
<i>Vfaba. Tiffany. R1. Ig357280.1</i>	VfWRKY56	509	56 540.89	6.14	56.59	56.72	-0.87
<i>Vfaba. Tiffany. R1. Ig365120.1</i>	VfWRKY57	354	40 464.19	7.21	48.52	63.53	-0.79
<i>Vfaba. Tiffany. R1. Ig369080.1</i>	VfWRKY58	523	56 235.22	6.97	43.70	64.42	-0.63
<i>Vfaba. Tiffany. R1. Ig381640.1</i>	VfWRKY59	214	23 399.96	5.07	54.93	64.21	-0.64
<i>Vfaba. Tiffany. R1. Ig115760.1</i>	VfWRKY60	212	23 976.10	4.84	47.44	56.93	-0.91
<i>Vfaba. Tiffany. R1. Ig141120.1</i>	VfWRKY61	153	17 729.49	4.97	29.87	47.06	-1.19
<i>Vfaba. Tiffany. R1. Ig145680.1</i>	VfWRKY62	254	28 678.23	8.76	47.93	68.62	-0.72
<i>Vfaba. Tiffany. R1. Ig145680.2</i>	VfWRKY63	257	28 976.62	8.88	47.16	69.34	-0.71
<i>Vfaba. Tiffany. R1. Ig153440.1</i>	VfWRKY64	317	35 362.78	9.30	53.58	60.91	-0.70
<i>Vfaba. Tiffany. R1. Ig178360.1</i>	VfWRKY65	413	47 138.74	8.74	47.77	62.76	-0.83
<i>Vfaba. Tiffany. R1. Ig211440.1</i>	VfWRKY66	312	34 624.84	8.31	47.65	64.68	-0.79
<i>Vfaba. Tiffany. R1. Ig051240.1</i>	VfWRKY67	359	40 764.10	9.83	49.16	57.02	-0.93
<i>Vfaba. Tiffany. R1. Ig064720.1</i>	VfWRKY68	545	59 171.88	4.90	51.88	52.57	-0.77
<i>Vfaba. Tiffany. R1. Ig078880.1</i>	VfWRKY69	361	40 569.08	6.02	54.12	62.55	-0.75

续表 1 Continued 1

转录本名称	WRKY 家族	氨基酸数目 (aa)	相对分子质量	等电点	不稳定指数	脂肪系数	亲疏水性
<i>Vfaba. Tiffany. R1.3g078920.1</i>	VfWRKY70	328	36 690.76	6.37	59.53	59.94	-0.73
<i>Vfaba. Tiffany. R1.3g106120.1</i>	VfWRKY71	296	33 177.29	5.47	52.26	47.43	-0.89
<i>Vfaba. Tiffany. R1.3g164760.1</i>	VfWRKY72	339	36 817.86	9.62	49.08	68.20	-0.56
<i>Vfaba. Tiffany. R1.3g171760.1</i>	VfWRKY73	242	27 808.50	8.67	42.56	71.28	-0.79
<i>Vfaba. Tiffany. R1.3g171800.1</i>	VfWRKY74	279	31 686.95	8.34	44.84	68.82	-0.67
<i>Vfaba. Tiffany. R1.4g003440.1</i>	VfWRKY75	228	25 772.33	7.61	49.71	49.52	-0.91
<i>Vfaba. Tiffany. R1.4g035440.1</i>	VfWRKY76	322	35 099.52	5.63	57.47	47.92	-0.79
<i>Vfaba. Tiffany. R1.4g097560.1</i>	VfWRKY77	330	35 997.69	9.75	47.57	65.67	-0.61
<i>Vfaba. Tiffany. R1.4g097560.2</i>	VfWRKY78	331	36 084.77	9.75	47.05	65.47	-0.61
<i>Vfaba. Tiffany. R1.4g122440.1</i>	VfWRKY79	613	67 328.29	6.12	52.66	59.35	-0.81
<i>Vfaba. Tiffany. R1.4g149920.1</i>	VfWRKY80	421	46 768.05	5.71	45.47	66.48	-0.61
<i>Vfaba. Tiffany. R1.4g153840.1</i>	VfWRKY81	367	40 534.09	8.91	46.68	77.30	-0.40
<i>Vfaba. Tiffany. R1.4g162960.1</i>	VfWRKY82	325	35 589.23	9.87	39.74	60.62	-0.65
<i>Vfaba. Tiffany. R1.4g186920.1</i>	VfWRKY83	498	54 491.21	5.75	39.57	68.59	-0.72
<i>Vfaba. Tiffany. R1.5g003240.1</i>	VfWRKY84	316	35 487.82	9.05	46.11	56.49	-0.90
<i>Vfaba. Tiffany. R1.5g015400.1</i>	VfWRKY85	164	18 829.15	9.27	40.31	58.17	-0.96
<i>Vfaba. Tiffany. R1.5g037520.1</i>	VfWRKY86	273	30 579.86	6.06	48.61	68.17	-0.79
<i>Vfaba. Tiffany. R1.5g059520.1</i>	VfWRKY87	207	23 854.91	8.59	44.35	68.21	-0.77
<i>Vfaba. Tiffany. R1.5g059640.1</i>	VfWRKY88	334	36 743.50	5.75	55.19	54.01	-0.80
<i>Vfaba. Tiffany. R1.5g132400.1</i>	VfWRKY89	292	33 013.79	7.18	52.23	55.34	-0.81
<i>Vfaba. Tiffany. R1.5g134880.1</i>	VfWRKY90	429	46 752.42	6.96	51.54	71.47	-0.52
<i>Vfaba. Tiffany. R1.5g134880.2</i>	VfWRKY91	424	46 416.17	7.67	47.58	72.78	-0.53
<i>Vfaba. Tiffany. R1.6g010360.1</i>	VfWRKY92	522	57 181.40	5.66	43.96	69.16	-0.68
<i>Vfaba. Tiffany. R1.6g074800.1</i>	VfWRKY93	318	35 744.19	6.67	55.06	50.85	-0.98
<i>Vfaba. Tiffany. R1.6g076880.1</i>	VfWRKY94	478	53 435.88	5.67	49.01	63.70	-0.86
<i>Vfaba. Tiffany. R1.6g104600.1</i>	VfWRKY95	235	26 500.86	6.02	65.87	59.74	-0.77
<i>Vfaba. Tiffany. R1.6g104600.2</i>	VfWRKY96	236	26 587.94	6.02	65.07	59.49	-0.77
<i>Vfaba. Tiffany. R1.6g116360.1</i>	VfWRKY97	185	21 439.62	6.54	41.57	54.70	-0.99
<i>Vfaba. Tiffany. R1.6g128160.1</i>	VfWRKY98	154	18 040.36	9.41	60.22	63.83	-0.94
<i>Vfaba. Tiffany. R1.6g134000.1</i>	VfWRKY99	276	31 795.78	5.47	36.87	67.83	-0.76
<i>Vfaba. Tiffany. R1.1g177040.1</i>	VfWRKY100	216	24 555.57	5.57	47.49	69.03	-0.68
<i>PB.14032.1</i>	VfWRKY101	193	22 013.06	9.36	42.66	62.12	-0.88
<i>PB.241.1</i>	VfWRKY102	211	23 563.33	4.97	46.05	67.44	-0.53
<i>Vfaba. Tiffany. R1.1g060160.1</i>	VfWRKY103	296	34 114.71	7.08	35.38	81.99	-0.56
<i>Vfaba. Tiffany. R1.1g060200.1</i>	VfWRKY104	303	34 661.46	8.44	54.66	81.35	-0.58
<i>Vfaba. Tiffany. R1.1g177080.1</i>	VfWRKY105	334	37 295.85	5.96	55.03	61.05	-0.65
<i>Vfaba. Tiffany. R1.1g177080.2</i>	VfWRKY106	333	37 208.77	5.96	55.57	61.23	-0.65
<i>Vfaba. Tiffany. R1.2g102760.1</i>	VfWRKY107	336	37 683.01	6.12	48.18	74.29	-0.55
<i>Vfaba. Tiffany. R1.2g102760.2</i>	VfWRKY108	330	36 998.21	6.00	48.76	73.27	-0.56
<i>Vfaba. Tiffany. R1.4g024720.1</i>	VfWRKY109	269	31 157.57	6.21	45.99	58.66	-0.78
<i>Vfaba. Tiffany. R1.4g037840.1</i>	VfWRKY110	296	33 989.97	5.37	52.47	59.90	-0.85
<i>Vfaba. Tiffany. R1.5g049280.1</i>	VfWRKY111	299	33 822.77	5.87	57.43	72.31	-0.71
<i>Vfaba. Tiffany. R1.5g049320.1</i>	VfWRKY112	299	34 006.76	5.59	60.05	70.37	-0.77
<i>Vfaba. Tiffany. R1.6g136040.1</i>	VfWRKY113	377	43 013.67	5.51	51.05	58.17	-0.81

2.3 VfWRKY 家族成员染色体分布及系统进化分析

为更好地了解蚕豆 WRKY 基因家族系统发育关

系,我们利用最大似然法构建了 71 个拟南芥 WRKY

蛋白和 113 个蚕豆 WRKY 蛋白的系统发育树,并根

据 Eulgem 等^[8]的分类方法将蚕豆 *WRKY* 家族成员分为 group 1~group 3, group 1 具有 2 个 WRKY 结构域, group 2 具有 1 个 WRKY 结构域和 C2H2 基序, group 3 具有 1 个 WRKY 结构域和 C2HC 基序。系统发育树分析结果显示, group 1、group 2、group 3 分别包含 38 个 (*VfWRKY1*~*VfWRKY38*)、61 个 (*VfWRKY39*~*VfWRKY99*) 和 14 个 (*VfWRKY100*~*Vf*

WRKY113) *VfWRKY* 基因, group 3 包含的基因数目最少, 且与拟南芥 13 个 *WRKY* 基因聚在同一亚家族中。同时我们发现除 group 1 中 *VfWRKY16*~*VfWRKY19* 与 group 2 中 *VfWRKY* 基因聚在同一分支上外, 其他 group 1 和 group 2 *WRKY* 家族成员与拟南芥该亚家族成员都能够很好地聚在同一分支上(图 2)。

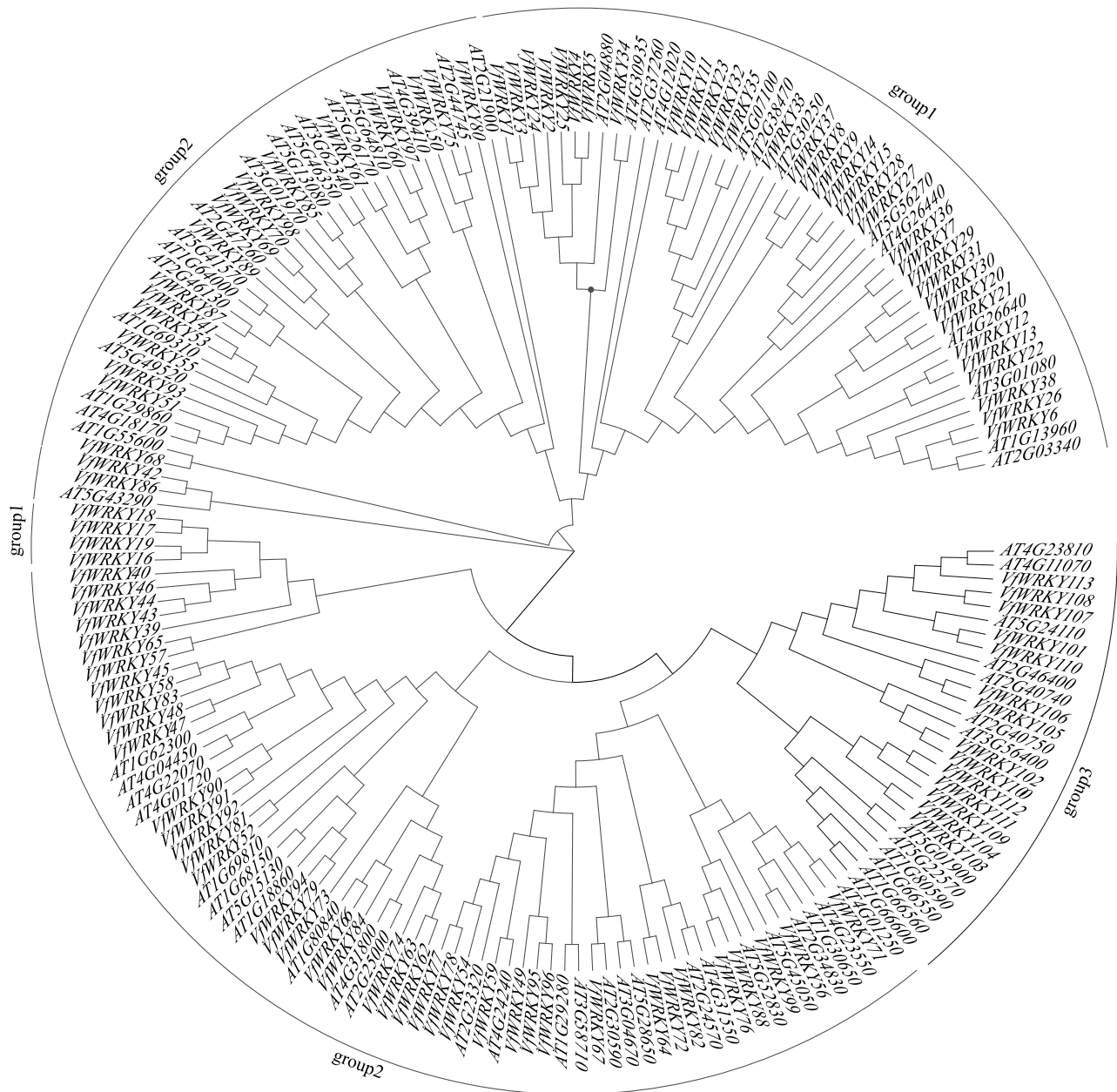


图 2 蚕豆和拟南芥 *WRKY* 基因系统发育分析

Fig.2 Phylogenetic analysis of *WRKY* genes in *Vicia faba* and *Arabidopsis thaliana*

我们将 *VjWRKY16*~*VjWRKY19* 与 *VjWRKY29* (5 个基因同为 *Vfaba. Tiffany. R1.2g038280* 的转录本) 进行多序列比对发现 5 个基因的第二个 WRKY 结构域都为 WRKYGQK-C2-HDH (完整), 但 *VfWRKY16*、*VfWRKY17*、*VfWRKY18* 和 *VfWRKY19* 的第二个 WRKY 结构域只有 WRKYGQK-(*VfWRKY29* 的第二个 WRKY 结构域具有完整的 WRKYGQK-C2-H2 结构), 所以这 4 个基因聚集在 group 2 分支是由于第二个结构域缺少 C2-H2 序列(图 3)。

<i>VfWRKY16</i>	NPSGNITGLQASQVDGRGSGLSVAADKVSDDGYNWRKYGQKLVKGCFFPRSY	159
<i>VfWRKY17</i>	FTSAAPVTTVCYNTNAVDGKSSFFFKPHNRSMVPANFDNHACEKSTQID	159
<i>VfWRKY18</i>	FTSAAPVTTVCYNTNAVDGKSSFFFKPHNRSMVPANFDNHACEKSTQID	159
<i>VfWRKY19</i>	NPSGNITGLQASQVDGRGSGLSVAADKVSDDGYNWRKYGQKLVKGCFFPRSY	159
<i>VfWRKY29</i>	FTSAAPVTTVCYNTNAVDGKSSFFFKPHNRSMVPANFDNHACEKSTQID	159
<i>VfWRKY16</i>	IKCTHNCVKKLFERSHIGQITEIVYKGTIDHPKPQSRYSGGNLSMVQED	212
<i>VfWRKY17</i>	GQKVKQTFVSLPLVKSEMAVPSDEISLSSSLQMTSGASAHVEVELDESNSP	212
<i>VfWRKY18</i>	GQKVKQTFVSLPLVKSEMAVPSDEISLSSSLQMTSGASAHVEVELDESNSP	212
<i>VfWRKY19</i>	IKCTHNCVKKLFERSHIGQITEIVYKGTIDHPKPQSRYSGGNLSMVQED	212
<i>VfWRKY29</i>	GQKVKQTFVSLPLVKSEMAVPSDEISLSSSLQMTSGASAHVEVELDESNSP	212
<i>VfWRKY16</i>	RSLTSRDAERDSTPELSPIATNDGSPGAGFLSNRNDEADEDDPFSSKRRKMD	265
<i>VfWRKY17</i>	GNITGLQASQVDGRGSGLSVAADKVSDDGYNWRKYGQKLVKGCFFPRSYK	265
<i>VfWRKY18</i>	GNITGLQASQVDGRGSGLSVAADKVSDDGYNWRKYGQKLVKGCFFPRSYK	265
<i>VfWRKY19</i>	RSLTSRDDKGFNNYQMSHAERDSTPELSPIATNDGSPGAGFLSNRNDEA	265
<i>VfWRKY29</i>	GNITGLQASQVDGRGSGLSVAADKVSDDGYNWRKYGQKLVKGCFFPRSYK	265
<i>VfWRKY16</i>	LDITPVVKPIREPRVVQTLSEVDILDDGYRWRKYGQKLVKGCFFPRSYK	318
<i>VfWRKY17</i>	THPNCVKKLFERSHIGQITEIVYKGTIDHPKPQSRYSGGNLSMVQEDRSL	318
<i>VfWRKY18</i>	THPNCVKKLFERSHIGQITEIVYKGTIDHPKPQSRYSGGNLSMVQEDRSL	318
<i>VfWRKY19</i>	DEDDPFSSKRRKMDLITPVVKPIREPRVVQTLSEVDILDDGYRWRKYGQKLV	318
<i>VfWRKY29</i>	THPNCVKKLFERSHIGQITEIVYKGTIDHPKPQSRYSGGNLSMVQEDRSL	318
<i>VfWRKY16</i>	TIQYKLI*	325
<i>VfWRKY17</i>	TSRDDKGFNNYQMSHAERDSTPELSPIATNDGSPGAGFLSNRNDEADE	371
<i>VfWRKY18</i>	TSRDDKGFNNYQMSHAERDSTPELSPIATNDGSPGAGFLSNRNDEADE	371
<i>VfWRKY19</i>	KGNNPRVVIITMTIQYKLI	338
<i>VfWRKY29</i>	ISRDDKGFNNYQMSHAERDSTPELSPIATNDGSPGAGFLSNRNDEADE	371
<i>VfWRKY16</i>		-
<i>VfWRKY17</i>	DPFSKRRKMDLITPVVKPIREPRVVQTLSEVDILDDGYRWRKYGQKLVKGC	424
<i>VfWRKY18</i>	DPFSKRRKMDLITPVVKPIREPRVVQTLSEVDILDDGYRWRKYGQKLVKGC	424
<i>VfWRKY19</i>		-
<i>VfWRKY29</i>	DPFSKRRKMDLITPVVKPIREPRVVQTLSEVDILDDGYRWRKYGQKLVKGC	424
<i>VfWRKY16</i>		-
<i>VfWRKY17</i>	ENPRVLTMTIQYKLI	441
<i>VfWRKY18</i>	ENPRVLTMTIQYKLI	441
<i>VfWRKY19</i>		-
<i>VfWRKY29</i>	ENPRSYKCTNAGCFFVKRIVERASHDPRKAVITTYEGKHNIDVPTARNNSHDM	477

标记处表示 WRKY 结构域氨基酸序列。

图 3 WRKY 基因编码的蛋白质序列比对

Fig.3 Sequences comparison of proteins encoded by WRKY genes

由于所使用的软件或数据本身具有局限性,导致所选参考基因组注释往往不够精确,这样就有必要对原有注释的基因结构进行优化。我们通过 gff-compare 软件与蚕豆 *Tiffany* 参考基因组注释进行比较,将已知基因和转录本保留原身份标识号(ID),将新基因和新转录本保留 PB 格式的 ID 添加到注释中,最终我们得到包含新基因和新转录本的优化参考基因组注释文件。利用优化参考基因组注释文件提取位置信息绘制 WRKY 家族成员在染色体上的物理分布(图 4)。蚕豆由 12 条染色体组成,每两条

染色体为 1 对,共 6 对,其中染色体 1 包含染色体 1L 和染色体 1S。113 个 WRKY 转录本不均匀地分布在 7 条染色体上,染色体 1 拥有的 WRKY 家族成员最多,为 38 个,其中染色体 1L 有 29 个,染色体 1S 有 9 个;染色体 2 拥有 WRKY 家族成员 24 个,染色体 5 拥有 WRKY 家族成员 19 个,染色体 4 拥有 WRKY 家族成员 13 个,染色体 6 拥有 WRKY 家族成员 11 个,染色体 3 拥有 WRKY 家族成员数量最少,为 8 个。

2.4 VfWRKY 家族保守结构域和保守基序鉴定

113 个 VfWRKY 蛋白包含 3 种保守结构域,分别为 WRKY、Plant_zn_clust、PTZ00265。其中 *VfWRKY54*、*VfWRKY64*、*VfWRKY67*、*VfWRKY72*、*VfWRKY77*、*VfWRKY78* 和 *VfWRKY82* 含有 WRKY 和 Plant_zn_clust 结构域,并且 7 个基因聚在同一分支上;*VfWRKY14*、*VfWRKY15*、*VfWRKY27* 和 *VfWRKY28* 含有 WRKY 和 PTZ00265 结构域,其余 102 WRKY 基因都只包含 WRKY 结构域(图 5)。同时我们发现 group 2 和 group 3 中的 61 个和 14 个 WRKY 转录本都只含有 1 个 WRKY 保守结构域,并且分别聚集在同一分支上;group 1 的 38 个 WRKY 转录本都含有 2 个 WRKY 保守结构域,但 *VfWRKY16*、*VfWRKY17*、*VfWRKY18* 和 *VfWRKY19* 的第二个 WRKY 结构域长度不完整(图 5)。

在一个基因家族中,基因的保守基序可以反映该基因家族成员之间的进化关系和功能。我们利用在线网站 MEME 分析了 VfWRKY 基因家族的保守基序,并使用 TBtools 进行了可视化分析(图 6)。VfWRKY 基因家族共含有 10 个 Motif,其中不同的 WRKY 基因含有 Motif 的数量为 2~10 个,*VfWRKY39* 和 *VfWRKY44* 只含有 2 个 Motif,*VfWRKY12*、*VfWRKY13*、*VfWRKY22*、*VfWRKY29*、*VfWRKY30* 和 *VfWRKY31* 含有的 Motif 最多,为 10 个。通过序列查询发现 Motif 1 和 Motif 3 与 WRKY 结构域相对应,group 2 和 group 3 中有 71 个 WRKY 转录本含有 Motif 1,只有 *VfWRKY39*、*VfWRKY43*、*VfWRKY44* 和 *VfWRKY46* 含有 Motif 3;group 1 中 37 个基因同时含有 Motif 1 和 Motif 3,只有 *VfWRKY34* 含有 2 个 Motif 1。每个 WRKY 亚家族都具有相对稳定的 Motif,其中 group 1 亚家族有 33 个基因(在 group 1 中占比 86.8%)同时含有 Motif 3、Motif 6、Motif 7、Motif 1、Motif 4、Motif 2 和 Motif 5;group 2 亚家族被分为 2 个

分支,分支 1 有 29 个基因(在 group 2 中占比 47.5%)同时含有 Motif 1、Motif 4、Motif 2 和 Motif 5,分支 2 有 27 个基因(在 group 2 中占比 44.3%)同时含有 Motif 7、Motif 1、Motif 4、Motif 2 和 Motif 5;group

3 亚家族所有基因都含有 Motif 1、Motif 4 和 Motif 2。*VjWRKY16*、*VjWRKY17*、*VjWRKY18* 和 *VjWRKY19* 的 Motif 数量与保守结构域长度相对应,相对 group 1 亚家族成员缺少了 Motif 4、Motif 2 和 Motif 5(图 6)。

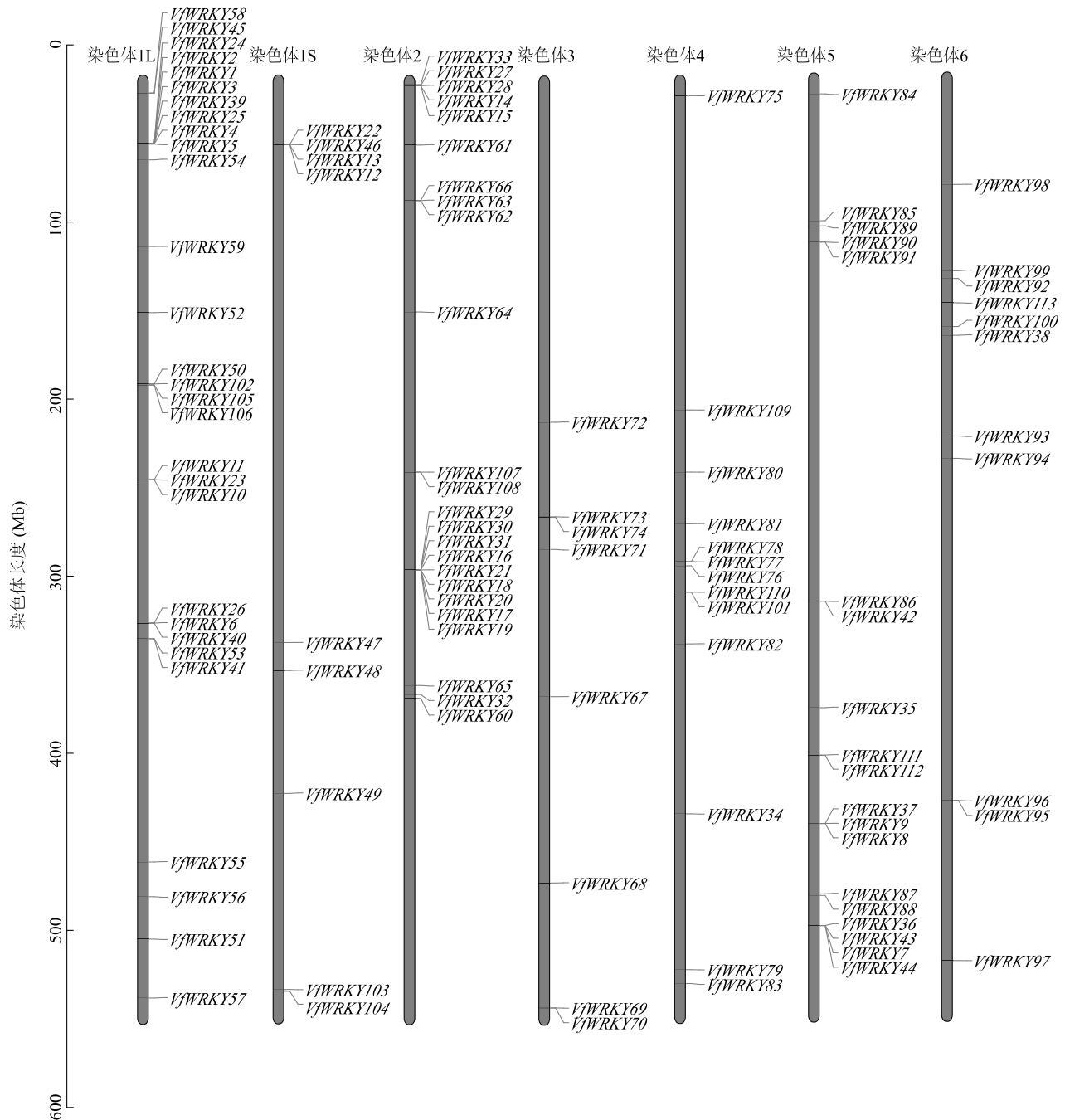


图 4 蚕豆 WRKY 基因在染色体上的分布

Fig.4 Distribution of *Vicia faba* WRKY genes on chromosomes

2.5 *VjWRKY* 家族成员 GO 功能注释及 KEGG 通路分析

将 *VjWRKY* 家族的 113 个 WRKY 转录本与 GO 数

据库进行比对分析,结果显示,113 个 WRKY 转录本都得到了注释(图 7a)。按照 GO 功能分类方式,将 113 个 WRKY 转录本分为生物过程 (Biological

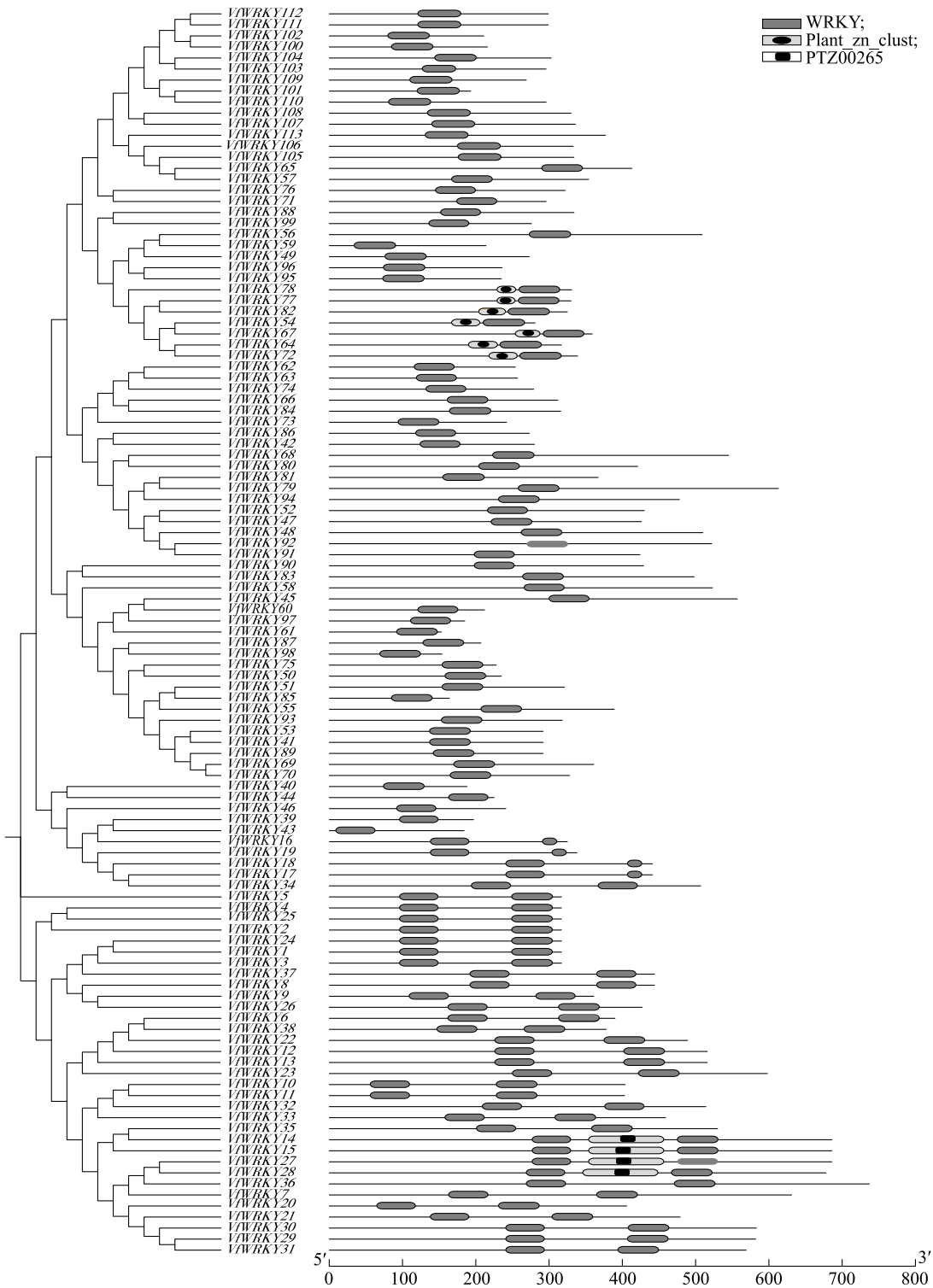


图 5 蚕豆 WRKY 家族成员保守结构域分析
Fig.5 Conserved domain analysis of WRKY family members in *Vicia faba*

process)、细胞组分 (Cellular component) 和分子功能 (Molecular function) 3 大类。3 个大类可细分为 20 个二级分类, 其中生物过程包含的二级分类最多, 为 16

个不同的亚类, 主要包括转录调控 (DNA 和 RNA 模板转录)、代谢过程的调控 (大分子、细胞、化合物和 RNA 等)、生物合成过程的调控 (细胞、大分子和 RNA

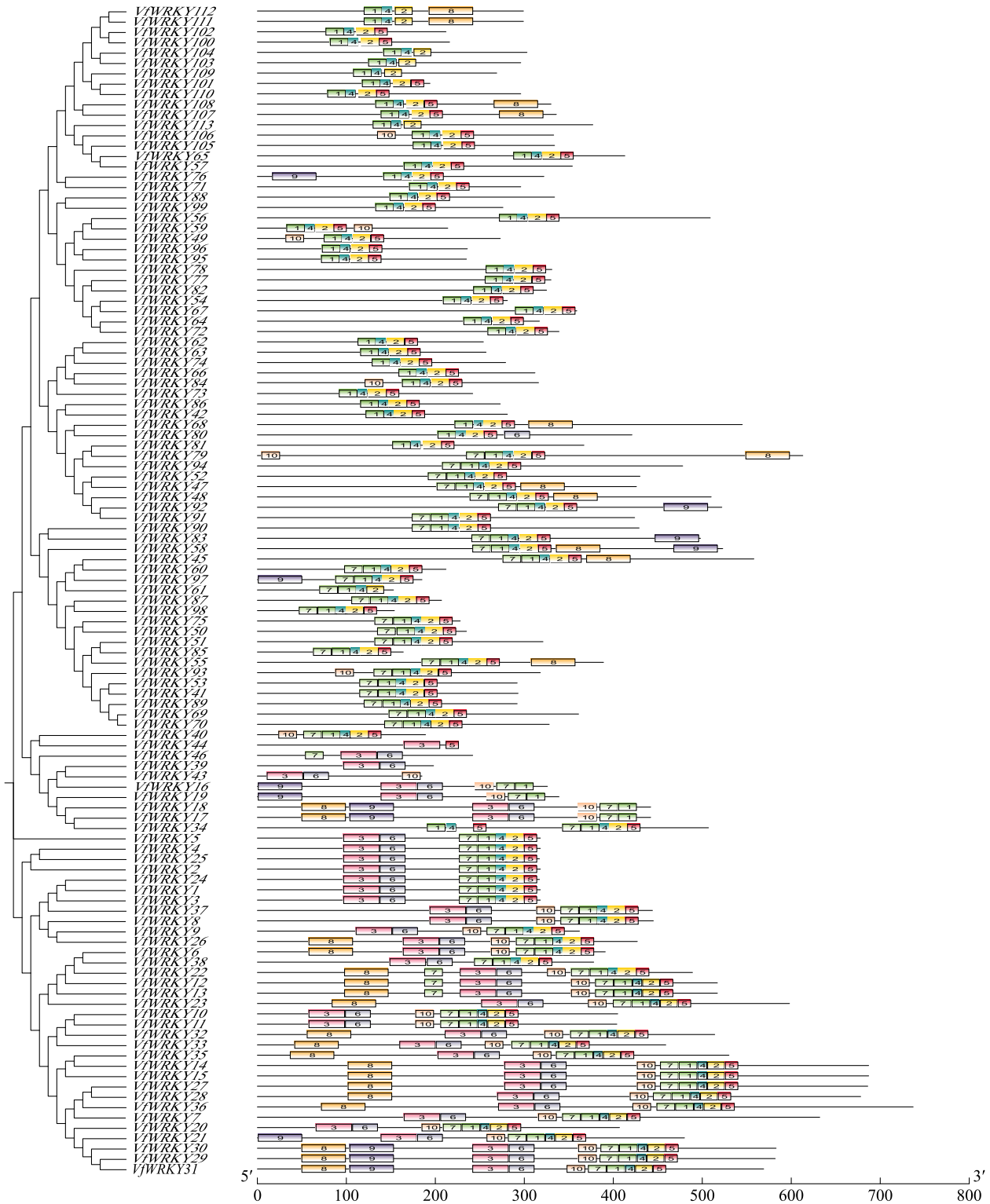


图 6 蚕豆 *WRKY* 家族成员保守基序分析

Fig.6 Conserved motif analysis of *WRKY* family members in *Vicia faba*

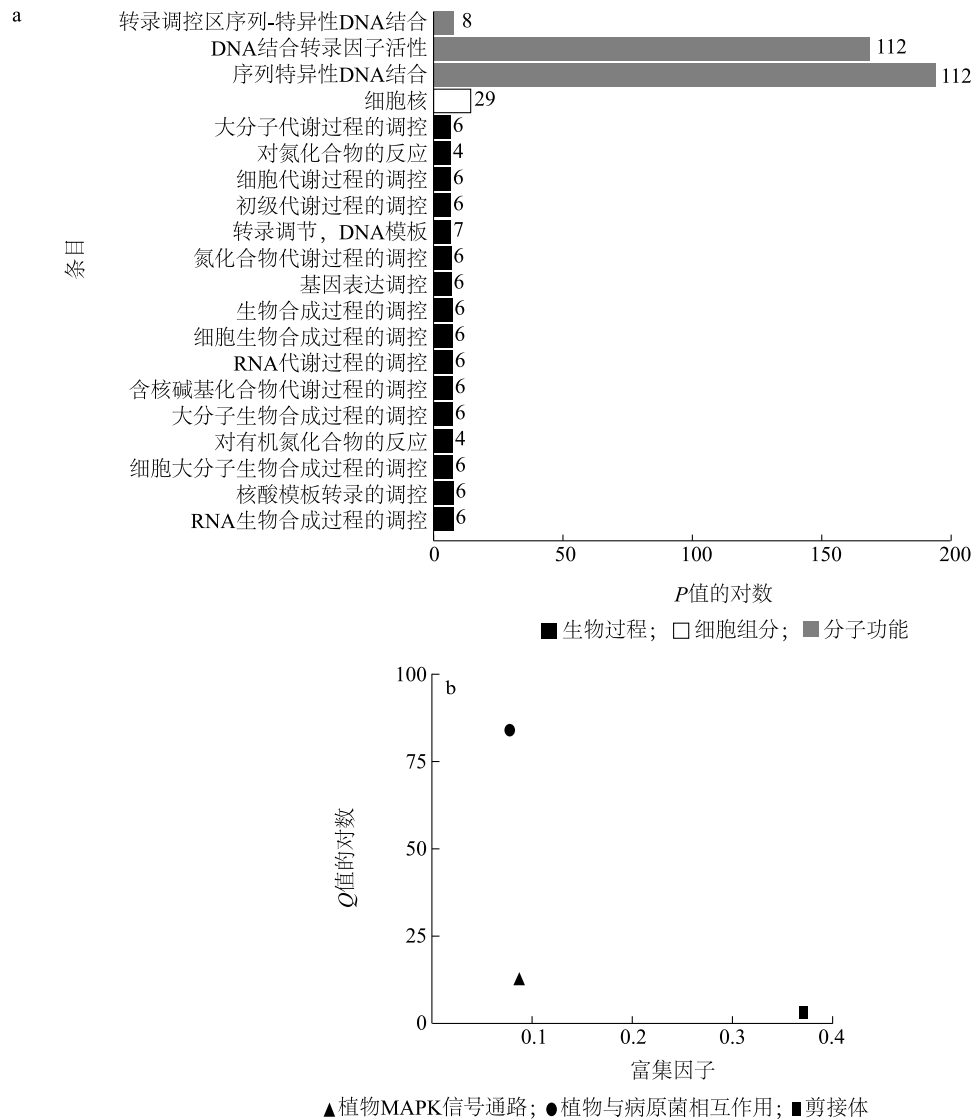
等);分子功能注释基因占比最多,为 112 个 *WRKY* 转录本,主要包括 DNA 结合转录因子活性和序列特异

性 DNA 结合 2 个功能;细胞组分包含的二级分类最少,主要为细胞核(29 个转录本)。

在 KEGG 数据库对 *VfWRKY* 家族的 113 个 *WRKY* 转录本进行通路富集分析,结果显示有 98 个 *VfWRKY* 转录本得到 KEGG 通路富集(图 7b)。*VfWRKY* 基因主要富集在植物 MAPK 信号通路、植物与病原菌相互作用和剪接体 3 个通路中,其中植物与病原菌相互作用通路 *WRKY* 基因最多,为 84 个(占比 85.71%),其次是植物 MAPK 信号通路,为 40 个(占比 40.82%),剪接体通路最少,为 13 个(占比 13.27%)。

为了探究蚕豆 *WRKY* 家族成员表达模式,我们利

用 *WRKY* 基因在根、茎叶、花、种子(开花后 20 d、30 d、40 d)、果皮(开花后 20 d、30 d、40 d)中的表达信息,构建了 *WRKY* 基因表达图谱(图 8)。分析表达图谱可知,*WRKY* 基因家族在 9 个样本中的表达模式可分为 5 组(I、II、III、IV、V):I 组主要在果皮中高表达;II 组主要在种子中高表达;III 组在 9 个样本中低表达或不表达;IV 组主要在根中表达且在其他组织中低表达或不表达,并且在根中高表达的 *WRKY* 基因占比较高(46.0%);V 组在多个组织中均有表达。



P 值、*Q* 值表示差异显著性检验指标。图 a 中,柱长度对应 *P* 值的对数,柱后的数字表示基因数。

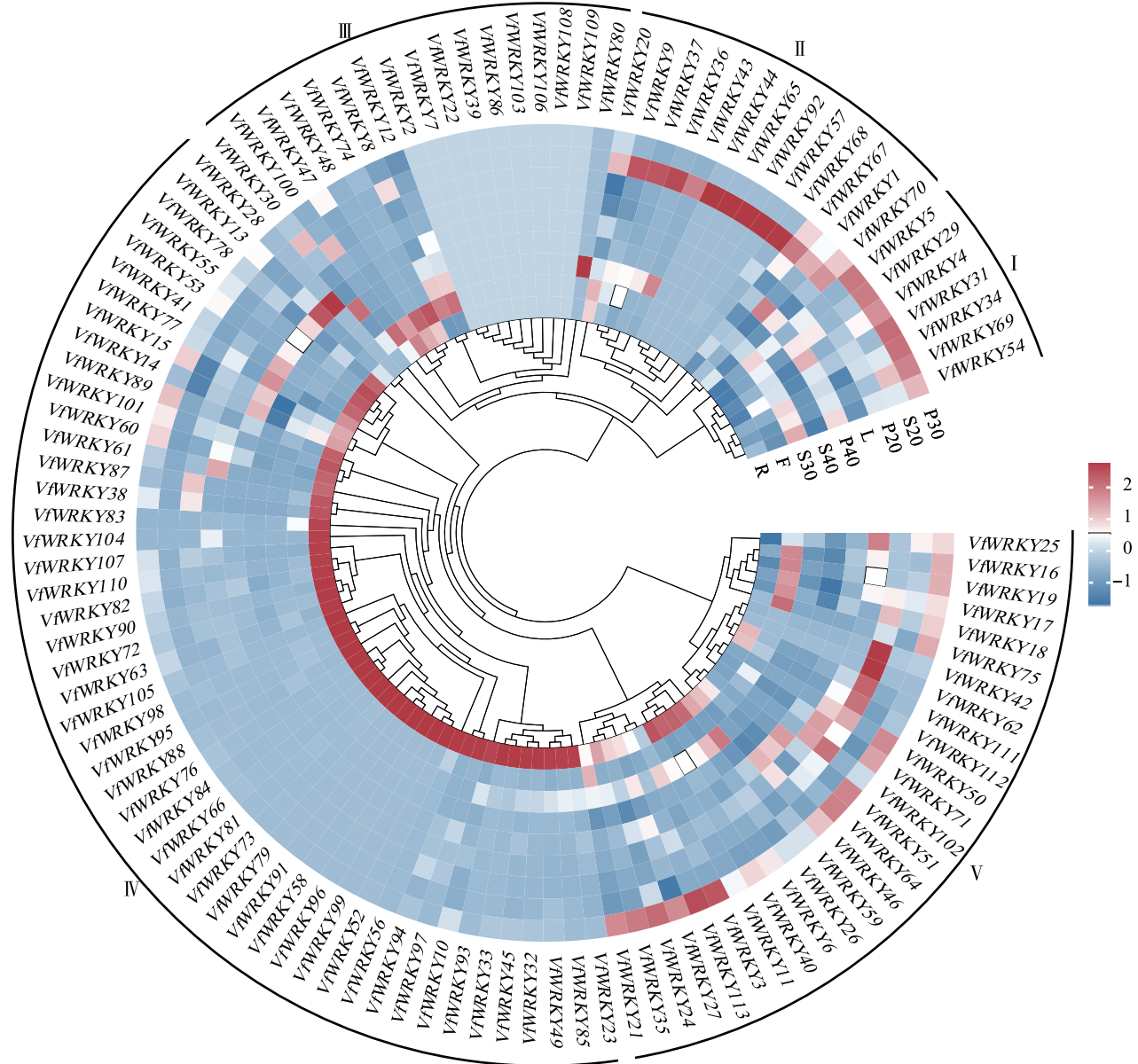
图 7 蚕豆 *WRKY* 家族成员 GO 富集(a)及 KEGG 通路分析(b)

Fig.7 GO enrichment (a) and KEGG pathway analysis (b) of *WRKY* family members in *Vicia faba*

2.6 蚕豆盐胁迫候选 *WRKY* 基因的挖掘

通过查阅已有的拟南芥 *WRKY* 基因家族报道,

我们发现 *AtWRKY1*、*AtWRKY3*、*AtWRKY4*、*AtWRKY8*、*AtWRKY9* 等 16 个基因与盐胁迫相关,其中 12 个基



R;根;L;茎叶;F;花;S20、S30、S40 表示开花后 20 d、30 d、40 d 的种子;P20、P30、P40 表示开花后 20 d、30 d、40 d 的果皮。

图 8 蚕豆 WRKY 基因表达模式分析

Fig.8 Expression pattern analysis of WRKY family members in *Vicia faba*

因为耐盐基因, *AtWRKY18*、*AtWRKY46*、*AtWRKY60*、*AtWRKY70* 为盐敏感基因(表 2)。将拟南芥中与盐胁迫相关的 16 个 WRKY 基因比对到蚕豆 WRKY 基因家族中,在蚕豆中发现了 14 个候选同源基因,分别为 *Vfaba. Tiffany. R1. 1g183400*、*Vfaba. Tiffany. R1. 2g038280*、*Vfaba. Tiffany. R1. 3g164760*、*Vfaba. Tiffany. R1. 2g048240*、*Vfaba. Tiffany. R1. 4g037840*、*Vfaba. Tiffany. R1. 6g074800*、*Vfaba. Tiffany. R1. 3g078920*、*Vfaba. Tiffany. R1. 1g060200*、*Vfaba. Tiffany. R1. 2g211440*、

Vfaba. Tiffany. R1. 5g003240、*Vfaba. Tiffany. R1. 1g006440*、*Vfaba. Tiffany. R1. 2g102760*、*Vfaba. Tiffany. R1. 6g076880* 和 *Vfaba. Tiffany. R1. 5g049280*,其中 *AtWRKY3* 和 *AtWRKY4* 共同比对到 *Vfaba. Tiffany. R1. 1g006440*,*AtWRKY11* 和 *AtWRKY17* 共同比对到 *Vfaba. Tiffany. R1. 3g164760*,且 *AtWRKY3*、*AtWRKY4*、*AtWRKY11* 和 *AtWRKY17* 都为耐盐基因。

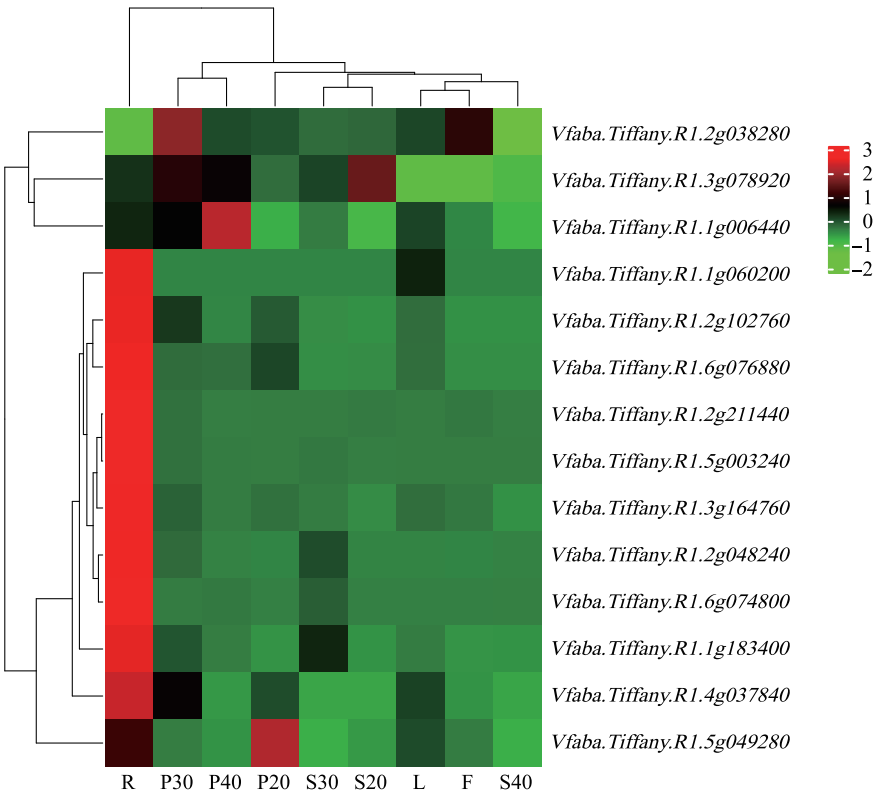
我们对启豆 2 号的根、茎叶、花、种子(开花后 20 d、30 d、40 d)、果皮(开花后 20 d、30 d、40 d)的

27 个样品进行了转录组分析,利用转录组数据对蚕豆 14 个盐胁迫相关候选 *WRKY* 基因进行了表达模式分析(图 9)。 *Vfaba. Tiffany. R1. 1g183400*、*Vfaba. Tiffany. R1. 3g164760*、*Vfaba. Tiffany. R1. 2g048240*、*Vfaba. Tiffany. R1. 4g037840*、*Vfaba. Tiffany. R1. 6g074800*、*Vfaba. Tiffany. R1. 2g211440*、*Vfaba. Tiffany. R1. 5g003240*、*Vfaba. Tiffany. R1. 2g102760*、*Vfaba. Tiffany. R1. 6g076880* 主要在根中高表达,在其他组织的表达量较低或不表达。 *Vfaba. Tiffany. R1. 1g006440* 在多个组织中均有表达,并在开花后 40 d 果皮中表达量最高; *Vfaba. Tiffany. R1. 1g060200* 在多个组织中表达量较低或不表达; *Vfaba. Tiffany. R1. 5g049280* 在开花后 20 d 果皮中表达量最高。 *Vfaba. Tiffany. R1. 1g183400*、*Vfaba. Tiffany. R1. 2g038280*、*Vfaba. Tiffany. R1. 2g048240*、*Vfaba. Tiffany. R1. 2g102760*、*Vfaba. Tiffany. R1. 2g211440*、*Vfaba. Tiffany. R1. 3g078920*、*Vfaba. Tiffany. R1. 3g164760*、*Vfaba. Tiffany. R1. 4g037840* 和 *Vfaba. Tiffany. R1. 5g003240* 在果皮中的表达量均先升高后下降(图 9)。

表 2 蚕豆盐胁迫候选 *WRKY* 基因

Table 2 Candidate *WRKY* genes of *Vicia faba* under salt stress

拟南芥耐盐基因	功能	蚕豆盐胁迫候选基因	参考文献
<i>AtWRKY33</i>	耐盐基因	<i>Vfaba. Tiffany. R1. 1g183400</i>	[18]
<i>AtWRKY1</i>	耐盐基因	<i>Vfaba. Tiffany. R1. 2g038280</i>	[20]
<i>AtWRKY11</i>	耐盐基因	<i>Vfaba. Tiffany. R1. 3g164760</i>	[26]
<i>AtWRKY17</i>	耐盐基因	<i>Vfaba. Tiffany. R1. 3g164760</i>	[26]
<i>AtWRKY25</i>	耐盐基因	<i>Vfaba. Tiffany. R1. 2g048240</i>	[18]
<i>AtWRKY46</i>	盐敏感基因	<i>Vfaba. Tiffany. R1. 4g037840</i>	[19]
<i>AtWRKY28</i>	耐盐基因	<i>Vfaba. Tiffany. R1. 6g074800</i>	[16]
<i>AtWRKY8</i>	耐盐基因	<i>Vfaba. Tiffany. R1. 3g078920</i>	[25]
<i>AtWRKY66</i>	耐盐基因	<i>Vfaba. Tiffany. R1. 1g060200</i>	[17]
<i>AtWRKY18</i>	盐敏感基因	<i>Vfaba. Tiffany. R1. 2g211440</i>	[24]
<i>AtWRKY60</i>	盐敏感基因	<i>Vfaba. Tiffany. R1. 5g003240</i>	[24]
<i>AtWRKY3</i>	耐盐基因	<i>Vfaba. Tiffany. R1. 1g006440</i>	[21]
<i>AtWRKY4</i>	耐盐基因	<i>Vfaba. Tiffany. R1. 1g006440</i>	[21]
<i>AtWRKY30</i>	耐盐基因	<i>Vfaba. Tiffany. R1. 2g102760</i>	[22]
<i>AtWRKY9</i>	耐盐基因	<i>Vfaba. Tiffany. R1. 6g076880</i>	[23]
<i>AtWRKY70</i>	盐敏感基因	<i>Vfaba. Tiffany. R1. 5g049280</i>	[19]



R;根;L;茎叶;F;花;S20、S30、S40 表示开花后 20 d、30 d、40 d 的种子;P20、P30、P40 表示开花后 20 d、30 d、40 d 的果皮。

图 9 蚕豆盐胁迫候选 *WRKY* 基因表达模式

Fig.9 Expression pattern of *Vicia faba* candidate *WRKY* genes under salt stress

3 讨论

蚕豆是具有较高营养价值和生物固氮作用的食用豆类,由于缺乏参考基因组序列,使得蚕豆的基础研究进展缓慢。2023 年 Jayakodi 等^[33]通过高通量测序,组装了第一个蚕豆高质量、染色体规模的基因组,并公布了 37 065 个基因的序列信息。全长转录组测序(三代高通量测序技术)能够直接获得高质量全长转录本,高质量全长转录本既可以为无参基因组的二代测序数据拼接提供参考,也可以完善基因组的基因数量。本研究首次对蚕豆的根、茎叶、花、种子(开花后 20 d、30 d、40 d)、果皮(开花后 20 d、30 d、40 d)等 9 个样本进行了三代转录组测序,获得 58 885 条转录本序列,与蚕豆基因组相比,发现新转录本 42 019 个。本研究构建的全长转录组为后续进一步开展蚕豆遗传研究提供了丰富的数据基础。

WRKY 转录因子基因是植物特有的转录因子基因,已有研究表明,*WRKY* 转录因子基因能够调控植物生长发育,例如,调控种子休眠、萌发^[34],调控幼苗的形态发生^[35],以及调节植物的开花时间等^[36]。除了这些功能外,*WRKY* 基因在响应非生物胁迫(干旱、高温、高盐等)^[37]和生物胁迫(病原菌^[38]和昆虫^[39]等)方面也具有非常重要的作用。本研究首次利用全长转录组测序对蚕豆 *WRKY* 家族进行研究,并鉴定出 113 个 *VjWRKY* 基因。不同物种间系统进化树的建立,有助于更加准确地研究物种中未知家族成员的功能,我们根据进化关系,将蚕豆 113 个 *WRKY* 基因分成 3 组,其中 group 2 成员最多,为 61 个。同时我们对 *VjWRKY* 家族进行保守结构域和保守基序分析发现,group 1 中的 4 个基因(*VjWRKY16*、*VjWRKY17*、*VjWRKY18*、*VjWRKY19*)具有 2 个 *WRKY* 保守结构域,但却与 group 2 亚家族聚集在同一分支,通过序列比对,发现是由于第二个 *WRKY* 结构域缺少 C2-H2 序列造成的,需要通过优化蚕豆参考基因组后来完善其序列。

VjWRKY 基因家族主要富集在植物 MAPK 信号通路、植物与病原菌相互作用和剪接体 3 个通路中。刘晨等^[40]通过分析前人的研究发现 MAPK 信号通路能够响应干旱胁迫、盐胁迫、极端温度及营养匮乏等非生物胁迫,并在植物抗逆过程中扮演重要角色,这与本研究的结果相符。同时 *WRKY* 基因在植物与病原菌相互作用通路中也具有多种功能,研究发现 *AtWRKY25* 和 *AtWRKY33* 既能调控植物与病原菌的

相互作用^[41-42],也能增强植物抵抗盐胁迫的能力^[24]。根据拟南芥 *WRKY* 基因家族已知基因功能,通过同源比对,在 *VjWRKY* 基因家族中发现 14 个候选基因可能与盐胁迫相关,其中部分基因同时富集在植物与病原菌相互作用和植物 MAPK 信号通路中。本研究的下一步工作将对 14 个候选基因进行功能验证,探索其与盐胁迫的相关性。

参考文献:

- [1] YAMAGUCHI-SHINOZAKI K, SHINOZAKI K. Transcriptional regulatory networks in cellular responses and tolerance to dehydration and cold stresses[J]. *Annual Review of Plant Biology*, 2006, 57: 781-803.
- [2] MUNNS R, TESTER M. Mechanisms of salinity tolerance[J]. *Annual Review of Plant Biology*, 2008, 59: 651-681.
- [3] FU L B, SHEN Q F, KUANG L H, et al. Metabolite profiling and gene expression of Na/K transporter analyses reveal mechanisms of the difference in salt tolerance between barley and rice[J]. *Plant Physiology and Biochemistry*, 2018, 130: 248-257.
- [4] 毕晨曦, 杨宇昕, 于月华, 等. 小麦 bZIP 家族转录因子的鉴定及其在盐胁迫条件下的表达分析[J]. *分子植物育种*, 2021, 19(15): 4887-4895.
- [5] ULLAH A, SUN H, YANG X, et al. A novel cotton *WRKY* gene, *GhWRKY6-like*, improves salt tolerance by activating the ABA signaling pathway and scavenging of reactive oxygen species[J]. *Physiologia Plantarum*, 2018, 162(4): 439-454.
- [6] DU B, LIU H, DONG K, et al. Over-expression of an *R2R3 MYB* gene, *MdMYB108L*, enhances tolerance to salt stress in transgenic plants[J]. *International Journal of Molecular Sciences*, 2022, 23(16): 9428.
- [7] 李 崑, 周春江. 植物 *WRKY* 转录因子的研究进展[J]. *植物生理学报*, 2014, 50(9): 1329-1335.
- [8] EULGEM T, RUSHTON P J, ROBATZEK S, et al. The *WRKY* superfamily of plant transcription factors[J]. *Trends in Plant Science*, 2000, 5(5): 199-206.
- [9] XIE T, CHEN C J, LI C H, et al. Genome-wide investigation of *WRKY* gene family in pineapple: evolution and expression profiles during development and stress[J]. *BMC Genomics*, 2018, 19: 1-18.
- [10] WU K L, GUO Z J, WANG H H, et al. The *WRKY* family of transcription factors in rice and *Arabidopsis* and their origins[J]. *DNA Research*, 2005, 12(1): 9-26.
- [11] SCHMUTZ J, CANNON S B, SCHLUETER J, et al. Genome sequence of the palaeopolyploid soybean[J]. *Nature*, 2010, 463(7278): 178-183.
- [12] 苏文娟, 曹瑞兰, 周增亮, 等. 油茶 *WRKY* 基因家族鉴定及逆境胁迫表达分析[J]. *中南林业科技大学学报*, 2023, 43(3): 155-166, 174.
- [13] 郝青婷, 高 伟, 闫虎斌, 等. 绿豆 *WRKY* 基因家族的全基因

- 组鉴定及生物信息学分析[J].西北农林科技大学学报(自然科学版),2023,51(5):59-71,81.
- [14] ZHOU L, WANG N N, GONG S Y, et al. Overexpression of a cotton (*Gossypium hirsutum*) *WRKY* gene, *GhWRKY34*, in *Arabidopsis* enhances salt-tolerance of the transgenic plants[J]. Plant Physiology and Biochemistry, 2015, 96:311-320.
- [15] BO C, CAI R H, FANG X, et al. Transcription factor *ZmWRKY20* interacts with *ZmWRKY115* to repress expression of *ZmZIP111* for salt tolerance in maize[J]. The Plant Journal, 2022, 111(6):1660-1675.
- [16] BABITHA K C, RAMU S V, PRUTHVI V, et al. Co-expression of *bHLH17* and *WRKY28* confers resistance to abiotic stress in *Arabidopsis*[J]. Transgenic Research, 2013, 22:327-341.
- [17] ZHANG Y Z, LI P, NIU Y Q, et al. Evolution of the *WRKY66* gene family and its mutations generated by the CRISPR/Cas9 system increase the sensitivity to salt stress in *Arabidopsis*[J]. International Journal of Molecular Sciences, 2023, 24(4):3071.
- [18] JIANG Y Q, DEYHOLOS M K. Functional characterization of *Arabidopsis* NaCl-inducible *WRKY25* and *WRKY33* transcription factors in abiotic stresses[J]. Plant Molecular Biology, 2009, 69:91-105.
- [19] SHEN Y, CHI Y H, LU S, et al. Involvement of JM15 in the dynamic change of genome-wide H3K4me3 in response to salt stress[J]. Frontiers in Plant Science, 2022, 13:1009723.
- [20] WU X, XU J N, MENG X N, et al. Linker histone variant HIS1-3 and *WRKY1* oppositely regulate salt stress tolerance in *Arabidopsis*[J]. Plant Physiology, 2022, 189(3):1833-1847.
- [21] LI P, LI X W, JIANG M. CRISPR/Cas9-mediated mutagenesis of *WRKY3* and *WRKY4* function decreases salt and Me-JA stress tolerance in *Arabidopsis thaliana*[J]. Molecular Biology Reports, 2021, 48(8):5821-5832.
- [22] SCARPECI T E, ZANOR M I, MUELLER-ROEBER B, et al. Overexpression of *AtWRKY30* enhances abiotic stress tolerance during early growth stages in *Arabidopsis thaliana*[J]. Plant Molecular Biology, 2013, 83:265-277.
- [23] KRISHNAMURTHY P, VISHAL B, BHAL A, et al. *WRKY9* transcription factor regulates cytochrome *P450* genes *CYP94B3* and *CYP86B1*, leading to increased root suberin and salt tolerance in *Arabidopsis*[J]. Physiologia Plantarum, 2021, 172(3):1673-1687.
- [24] CHEN H, LAI Z B, SHI J W, et al. Roles of *Arabidopsis* *WRKY18*, *WRKY40* and *WRKY60* transcription factors in plant responses to abscisic acid and abiotic stress[J]. BMC Plant Biology, 2010, 10(1):1-15.
- [25] HU Y, CHEN L, WANG H P, et al. *Arabidopsis* transcription factor *WRKY8* functions antagonistically with its interacting partner VQ 9 to modulate salinity stress tolerance[J]. The Plant Journal, 2013, 74(5):730-745.
- [26] ALI M A, AZEEM F, NAWAZ M A, et al. Transcription factors *WRKY11* and *WRKY17* are involved in abiotic stress responses in *Arabidopsis*[J]. Journal of Plant Physiology, 2018, 226:12-21.
- [27] 杨俊品, 罗菊枝. 中国蚕、豌豆育种进展[J]. 西南农业学报 1996, 9(增刊1):142-146.
- [28] 田莹莹. 蚕豆籽粒大小的 QTL 分析[D]. 西宁:青海大学, 2018.
- [29] PUNTA M, COGGILL P C, EBERHARDT R Y, et al. The Pfam protein families database[J]. Nucleic Acids Research, 2012, 40(1):290-301.
- [30] SUBRAMANIAN B, GAO S H, LERCHER M J, et al. Evolvview v3: a webserver for visualization, annotation, and management of phylogenetic trees[J]. Nucleic Acids Research, 2019, 47(1):270-275.
- [31] CHEN C J, CHEN H, ZHANG Y, et al. TBtools: an integrative toolkit developed for interactive analyses of big biological data[J]. Molecular Plant, 2020, 13(8):1194-1202.
- [32] BAILEY T L, BODEN M, BUSKE F A, et al. MEME SUITE: tools for motif discovery and searching[J]. Nucleic Acids Research, 2009, 37(S2):202-208.
- [33] JAYAKODI M, GOLICZ A A, KREPLAK J, et al. The giant diploid faba genome unlocks variation in a global protein crop[J]. Nature, 2023, 615(7953):652-659.
- [34] WANG H P, CHEN W Q, XU Z Y, et al. Functions of *WRKYs* in plant growth and development[J]. Trends in Plant Science, 2023, 28(6):630-645.
- [35] ZHOU H, ZHU W, WANG X C, et al. A missense mutation in *WRKY32* converts its function from a positive regulator to a repressor of photomorphogenesis[J]. The New Phytologist, 2021, 235(1):111-125.
- [36] HUNG F Y, SHIH Y H, LIN P Y, et al. *WRKY63* transcriptional activation of *COOLAIR* and *COLDAR* regulates vernalization-induced flowering[J]. Plant Physiology, 2022, 190(1):532-547.
- [37] LI W X, PANG S Y, LU Z G, et al. Function and mechanism of *WRKY* transcription factors in abiotic stress responses of plants[J]. Plants, 2020, 9(11):1515.
- [38] WANG L J, GUO D Z, ZHAO G D, et al. Group IIc *WRKY* transcription factors regulate cotton resistance to *Fusarium oxysporum* by promoting GhMKK2-mediated flavonoid biosynthesis[J]. New Phytologist, 2022, 236(1):249-265.
- [39] YIN M, SONG N, CHEN S Y, et al. NaKT12, a Kunitz trypsin inhibitor transcriptionally regulated by *NaWRKY3* and *NaWRKY6*, is required for herbivore resistance in *Nicotiana attenuata*[J]. Plant Cell Reports, 2021, 40:97-109.
- [40] 刘晨, 曹小汉, 殷丹丹, 等. MAPK 信号通路调控植物响应非生物胁迫的研究进展[J]. 安徽农业科学, 2022, 50(18):9-16.
- [41] ZHENG Z Y, MOSHER S L, FAN B, et al. Functional analysis of *Arabidopsis* *WRKY25* transcription factor in plant defense against *Pseudomonas syringae*[J]. BMC Plant Biology, 2007, 7(1):1-13.
- [42] ANDREASSON E, JENKINS T, BRODERSEN P, et al. The MAP kinase substrate MKS1 is a regulator of plant defense responses[J]. The EMBO Journal, 2005, 24(14):2579-2589.

(责任编辑:陈海霞)