

• 药材与资源 •

金线莲中调控胚胎发育 WRKY 转录因子筛选及克隆分析

邢丙聪^{1,2}, 苏立样^{1,2}, 万思琦^{1,2}, 邵清松^{1,2*}

1. 浙江省特色中药资源保护与创新利用重点实验室, 浙江 杭州 311300

2. 浙江农林大学中药学科, 浙江 杭州 311300

摘要:目的 探究金线莲 *Anoectochilus roxburghii* 中 WRKY 转录因子时空表达模式并筛选调控金线莲胚胎发育的 *ArWRKY* 基因。方法 基于金线莲转录组数据库进行 *ArWRKY* 转录因子筛选, 利用软件进行蛋白质结构域预测、序列比对、理化性质预测与进化树构建等分析。采用实时荧光定量 PCR (qRT-PCR) 检测 *ArWRKY* 基因时空表达模式。根据转录组数据 Unigenes 克隆 *ArWRKY* 的 CDS 序列。结果 筛选得到金线莲中 23 条 WRKY 转录因子, 分别命名为 *ArWRKY1*~*ArWRKY23*; qRT-PCR 结果表明 *ArWRKY5* 与 *ArWRKY20* 基因在金线莲叶芽期、花芽期及开花期的花中表达量都显著高于根、茎、叶; 克隆得到 *ArWRKY5* 的 CDS 为 600 bp, *ArWRKY20* 的 CDS 为 1149 bp。结论 筛选得到 23 个 *ArWRKY* 转录因子并成功克隆获得 *ArWRKY5* 与 *ArWRKY20* 的 CDS 序列, 对探究 *ArWRKY* 转录因子调控金线莲生长发育及验证 *ArWRKY5* 与 *ArWRKY20* 基因功能奠定基础。

关键词: 金线莲; 胚胎发育; WRKY 转录因子; 生物信息学分析; 表达模式

中图分类号: R286.12 文献标志码: A 文章编号: 0253-2670(2022)12-3745-10

DOI: 10.7501/j.issn.0253-2670.2022.12.021

Screening and cloning analysis of WRKY transcription factor regulating embryo development from *Anoectochilus roxburghii*XING Bing-cong^{1,2}, SU Li-yang^{1,2}, Wan Si-qi^{1,2}, SHAO Qing-song^{1,2}

1. Zhejiang Provincial Key Laboratory of Resources Protection and Innovation of Traditional Chinese Medicine, Hangzhou 311300, China

2. Department of Traditional Chinese Medicine, Zhejiang Agriculture and Forestry University, Hangzhou 311300, China

Abstract: **Objective** In order to lay a foundation for further study of *EuTCP* genes function, TCP (Teosinte branched 1/Cycloidea/Proliferating) gene family of *Eucommia ulmoides* (*E. ulmoides*) is screened and analyzed in this study. **Method** Based on the *E. ulmoides* genome database, the physicochemical properties, phylogeny, gene structure, promoter *cis*-acting elements and their expression levels in the leaf development and eu-rubber formation of the TCP gene family were analyzed by bioinformatics system. **Result** In this study, a total of 14 *EuTCPs* were identified from *E. ulmoides* genome, *EuTCPs* encoded 139—492 amino acids, and theoretical isoelectric points distribution were 5.54—9.72, molecular weight between 18 880 to 53 620 by physicochemical analysis, all of *EuTCPs* proteins were hydrophilic proteins and expressed in nucleus. *EuTCP* gene family was divided into three subfamilies: PCF, CIN and CYC/TB1, which contained 5, 6 and 3 *EuTCPs* proteins respectively. The expression pattern analysis showed that *EuTCPs* were significantly different at different developmental stages of *E. ulmoides* leaves, and the expression levels of *EuTCPs* were different at each developmental stage. *EuTCPs* positively regulated the formation of eu-rubber. **Conclusion** There are 14 members of TCP family identified in *E. ulmoides*, and their molecular characteristics and tissue expression specificity are different, which may be involved in the development of leaves and eu-rubber formation.

Key words: *Anoectochilus roxburghii* (Wall.) Lindl.; embryonic development; WRKY transcription factor; bioinformatics analysis; expression pattern

金线莲 *Anoectochilus roxburghii* (Wall.) Lindl. 为兰科开唇兰属多年生草本植物, 是我国传统珍贵药材, 全草可入药, 具有清热凉血、祛风利湿等功效^[1]。在

自然条件下, 金线莲生长缓慢、结实率低、种子败育现象严重、幼苗存活率低等原因导致野生金线莲种群规模急剧收缩, 濒临灭绝^[2-3]。金线莲种子自然

收稿日期: 2021-11-06

基金项目: 国家自然科学基金项目 (82173916); 浙江省农业新品种选育重大科技专项 (2021C02074); 浙江省特色中药资源保护与创新利用重点实验室 (2021E10013)

作者简介: 邢丙聪 (1989—), 讲师, 研究方向为药用植物次生代谢调控。E-mail: xingbingcong@163.com

*通信作者: 邵清松 (1980—), 男, 博士, 教授, 主要从事药用植物栽培与育种研究。E-mail: sqszjfc@126.com

繁殖率极低,种子细小且种胚形态结构不健全,并且种子必须与菌根真菌共生,将种子胚细胞中的淀粉转化为糖,才能萌芽生长^[4-6]。近年来,许多学者对金线莲人工栽培进行研究,在种苗培育、组培苗移栽、设施栽培等关键技术取得突破性进展,但是关于生殖发育分子生物学方面研究较少报道。

WRKY 转录因子是植物中最大的转录因子家族之一。WRKY 转录因子因具有保守的“WRKYGQK”序列而被简称为“WRKY”^[7]。WRKY 已经被证实在调控植物种子的形成、发育及萌发等方面具有重要作用。研究发现,AtWRKY34 与其同源蛋白 AtWRKY2 是拟南芥花粉发育所必需的。若 WRKY34 在花粉发育的早期阶段没有被蛋白激酶 MPK3 和 MPK6 磷酸化,拟南芥花粉发育和花粉管生长等功能都会受到损害^[8]。编码 AtWRKY10 的 *MINI3* 和编码 LRR 受体激酶的 *IKU2* 共同正调控了拟南芥种子的大小和胚乳的发育^[9]。AtWRKY43 能促进拟南芥种子内脂肪物质转化为糖类从而供给发芽所需的能量消耗,增强种子的发芽能力^[10]。AtTTG2 则能调控拟南芥种皮的发育^[11]。葡萄中的 VvWRKY20 对葡萄胚珠的发育至关重要,VvWRKY20 的缺失会导致种子前体——胚珠的败育^[12]。Xiang 等^[13]研究发现,在谷子中编码一种第一亚族 WRKY 转录因子的 *lp1* 基因突变会导致谷子种子变大。水稻中的 OsWRKY78 也与水稻种子大小的发育有关,抑制 OsWRKY78 的表达会引起细胞长度减小从而使种子变小^[14]。棉花中 GhWRKY22 能通过调节 *JAZ* 基因的表达来影响雄性花粉的发育^[15];而 GhWRKY17 调控了棉花种子的发芽能力^[16]。人参中的 PgWRKY3 参与种子的后熟过程^[17]。金钗石斛中的 DnWRKY11 则可增强其种子的抗逆性^[18]。目前,在金线莲中 WRKY 转录因子的调控作用尚未见报道。

本研究利用 Illumina 结合 PacBio 技术对金线莲转录组进行测序,首次筛选得到 23 条 ArWRKY 序列,分别命名为 ArWRKY1~ArWRKY23。通过 qRT-PCR 检测分析,发现 *ArWRKY5* 与 *ArWRKY20* 在金线莲中显著高表达,推测两者在花的发育过程中扮演了重要的角色。最后,本课题组克隆得到 2 条基因的 CDS 序列,为其功能进一步验证奠定了基础。

1 材料与试剂

材料购自浙江金华市农业科学研究院,经浙江农林大学胡润淮教授鉴定为金线莲 *A. roxburghii*

(Wall.) Lindl. 组培苗。

RNAprep Pure 多糖多酚植物总 RNA 提取试剂盒购自天根生化科技(北京)有限公司;AxyPrep™ DNA 凝胶回收试剂盒购自爱思进生物技术(杭州)有限公司;SYBR® Premix Ex Taq™ II (Tli RNaseH Plus) 荧光定量试剂盒、反转录 PrimeScript™ II 1st strand cDNA Synthesis Kit 试剂盒、pMD19-T 载体、PrimeSTAR Max DNA Polymerase 试剂盒、DNA Marker 均购自 TaKaRa 公司(日本);胰蛋白胨、酵母提取物、氯化钠、氨苄青霉素均购自生工生物工程(上海)股份有限公司;琼脂粉购自北京鼎国昌盛生物技术有限责任公司。

2 方法

2.1 样品处理

将已驯化过的金线莲组培苗洗净培养基,移栽至浙江农林大学百草园试验基地,栽培基质为泥炭-河沙-花生壳 4:2:2(体积比)^[19]。待生长稳定后随机采摘 5~6 株分为根、茎、叶 3 部分迅速放入液氮速冻,并送往上海派森诺生物科技股份有限公司进行转录组测序,3 次生物学重复。

2.2 ArWRKY 的生物信息学分析

利用 MEME 在线网站(<http://meme-suite.org/tools/meme>)预测蛋白功能结构域,得到结构域蛋白序列,并拼接得到 ArWRKY 核心序列;利用 Clustal W 对核心序列进行多序列比对;使用 MEGA7 NJ 法(bootstrap 1000)构建系统进化树。利用 ProtParam(<http://web.expasy.org/protparam/>)预测编码蛋白的相对分子质量、等电点等理化特性。

2.3 RNA 提取与反转录

按照 TIANGEN 的 RNAprep Pure 多糖多酚植物总 RNA 提取试剂盒的说明书提取金线莲总 RNA,并检测 RNA 的完整性与浓度,合格后用于后续实验。按照 PrimeScript™ II 1st strand cDNA Synthesis Kit 试剂盒的说明书合成 cDNA 第 1 链。

2.4 ArWRKY 表达模式分析

分别收集金线莲叶芽期和花芽期的花及开花期的花、根、茎和叶提取金线莲总 RNA,反转录后用做测量基因表达量的模板。以金线莲 *Actin*(*ArActin*)为内参基因,利用 qRT-PCR 检测 *ArWRKY1*、*ArWRKY5*、*ArWRKY6*、*ArWRKY8*、*ArWRKY9*、*ArWRKY11*、*ArWRKY13*、*ArWRKY19*、*ArWRKY20* 与 *ArWRKY23* 的时空表达模式(引物序列见表 1)。qRT-PCR 反应体系见表 2。

表 1 金线莲 *ArWRKY* 和 *ArActin* 的 qRT-PCR 引物序列
Table 1 qRT-PCR primer sequences of *ArWRKY* and *ArActin*

基因名称	引物序列 (5'-3')
<i>ArActin</i>	F: GCTAGTGGCCGTACAACCTGG R: GCCAGCAAGGTCCAATCGAA
<i>ArWRKY1</i>	F: TTCAGCAGTGGACCAAAGCC R: CTCGGTTCCCTCACCTTCCT
<i>ArWRKY5</i>	F: CGCAGCAACAACAACAGCAA R: AGGGTTTCCCTTCGCCATCT
<i>ArWRKY6</i>	F: GAACCATCGGGAGGAGGTGA R: GGGCGGAATCAACAACCTCA
<i>ArWRKY8</i>	F: GTGCCGAAAGCGAGAGAGAC R: ATTGCTTCCATGCTGCTGCT
<i>ArWRKY9</i>	F: TGGCCATCATCTTCCCAGGT R: CAAACACTTCGGGAGTGGCA
<i>ArWRKY11</i>	F: AGCACGTCGAGAGGTCTTCA R: AAGAGAAGGCGGGTCATGGT
<i>ArWRKY13</i>	F: CCGCCTTCTCCTTCTCCTCA R: CGGCGTTGACTTTAGCCCTT
<i>ArWRKY19</i>	F: GGAAAGCACAGCCACCCAAT R: CCGACCTGCCGAAGTTTGAA
<i>ArWRKY20</i>	F: GGTGGAGGCAGAGTTCAGGA R: CACAGCCCAGCTTGAACAT
<i>ArWRKY23</i>	F: CTTATCGCGGAGCTCACACC R: ACTCGAGAAGCGGTTGAGGA

表 2 荧光定量 PCR 反应液
Table 2 RT-qPCR reaction liquid

反应成分	用量/ μL
2 $\times$ SYBR [®] Premix Ex Taq II	5.0
Forward primer (10 $\mu\text{mol}\cdot\text{L}^{-1}$)	0.4
Reverse primer (10 $\mu\text{mol}\cdot\text{L}^{-1}$)	0.4
cDNA	1.0
RNase Free ddH ₂ O	3.2

以开花期的根为对照, 采用相对定量法 $2^{-\Delta\Delta C_t}$ 计算基因相对表达量。使用 SPSS 23 软件对数据进行处理, Duncan 与 LSD 分析法检测进行显著性分析, 字母标记法进行标记。

2.5 *ArWRKY5* 与 *ArWRKY20* 基因扩增与测序

根据转录组数据 Unigenes 序列设计上下游引物 (序列见表 3)。以金线莲 cDNA 为模板进行序列扩增。*ArWRKY5* 序列扩增反应程序为 98 $^{\circ}\text{C}$ 、3

表 3 基因克隆引物
Table 3 Cloning primers of genes

基因名称	引物序列 (5'-3')
<i>ArWRKY5</i>	F: ATGCAGGTAATGGATGCAAAAATGG R: CTACTTGGTTTCACCGCTCCCACTT
<i>ArWRKY20</i>	F: ATGGCGGCTCTCATGGCGGCTCT R: CTATGAAGGCGGCGCAAACTGAAAG

min, 98 $^{\circ}\text{C}$ 、10 s, 55 $^{\circ}\text{C}$ 、5 s, 72 $^{\circ}\text{C}$ 、1.2 min, 30 个循环。*ArWRKY20* 序列扩增反应程序为 98 $^{\circ}\text{C}$ 、3 min, 98 $^{\circ}\text{C}$ 、10 s, 55 $^{\circ}\text{C}$ 、5 s, 72 $^{\circ}\text{C}$ 、0.7 min, 30 个循环。用 1% 琼脂糖凝胶电泳检测 PCR 产物, 选择正确大小的凝胶片段进行切胶回收。

将胶回收产物连接为 pMD19-T 载体, 转化大肠杆菌 DH5 α 中, 并在含有氨苄青霉素 (Amp) 的 LB 固体培养基平板上倒置培养 12~16 h, 挑选阳性菌落送往浙江有康生物科技有限公司测序。

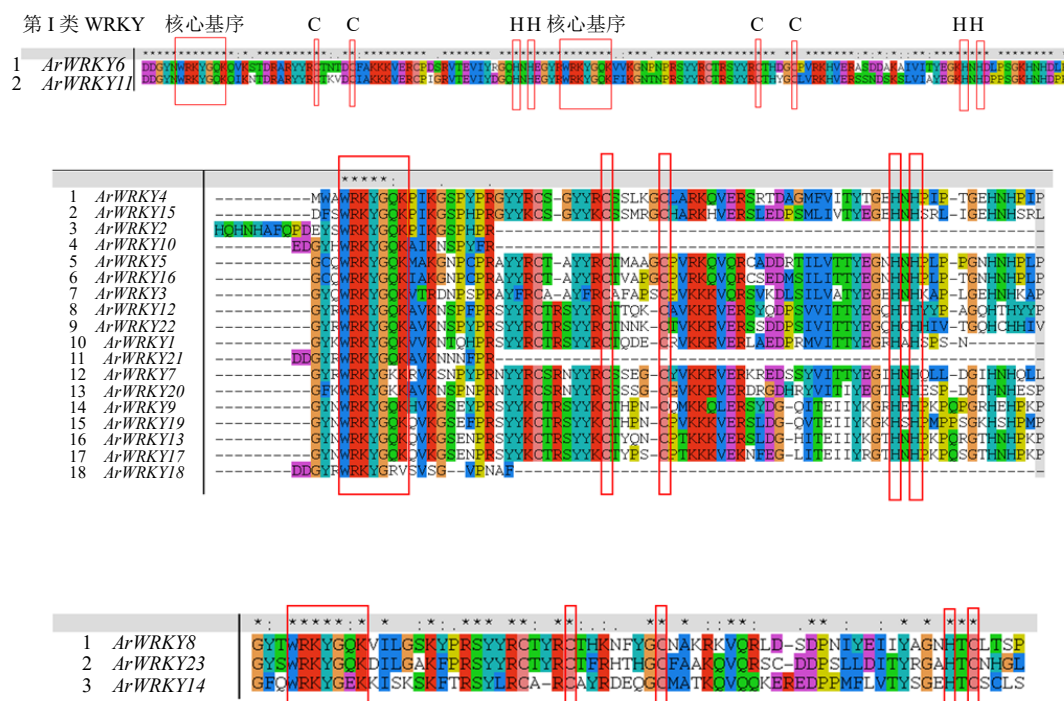
2.6 *ArWRKY5* 与 *ArWRKY20* 的生物信息学分析

利用在线网站 (<https://www.detaibio.com/tools/signal-peptide.html>) 预测信号肽; 利用 NetPhos 3.1 Server (<http://www.cbs.dtu.dk/services/NetNGlyc/>) 预测蛋白的糖基化位点、磷酸化位点和蛋白激酶结合位点; 利用 ExPasy-ProtScale (<https://web.expasy.org/protscale/>) 分析蛋白亲/疏水性。分别采用 GOR4 (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_gor4.html) 和 SWISSMODEL (<https://swissmodel.expasy.org/>) 预测蛋白的二、三级结构。

3 结果与分析

3.1 *ArWRKY* 的生物信息学分析

从金线莲中筛选得到 23 条 *ArWRKY* 序列, 分别命名为 *ArWRKY1*~*ArWRKY23*。根据 WRKY 转录因子的蛋白结构特征将 23 条 *ArWRKY* 分为 3 类: 其中 *ArWRKY6* 和 *ArWRKY11* 具有 2 个 WRKY 域, 属于第 I 类; *ArWRKY8*、*ArWRKY14* 和 *ArWRKY23* 具有 1 个 WRKY 域, 所含的锌指结构为 C₂-HC (C-X₇-C-X₂₃-H-X-C), 属于第 III 类; 其他 18 个家族成员具有一个 WRKY 域, 所含锌指结构为 C₂-H₂ (C-X_{4.5}-C-X₂₂₋₂₃-HXH), 属于第 II 类 (图 1)。*ArWRKY* 蛋白序列氨基酸长度从 100 个到 623 个不等, 蛋白大小最小为 11 445.87, 最大为 66 584.31。*ArWRKY* 所含的保守结构域基序除了常见的 WRKYGQK, 还含有 WRKYGKK、WRKYGRV 或 WRKYGEK (表 4)。



其中红色框内表示 WRKY 的保守基序与锌指结构域
the red box indicates the conserved motif and zinc finger domain of WRKY

图 1 *ArWRKY* 核心序列比对图

Fig. 1 *ArWRKY* core sequence alignment diagram

表 4 *ArWRKY* 蛋白信息

Table 4 *ArWRKY* protein information

转录因子	氨基酸数目	等电点	蛋白大小	家族	保守域	锌指结构	锌指结构类型
ArWRKY1	222	9.83	25 185.29	第 II 类	WRKYGQK	C-X4-C-X23-HXH	C ₂ -H ₂
ArWRKY2	197	10.27	21 985.05	第 II 类	WRKYGQK	无	无
ArWRKY3	222	9.34	24 561.83	第 II 类	WRKYGQK	C-X5-C-X23-HXH	C ₂ -H ₂
ArWRKY4	176	4.58	19 497.14	第 II 类	WRKYGQK	C-X5-C-X23-HXH	C ₂ -H ₂
ArWRKY5	382	7.07	40 216.17	第 II 类	WRKYGQK	C-X5-C-X23-HXH	C ₂ -H ₂
ArWRKY6	514	6.29	55 377.25	第 I 类	WRKYGQK	C-X4-C-X22-23-HXH	C ₂ -H ₂
ArWRKY7	133	9.79	15 088.88	第 II 类	WRKYGKK	C-X4-C-X23-HXH	C ₂ -H ₂
ArWRKY8	265	5.06	28 827.02	第 III 类	WRKYGQK	C-X7-C-X23-H-X-C	C ₂ -HC
ArWRKY9	318	6.83	35 173.68	第 II 类	WRKYGQK	C-X4-C-X22-HXH	C ₂ -H ₂
ArWRKY10	203	4.88	22 959.65	第 II 类	WRKYGQK	无	无
ArWRKY11	368	8.79	41 695.80	第 I 类	WRKYGQK	C-X4-C-X22-23-HXH	C ₂ -H ₂
ArWRKY12	131	9.12	15 048.89	第 II 类	WRKYGQK	C-X4-C-X23-HXH	C ₂ -H ₂
ArWRKY13	339	6.80	37 058.17	第 II 类	WRKYGQK	C-X4-C-X22-HXH	C ₂ -H ₂
ArWRKY14	172	8.70	19 374.75	第 III 类	WRKYGEK	C-X7-C-X24-H-X-C	C ₂ -HC
ArWRKY15	322	9.71	35 640.63	第 II 类	WRKYGQK	C-X5-C-X23-HXH	C ₂ -H ₂
ArWRKY16	623	9.09	66 584.31	第 II 类	WRKYGQK	C-X5-C-X23-HXH	C ₂ -H ₂
ArWRKY17	316	6.25	34 067.18	第 II 类	WRKYGQK	C-X4-C-X22-HXH	C ₂ -H ₂
ArWRKY18	100	9.75	11 445.87	第 II 类	WRKYGRV	无	无
ArWRKY19	269	8.66	29 993.20	第 II 类	WRKYGQK	C-X4-C-X22-HXH	C ₂ -H ₂
ArWRKY20	223	4.72	24 833.40	第 II 类	WRKYGKK	C-X4-C-X23-HXH	C ₂ -H ₂
ArWRKY21	171	7.10	19 259.41	第 II 类	WRKYGQK	无	无
ArWRKY22	142	8.81	15 984.06	第 II 类	WRKYGQK	C-X4-C-X23-HXH	C ₂ -H ₂
ArWRKY23	338	6.27	37 459.57	第 III 类	WRKYGQK	C-X7-C-X23-H-X-C	C ₂ -HC

为了解 ArWRKY 的保守结构域信息,运用 MEME 在线网站预测保守基序,设置了 5 个保守域 (Motif) 为条件进行分析,结果见图 2。Motif 1 与 Motif 3 虽属于不同的保守基序,但均含有“WRKY”4 个氨基酸残基,且所有 ArWRKY 中均含 Motif 1 或 Motif 3,说明在 ArWRKY 中,色氨酸 (W)、精氨酸 (R)、赖氨酸 (K) 和酪氨酸 (Y) 是核心序列中必不可少的,这也符合 WRKY 转录因子的结构特征与命名规则。Motif 1 和 Motif 3 属于不同结构域暗示着在含有 2 个 WRKY 域的第一类 ArWRKY 中,2 者可能具有不同的功能。通过分析可以发现 Motif 2 与 Motif 4 属于锌指结构,大部分

ArWRKY 中 Motif 2 与 Motif 4 共同存在,且 Motif 4 在 Motif 2 的下游,也符合 WRKY 转录因子中锌指结构的构造。Motif 5 同样是锌指结构,但只在 ArWRKY6 与 ArWRKY11 这 2 个第 I 类 WRKY 中出现,且分布在 Motif 3 的下游,说明在 ArWRKY 中每一段保守序列 WRKY 的下游可能都存在一个锌指结构与之对应。但在 ArWRKY2、ArWRKY10、ArWRKY18 和 ArWRKY21 中只存在 WRKY 保守序列却未出现锌指结构,可能是由于蛋白序列不完整造成的。另外,ArWRKY14 所含有的锌指结构为 C-X₇-C-X₂₄-H-X-C,与文献报道的 C₂-HC 类型为 C-X₇-C-X₂₃-H-X-C 构造相差 1 个氨基酸残基。

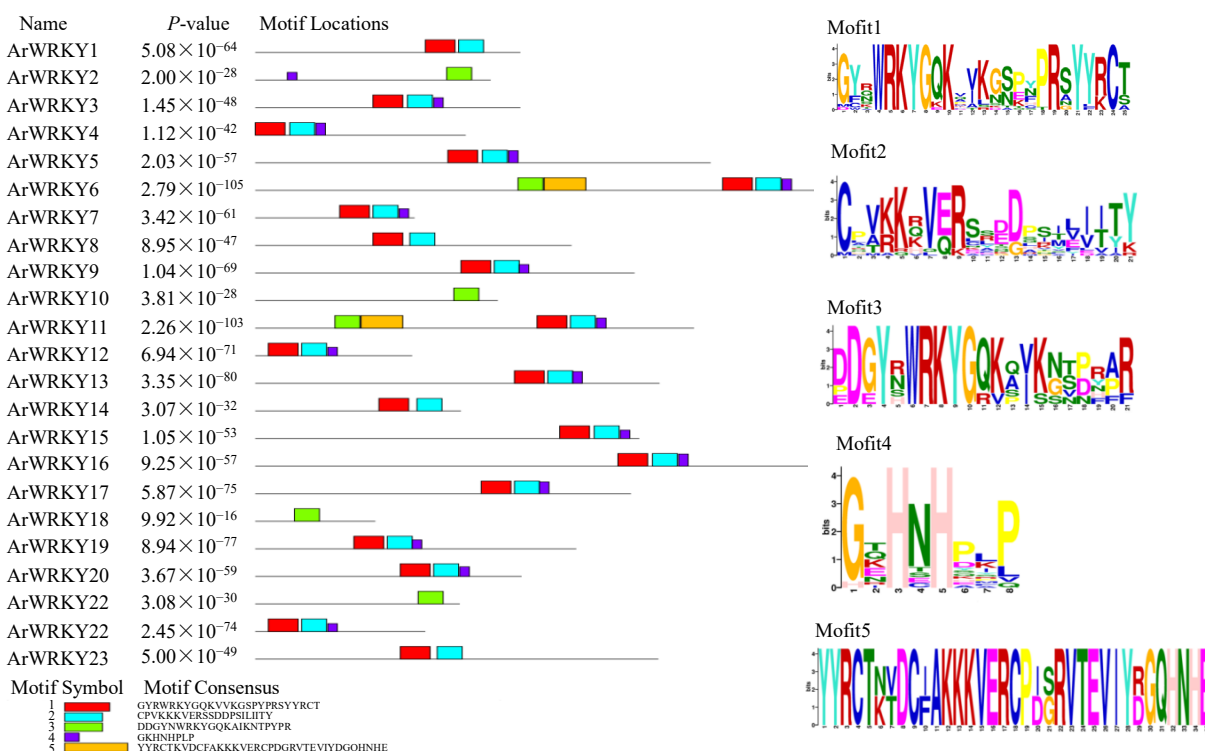


图 2 ArWRKY 保守域预测

Fig. 2 ArWRKY conservative domain prediction

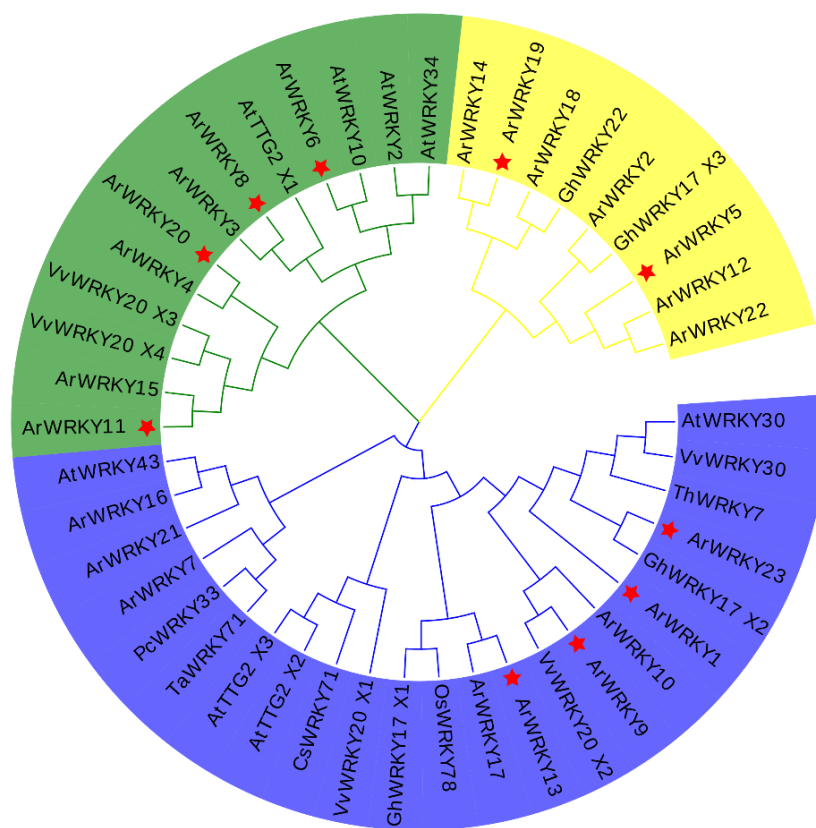
为了进一步筛选可能参与金线莲胚胎发育的 ArWRKY,利用 NCBI 数据库下载得到其他物种中已经报道的对种子形成、发育或萌发等方面具有调控作用的 WRKY 蛋白: AtWRKY2 (NP_200438.1)、AtWRKY34 (NP_194374.1)、AtWRKY43 (AEC10646.1)、AtWRKY30 (NP_568439.1)、AtWRKY10 (NP_175956.1)、TTG_X1 (NP_181263.2)、TTG_X2 (NP_001078015.1)、TTG_X3 (NP_001323504.1); GhWRKY22 (AIE43820.1)、GhWRKY17_X1 (AIE43818.1)、GhWRKY17_X2

(AJT43302.1)、GhWRKY17_X3 (ADW82100.1); VvWRKY20_X1 (XP_010647039.1)、VvWRKY20_X2 (XP_019074038.1)、VvWRKY20_X3 (XP_019074039.1)、VvWRKY20_X4 (XP_010647041.1)、VvWRKY30 (ALM96663.1); OsWRKY78 (DAA05640.1); ThWRKY7 (AFS64072.1); CsWRKY71 (NP_001292668.1); TaWRKY71 (ABN43177.1); PcWRKY33 (AYN74370.1)。

采用 MAGA 7.0 软件 NJ 法构建系统进化树(图

3)。结果发现 *ArWRKY* 可以聚为 3 大类, 第 1 大类为 *ArWRKY2*、*ArWRKY12*、*ArWRKY14*、*ArWRKY18* 与 *ArWRKY22*; 第 2 大类为 *ArWRKY3*、*ArWRKY4*、*ArWRKY6*、*ArWRKY8*、*ArWRKY11*、*ArWRKY15* 与 *ArWRKY20*, 其余 *ArWRKY* 为第 3 类。同时大部分 *ArWRKY* 与已报道的 *WRKY* 转录因子在基因家族的进化起源和基因重复上比较相

近, 具有相近起源的蛋白往往具有较为相似的结构与较为相近的功能作用, 因此, 综合考虑了系统进化树的分支情况及 *ArWRKY* 基因 CDS 序列的完整性, 筛选得到 *ArWRKY1*、*ArWRKY5*、*ArWRKY6*、*ArWRKY8*、*ArWRKY9*、*ArWRKY11*、*ArWRKY13*、*ArWRKY19*、*ArWRKY20* 与 *ArWRKY23* 用于后续的研究。



加红色星号的为筛选得到的 *ArWRKY*

Those with red asterisk are *ArWRKY* obtained by screening

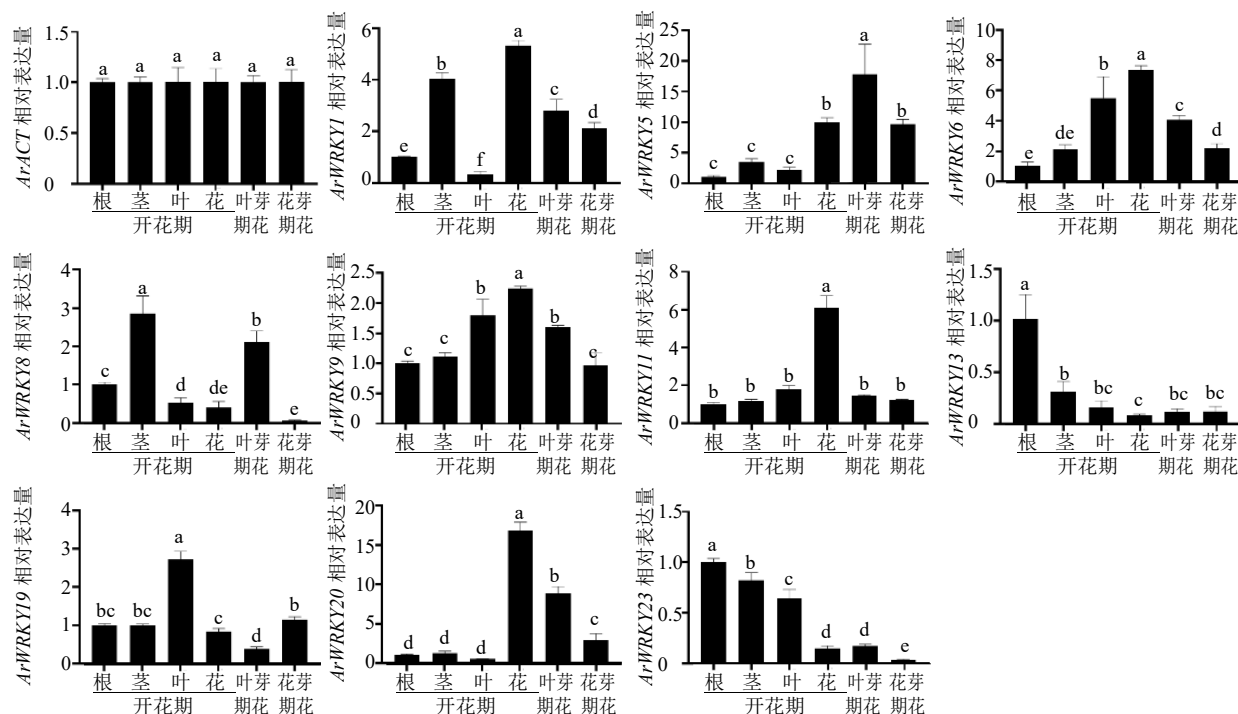
图 3 *ArWRKY* 系统进化树

Fig. 3 *ArWRKY* phylogenetic tree

3.2 *ArWRKY* 的时空表达分析

利用 qRT-PCR 对 *ArWRKY1*、*ArWRKY5*、*ArWRKY6*、*ArWRKY8*、*ArWRKY9*、*ArWRKY11*、*ArWRKY13*、*ArWRKY19*、*ArWRKY20* 与 *ArWRKY23* 基因在金线莲不同发育时期、不同组织部位的表达量进行检测 (图 4), 进一步确定 *ArWRKY* 可能存在的功能, 以筛选与金线莲胚胎发育相关的基因。通过时空特异性表达分析可以发现, 10 条 *ArWRKY* 在金线莲不同部位及不同时期的花中均有表达, 但表达模式不尽相同。*ArWRKY1*、*ArWRKY6*、*ArWRKY9*、*ArWRKY11* 与 *ArWRKY20* 的表达量随着

金线莲的发育呈下降趋势; *ArWRKY5*、*ArWRKY8* 与 *ArWRKY23* 的表达量随着花的发育先上升后下降; *ArWRKY13* 呈上升趋势; *ArWRKY19* 先下降再上升。结合金线莲开花期不同组织表达分析显示, *ArWRKY5* 与 *ArWRKY20* 在不同时期的花中表达量均显著高于开花期的根、茎、叶, 而 *ArWRKY13* 与 *ArWRKY23* 在不同时期的花中表达量均低于开花期的根、茎、叶。说明 *ArWRKY5* 与 *ArWRKY20* 对金线莲的花的形成及发育可能具有重要的调控作用。*ArWRKY13* 与 *ArWRKY23* 与金线莲开花呈负相关, 说明它们也可能参与了金线莲花的形成与发育。



不同小写字母表示存在显著差异 ($P < 0.05$)

Different lowercase letters indicate significant differences ($P < 0.05$)

图 4 *ArWRKY* 基因在不同组织的表达

Fig. 4 Expression of *ArWRKY* gene in different tissues

3.3 *ArWRKY5* 与 *ArWRKY20* 基因克隆

时空表达模式结果显示, *ArWRKY5* 与 *ArWRKY20* 在金线莲花中的表达量显著高于根、茎和叶, 暗示它们可能参与了金线莲的花形成及发育过程。为进一步研究 *ArWRKY5* 与 *ArWRKY20* 的功能, 本课题组以金线莲转录组测序拼接得到的 CDS 序列设计引物, 成功从金线莲中克隆得到 *ArWRKY5* 与 *ArWRKY20* (图 5)。测序结果显示 *ArWRKY5* 为 1149 bp (图 6), 与转录组数据中的序列一致; 而 *ArWRKY20* 的 CDS 序列为 600 bp (图 7), 虽然编码的蛋白包含了 WRKY 基本的结构域, 但是与转录组数据的序列相比中间部分丢失了一段 60 bp 的核苷酸序列, 猜测可能是转录组测序的结果拼接过程出错或者发生 mRNA 可变剪接。

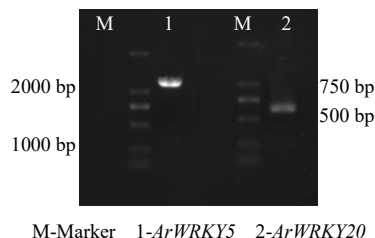


图 5 *ArWRKY5* 和 *ArWRKY20* PCR 产物电泳图

Fig. 5 Electrophoresis of *ArWRKY5* and *ArWRKY20* PCR products

```

1  ATGCAGGTAATGGATGCAAAAATGGATTCTACTAAGCAACAGGCAAGAGAAACAATAGTG
M  Q V M D A K M D S T K Q Q A R E T I V
61  CCTAGGCGATCTTGGAGCTTGGGCTTGGATTATGGCGGCCGCGATCTGGACAAAT
P  R Q F L E L G P G L M A A P A D L D N
121 GATTGCTTGGTGGATGGAGAGAGCAAGGAGCAGTCGATATCAATCAACCAATCTTATT
D  S L S D G E S K E Q S I S Q S T N L I
181 AGGCGGCTGCTCAAGATGACTGTTTCTCAAGATTAACTATGGTGATCTGATCATCAT
R  P A A Q D D C F L K I N Y G A S D H H
241 TCGGAAAGGATGTGTTCCTTCGGCATCCATGATCTGAGTTGCGGGAAGATGAC
S  G K D V V P S G I H D L R F A G R S D
301 CGAGGACTGTTGACAACTGAGCTGGAGTGGCTGGAGCCCAAGCAATAAGTTCGCT
R  G L V R Q L S L E S G W S P G N K V A
361 GAATTATCTCCCTCCAGAACTAGTGGCGGCAAGCAACAGCAACAAAAGTGCAA
E  L S P S Q N T E P Q Q Q Q Q Q Q K V Q
421 GAGGCTACTATGAGAAAACCTGTTGTTCTGTTGAGCTGTTCTGAAGCTTCATGATT
E  A T M R K T R V S V R A R S E A S M I
481 ACCGACGGATGCCAGTGGAGGAGTACGGCGAGAAGATGGCGAAGGAAACCCCTTGCCCT
T  D G C Q W R K Y G Q K M A K G N P C P
541 CGAGCATACTATGCTGCACTATGCTGCGGGCTGCCGGTTGCAAAACAGGTCACGGA
R  A Y Y R C T M A A G C P V R K Q V Q R
601 TGCGCGACGACCGCACTATCTGTTGAGAGCTACGAAAGGAAACCAACCAACCGCTC
C  A D D R T I L V T T Y E G N H N H P L
661 CGCGCGCGCCATGGCCATGGCTCCCACTTCCGCGCGCCCTCTATGCTCTCTCT
P  P A A M A M A S T T S A A A S M L L S
721 GGCTCCATGCCAGCGCGACGGCTCTGAGGCTCCCACTTTTATGACACAGGCGCTC
G  S M P S A D G L L S A P N F L A Q A V
781 CTCCTCTGCTCTCCAGCTCGCCACCATATCGGCTCTCTACCCCTTCCCAACCGTACCC
L  P C S S S L A T I S A S S P F P T V T
841 CTCGACCTCAACCACTCCCAACCCCTCCAGCTACATAAATCCCACTCTCTCGTTTC
L  D L T H S P N P L Q L H K L P H S P F
901 TGCGTCCCAAAAATATCTTAGGCAATCCAGGTTCTCGGTTCTCACTGCTGCTGCGGA
C  V P Q N I L G Q S R F S G L Q L P A G
961 GACCAGAAAGCGCAGTCGCTTGCAGATACGGTGAGCGCGCAACGCGGCCATACAGCT
D  Q K P Q S L A D T V S A A T A I T A
1021 GATCCGAACCTTACGCGAGCGCTGGCGGCGAGCAATACGTCATCATTTGGGGTGGAGTC
D  P N F T A A L A A A I T S I I G G G V
1081 GGAGGAGGGATGAAACCGCGGCTGTTTACGGGTGGAGTCAAGCAAGTGGGAGCGGTGAA
G  G G D E T G G V Q G G S Q A S G S G E
1141 ACCAAGTAG
T  K *
    
```

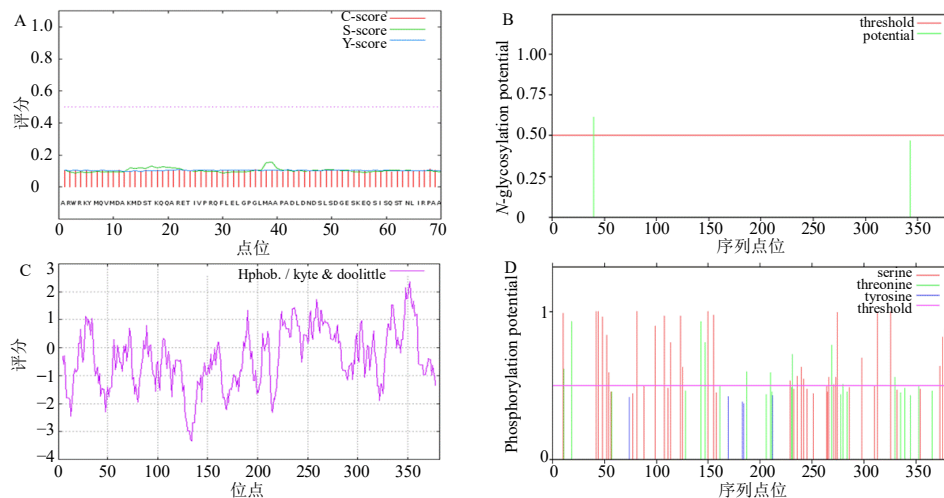
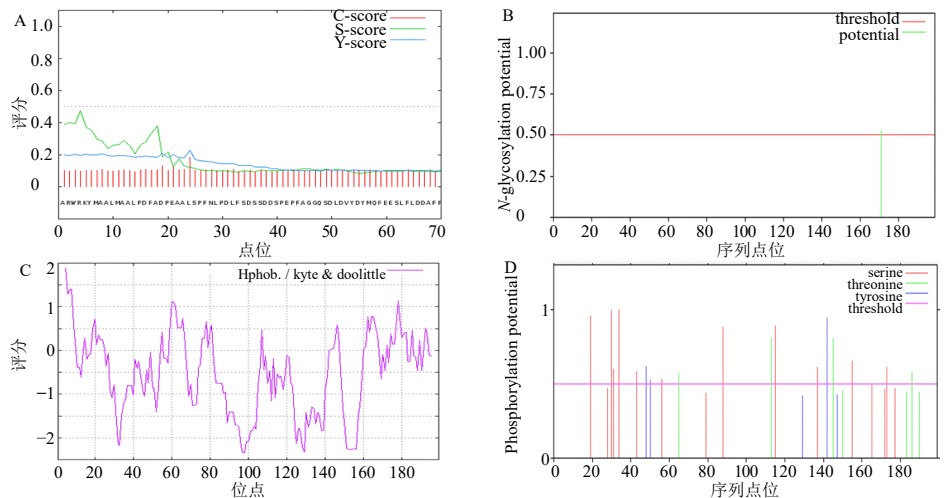
图 6 *ArWRKY5* CDS 序列及其对应的氨基酸图

Fig. 6 *ArWRKY5* CDS sequence and its corresponding amino acid map

```

1  ATGGGGCTCTCATGGCGGCTCTCCCGACTTGGCGGATCCAGAAGCGCGCTCTCTCC
   M A A L M A A L P D F A D P E A A L S P
61  TTCAACTTACGAGATCTCTTCCGACAGCTCCGATGACTCTCCGAGCCTTTCGCGGC
   F N L P D L F S D S S D D S P E P F A G
121  GGGCAGTCTGACCTCGAGCTCTACGACTATATGAGTGAAGATCCCTGTTCTCGAC
   G Q S D L D V Y D Y M Q F E E S L F L D
181  GACGCTTCTTACCTTCCACTAGAACGAGAACCGCGCGCGGTTCTGTGGCGG
   D A F F T F P L E P E E P A P P V L S A
241  AACGCGAACCGCGCGAGGAGCGCGAAGGCGATCGGATGAATGGGGAGGCGAAGAG
   N A N A G Q G S A E G H R M N G E A K R
301  AGAAGGGTGGAGGCGAGTTCAGGATAGGATTCAAGACTAAATCGACGTTGGAATAATG
   R R V E A E F R I G F K T K S D V E I M
361  GACGATGGCTTCAAGTGAAGATGGAAGAGCGGTTGAAGAACGCCAATCCA
   D D G F K W R K Y G K K A V K N S P N P
421  AGGTATGTGATAACGATATACGAAAGAACGACAAATCAGAGAGCCAGATGGCGGTAA
   R Y V I T I I Y E G T H N H E S P D G G N
481  CTTCCATTTATTCATGTGAATTGAAGCAATCTCTGCTGGGCTTCTCGCGGCCAA
   L P F I S C E F E A N S S V W A S C G Q
541  ATATTCACGACTTTCGCAAGGCTTTCGTAACCTTCAGTTTGGCGCGCTTCATAG
   I F T D F T Q G F T C N F Q F A P P S *

```

图 7 *ArWRKY20* CDS 序列及其对应的氨基酸图Fig. 7 *ArWRKY20* CDS sequence and its corresponding amino acid map图 8 *ArWRKY5* 蛋白信号肽 (A)、糖基化位点 (B)、亲水性/疏水性 (C) 及磷酸化位点 (D) 预测Fig. 8 *ArWRKY5* protein signal peptide (A), glycosylation site (B), hydrophilicity/hydrophobicity (C) and phosphorylation site (D) prediction图 9 *ArWRKY20* 蛋白信号肽 (A)、糖基化位点 (B)、亲水性/疏水性 (C) 及磷酸化位点 (D) 预测Fig. 9 *ArWRKY20* protein signal peptide (A), glycosylation site (B), hydrophilicity/hydrophobicity (C) and phosphorylation site (D) prediction

3.4 *ArWRKY5* 与 *ArWRKY20* 的生物信息学分析

利用在线网站对 *ArWRKY5* 与 *ArWRKY20* 信号肽、蛋白的糖基化位点、磷酸化位点、蛋白激酶结合位点和蛋白亲/疏水性进行预测 (图 8、9)。发现 *ArWRKY5* 蛋白不存在信号肽, 虽然预测图上出现可能存在的 2 个糖基化位点, 但是因为没有信号肽的蛋白质不太可能暴露于 *N*-糖基化机制下, 因此 *ArWRKY5* 应不存在糖基化位点; *ArWRKY5* 呈亲水性, 存在 37 个磷酸化位点, 其中 27 个是丝氨酸, 10 个是酪氨酸。预测结果发现 *ArWRKY20* 同样不存在信号肽和糖基化位点, 且为亲水性蛋白, 存在 18 个磷酸化位点, 其中 11 丝氨酸、3 个酪氨酸、4 个苏氨酸。

利用 GOR4 进行蛋白的二级结构预测, 结果如图 10、11 所示, 可以发现 ArWRKY5 和 ArWRKY20 均由 α -螺旋(alpha helix)、延伸链(extended strand)、无规则卷曲(random coil)所组成的, 并且分布于

整个蛋白。其中 ArWRKY5 中 α -螺旋占 30.89%, 延伸链占 14.14%, 无规则卷曲占 54.97%; ArWRKY20 中 α -螺旋占 22.16%, 延伸链占 13.07%, 无规则卷曲占 64.32%。蛋白三级结构预测见图 11。

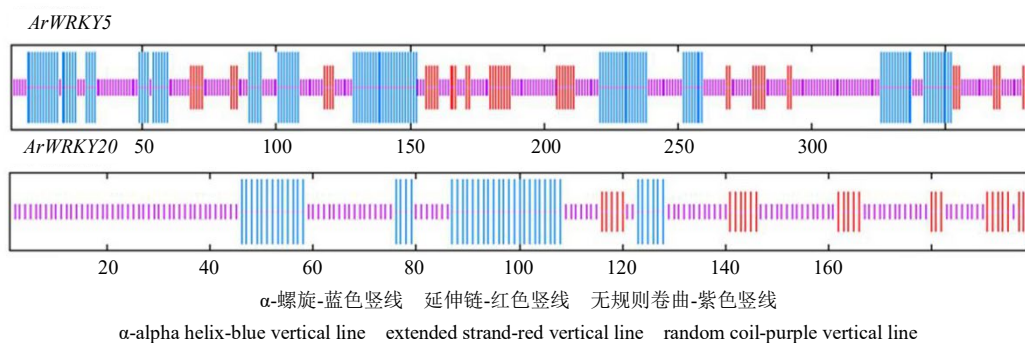


图 10 ArWRKY5 和 ArWRKY20 二级结构预测

Fig. 10 Predicted secondary structure of ArWRKY5 and ArWRKY20

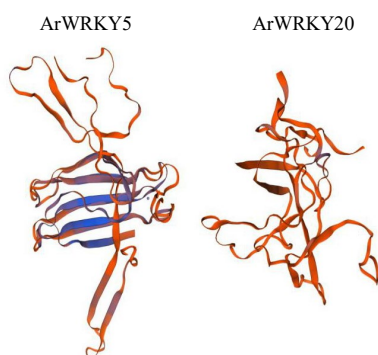


图 11 ArWRKY5 和 ArWRKY20 三级结构预测

Fig. 11 Prediction of tertiary structure of ArWRKY5 and ArWRKY20

4 讨论

金线莲自然结实率低、种子败育率高和幼苗死亡率高原因是导致野生金线莲资源匮乏的重要因素。金线莲的种胚结构不健全且不含胚乳, 由于长期的无性繁殖, 引起了金线莲种质退化、抗逆性差等现象, 导致产量和品质下降, 最终影响金线莲的药材质量^[5]。因此解决金线莲胚胎发育问题是金线莲优良品种选育的目标之一。筛选对金线莲胚胎发育具有调控作用的 ArWRKY 转录因子对于明确金线莲种子胚胎发育的分子机制具有重要意义。

为从金线莲中筛选调控胚胎发育相关的 WRKY 转录因子, 本研究利用 Illumina 结合 PacBio 技术对金线莲进行转录组测序, 首次从金线莲转录组数据库中筛选得到 23 条 ArWRKY 序列。根据 ArWRKY 核心序列比对结果与 WRKY 蛋白结构类型将 ArWRKY 分为 3 大类^[20]。将 ArWRKY 与其他

物种中已知功能的 WRKY 一同构建进化树, 发现: ArWRKY6 与 AtWRKY10 和 AtTTG2_X1 相距最近; ArWRKY8 与 AtTTG2_X1 相距最近。已报道 AtWRKY10 能促进拟南芥种子的胚乳发育, 并且通过调控种子腔的发育来影响拟南芥种子的大小^[21]; AtTTG2 则能调控拟南芥种子种皮中单宁物质和粘液的生成, 从而影响种皮发育^[11]。这说明 ArWRKY6 和 ArWRKY8 可能与金线莲种子大小、胚乳发育和种皮发育相关。除此之外, ArWRKY9 与 ArWRKY13 和 VvWRKY20_X2 相距最近; ArWRKY11 与 AtWRKY43 相距最近。VvWRKY20 参与了葡萄胚珠发育过程, 缺少 VvWRKY20 的调控作用会导致葡萄胚珠败育^[12]; AtWRKY43 可以调控拟南芥种子中的脂肪动员过程, 促进糖类的积累, 从而增强种子的发芽能力^[10]。这些结果也暗示 ArWRKY9、ArWRKY11 和 ArWRKY13 可能具有同样或类似的功能。另外, ArWRKY1 与 GhWRKY17_X2 的分支相距最近; ArWRKY5 与 GhWRKY17_X3 相距最近; ArWRKY19 与 AtWRKY34 和 GhWRKY22 相距较近; ArWRKY20 与 VvWRKY20_X3 相距较近; ArWRKY23 与 GhWRKY17_X2 和 ThWRKY7 相距最近。聚类越近的蛋白可能在遗传关系上越密切, 进化树的聚类分析结果为进一步筛选调控金线莲胚胎发育的 ArWRKY 提供了一定参考依据。

本研究对 ArWRKY1、ArWRKY5、ArWRKY6、ArWRKY8、ArWRKY9、ArWRKY11、ArWRKY13、ArWRKY19、ArWRKY20 与 ArWRKY23 在金线莲的不同组织部位及金线莲不同时期花中的表达量进行

检测。发现 23 条 *ArWRKY* 在金线莲的不同组织与不同花期中表达模式不尽相同。其中, *ArWRKY5* 和 *ArWRKY20* 在金线莲不同时期的花中都显著高于根、茎、叶, 因此进一步确定了 *ArWRKY5* 和 *ArWRKY20* 可能存在调控金线莲花发育的功能, 进而影响金线莲的胚胎形成及种子发育。为验证 *ArWRKY5* 和 *ArWRKY20* 的功能, 本课题组利用金线莲转录组数据库的 Unigenes 克隆得到 *ArWRKY5* 和 *ArWRKY20* 的 CDS 序列, 其中 *ArWRKY20* 与转录组数据库相比中间减少了 60 bp 大小的序列, 猜测是由于转录组数据拼接时出错, 功能上是否存在差异还有待进一步研究。本研究首次对金线莲 *ArWRKY* 蛋白家族的组成、分类、进化、基序特点及结构特征进行了分析, 并筛选到在花中显著高表达的 *ArWRKY5* 和 *ArWRKY20*, 成功克隆得到 *ArWRKY5* 与 *ArWRKY20* 的 CDS 序列, 为后续的功能研究奠定基础。

利益冲突 所有作者均声明不存在利益冲突

参考文献

- [1] 沈廷明, 黄春情, 刘知远, 等. 林下仿生态种植金线莲的质量标准研究 [J]. 中草药, 2018, 49(2): 450-454.
- [2] 杨帆, 廉美兰, 李美兰, 等. 金线莲丛生芽培养及有效物质生产研究 [J]. 中草药, 2016, 47(18): 3284-3288.
- [3] 许梦洁, 叶申怡, 吴梅, 等. 不同种质金线莲氨基酸和矿物质元素量的比较 [J]. 中草药, 2017, 48(2): 368-372.
- [4] 邵清松, 叶申怡, 周爱存, 等. 金线莲种苗繁育及栽培模式研究现状与展望 [J]. 中国中药杂志, 2016, 41(2): 160-166.
- [5] 邵清松. 珍稀名贵药材金线莲 [M]. 北京: 中国农业出版社, 2018: 263.
- [6] 叶冰竹. 金线莲内生真菌促进宿主生长代谢及增强抗病性作用研究 [D]. 上海: 中国人民解放军海军军医大学, 2019.
- [7] 张凡, 尹俊龙, 郭瑛琪, 等. WRKY 转录因子的研究进展 [J]. 生物技术通报, 2018, 34(1): 40-48.
- [8] Guan Y F, Meng X Z, Khanna R, et al. Phosphorylation of a WRKY transcription factor by MAPKs is required for pollen development and function in *Arabidopsis* [J]. *PLoS Genet*, 2014, 10(5): e1004384.
- [9] Meng L S, Wang Y B, Loake G J, et al. Seed embryo development is regulated via an *AN3-MINI3* gene cascade [J]. *Front Plant Sci*, 2016, 7: 1645.
- [10] Geilen K, Heilmann M, Hillmer S, et al. WRKY43 regulates polyunsaturated fatty acid content and seed germination under unfavourable growth conditions [J]. *Sci Rep*, 2017, 7(1): 14235.
- [11] Johnson C S, Kolevski B, Smyth D R. TRANSPARENT TESTA GLABRA2, a trichome and seed coat development gene of *Arabidopsis*, encodes a WRKY transcription factor [J]. *Plant Cell*, 2002, 14(6): 1359-1375.
- [12] 李树秀, 张朝红, 巩培杰, 等. 无核葡萄 VvWRKY20 基因的克隆及其在胚珠败育过程中的差异表达分析 [J]. 农业生物技术学报, 2012, 20(10): 1143-1149.
- [13] Xiang J S, Tang S, Zhi H, et al. Loose Panicle1 encoding a novel WRKY transcription factor, regulates panicle development, stem elongation, and seed size in foxtail millet [*Setaria italica* (L.) P. Beauv.] [J]. *PLoS One*, 2017, 12(6): e0178730.
- [14] Zhang C Q, Xu Y, Lu Y, et al. The WRKY transcription factor OsWRKY78 regulates stem elongation and seed development in rice [J]. *Planta*, 2011, 234(3): 541-554.
- [15] Wang Y, Li Y, He S P, et al. A cotton (*Gossypium hirsutum*) WRKY transcription factor (GhWRKY22) participates in regulating anther/pollen development [J]. *Plant Physiol Biochem*, 2019, 141: 231-239.
- [16] Yan H R, Jia H H, Chen X B, et al. The cotton WRKY transcription factor GhWRKY17 functions in drought and salt stress in transgenic *Nicotiana benthamiana* through ABA signaling and the modulation of reactive oxygen species production [J]. *Plant Cell Physiol*, 2014, 55(12): 2060-2076.
- [17] 曲庆玲. 人参种子后熟分子机理的研究 [D]. 长春: 吉林大学, 2020.
- [18] 潘园园. DnWRKY11 转录因子克隆及其功能分析 [D]. 杭州: 杭州师范大学, 2013.
- [19] 朱建军, 黄雨佳, 金建红, 等. 不同栽培基质对金线莲 3 种基原植物生长及其活性成分含量的影响 [J]. 中国中药杂志, 2019, 44(12): 2467-2471.
- [20] Li H E, Xu Y, Xiao Y, et al. Expression and functional analysis of two genes encoding transcription factors, VpWRKY1 and VpWRKY2, isolated from Chinese wild *Vitis pseudoreticulata* [J]. *Planta*, 2010, 232(6): 1325-1337.
- [21] Kang X, Li W, Zhou Y, et al. A WRKY transcription factor recruits the SYG1-like protein SHB1 to activate gene expression and seed cavity enlargement [J]. *PLoS Genetics*, 2013, 9(3): e1003347.

[责任编辑 时圣明]