

doi:10.11937/bfyy.20170279

银杏两个 WD40 转录因子基因克隆及序列分析

冯 如, 李泽宏, 袁红慧, 程 华, 程水源

(武汉轻工大学 生物与制药工程学院, 湖北 武汉 430023)

摘 要:WD40 转录因子也称 WDR(WD repeat)蛋白,是一种古老蛋白家族,参与调节类黄酮次生代谢途径中多种酶基因的表达,从而影响类黄酮的积累。以三年生银杏苗为试材,采用 RT-PCR 和 RACE-PCR 方法,研究了银杏 WD40 基因序列,以期深入了解银杏叶类黄酮合成代谢的分子机理。结果表明:*GbWD401* 序列包含一段 2 313 bp 的开放阅读框,编码一个 770 aa 的肽链;*GbWD402* 序列包含一段 1 605 bp 的开放阅读框,编码一个 534 aa 的肽链。银杏 *GbWD401* 和 *GbWD402* 基因推测编码蛋白序列均含有典型 WD40 重复基序。*GbWD401* 蛋白序列与无油樟 *AtWD401* 相似性为 49%,*GbWD402* 蛋白序列与无油樟 *AtWD402* 相似性为 78%。系统进化树发现,*GbWD401* 与无油樟 *AtWD401*、荷花 *NnWD40* 划归为同一分支;而 *GbWD402* 与无油樟 *AtWD402* 划归为另一分支。

关键词:银杏;类黄酮;WD40 转录因子;基因克隆

中图分类号:S 662.203.6 **文献标识码:**A **文章编号:**1001-0009(2017)18-0100-09

银杏(*Ginkgo biloba* L.)属裸子植物银杏科银杏属落叶乔木,又名白果、公孙树^[1]。银杏叶具有很高药用价值,其主要有效成分为类黄酮和萜内酯。类黄酮是一类以 C₆-C₃-C₆ 为基本结构或具有苯基苯并吡喃结构的植物次生代谢物^[2],银杏叶中类黄酮主要包括单黄酮、双黄酮、儿茶素等三大类^[3],具有抗氧化、清除自由基、消炎、抗病毒、抗癌和治疗免疫缺陷等保护人类健康的多种保健与药用功能^[4-13]。

在植物体内类黄酮生物合成是由多种生物酶共同参与调控的复杂过程,其产物积累也不是由 1~2 个关键酶所能控制^[14]。研究转录因子对黄酮合成调控作用,是提高类黄酮代谢通路的另一个途径。WD40 转录因子可通过激活类黄酮等次生代谢途径中多种酶基因的表达,有效启动或关闭次生代谢流量,从而调节类黄酮的形成^[15]。WD40 是一类古老的蛋白家族,其结构高度保守,一般含有 4~16 个串联重复的 WD 基元,每个 WD 基元由 40 个氨基酸残基组成^[16]。目前多个与类黄酮合成相关的 WD40 转录因子基因得到分离与鉴定,如紫苏 *PFWD*、玉米 *ZmPAC1*、马铃薯 *STWD40*、苜蓿 *WtWD40-1*、苹果 *MdTTG1* 以及葡萄 *WDR1* 和 *WDR2*,但在银杏转录因子的研究中尚鲜见 WD40 类转录因子的报道^[17-22]。

该研究从已获得的银杏转录组数据中筛选出 WD40 基因核苷酸拼接片段,采用 RACE-PCR 获得了 2 个 WD40 基因 *GbWD401* 和 *GbWD402* 全

第一作者简介:冯如(1992-),男,硕士研究生,研究方向为生药学。E-mail:799809864@qq.com.

责任作者:程水源(1965-),男,博士,教授,研究方向为经济林种质资源评价与利用。E-mail:s_y_cheng@sina.com.

基金项目:国家自然科学基金资助项目(31270717,31400556);湖北省自然科学基金重点资助项目(2013BHE029)。

收稿日期:2017-05-16

长 cDNA 序列,对其序列进行生物信息学分析,探讨该转录因子序列及结构特点,以期为进一步阐明该转录因子在银杏类黄酮生物合成途径中的分子作用机制奠定基础。

1 材料与方法

1.1 试验材料

供试植物采于武汉轻工大学银杏种植园,在相同生长环境下取新鲜银杏叶片,经无菌水洗净后,由液氮快速冷却后置于 $-80\text{ }^{\circ}\text{C}$ 保存。

植物总 RNA 快速提取试剂盒、琼脂糖凝胶 DNA 回收纯化试剂盒购自北京艾德莱(Aidlab)生物科技有限公司;PrimeScriptTM RT Reagent Kit with gDNA Eraser 反转录试剂盒、TaKaRa LA Taq[®]试剂盒、PMD18-T Vector、反转录试剂盒 SMARTTM RACE cDNA Amplification kit 购自 Clontech。

1.2 试验方法

1.2.1 总 RNA 提取与反转录

取银杏叶片样本 0.45 g 于液氮中充分研磨,利用植物总 RNA 快速提取试剂盒提取 RNA,采用超微量紫外分光光度计检测 RNA 的浓度和纯度,1%琼脂糖凝胶电泳检测 RNA 完整性。取银杏叶片总 RNA 1 μg ,按照 TaKaRa 公司 PrimeScriptTM RT Reagent Kit with gDNA Eraser 试剂盒说明书进行反转录反应,分别合成 cDNA 第一链,于 $-20\text{ }^{\circ}\text{C}$ 保存。

1.2.2 *GbWD401* 和 *GbWD402* 基因的克隆

分别设计 *GbWD40* 基因特异性引物 GbWD40s 和 GbWD40a(表 1),以银杏叶 cDNA 为模板,扩增 *GbWD40* 基因 cDNA 中间片段。根据已知中间片段序列设计 *GbWD40* 基因 3'端特异性引物 GbWD403、GbWD403N,通过 RT-PCR 以及巢式 PCR,从银杏 cDNA 中克隆出 *GbWD40* 基因 3'端。第一轮扩增用引物 GbWD403 和 RACE kit 通用引物(UPM),3'-RACE PCR 第二轮巢式扩增以第一轮产物稀释 10 倍为模板,利用引物 GbWD403N,结合 RACE kit 的巢式通用引物(NUP)完成。PCR 扩增第一轮反应体系:3'-RACE-Ready cDNA 2.5 μL 、UPM (10 \times) 5 μL 、

GbWD403 (10 mmol \cdot L⁻¹) 1 μL 、Advantage 2 PCR buffer (10 \times) 5 μL 、dNTP (10 mmol \cdot L⁻¹) 1 μL 、Advantage 2 Polymerase Mix 1 μL ,补水至 50 μL 。反应条件为:94 $^{\circ}\text{C}$ 3 min;94 $^{\circ}\text{C}$ 30 s,70 $^{\circ}\text{C}$ 3 min,5 个循环;94 $^{\circ}\text{C}$ 30 s,70 $^{\circ}\text{C}$ 30 s,72 $^{\circ}\text{C}$ 3 min,5 个循环;94 $^{\circ}\text{C}$ 30 s,68 $^{\circ}\text{C}$ 30 s,72 $^{\circ}\text{C}$ 3 min,25 个循环;72 $^{\circ}\text{C}$,10 min。PCR 扩增第二轮反应体系:第一轮 PCR 产物 2.5 μL 、NUP (10 mmol \cdot L⁻¹) 1 μL 、GbWD403N (10 mmol \cdot L⁻¹) 1 μL 、Advantage 2 PCR buffer (10 \times) 5 μL 、dNTP(10 mmol \cdot L⁻¹) 1 μL 、Advantage 2 Polymerase Mix 1 μL ,补水至 50 μL 。第二轮反应条件为:94 $^{\circ}\text{C}$ 3 min;94 $^{\circ}\text{C}$ 30 s,68 $^{\circ}\text{C}$ 30 s,72 $^{\circ}\text{C}$ 3 min,35 个循环;72 $^{\circ}\text{C}$ 10 min。按照同样方法利用 5'端特异性引物 GbWD405、GbWD405N,从银杏 cDNA 中克隆出 *GbWD40* 基因 5'端。拼接比对 *GbWD40* 基因 3'-RACE 和 5'-RACE 测序结果,根据拼接序列设计特异性引物 GbWD40UR 和 GbWD40TR 进行全长 cDNA PCR 扩增。1%琼脂糖凝胶电泳对扩增产物进行检测,对 PCR 扩增所得特异目的片段回收纯化。将回收所得 DNA 片段与 PMD18-T 载体连接后,送往公司测序。

表 1 RACE-PCR 所用引物

Table 1 Primers used in RACE-PCR

引物名称 Primer name	序列 Sequence(5'-3')
GbWD401s	AAGAGATTCAAGCTCACSAKGGTT
GbWD401a	CATCCTACTGGACTTCTTCACTACA
GbWD4013	AAAGTACGAATTTGGAGCATCCCCGA
GbWD4013N	AGTCGTTGATTGGAGCGATGTA
GbWD4015	TGGCTTATCTGCGTGACCTTCAACATT
GbWD4015N	GATTCTATCACTTCCGACACAT
GbWD401UR	TGTTGCGAATACTCCAAAACCAGG
GbWD401TR	GCTGATTCAAACATGGATTGCTAGG
GbWD402s	GCAAGTGTTCCGAAAACCTTTATAC
GbWD402a	TGAAAATCAAGCCGATAGAGGTT
GbWD4023	CTTTTGACTCATTCTGGACCTCGCCTT
GbWD4023N	TCTGTGGGACTTGGTAATGGAT
GbWD4025	AAGGCAAAGGTTGTTGGACTGGCTTAT
GbWD4025N	TTCAAATGAGCCAGTGTAGCCT
GbWD402UR	TGTGGTATTTAGTGTAGCAAACGAG
GbWD402TR	ATAATGAAATGACAACCAGCAGCAC

1.2.3 *GbWD401* 和 *GbWD402* 基因的生物信息学分析

依据 *GbWD401* 和 *GbWD402* 基因测序结果,利用 DNAstar 中 EditSeq 软件获取基因序列的开放阅读框(open reading frame, ORF)和其所编码氨基酸序列;使用在线工具 BLAST(<http://blast.ncbi.nlm.nih.gov/Blast.cgi>),将获得的核苷酸和氨基酸序列与 NCBI 中已登录的其它物种进行序列比对和同源性分析,并利用 MEGA 7.0 软件构建系统进化树;借助 ExPASy 网站中在线工具 ProtParam(<http://web.expasy.org/prot-param/>)进行蛋白质基本理化性质预测;采用(https://npsa-prabi.ibcp.fr/cgi-bin/npsa_auto-mat.pl?page=/NPSA/npsa_sopma.html)预测蛋白质二级结构;采用在线工具 SWISS-MODEL(<http://swissmodel.expasy.org/>)对蛋白质三维结构建模;蛋白质亚细胞定位运用在线程序 WOLF PSORT(<http://wolfsort.org/>);蛋白质序列信号位点使用 Signal P4.1 Server(<http://www.cbs.dtu.dk/service/SignalP/>);运用在线工具 TMHMM 2.0(<http://www.cbs.dtu.dk/service/TMHMM-2.0/>)预测蛋白质跨膜区结构域。

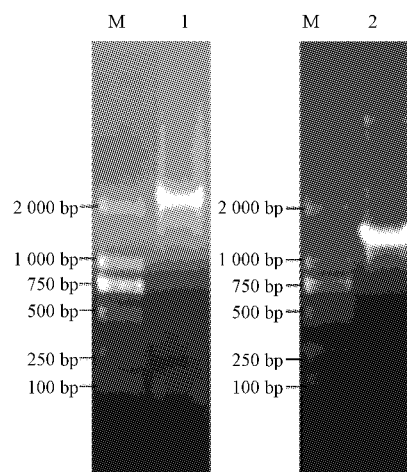
2 结果与分析

2.1 *GbWD401* 和 *GbWD402* 转录因子全长基因克隆

以银杏叶片 cDNA 为模板,利用特异性引物对 *GbWD401* 和 *GbWD402* 基因 ORF 序列扩增,扩增产物经 1% 琼脂糖凝胶电泳检测,结果显示(图 1),分别获得长度约为 2 300 bp 和 1 600 bp 的明亮条带,与预期扩增片段大小基本一致。初步判定为银杏 WD40 转录因子 cDNA 全长片段。

2.2 *GbWD401* 和 *GbWD402* 基因测序结果及生物信息学分析

经软件分析, *GbWD401* 包含一段 2 313 bp 的开放阅读框,从翻译起始密码子 ATG 到翻译终止密码子 TGA 共 770 个氨基酸,分子式为 $C_{3745}H_{5879}N_{1059}O_{1185}S_{30}$,相对分子质量为 85.66 kD,理论等电点(pI)为 5.88。 *GbWD402* 包含一段



注: M, DNA Marker(DL 2 000); 1, *GbWD401* 基因 PCR 产物; 2, *GbWD402* 基因 PCR 产物。

Note: M, DNA Marker (DL 2 000); 1, PCR production of *GbWD401*; 2, PCR production of *GbWD402*.

图 1 RACE 扩增产物电泳结果

Fig. 1 Electrophoresis results of RACE

1 605 bp 的开放阅读框,从翻译起始密码子 ATG 到翻译终止密码子 TAA 共 534 个氨基酸,分子式为 $C_{2518}H_{3879}N_{705}O_{786}S_{18}$,相对分子质量为 57.18 kD,理论等电点(pI)为 6.8。

氨基酸序列分析发现,转录因子 *GbWD401* 包含 6 个 WD40 重复基序, *GbWD402* 包含 5 个 WD40 重复基序,符合 WD40 转录因子的结构特征,因此可推断所获得的 cDNA 序列为转录因子 WD40 基因。将推导的氨基酸序列在 NCBI 中进行 Blast 比对,结果发现 *GbWD401* 编码的氨基酸序列与无油樟 *Amborella trichopoda* (XP_006857843.1)、毛果杨 *Populus trichocarpa* (XP_002317512.1)、胡桃 *Juglans regia* (XP_018851091.1) 编码的氨基酸相似性分别为 49%、51%、52%(图 2)。 *GbWD402* 编码的氨基酸序列与无油樟 *Amborella trichopoda* (XP_006854341.2)、荷花 *Nelumbo nucifera* (XP_010264714.1)、梅 *Prunus mume* (XP_008243924.1) 编码的氨基酸相似性分别为 78%、76%、74%(图 3)。而 *GbWD401* 和 *GbWD402* 核苷酸序列相似性均在 72% 以上。

GbWD401	MDGFANEEDQFFETLEDIVSVSDSGSDVAETDLDTVKTDLYNASSKYEIWKSEPGSSRE	60
AtWD401	-----SRD	3
PtWD40	-----	1
JrWD40	-----	1
Chantal Coneseans		
GbWD401	VEEIRDVPFSPADSSGEFVCKIKNLDGKEFVVDLGGDGMWNKLEIGTDROLTDEEFE	120
AtWD401	LEEIN---SIDDDRENMVCRIRNLDNGIEFVLDNPSGDGVLSSLRVGSRLATIDEFQ	59
PtWD40	--SPKDGSSSVAAAGEVEVCTIKNLDNGKEFVVVEIREDCMWNKLEVGTGROLTMEFE	58
JrWD40	-----EKVCLIKNLDNGKEFVVVEIREDCMWNKLEVGTGROLTMEFE	44
Chantal Coneseans	*****	
GbWD401	RSLGVSPIVOELMRRSSAVELAEIKVED---HDVDCSK---RKKRGWLSLEYVHAMRR	174
AtWD401	RDIGLSPVQHFMR---DLVESGIS---EGLK---RKM-GWLSLRRAASIVER	105
PtWD40	MSVGHSPIVOELMRRONVEDGTRGNLDSNANGCIGGGVTKFAKNGSWFGSIRSVANSVTR	118
JrWD40	MSVGHSPIVQELMRRONVEEGNKDSSDSNANG-VDGSGAKMKKGSWFGSIRSVASSVTG	103
Chantal Coneseans	*****	
GbWD401	HKDDRIIND---SENSRRSGSFT---DVS---SDRFORHVRHYKESCKEFTALYKGOEI	225
AtWD401	KKGGGSSSR---SCNSENINGSN---SVNPIHQMPORVRVHCKKCKEFTALYLGQEI	159
PtWD40	HKERRSSDERDTGSEKGGRRSSSATDSSQDVSTHGPHERVVRQVGSSEKELALYKGOEI	178
JrWD40	QKERRSSDERDTSSSEKGGRRSSSATDSSQDVSTHGPHERVVRQVGSSEKELALYKGOEI	163
Chantal Coneseans	*****	75
GbWD401	QAHEGSIWIMKFSLDGRYLASAGCDRVIHVWQVSESERK---FGAGVLDDNASNTVFIVNG	284
AtWD401	QAHEGSIWIMKFSLDGRYLASAGGCDRVIHVWQVSESAGT---ELGDTLEIDRSCTVFIVNE	217
PtWD40	QAHNGSIWISIKFSLDGRYLASAGEDCVIHVWQVSESERKCELLMKFPDDGGLNLLLIANG	238
JrWD40	QAHNGSIWISIKFSLDGRYLASAGEDCVIHVWQVSESERKCELLMKFPDDGGLNLLLIANG	223
Chantal Coneseans	*****	110
GbWD401	SSE--LVPLNVEGHADKPERGKIKGSKSLNTDCMVLDPDKVFLQSEKPLCSFGHLLDDVL	342
AtWD401	SSE--LVPL---VANKEQCGFRGLRSESGACVVIPIKIPKLSKPLHEFHCHTGDVL	271
PtWD40	SPEPNLLSPLVDSHLEKRRGRSSISRSLSLDHIFVPIVPSLIDKPLCSFGHLLDDVL	298
JrWD40	SPEPTLLSPSMDNHPEKRRGRSSISRSLSLDHIFVPIVPSLIDKPLCSFGHLLDDVL	283
Chantal Coneseans	*****	141
GbWD401	DLWSRSQVLLSSSMDKTVRLWHVGCCTCLKTFSHNDVTCIOFNPDDRVFISGSLDGK	402
AtWD401	DLWSQVLLSSSMDKTVRLWVGKASCKEFPSHNDVTCIOFNPVDDRVFISGSLDGK	331
PtWD40	DLWSRSQVLLSSSMDKTVRLWHVGCCTCLKTFSHNDVTCIOFNPVDDRVFISGSLDAK	358
JrWD40	DLWSRSQVLLSSSMDKTVRLWHVGCCTCLKTFSHNDVTCIOFNPVDDRVFISGSLDAK	343
Chantal Coneseans	*****	193
GbWD401	VRIWSIPDRQVVDWSDVREMTAACTPDGCGGLVGSINGNCRFYNTSGNKLQLETOGV	462
AtWD401	VRIWIPDRQVVDWSDIREMTAACTPDGCGGLVGSINGNCRFYNTSGNKLQLETOGV	391
PtWD40	VRIWSIPDRQVVDWSDVREMTAACTPDGCGGLVGSINGNCRFYNTSGNKLQLETOGV	418
JrWD40	VRIWSIPDRQVVDWSDVREMTAACTPDGCGGLVGSINGNCRFYNTSGNKLQLETOGV	403
Chantal Coneseans	*****	238
GbWD401	HSEKKKSSSGKEITGFOFAPGDSQVLTISADSEIRVVDGVLDVCOYEGFRNTCSQISASF	523
AtWD401	QGGKA--SSCKEITGFOFAPEDPRVMTISADSEIRVVDGVLDVCOYEGFRNTCSQISASF	450
PtWD40	QKKKA--AHLKEITGFOFAPGSSSEVLITISADSEIRVIDGEDLVHEEGFRNTCSQISASF	477
JrWD40	QKKKA--HKEITGFOFAPGSSSEVLITISADSEIRVIDGVLDVHEEGFRNTCSQISASF	453
Chantal Coneseans	*****	234
GbWD401	TPNGKHVISAEDSQVVIWNVETIPIGP-AVRRKSSVWGHEHFLSQSVSAIPWFGAKLEP	581
AtWD401	TNGKHVISAEDSHVYVWVETIPEGR-TSEQAKSVWGCERFFSNSSVAITWGLCKSEA	509
PtWD40	TINGKHVISAEDSYVYVWVHEADSRLSRSEGYTITRSEHFLCQDYSVAIPWFGADTW	537
JrWD40	TNGKHVISAEDSCVYVWVHEADSRPNRSECTVTITRSEHFLCQDYSVAIPWGLGDSW	523
Chantal Coneseans	*****	514
GbWD401	LALSSQEMMPGQRDNQPLQVGVDAWETNGFLREPAGGQELLETALQYVVSQRSQLCESQIS	541
AtWD401	RPSSEVSLKTNERISDELTE---SEENGPHS---	538
PtWD40	GLQDTLSGQNGLDNHLDEV---SIVNHPTP---	555
JrWD40	GLQDNYCREQNGLDNHLDEV---STANHPPTP---	552
Chantal Coneseans		
GbWD401	ACMVERTDLDDLSSGPCIPDSCQATFTALSASPRNLDHVNSSNVFFSIDGVSKGSATWPE	701
AtWD401	-----WDSRETPSVSSSP---DRFSLGHGFPADSLARGSATWPE	575
PtWD40	-----VEEASNEGSSSLTGCTNSPMANG-ISSATNGVFFDRIS---ATWPE	605
JrWD40	-----VEEING-GSQSAGFTNSSELGCTISSATNSTEFFDRIS---ATWPE	593
Chantal Coneseans	*****	
GbWD401	EKLPPSSNLSPSKLVASKTHYNILKTIGDITQCGSTVAAAANGLVIVTAGCGGEIRSF	761
AtWD401	EKLPTDALSSAQSTCKSQFKFLKT---ASEGMFASPHTWGLVIVTAGWDGEIRSF	629
PtWD40	EKLDLATRTRSHHASVDISN-----GLSQNVSAAGLVIVTAGLRGEIRTF	653
JrWD40	EKLLVPTNRNRSPPRVSDFFN-----GVNQNVSAAGLVIVTAGLRGEIRTF	639
Chantal Coneseans	*****	352
GbWD401	QNYGLPVRL	770
AtWD401	HNVGLPVRL	633
PtWD40	QNEGLPVRI	662
JrWD40	QNEGLPVRI	643
Chantal Coneseans	*****	361

图 2 GbWD401 基因编码蛋白的多序列比较

Fig. 2 Multiple sequence alignments of the GbWD401 protein from *Ginkgo biloba* L.

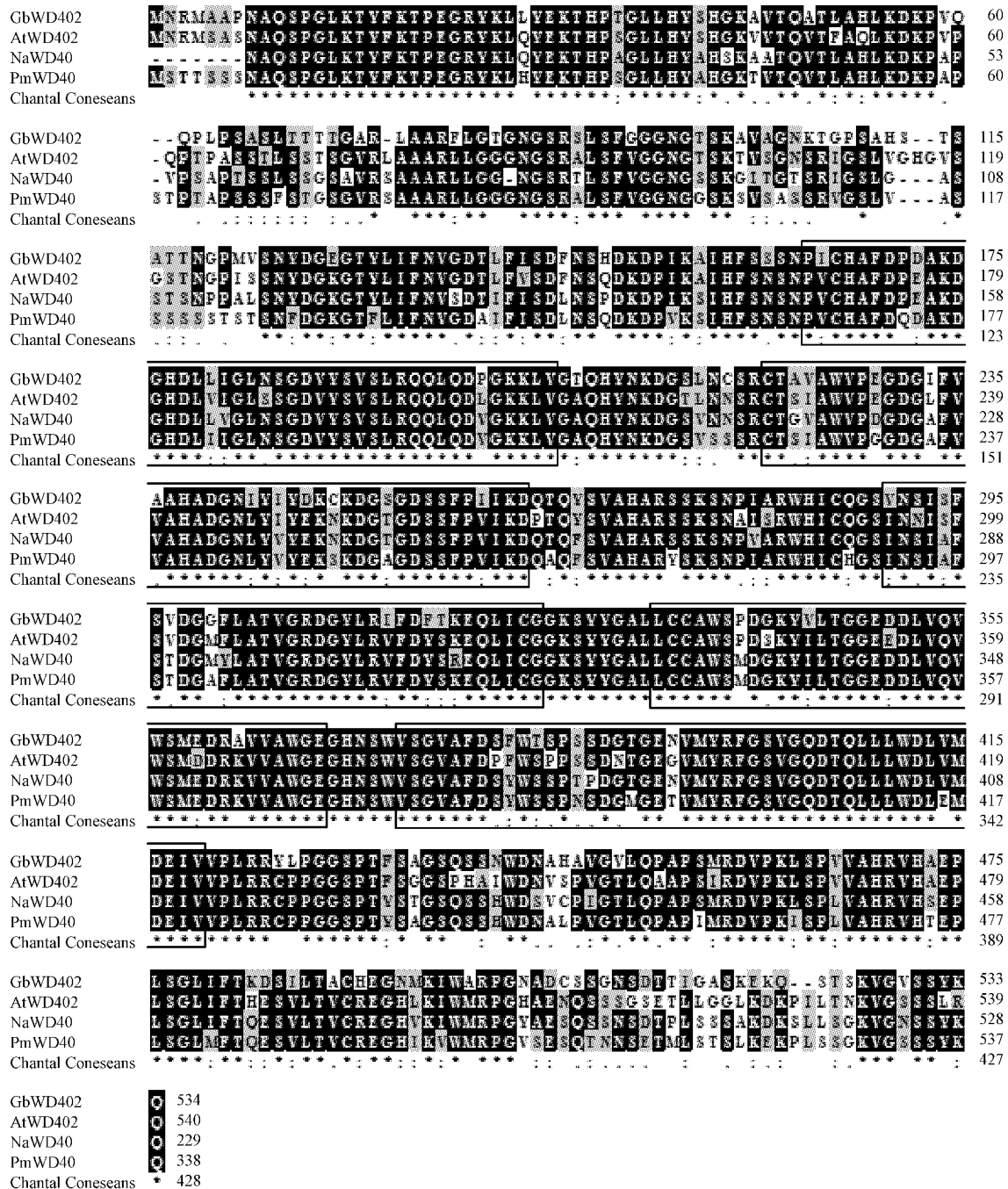


图3 *GbWD402* 基因编码蛋白的多序列比较(方框代表 WD40 重复基序)

Fig. 3 Multiple sequence alignments of the GbWD402 protein from *Ginkgo biloba* L.

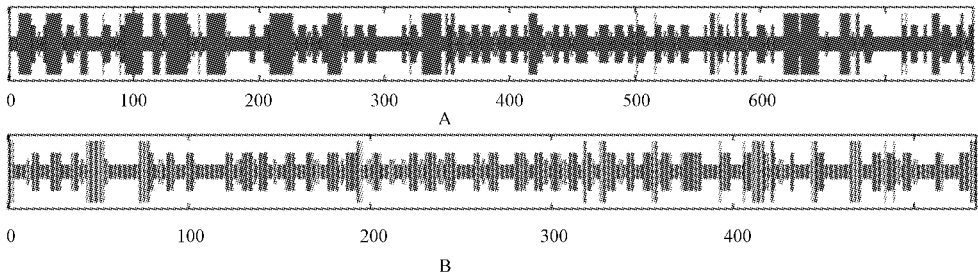
2.3 *GbWD401* 和 *GbWD402* 基因编码蛋白的二级结构和三级结构预测

利用 SOPMA 在线软件对 *GbWD401* 和 *GbWD402* 基因编码蛋白的二级结构进行预测,结果显示(图 4),*GbWD401* 转录因子蛋白由 27.66%

α -螺旋,25.58%延伸链,10.91% β 转角和 35.84% 无规则卷曲构成,表明 α -螺旋、延伸链和无规则卷曲是主要的二级结构元件。*GbWD402* 转录因子蛋白由 10.67% α -螺旋、34.46% 延伸链、11.42% β 转角和 43.45% 无规则卷曲构成,表

明延伸链和无规则卷曲是主要的二级结构元件。采用 SWISS-MODEL 工具,运用同源建模法对

GbWD401 和 *GbWD402* 基因编码蛋白进行三级结构建模(图 5)。



注:A. *GbWD401*;B. *GbWD402*。下同。

Note:A. *GbWD401*;B. *GbWD402*. The same below.

图 4 *GbWD40* 基因编码蛋白的二级结构

Fig. 4 Constructed 2D structure of *GbWD40* protein

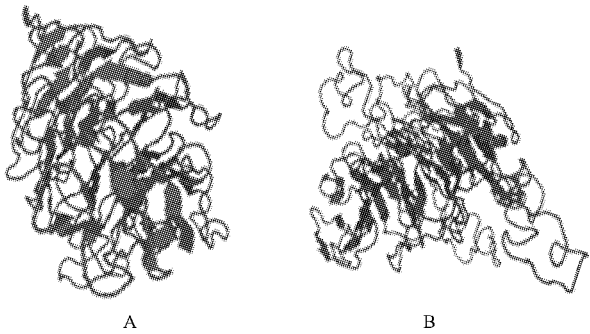


图 5 *GbWD40* 基因编码蛋白的三级结构

Fig. 5 Constructed 3D structure of *GbWD40* protein

2.4 *GbWD401* 和 *GbWD402* 基因编码蛋白的亚细胞定位、信号肽和跨膜区分析

使用 GenScript 程序预测 *GbWD40* 基因编码蛋白的亚细胞定位情况,结果如下:*GbWD401* 在细胞核占 30.4%,细胞质占 26.1%,高尔基体占 13.0%,线粒体占 8.7%,内质网占 8.7%,囊泡分泌系统占 8.7%,过氧化物酶体占 4.3%。*Gb-*

WD402 在线粒体占 43.5%,细胞核占 30.4%,细胞质占 17.4%,液泡占 4.3%,过氧化物酶体占 4.3%。Signal P4.1 在线软件预测 *GbWD401*、*GbWD402* 蛋白均无信号肽,属非分泌蛋白。在线工具 TMHMM2.0 预测 *GbWD401*、*GbWD402* 蛋白几乎不含跨膜区,为非跨膜蛋白(图 6)。

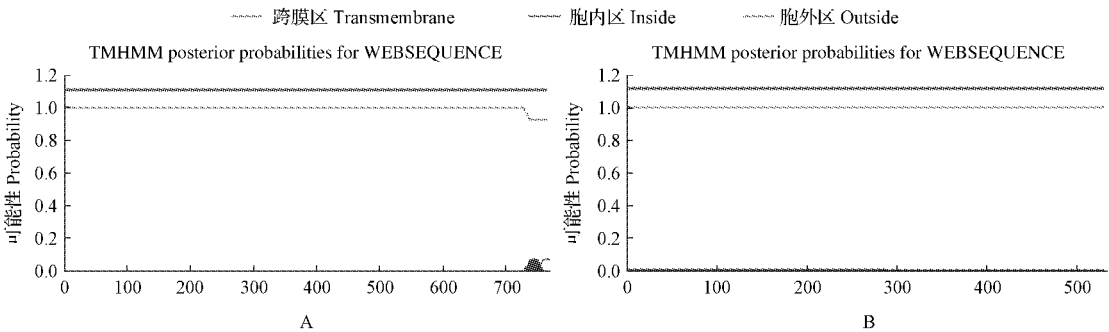


图 6 *GbWD40* 基因编码蛋白跨膜区域

Fig. 6 Transmembrane domain of *GbWD40* protein

2.5 GbWD401 和 GbWD402 基因编码蛋白的系统进化分析

为了解 GbWD401 和 GbWD402 转录因子在自然界中大致的进化关系,从 NCBI 数据库中筛选出与 GbWD401 和 GbWD402 基因编码蛋白同源性较高的 10 种植物的 WD40 蛋白序列,分别为无油樟 *Amborella trichopoda* (XP_006857843.1, AtWD401)、野草莓 *Fragaria vesca* subsp. *vesca* (XP_004293101.1, FvWD40)、枣 *Ziziphus jujuba* (XP_015899081.1, ZjWD40)、烟草 *Nicotiana sylvestris* (XP_009767447.1, NsWD40)、荷花 *Nelumbo nucifera* (XP_010252806.1, NnWD401)、马铃薯 *Solanum tuberosum* (XP_006348232.1, StWD40)、无油樟 *Amborella trichopoda* (XP_006854341.2, AtWD402)、荷花 *Nelumbo nucifera* (XP_

010264714.1, NnWD402)、葡萄 *Vitis vinifera* (XP_002280841.1, VvWD40)、甜椒 *Capsicum annuum* (XP_016542723.1, CaWD40)、胡萝卜 *Daucus carota* subsp. *sativus* (XP_017252788.1, DcWD40)、芝麻 *Sesamum indicum* (XP_011094676.1, SiWD40)、烟草 *Nicotiana tomentosiformis* (XP_009590375.1, NtWD40); 由 ClustalW 序列比对后,通过 MEGA 7.0 软件采用 NJ 法,构建进化树,结果显示(图 7),GbWD401 基因与无油樟、荷花 WD40 基因亲缘性较近,划归为同一分支;GbWD402 基因与无油樟 WD40 基因亲缘性最近,划归为同一分支,与荷花在进化水平上遗传距离较近,GbWD401 和 GbWD402 均与胡萝卜、芝麻、甜椒、烟草遗传距离较远。

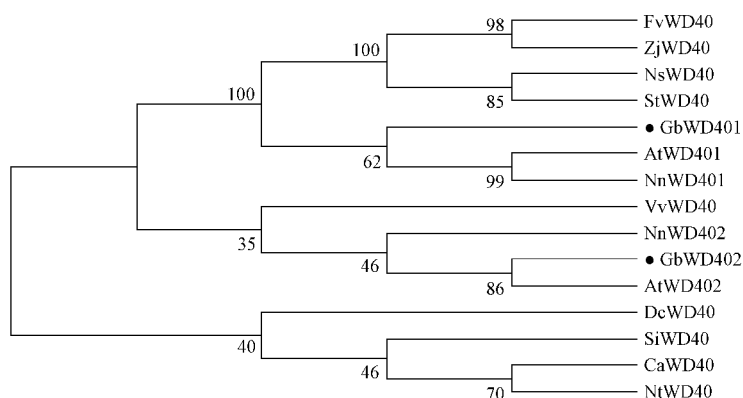


图 7 15 种植物 WD40 基因的系统进化树

Fig. 7 Phylogenetic tree analysis of 15 plant WD40 transcription factors

3 讨论

银杏叶类黄酮具有重要的药理活性及药用价值,然而大多数银杏叶中类黄酮含量极低,国内外学者围绕银杏栽培方式、体外合成、分子水平等方面进行大量研究,效果并不理想。近年来研究发现,参与类黄酮生物合成途径中的调节基因,它们编码的转录因子蛋白能结合酶基因启动子中顺式作用元件,进而激活类黄酮合成途径中多种基因的表达,有效地积累类黄酮,如 WD40、bHLH、MYB 类调节基因^[23-24]。

WD40 基因家族广泛参与植物类黄酮生物合

成,其具有 β 螺旋桨蛋白组结构,核心区域由约 40 个氨基酸组成,该区域包含甘氨酸-组氨酸二肽和色氨酸-天门冬氨酸二肽,此基序在同一蛋白中一般可串联 4~16 次。首个报道参与花青素代谢调控的 WD40 蛋白是矮牵牛 PhAN11,该蛋白能调控矮牵牛花瓣中色素的表达,在无色素形成的植物体内也有表达^[25]。目前已经在紫苏、玉米、马铃薯、苜蓿、苹果以及葡萄等植物中鉴定了类黄酮合成相关的 WD40 类转录因子。该研究克隆了银杏叶片中调节类黄酮合成的 2 个转录因子 GbWD401 和 GbWD402,并对该基因序列进行深入分析。2 个转录因子基因在银杏叶片不同发

育阶段的表达,与类黄酮含量变化关系有待进一步验证。其中将 GbWD401 氨基酸序列在 NCBI 中进行 Blast 比对,氨基酸相似性在 50%左右, GbWD402 氨基酸相似性在 74%左右。而核苷酸序列的相似性也较低,均在 72%左右。课题组认为这些比对数据较低是因为银杏是银杏纲下唯一的一个物种,与其它物种亲缘关系较远较正常。

另外,拟南芥 AtTTG1 能与 bHLH 类转录因子及 MYB 类转录因子形成 MBW 结合物,来调节 DFR, LDOX 和 ANR 基因的表达,从而影响种子中类黄酮及组织内花青素含量^[26]。AN 等^[27]研究发现,苹果 WD40 转录因子基因 *MdTTG1* 在拟南芥中过表达时能与其它转录因子结合,形成转录激活体,刺激花青素合成途径中关键基因的表达,从而提高花青素含量。因此后续试验研究可以将构建 *GbWD401* 和 *GbWD402* 过表达载体导入到烟草中,分析烟草花青素合成关键酶基因的表达,且与 *GbWD401* 和 *GbWD402* 基因表达的相关性。以进一步表明转录因子 GbWD401 和 GbWD402 参与调节类黄酮合成,为提高银杏类黄酮含量技术研究提供一定的理论基础。

4 结论

该研究成功克隆了银杏 GbWD401 和 GbWD402 转录因子 cDNA 序列,明确了 2 个转录因子基因序列结构。*GbWD401* 编码 770 个氨基酸, *GbWD402* 编码 534 个氨基酸,均含有 WD40 重复基序,进化树分析 *GbWD401* 基因与无油樟、荷花 WD40 基因亲缘性较近,划归为同一分支; *GbWD402* 基因与无油樟 WD40 基因亲缘性最近,划归为同一分支,与荷花在进化水平上遗传距离较近, *GbWD401* 和 *GbWD402* 均与胡萝卜、芝麻、甜椒、烟草遗传距离较远。

参考文献

- [1] 曹福亮. 中国银杏志[M]. 北京:中国林业出版社,2007.
- [2] 陈西娟,王成章,叶建中. 银杏叶化学成分及其应用研究进展[J]. 生物质化学工程,2008(4):57-62.
- [3] 张传丽,陈鹏. 银杏类黄酮研究进展[J]. 北方园艺,2014(3):177-181.
- [4] van BEEK T A, MONTORO P. Chemical analysis and quality control of *Ginkgo biloba* leaves, extracts, and phytopharmaceu-

tics[J]. Journal of Chromatography A, 2009, 1216(11): 2002-2032.

- [5] 陈梅霞,苏洁,张玥莉,等. 银杏叶提取物与银杏黄酮对高脂血症大鼠血脂水平影响的比较研究[J]. 中国新药杂志,2014(7): 833-838.
- [6] 陈晓珏,季斌,王燕,等. 银杏黄酮对大鼠放射性脑损伤的保护作用[J]. 江苏医药,2010(11):1328-1329.
- [7] 邱耀宇,林密,黄荣晖,等. 银杏黄酮对力竭小鼠心肌抗氧化的效用[J]. 中国医药导报,2010(15):23-24.
- [8] 张静,张晓鸣,佟建明,等. 银杏黄酮抑制脂质氧化的研究[J]. 食品与生物技术学报,2010(6):842-846.
- [9] 朱红,高尔,王德伟,等. 银杏黄酮的药学研究进展[J]. 食品与药品,2005(2):4-7.
- [10] 陈健康,卫张蕊,刘水冰,等. 银杏黄酮对急性心肌缺血/再灌注的药理作用[J]. 心脏杂志,2007(3):293-294,301.
- [11] 杜辉,张琰,王惠兰,等. 银杏黄酮对人卵巢恶性肿瘤细胞株 PA-1 及人卵巢颗粒细胞的影响[J]. 中国实用妇科与产科杂志,2013(8):633-637.
- [12] 程子翠,孙保亮,杨明峰,等. 银杏内酯及银杏黄酮对脑淋巴阻滞大鼠脑实质微循环血流量变化的影响[J]. 中国微循环,2009(5):355-358.
- [13] 李娜,常征,乔宗旭,等. 银杏黄酮对恶性卵巢肿瘤化疗所致功能损伤的保护作用机制[J]. 西部医学,2015(9):1306-1308.
- [14] HICHERI I, BARRIEU F, BOGS J, et al. Recent advances in the transcriptional regulation of the flavonoid biosynthetic pathway[J]. Journal of Experimental Botany, 2011, 62(8): 2465-2483.
- [15] 沈忠伟,许昱,夏彝,等. 植物类黄酮次生代谢生物合成相关转录因子及其在基因工程中的应用[J]. 分子植物育种,2008(3): 542-548.
- [16] XU C, MIN J. Structure and function of WD40 domain proteins[J]. Protein & Cell, 2011, 2(3): 202-214.
- [17] SOMPORNPAILIN K, MAKITA Y, YAMAZAKI M, et al. A WD-repeat-containing putative regulatory protein in anthocyanin biosynthesis in *Perilla frutescens*[J]. Plant Molecular Biology, 2002, 50(3): 485-495.
- [18] CAREY C C, STRAHLE J T, SELINGER D A, et al. Mutations in the pale aleurone color1 regulatory gene of the *Zea mays* anthocyanin pathway have distinct phenotypes relative to the functionally similar TRANSPARENT TESTA GLABRA1 gene in *Arabidopsis thaliana* [J]. The Plant Cell, 2004, 16(2): 450-464.
- [19] 罗遵喜,刘仕芸,张树珍,等. 马铃薯花青素转录激活基因 *stwd40* 的克隆与表达分析[J]. 园艺学报,2008(9):1317-1322.
- [20] PANG Y, WENGER J P, SAATHOFF K, et al. A WD40 repeat protein from *Medicago truncatula* is necessary for tissue-specific anthocyanin and proanthocyanidin biosynthesis but not for trichome development[J]. Plant Physiology, 2009, 151(3): 1114-1129.
- [21] BRUEGGEMANN J, WEISSHAAR B, SAGASSER M. A WD40-repeat gene from *Malus domestica* is a functional homo-

logue of *Arabidopsis thaliana* TRANSPARENT TESTA GLABRA1[J]. Plant Cell Reports, 2010, 29(3): 285-294.

[22] MATUS J T, POUPIN M J, CAÑÓN P, et al. Isolation of *WDR* and *bHLH* genes related to flavonoid synthesis in grapevine (*Vitis vinifera* L.) [J]. Plant Molecular Biology, 2010, 72(6): 607-620.

[23] ESPLEY R V, HELLENS R P, PUTTERILL J, et al. Red colouration in apple fruit is due to the activity of the MYB transcription factor, MdMYB10[J]. The Plant Journal, 2007, 49(3): 414-427.

[24] PIRES N, DOLAN L. Origin and diversification of basic-helix-loop-helix proteins in plants[J]. Molecular Biology and Evolution, 2010, 27(4): 862-874.

[25] de VETTEN N, QUATTROCCHIO F, MOL J, et al. The *anll* locus controlling flower pigmentation in petunia encodes a novel WD-repeat protein conserved in yeast, plants, and animals[J]. Genes & Development, 1997, 11(11): 1422-1434.

[26] APPELHAGEN I, JAHNS O, BARTELNEUWOEHNER L, et al. Leucoanthocyanidin dioxygenase in *Arabidopsis thaliana*; Characterization of mutant alleles and regulation by MYB-BHLH-TTG1 transcription factor complexes[J]. Gene, 2011, 484(1): 61-68.

[27] AN X H, TIAN Y, CHEN K Q, et al. The apple WD40 protein MdTTG1 interacts with bHLH but not MYB proteins to regulate anthocyanin accumulation[J]. Journal of Plant Physiology, 2012, 169(7): 710-717.

Cloning and Sequence Analysis of Two WD40 Transcription Factors in *Ginkgo biloba* L.

FENG Ru, LI Zehong, YUAN Honghui, CHENG Hua, CHENG Shuiyuan

(School of Biology and Pharmaceutical Engineering, Wuhan Polytechnic University, Wuhan, Hubei 430023)

Abstract: WD40 transcription factor, also known as WDR (WD repeat), was an ancient family of proteins. It could participate in a few genes' co-expression in the plant flavonoid secondary metabolic pathway, effectively improved the contents of flavonoids. This study used triennial ginkgo trees as materials, used RT-PCR and RACE-PCR technology explored the sequences of WD40, in order to understand the molecular mechanism and transcriptional regulation mechanisms of flavonoid metabolism in *Ginkgo*. The results showed that *GbWD401* contained a 2 313 bp ORF encoding 770 amino acid residues, and *GbWD402* contained a 1 605 bp ORF encoding 534 amino acid residues. The predicted protein sequences of *GbWD401* and *GbWD402* genes both contained the typical WD40 structural domain. *GbWD401* shared 49% similarity with *AtWD401* of *Amborella trichopoda*, *GbWD402* shared 78% similarity with *AtWD402* of *Amborella trichopoda*. Phylogenetic tree analysis showed that *GbWD401* was grouped into the same clade with *AtWD401* of *Amborella trichopoda* and *NnWD40* of *Nelumbo nucifera*, while *GbWD402* was grouped into the other clade with *AtWD402* of *Amborella trichopoda*.

Keywords: *Ginkgo biloba* L.; flavonoid; WD40 transcription factor; gene cloning