

巨龙竹木质素合成关键基因 *CCoAOMT* 的克隆及表达分析

陈凌娜, 郭晓娟, 杨汉奇

(中国林业科学研究院资源昆虫研究所, 昆明 650224)

摘要: 咖啡酰辅酶 A-O-甲基转移酶 (*CCoAOMT*, caffeoyl-CoA O-methyltransferase) 是木质素合成过程中关键酶之一。研究通过逆转录 PCR (RT-PCR) 从巨龙竹 (*Dendrocalamus sinicus* L.C.Chia & J.L.Sun) 发育早期的竹笋中克隆了 *CCoAOMT* 基因 (*DsCCoAOMT*), 基因全长 cDNA 序列为 777 bp, 编码 258 个氨基酸, 理论分子量 28.85 kD, 等电点为 5.35, 蛋白质结构预测的结果显示该蛋白含有丰富的 α -螺旋和无规则卷曲。系统进化分析表明, 巨龙竹 *CCoAOMT* 基因与禾本科植物 *CCoAOMT* 基因同源性最高, 亲缘关系最近; 与非禾本科单子叶植物的亲缘关系次之, 然后是裸子植物, 与双子叶植物的亲缘关系最远, 说明 *CCoAOMT* 基因在双子叶植物和单子叶植物之间的差异可能在被子植物进化之前就已经存在。 *DsCCoAOMT* 基因密码子偏好性分析的结果显示, 该基因选择偏性较强且偏好以 G/C 结尾; 密码子使用频率比较结果发现, *DsCCoAOMT* 基因在微生物中异源表达分析时采用酵母表达系统更为合适, 而模式植物拟南芥、烟草、番茄与巨龙竹 *CCoAOMT* 密码子使用频率差异较小, 尤其烟草、番茄可能是该基因进行转基因功能验证时最为理想的受体。实时荧光定量 PCR 分析巨龙竹不同组织及竹笋发育过程中 *DsCCoAOMT* 基因表达水平的结果显示, 该基因在巨龙竹不同组织中的表达具有明显差异, 竹笋中表达量最高, 茎秆中次之, 叶片中最少; 随着竹笋的发育, *DsCCoAOMT* 基因的表达量上调并在竹笋发育中期维持较高水平, 表明 *DsCCoAOMT* 基因可能在巨龙竹快速生长中发挥重要调节作用。本研究的结果为进一步阐明 *DsCCoAOMT* 基因的功能和作用机制奠定了基础。

关键词: 巨龙竹; *CCoAOMT*; 基因克隆; 生物信息学分析; 基因表达

Cloning and Expression Analysis of *CCoAOMT*, a Key Gene in Lignin Biosynthesis of *Dendrocalamus sinicus*

CHEN Ling-na, GUO Xiao-juan, YANG Han-qi

(Research Institute of Resources Insects, Chinese Academy of Forestry, Kunming 650224)

Abstract: Caffeoyl-CoA O-methyltransferase (*CCoAOMT*) is one of the key enzymes in lignin biosynthesis. By reversible transcription PCR (RT-PCR) technique, a *CCoAOMT* gene (*DsCCoAOMT*) was cloned from the bamboo shoot of *Dendrocalamus sinicus* L.C.Chia & J.L.Sun. The full-length cDNA of *DsCCoAOMT* was 777 base pairs that encoded a protein of 258 amino acids with predicted molecular mass of 28.85 kD and a basic isoelectric point of 5.35. The deduced protein was rich in helix and coin. The phylogenetic analysis revealed the highest sequence homology of *DsCCoAOMT* in relative to *CCoAOMT* from the poaceous plants, followed by the monocotyledon of non-graminaceae and gymnosperms and then some dicotyledons. That suggested a differentiation of *CCoAOMT* between monocotyledons and dicotyledons before angiosperms evolution. According to the analysis of codon bias, the *DsCCoAOMT* had high bias toward the synonymous codons with G and C at third codon position. The results of codon usage frequency showed that heterologous expression of *DsCCoAOMT* might be applicable in the microorganism yeast expression system and in plants *Nicotiana tabacum* L. and

收稿日期: 2018-08-01 修回日期: 2018-09-06 网络出版日期: 2018-10-18

URL: <http://kns.cnki.net/kcms/detail/11.4996.S.20181017.1612.008.html>

第一作者研究方向为竹类植物生长发育的分子机制, E-mail: chenln@caf.ac.cn

通信作者: 杨汉奇, 研究方向为竹类种质资源遗传改良, E-mail: yanghanqikm@aliyun.com

基金项目: 中央级公益性科研院所基本科研业务费专项 (CAFYBB2017QA014, CAFYBB2017ZX001-8); 云南省科学技术厅 (2014HB041)

Fundation project: This research was funded by the Fundamental Research Funds of the Chinese Academy of Forestry (CAFYBB2017QA014, CAFYBB2017ZX001-8), Department of Sciences and Technology of Yunnan Province (2014HB041)

Solanum lycopersicum L..By real-time quantitative PCR, the *DsCCoAOMT* gene was highly expressed in the shoots, followed by the culms and the leaves at last.As the shoots developed, *DsCCoAOMT* is up-regulated at the higher level, suggesting that the gene might play an important regulation role in fast-growing *D.sinicus* L.C.Chia & J.L.Sun.Thus, this study might be helpful for further understanding the function and molecular mechanism of the *DsCCoAOMT* gene.

Key words: *Dendrocalamus sinicus* L.C.Chia & J.L.Sun; *CCoAOMT* gene; gene cloning; bioinformatics analysis; gene expression

木质素是维管植物细胞壁的主要组分,也是苯丙烷代谢途径中最重要的产物之一,在提高植物体的机械强度,保护植物抵御病害侵袭中起重要作用^[1]。而在纸品加工、纤维素生物燃料生产及饲料的高效利用中,植物细胞壁中的木质素又起负面作用^[2]。因此,根据研究目标对植物体木质素单体合成途径调控成为研究热点,且已从植物中成功分离了多个木质素合成途径中的关键酶基因。咖啡酰辅酶 A-O- 甲基转移酶(*CCoAOMT*, caffeoyl-CoA O-methyltransferase)是木质素单体前体甲基化过程中两个关键酶之一,催化咖啡酰辅酶 A(caffeoyl-CoA)甲基化为阿魏酰辅酶 A(feruloyl-CoA)^[1]。*CCoAOMT* 基因的调节能引起愈创木基/紫丁香基(S/G)木质素单体比例的改变,使木材的材质、饲草的消化率得以提升;不仅如此,*CCoAOMT* 基因的过表达还能增加植物对不同病原的抗性,因此越发受到研究者关注^[1,3-4]。目前,已经从拟南芥(*Arabidopsis thaliana* (L.) Heynh.)^[5]、慈竹(*Neosinocalamus affinis* (Rendle) Keng f.)^[6]、玉米(*Zea mays* L.)^[2]、赤桉(*Eucalyptus camaldulensis* Dehnh.)^[3]等多种植物中分离了 *CCoAOMT* 基因。

巨龙竹(*Dendrocalamus sinicus* L.C.Chia & J.L.Sun)是云南西南地区特有大型木本丛生竹,为目前文献报道的世界最大竹种^[7],其秆高可达 30 m,秆径可达 30 cm,单位面积产材量比我国主要经济竹种毛竹高 5~8 倍^[8]。巨龙竹材积大、生长快、竹壁厚、抗压力强,是其分布区家具、建筑、引水管道及工业纸浆生产的重要原料^[9],木质素含量、组成和类型在竹类茎秆发育和伸长、结构成分的形成中起重要作用^[10]。通过对巨龙竹进行无参考基因组的转录组测序发现,巨龙竹不同发育时期 *CCoAOMT* 基因 FPKM 值(Reads Per Kilobase of exon model per Million mapped reads)在 100~700 之间,表达量较高,可能在巨龙竹茎秆的发育中作用显著^[11]。为了探明 *CCoAOMT* 基因在巨龙竹茎秆发育中对木质素单体合成的作用,本研究从其发育早期竹笋中克

隆了 *CCoAOMT* 基因(命名为 *DsCCoAOMT*),对该基因进行生物信息学分析,包括 *CCoAOMT* 基因的系统进化、*DsCCoAOMT* 编码蛋白质的理化性质和结构、*DsCCoAOMT* 密码子偏好性进行分析,根据密码子使用频率比值对 *DsCCoAOMT* 遗传转化功能验证的受体系统进行了预测,并检测了巨龙竹不同组织及不同发育阶段竹笋中 *DsCCoAOMT* 的表达,为巨龙竹 *CCoAOMT* 功能的深入研究提供理论依据。

1 材料与方法

1.1 材料

巨龙竹取样于云南省孟连县,将出土 20 d 左右的竹笋剥除笋箨,切取基部 1/3 处组织并切成小块后置于液氮中,带回实验室于 -80 ℃ 冰箱长期保存。

RNA 提取、反转录试剂盒及基因克隆所需分子生物学试剂均购自全式金生物技术有限公司。

1.2 巨龙竹总 RNA 提取及 *DsCCoAOMT* 基因克隆

用 TransZol Up Plus RNA Kit 按说明书提取总 RNA,经 1.0% 琼脂糖电泳检测 RNA 完整性,然后用 NanoDrop 2000 分光光度计(Thermo Fisher Scientific)检测总 RNA 的质量并进行定量分析。取 10 μg 总 RNA,用 TransScript II One-Step gDNA Removal and cDNA Synthesis SuperMix 按说明书合成 cDNA。

根据巨龙竹转录组测序拼接的 *CCoAOMT* 基因序列,结合 NCBI 在线数据库(<https://www.ncbi.nlm.nih.gov/>)中水稻(*Oryza sativa* L.)等物种 *CCoAOMT* 序列,设计上游引物 *DsCCoCOMT-F* 为: GCA ATG GCC ACC GCG ACA GCC CAG A,下游引物 *DsCCoCOMT-R* 为: GGC TCA CTT GGC GCG GCG GCA GAG G。以巨龙竹笋 cDNA 为模板,以 *DsCCoCOMT-F* 和 *DsCCoCOMT-R* 为引物进行 PCR 扩增,反应条件为: 95 ℃ 预变性 5 min; 95 ℃ 变性 30 s, 59 ℃ 退火 30 s, 72 ℃ 延伸 1 min, 30 个循环; 72 ℃ 延伸 7 min。PCR 扩增产物经琼脂糖电泳分离

回收目的片段后连接至 pEASY-T5 载体,转化大肠杆菌 (*Escherichia coli*) DH5 α 菌株后选阳性克隆进行测序验证。

以 caffeoyl-CoA O-methyltransferase 为关键词从 NCBI 数据库中检索并下载拟南芥、慈竹、玉米等 15 个科共 25 个物种的 *CCoAOMT* 基因 CDS 序列,与巨龙竹 *CCoAOMT* 基因序列进行对比分析。

1.3 *CCoAOMT* 基因生物信息学分析

1.3.1 *CCoAOMT* 基因的确认 将 *CCoAOMT* 基因序列编码氨基酸进行 SMART (<http://smart.embl-heidelberg.de/>) 在线分析,鉴定 *CCoAOMT* 编码蛋白质的结构域。用 ExPaSy 在线工具 (<http://www.expasy.ch/tools/>) 中 InterProScan 程序 (<http://www.ebi.ac.uk/interpro/search/sequence-search>) 进行蛋白信号匹配分析,并使用 MEGA7 中的 ClustalW 程序将其与其他物种 *CCoAOMT* 基因进行多序列比对,确认克隆的基因序列是否属于 *CCoAOMT* 基因家族。

1.3.2 *DsCCoAOMT* 编码蛋白质的理化性质 使用 ProtParam 在线工具 (<http://web.expasy.org/cgi-bin/protparam/protparam>) 估计鉴定过的 *DsCCoAOMT* 蛋白序列的分子量、稳定指数和亲水指数等基本理化参数。

1.3.3 *DsCCoAOMT* 编码蛋白质的二级和三级结构预测 用 ExPaSy 在线工具中 PredictProtein 程序 (<https://www.predictprotein.org/>) 预测 *DsCCoAOMT* 编码的蛋白质结构,包括二级结构中 α -螺旋、 β -转角及无规则卷曲等,SignalP 4.1 程序 (<http://www.cbs.dtu.dk/services/SignalP/>) 预测蛋白是否存在信号肽,TMHMM 程序 (<http://www.cbs.dtu.dk/services/TMHMM-2.0/>) 预测是否存在跨膜结构,用 SWISS-MODEL (<https://swissmodel.expasy.org/>) 进行蛋白质同源三级结构建模。

1.3.4 *CCoAOMT* 基因的密码子偏好性 使用 CHIPS 在线程序 (<http://www.bioinformatics.nl/cgi-bin/emboss/chips>) 分析 *CCoAOMT* 基因的有效密码子数 (ENc, effective number of codons),CUSP 在线程序 (<http://www.bioinformatics.nl/emboss-explorer/>) 分析密码子 G+C 含量 (GC)、同义密码子第 3 位中 G/C 含量 (GC3s)。大肠杆菌、酵母 (*Saccharomyces cerevisiae*)、拟南芥、烟草和番茄 (*Solanum lycopersicum* L.) 的基因组密码子使用偏性数据来源于 Codon Usage Database 在线数据库 (<http://www.kazusa.or.jp/codon/>)。

1.3.5 *CCoAOMT* 系统进化树构建 使用 MEGA7 中的 ClustalW 程序,以 Neighbor-Joining method,采用 Bootstrap 方法重复 1000 次检验构建巨龙竹与其他 25 个物种 *CCoAOMT* 基因的进化树。

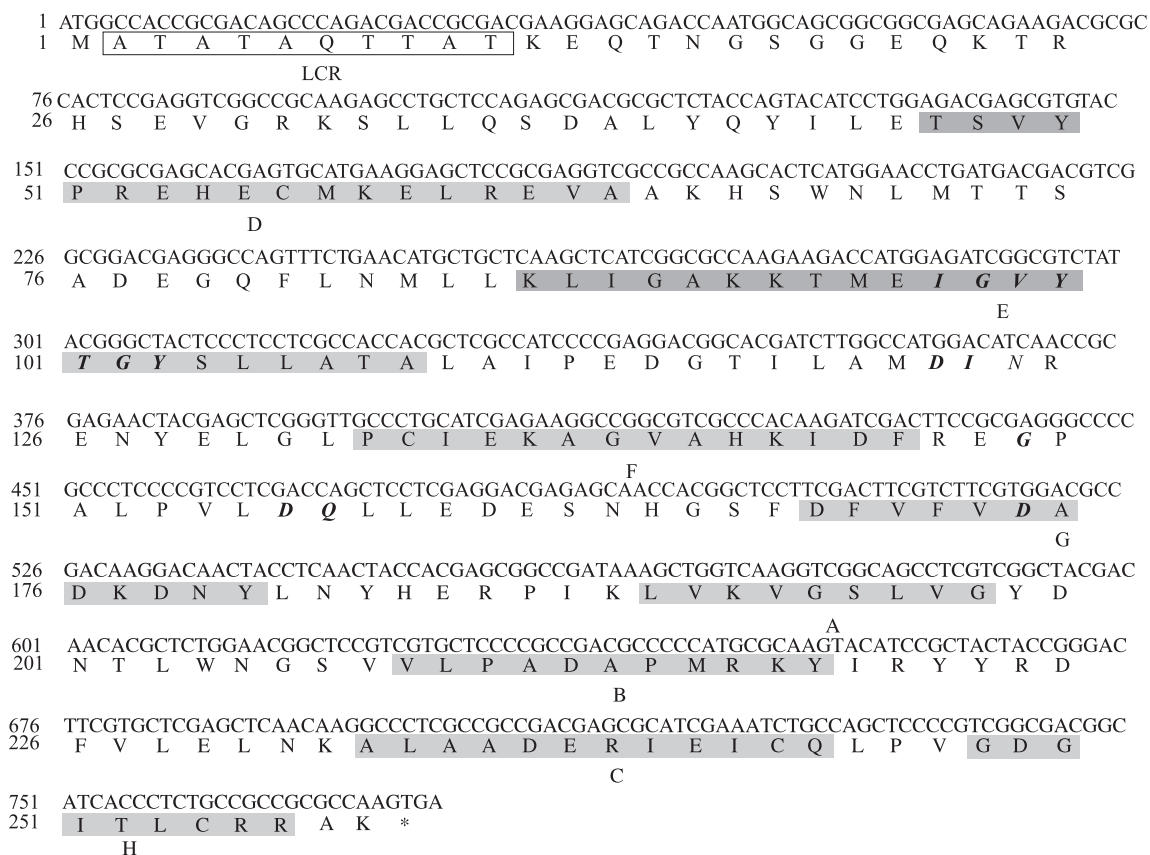
1.3.6 实时荧光定量 PCR (qRT-PCR) 检测巨龙竹 *DsCCoAOMT* 基因的表达 分别提取巨龙竹笋、笋箨、秆、小枝、叶、叶鞘,以及巨龙竹笋发育第 5、10、15、20、30、45 天竹笋的总 RNA,并反转录为 cDNA,以巨龙竹 *Actin* 基因为内参,分别以引物 DsActin-F: TGC CTA CAT TGC CCT GGA CTA 和 DsActin-R: TCC ATT CCA ATG AAG GAT GGC T, qDsCCoCOMT-F: GTA CAT CCT GGA GAC GAG CG 和 qDsCCoCOMT-R: GAG GAG GGA GTA GCC CGT AT,使用 TransStart Top Green qPCR SuperMix 按照产品说明书对各样品中 *DsActin* 和 *DsCCoAOMT* 进行扩增,每样品 3 次重复。数据由仪器自动采集后,采用 $2^{-\Delta\Delta Ct}$ 法计算各样品中 *DsCCoAOMT* 基因的相对表达量。

2 结果与分析

2.1 *DsCCoAOMT* 基因克隆及编码氨基酸序列分析

通过 RT-PCR 技术克隆获得 *DsCCoAOMT* 基因 cDNA 全长为 777 bp,编码 258 个氨基酸,包含完整的开放阅读框。将该氨基酸序列进行 SMART 在线分析的结果表明,第 2~11 氨基酸区域属于低复杂性区域 (LCR, low complexity region),第 45~257 氨基酸区域是 Methyltransf_3 (PF01596)。通过 InterProScan 在线分析,该序列包含 S-腺苷甲硫氨酸 (SAM, S-adenosyl-L-methionine) 结合位点 (第 97~103, 122~123, 149, 156~157, 174 位氨基酸),属于 S-腺苷甲硫氨酸甲基转移酶基因家族 (SAM-dependent O-methyltransferase)。经同源性比对发现,该序列也包括 5 个酪蛋白激酶 II 磷酸化位点 (casein kinase II phosphorylation site, D、E、F、G、H) 和 3 个 N-豆蔻酰化位点 (N-myristoylation site, A、B、C),因此确定是 *CCoAOMT* 基因家族中的成员 (图 1)。

使用 MEGA7 比对巨龙竹与 NCBI 下载的拟南芥、水稻、番茄等 25 个物种 *CCoAOMT* 基因的 CDS 序列,26 个物种 *CCoAOMT* 基因的序列相似性为 77.89%;其中与慈竹 *CCoAOMT* 基因 (NCBI 登录号: JF742462.1) 的相似性最高,为 98.46%;与烟草 *CCoAOMT* 基因 (NCBI 登录号: U62734.1) 的相似性最低,为 63.06%。



方框中表示 LCR 区域,阴影部分表示 8 个保守区域,粗斜体显示 SAM 结合位点

LCR region and 8 conserved regions are marked by the box and shadow, respectively. And letters in bold italics display SAM binding sites

图 1 *DsCCoAOMT* 基因 CDS 序列及其编码的氨基酸序列

Fig.1 cDNA sequence of *DsCCoAOMT* gene and its putative amino acid sequence

2.2 *DsCCoAOMT* 的理化性质

用 ProtParam 工具估计 *DsCCoAOMT* 蛋白的基本理化参数的结果表明,该蛋白的分子量为 28.85 kD,理论等电点为 5.35,包含 37 个酸性氨基酸残基 (Asp+Glu) 和 29 个碱性氨基酸残基 (Arg+Lys),为酸性蛋白质;在组成 *DsCCoAOMT* 的 20 种氨基酸中,亮氨酸 (Leu) 含量最高为 11.6%,色氨酸 (Trp) 含量最少,仅为 0.8%,肽链 N 端为甲硫氨酸 (Met);蛋白不稳定指数为 29.82,为稳定蛋白;亲水性平均系数为 -0.306,为亲水性蛋白。

2.3 *DsCCoAOMT* 的二级和三级结构

使用 PredictProtein 预测 *DsCCoAOMT* 的二级结构的结果显示,在肽链 N 端有一个 21 个氨基酸组成的无序化区域,并在二级结构中包含 12 个蛋白质结合位点; α -螺旋、 β -折叠和无规则卷曲分别为 36.43%、13.57% 和 50.00%;暴露 (exposed)、包埋 (buried) 和中间状态的氨基酸残基分别为 42.64%、48.84% 和 8.53% (图 2)。*DsCCoAOMT* 不包含信

号肽,且无跨膜区域,在真核生物中定位于细胞核中。利用 SWISS-MODLE 工具对 *DsCCoAOMT* 进行同源建模,构建三维模型 (图 3),可见 *DsCCoAOMT* 有两个亚基组成,含有丰富的 α -螺旋和无规则卷曲,并存在两个 SAM 结合区和两个 Ca^{2+} 的配位 (图 3b)。

2.4 *DsCCoAOMT* 基因的密码子偏好性及受体系统选择

根据 *DsCCoAOMT* 基因编码氨基酸密码子数量的差异,使用 CHIPS 程序预测其有效密码子数 (ENc) 值为 26.579,大于 20 且远小于 61,说明该基因对密码子选择偏性很强且稀有密码子较少,表达量较高。GC 和 GC3s 含量分别为 0.651 和 0.969,均大于 0.500,说明 *DsCCoAOMT* 基因编码区对碱基的选择倾向于 G+C,且密码子偏好于以 G/C 结尾。

根据表 1 中不同物种 *CCoAOMT* 基因的密码子偏好分析显示,包括巨龙竹在内的 26 个物种的 ENc 值在 26.302~53.471 之间,平均值为 44.068,

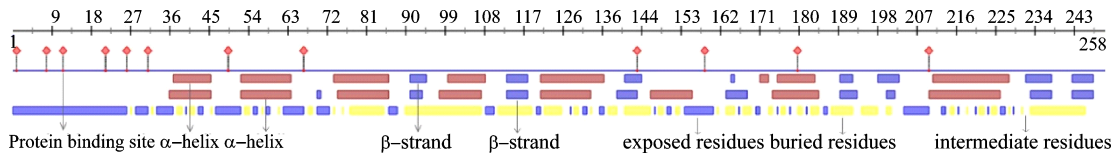
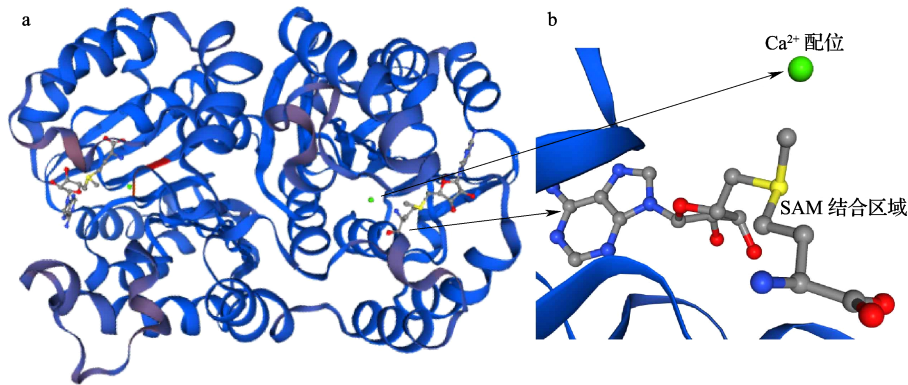


图 2 DsCCoAOMT 蛋白质的二级结构
Fig.2 The predicted secondary structure of DsCCoAOMT protein



a: DsCCoAOMT 的整体结构; b: DsCCoAOMT 的 SAM 结合区域
a: Global structure of DsCCoAOMT, b: SAM-binding region of DsCCoAOMT

图 3 DsCCoAOMT 蛋白质的三级结构
Fig.3 The predicted tertiary structure of DsCCoAOMT protein

表 1 不同物种 CCoAOMT 密码子使用偏性分析
Table 1 Analysis of codon usage bias of CCoAOMT in different species

物种 Species	拉丁文 Latin	科 Family	登录号 GenBank accession number	有效密码 子数 ENc	G+C 含量 GC	密码子第 3 位 G/C 含量 GC3s
金合欢	<i>Acacia koa</i> (A.Gray)	豆科	KX784941.1	51.861	0.479	0.555
拟南芥	<i>Arabidopsis thaliana</i> (L.) Heynh.	十字花科	AY088577.1	51.209	0.458	0.477
光皮桦	<i>Betula luminifera</i> H.Winkl	桦木科	FJ410449.1	46.878	0.507	0.641
象草	<i>Cenchrus purpureus</i> (Schumach.) Morrone	禾本科	KY347028.1	33.859	0.657	0.923
红桉	<i>Chamaecyparis formosensis</i> Matsum.	柏科	DQ305976.1	53.471	0.463	0.500
龙眼	<i>Dimocarpus longan</i> Lour.	无患子科	JN093023.1	49.603	0.475	0.565
荸荠	<i>Eleocharis dulcis</i> (Burm.f.) Trin.ex Hensch.	莎草科	MG719240.1	48.987	0.504	0.561
枇杷	<i>Eriobotrya japonica</i> (Thunb.) Lindl.	蔷薇科	JX885583.1	48.602	0.508	0.633
桉树	<i>Eucalyptus camaldulensis</i> Dehnh.	桃金娘科	HM106291.1	48.481	0.524	0.648
麻风树	<i>Jatropha curcas</i> L.	大戟科	FJ619046.1	44.575	0.449	0.466
百合	<i>Lilium regale</i> E.H.Wilson	百合科	KX842497.1	44.534	0.583	0.769
南荻	<i>Miscanthus lutarioriparius</i> L.Liu ex S.L.Chen & Renvoizo	禾本科	JF965572.2	28.914	0.657	0.966
慈竹	<i>Neosinocalamus affinis</i> (Rendle) Keng f.	禾本科	JF742462.1	26.496	0.651	0.969
烟草	<i>Nicotiana tabacum</i> L.	茄科	U62734.1	52.246	0.425	0.391
水稻	<i>Oryza sativa</i> L.	禾本科	AY644636.1	30.541	0.659	0.969
芍药	<i>Paeonia lactiflora</i> Pall.	毛茛科	JQ684014.1	50.942	0.466	0.496
柳枝稷	<i>Panicum virgatum</i> L.	禾本科	KF041775.1	26.302	0.659	0.992
挪威云杉	<i>Picea abies</i> (L.) H. Karst	松科	AM262870.1	53.166	0.521	0.600
马尾松	<i>Pinus massoniana</i> Lamb.	松科	KF419292.1	51.330	0.509	0.581
梅	<i>Prunus mume</i> (Siebold) Siebold & Zucc.	蔷薇科	KP202399.1	50.860	0.501	0.613
青杨	<i>Populus cathayana</i> Rehder	杨柳科	KJ632455.1	47.043	0.465	0.518
鸭梨	<i>Pyrus x bretschneideri</i> Rehder	蔷薇科	JX457336.1	51.615	0.507	0.625
番茄	<i>Solanum lycopersicum</i> L.	茄科	EU161983.1	50.519	0.428	0.395
马铃薯	<i>Solanum tuberosum</i> L.	茄科	AB061268.1	49.139	0.425	0.395
玉米	<i>Zea mays</i> L.	禾本科	AJ242980.1	28.004	0.658	0.988
巨龙竹	<i>Dendrocalamus sinicus</i> L.C.Chia & J.L.Sun	禾本科		26.579	0.651	0.969
均值 Mean				44.068	0.530	0.662

可见不同物种的 *CCoAOMT* 基因差异较大。其中包括巨龙竹在内的 7 个禾本科物种的 ENc 值均较低, 在 26.302~33.859 之间, 远低于平均水平, 表达水平较高; 其余 19 个物种的 ENc 值在 44~55 之间, 表达水平一般。通过 GC 含量分析发现, 禾本科物种的 GC 含量较高, 均大于 0.6, 另有 9 个物种的 GC 含量大于 0.5 但小于 0.6; 而 GC3s 值大于等于 0.5 的为 20 个, 占总体的 76.9%, 其中禾本科 7 个物种的 GC3s 值均大于 0.9, 茄科 3 个物种 GC3s 值均小于 0.4。这些结果说明, 不同物种 *CCoAOMT* 基因密码子对 GC 含量和 G/C 结尾的偏好性不一致,

多数物种 *CCoAOMT* 基因密码子偏好 G/C 结尾, 禾本科与其他科物种 *CCoAOMT* 密码子使用偏好差异较大。

将 *DsCCoAOMT* 密码子使用频率与大肠杆菌 (*Ec*)、酵母菌 (*Sc*)、拟南芥 (*At*)、烟草 (*Nt*) 和番茄 (*Sl*) 基因组密码子使用频率进行比较, 以密码子使用频率比值来衡量物种间密码子使用偏性的差异, 比值介于 0.5~2.0 之间时物种间密码子使用特性较为一致, 否则说明密码子偏小差异较大^[12]。巨龙竹 *DsCCoAOMT* 与部分模式生物基因组密码子使用偏性比较的结果显示 (表 2), 巨龙竹 *DsCCoAOMT*

表 2 *DsCCoAOMT* 与部分模式生物基因组密码子使用偏性比较

Table 2 Comparison of codon usage preference between *CCoAOMT* in *Dendrocalamus sinicus* L.C.Chia & J.L.Sun and other representative organisms

密码子 Codon	氨基酸 Amimmo acid	<i>DsCCoAOMT</i>	大肠 杆菌 <i>Ec</i>	酵母 菌 <i>Sc</i>	拟南 芥 <i>At</i>	烟草 <i>Nt</i>	番茄 <i>Sl</i>	<i>DsCCoAOMT</i> / <i>Ec</i> *	<i>DsCCoAOMT</i> / <i>Sc</i> *	<i>DsCCoAOMT</i> / <i>At</i> *	<i>DsCCoAOMT</i> / <i>Nt</i> *	<i>DsCCoAOMT</i> / <i>Sl</i> *
GCC	Ala	69.5	21.6	12.6	10.3	12.5	10.1	<u>3.22</u>	1.71	1.22	0.82	1.24
GCG	Ala	19.3	21.1	6.2	9	5.8	5.2	0.91	<u>3.40</u>	0.69	1.55	1.12
TGC	Cys	15.4	5.5	4.8	7.2	7.2	6.7	<u>2.80</u>	1.15	0.67	1.00	1.07
GAC	Asp	61.8	17.9	20.2	17.2	16.9	15	<u>3.45</u>	0.89	1.17	1.02	1.13
GAA	Glt	3.8	35.1	45.6	34.3	36	34.8	<u>0.11</u>	0.77	1.33	0.95	1.03
GAG	Glt	77.2	19.4	19.2	32.2	29.4	26.6	<u>3.98</u>	1.01	0.60	1.10	1.11
TTC	Phe	19.3	13.9	18.4	20.7	18	17.5	1.39	0.76	0.89	1.15	1.03
TTT	Phe	3.8	24.4	26.1	21.8	25.1	26	<u>0.16</u>	0.93	1.20	0.87	0.97
GGC	Gly	65.6	20.6	9.8	9.2	11.2	9.7	<u>3.18</u>	<u>2.10</u>	1.07	0.82	1.15
GGG	Gly	3.8	12.3	6	10.2	10.5	10.8	<u>0.31</u>	<u>2.05</u>	0.59	0.97	0.97
CAC	His	23.1	7.3	7.8	8.7	8.7	7.8	<u>3.16</u>	0.94	0.90	1.00	1.12
ATA	Ile	3.8	13.3	17.8	12.6	14	14	<u>0.29</u>	0.75	1.41	0.90	1.00
ATC	Ile	46.3	19.4	17.2	18.5	13.9	14	<u>2.39</u>	1.13	0.93	1.33	0.99
AAG	Lys	61.8	15.3	30.8	32.7	33.5	31	<u>4.04</u>	0.50	0.94	0.98	1.08
CTC	Leu	84.9	9.5	5.4	16.1	12.3	11.2	<u>8.94</u>	1.76	<u>0.34</u>	1.31	1.10
CTG	Leu	23.2	37.4	10.5	9.8	10.2	10.5	0.62	<u>3.56</u>	1.07	0.96	0.97
TTG	Leu	7.7	12.9	27.2	20.9	22.3	24.2	0.60	<u>0.47</u>	1.30	0.94	0.92
ATG	Met	27.0	23.7	20.9	24.5	25	24.7	1.14	1.13	0.85	0.98	1.01
AAC	Asn	38.6	20.3	24.8	20.9	17.9	17.3	1.90	0.82	1.19	1.17	1.03
AAT	Asn	3.8	29.3	35.7	22.3	28	30.5	<u>0.13</u>	0.82	1.60	0.80	0.92
CCC	Pro	27.0	6.2	6.8	5.3	6.6	5.7	<u>4.35</u>	0.91	1.28	0.80	1.16
CCG	Pro	7.7	14.5	5.3	8.6	5	4.6	0.53	<u>2.74</u>	0.62	1.72	1.09
CAG	Glu	30.9	26.7	12.1	15.2	15	14	1.16	<u>2.21</u>	0.80	1.01	1.07
CGC	Arg	42.5	14	2.6	3.8	3.9	3.1	<u>3.04</u>	<u>5.38</u>	0.68	0.97	1.26
CGG	Arg	7.7	7.9	1.7	4.9	3.7	3.1	0.97	<u>4.65</u>	<u>0.35</u>	1.32	1.19
AGC	Ser	23.2	14.3	9.8	11.3	10	9.3	1.62	1.46	0.87	1.13	1.08
TCA	Ser	3.9	13.1	18.7	18	17.6	20.7	<u>0.30</u>	0.70	1.04	1.02	0.85
TCC	Ser	15.4	9.7	14.2	11.2	10.2	9.9	1.59	0.68	1.27	1.10	1.03
TCG	Ser	3.9	8.2	8.6	9.3	5.3	5.6	<u>0.48</u>	0.95	0.92	1.75	0.95
ACA	Thr	3.9	15.1	17.8	15.7	17.4	17.9	<u>0.26</u>	0.85	1.13	0.90	0.97
ACC	Thr	23.2	18.9	12.7	10.3	9.7	8.6	1.23	1.49	1.23	1.06	1.13
ACG	Thr	34.7	13.6	8	7.7	4.5	4.6	<u>2.55</u>	1.70	1.04	1.71	0.98
GTC	Val	42.5	13.1	11.8	12.8	11.1	10.1	<u>3.24</u>	1.11	0.92	1.15	1.10
GTG	Val	15.4	19.9	10.8	17.4	16.7	16	0.77	1.84	0.62	1.04	1.04
TGG	Trp	7.7	13.4	10.4	12.5	12.2	13.5	0.57	1.29	0.83	1.02	0.90
TAC	Tyr	42.5	11.7	14.8	13.7	13.5	12.4	<u>3.63</u>	0.79	1.08	1.01	1.09
TAT	Tyr	3.9	21.6	18.8	14.6	17.8	18.6	<u>0.18</u>	1.15	1.29	0.82	0.96

* 标注的是巨龙竹 *CCoAOMT* 基因与该物种基因组密码子使用差异, 下划线表示巨龙竹 *CCoAOMT* 基因与该物种基因组密码子使用偏差较大

The asterisks (*) labels differences between *DsCCoAOMT* the codons of two species. The underlines mean there are obvious differences between the codons of two species

与大肠杆菌使用偏性差异大的密码子有 23 个,与酵母菌使用偏性差异较大的密码子有 9 个,可见 *DsCCoAOMT* 在微生物中表达时采用酵母表达系统更为合适,但二者密码子使用频率仍具有差异,需要在表达过程中进行密码子的优化和改造。*DsCCoAOMT* 与拟南芥、烟草、番茄基因组密码子使用特性差异较小,差异频率较大的密码子个数分别为 2、0 和 0 个,说明这 3 种模式植物都适合作为 *DsCCoAOMT* 基因遗传转化功能验证的受体,其中烟草和番茄可能是更为理想的受体。

2.5 CCoAOMT 基因系统进化关系

应用 MEGA7.0 软件构建不同物种 *CCoAOMT* 基因 CDS 序列的系统进化树(图 4),26 个物种划分为 5 个分支:A 分支为双子叶植物,B 分支为裸子植物,C、D 和 E 3 个分支均为单子叶植物,而且同一科的物种都聚在一起,如:A 分支中的蔷薇科、茄科,B 分支中的松科,E 分支中的禾本科,反映出 *CCoAOMT* 基因在物种内的相似性和物种间的差异。巨龙竹与 5 个禾本科植物亲缘关系较近,其中与慈竹的 *CCoAOMT* 基因亲缘关系最近,其他次之;与拟南芥、烟草、杨树等双子叶植物亲缘关系较远,说明基因和植物的进化关系密切有关;但裸子植物在进化树中位于双子叶植物与单子叶植物之间,说明 *CCoAOMT* 基因在单子叶植物和双子叶植

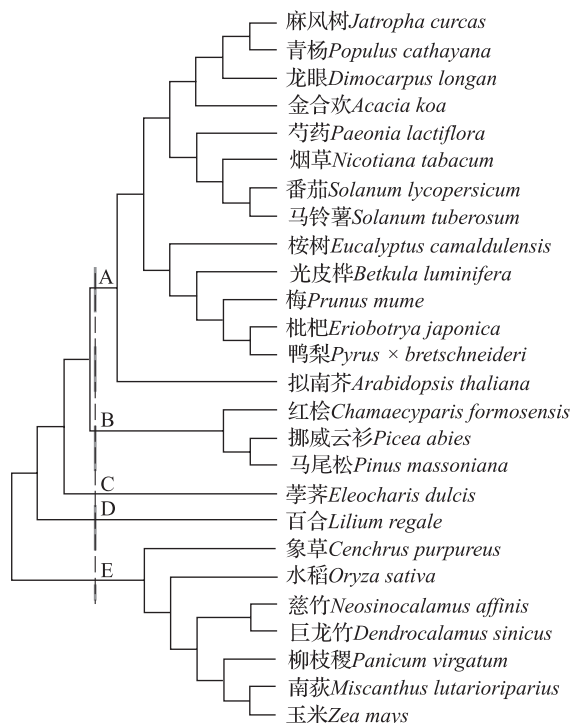


图 4 26 个物种 *CCoAOMT* 基因的系统进化树

Fig.4 The phylogenetic tree of *CCoAOMT* from 26 species

物之间的差异,可能在裸子植物向被子植物进化之前已经形成,造成 *DsCCoAOMT* 基因与裸子植物的亲缘关系相比双子叶植物更近。

2.6 DsCCoAOMT 在巨龙竹不同组织及竹笋不同发育阶段的表达分析

采用 qRT-PCR 方法检测巨龙竹不同组织及不同发育阶段竹笋中 *DsCCoAOMT* 基因的表达情况,结果如图 5 所示。*DsCCoAOMT* 基因在巨龙竹不同组织中的表达量有所不同,在竹笋中的最高,秆中次之,笋箨、小枝、叶鞘中随后,在叶片中的最低,笋中的表达量是其在叶片中的 150 倍以上,说明 *DsCCoAOMT* 基因在巨龙竹的表达具有组织表达差异(图 5a)。在竹笋发育的不同时期,*DsCCoAOMT* 基因的表达量也具有明显差异(图 5b),在竹笋发育早期,*DsCCoAOMT* 的表达量持续升高,并在笋发育中期保持较高的表达水平(发育第 30 天),而此时正处于巨龙竹快速生长期,*DsCCoAOMT*

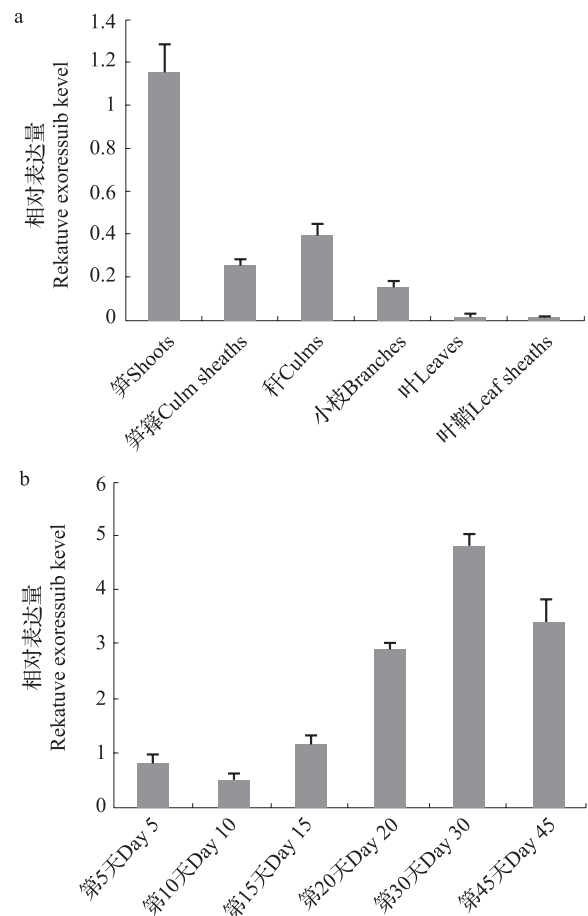


图 5 巨龙竹不同组织 (a) 及笋期不同发育阶段 (b) *DsCCoAOMT* 基因的表达

Fig.5 Expression levels of *DsCCoAOMT* in different tissues (a) and different development stages of the bamboo shoots (b)

的表达量远高于成熟茎秆中的表达量,说明巨龙竹生长过程中,木质素在笋期就开始大量合成,这可能会对巨龙竹的快速生长是必需的。

3 讨论

植物中木质素含量可以通过改变木质素合成途径中相关基因的表达来调节,包括苯丙氨酸解氨酶(PAL, Phenylalanine ammonia-lyase)、肉桂酸4-羟化酶(C4H, Cinnamate 4-hydroxylase)、4-香豆酸辅酶A连接酶(4CL, 4-coumarate: CoA ligase)、莽草酸羟基肉桂酰转移酶(HCT, Shikimate O-hydroxycinnamoyltransferase)、肉桂酰辅酶A还原酶(CCR, Cinnamoyl-CoA reductase)、肉桂醇脱氢酶(CAD, Cinnamyl-alcohol dehydrogenase)、CCoAOMT等^[2,13]。其中,转CCoAOMT基因的研究显示CCoAOMT下调造成的木质素含量降低,以及G木质素单体降低伴随的S/G木质素单体比例增加^[2,14],这对造纸中木质素的去除难易和饲料中的可消化性具有重要影响。反义CCoAOMT基因调控的转基因玉米与野生型对照相比,木质素平均含量降低了22.4%,纤维素含量增加了23.3%,植株表型正常^[2]。而CCoAOMT基因过表达的玉米植株比野生型植株抗枯叶病能力增强^[4],窄叶野豌豆(*Vicia sativa* L.) CCoAOMT基因在拟南芥中过表达还能使转基因植株对重金属镉的吸收能力提高^[15]。可见,利用遗传工程手段对植物中CCoAOMT基因进行调节,在不影响植物正常生长的条件下,结合研究目标对木质素含量和质量进行调控有重要意义。

通过15个科26个物种的CCoAOMT基因的密码子偏好性结果显示,CCoAOMT基因在禾本科物种中的表达量普遍较高,GC含量高且密码子偏性较强;在该基因的系统进化关系中,单子叶植物与双子叶植物的亲缘关系比裸子植物还要远。这些结果说明CCoAOMT基因可能在裸子植物向被子植物进化之前就已经分化成完全不同的类群。而不同物种CCoAOMT基因密码子偏好性和基因结构的差异因为可能引起蛋白质结合构象的改变,因此对基因功能有巨大的影响,引起植物中木质素和苯丙烷合成途径中其他代谢产物水平的差异^[4,16]。

巨龙竹秆材优异,是滇西南和滇南地区最具发展前景的优良竹种。但巨龙竹的分子生物学研究发展较晚,有关巨龙竹木质素生物合成及细

胞壁形成机理的研究更是十分有限。分析前期巨龙竹转录组测序数据注释结果,巨龙竹中至少存在CCoAOMT基因家族的2个成员,二者之间的表达量相差超过100倍,选取其中表达量高的CCoAOMT基因进行克隆。通过对比,所克隆巨龙竹的CCoAOMT基因密码子偏好性与慈竹、玉米、水稻、南荻等类似,禾本科植物CCoAOMT基因的高表达,可能与禾本科植物纤维和导管的形成密切相关,巨龙竹CCoAOMT基因在竹笋和茎秆中的高表达也印证了这一结果(图5)。由此可见,在基因的功能研究中,禾本科可能提供优良的CCoAOMT基因资源。但巨龙竹的组织培养体系尚未建立,限制了内源基因在它本身进行的功能验证。为了筛选基因功能验证中有效的受体植物,通过密码子使用频率的对比发现龙眼生长素受体基因TIR1密码子与3种模式植物拟南芥、烟草、番茄分别有7、6和6个密码子使用频率差异^[12],而巨龙竹CCoAOMT基因与拟南芥有2个密码子使用频率差异较大,与烟草、番茄没有密码子频率使用差异,说明CCoAOMT基因可在拟南芥、烟草、番茄中进行有效的功能验证。但目标基因的异源表达还受miRNA转录后调控、mRNA结构稳定性、基因甲基化等各种因素的影响,如何实现巨龙竹CCoAOMT基因有效的功能表达有待深入研究。

参考文献

- [1] Wang G F, Balint-Kurti P J. Maize homologs of CCoAOMT and HCT, two key enzymes in lignin biosynthesis, form complexes with the NLR Rpl protein to modulate the defense response. *Plant Physiology*, 2016, 171(3): 2166-2177
- [2] Li X, Chen W, Zhao Y, Xiang Y, Jiang H, Zhu S, Cheng B. Downregulation of caffeoyl-CoA O-methyltransferase (CCoAOMT) by RNA interference leads to reduced lignin production in maize straw. *Genetics and Molecular Biology*, 2013, 36(4): 540
- [3] 谷振军, 章怀云, 张党权, 谢耀坚, 何含杰, 陈丽莉, 彭宽, 刘果, 杨丹. 赤桉CCoAOMT亚家族的基因克隆及可变剪切分析. *林业科学*, 2014, 50(5): 62-69
Gu Z J, Zhang H Y, Zhang D Q, Xie Y J, He H J, Chen L L, Peng K, Liu G, Yang D. Gene cloning and alternative splicing of CCoAOMT subfamily genes from *Eucalyptus camaldulensis*. *Scientia Silvae Sinicae*, 2014, 50(5): 62-69
- [4] Yang Q, He Y, Kabahuma M, Chaya T, Kelly A, Borrego E, Bian Y, Kasmi F E, Yang L, Teixeira P, Kolkman J, Nelson R, Kolomiets M, Dangl J L, Wissner R, Caplan J, Li X, Lauter N, Balint-Kurti P. A gene encoding maize caffeoyl-CoA O-methyltransferase confers quantitative resistance to multiple pathogens. *Nature Genetics*, 2017, 49(9): 1364-1372

- [5] Do C T, Pollet B, Thévenin J, Sibout R, Denoue D, Barrière Y, Lapierre C, Jouanin L. Both caffeoyl Coenzyme A 3-*O*-methyltransferase 1 and caffeic acid *O*-methyltransferase 1 are involved in redundant functions for lignin, flavonoids and sinapoyl malate biosynthesis in *Arabidopsis*. *Planta*, 2007, 226 (5): 1117-1129
- [6] 吴晓宇, 胡尚连, 曹颖, 卢学琴, 任鹏, 周美娟, 李晓瑞. 慈竹 *CCoAOMT* 基因的克隆及生物信息学分析. 南京林业大学学报: 自然科学版, 2012, 36 (3): 17-22
Wu X Y, Hu S L, Cao Y, Lu X Q, Ren P, Zhou M J, Li X R. Cloning of *CCoAOMT* gene in *Neosinocalamus affinis* and its bioinformatics analysis. *Journal of Nanjing Forestry University: Natural Science Edition*, 2012, 36 (3): 17-22
- [7] Hong D Y, Blackmore S. *Plants of China//A Companion to the Flora of China*. Beijing: Science Press, 2005: 319-320
- [8] 辉朝茂, 张国学, 李在留, 谷中明, 杨宇明. 珍稀竹种巨龙竹种群特性及其保护研究. 竹子学报, 2004, 23 (4): 4-9
Hui C M, Zhang G X, Li Z L, Gu Z M, Yang Y M. A study on the population characteristics of the valuable and rare bamboo species of *Dendrocalamus sinicus* and its conservation in Yunnan, China. *Journal of Bamboo Research*, 2004, 23 (4): 4-9
- [9] 孙金金, 朱金鑫, 王娟, 杨宇明, 王毅. 巨龙竹肉桂醇脱氢酶基因 (*DsCAD*) 的克隆及功能分析. 竹子研究汇刊, 2016, 35 (4): 1-7
Sun J J, Zhu J X, Wang J, Yang Y M, Wang Y. Cloning and functional analysis of cinnamyl alcohol dehydrogenase gene from *Dendrocalamus sinicus*. *Journal of Bamboo Research*, 2016, 35 (4): 1-7
- [10] Gamuyao R, Nagai K, Ayano M, Mori Y, Minami A, Kojima M, Suzuki T, Sakakibara H, Higashiyama T, Ashikari M, Reuscher S. Hormone distribution and transcriptome profiles in bamboo shoots provide insights on bamboo stem emergence and growth. *Plant and Cell Physiology*, 2017, 58: 702-716
- [11] Chen L, Guo X, Cui Y, Zheng X, Yang H. Comparative transcriptome analysis reveals hormone signaling genes involved in the launch of culm-shape differentiation in *Dendrocalamus sinicus*. *Genes*, 2018, 9 (1): 4
- [12] 赖瑞联, 林玉玲, 钟春水, 赖钟雄. 龙眼生长素受体基因 *TIR1* 密码子偏好性分析. 园艺学报, 2016, 43 (4): 771-780
Lai R L, Lin Y L, Zhong C S, Lai Z X. Analysis of codon bias of auxin receptor gene *TIR1* in *Dimocarpus longan*. *Acta Horticulturae Sinica*, 2016, 43 (4): 771-780
- [13] 裴艳梅, 王金鑫, 彭建营. 无核小枣木质素合成基因 *Zj4CL* 的克隆和表达分析. 植物遗传资源学报, 2016, 17 (1): 147-152
Pei Y M, Wang J X, Peng J Y. Molecular cloning and expression analysis of *4CL* gene in *Ziziphus jujuba* Mill. 'Wuhexiaozao'. *Journal of Plant Genetic Resource*, 2016, 17 (1): 147-152
- [14] Boerjan W, Ralph J, Baucher M. Lignin biosynthesis. *Annual Review of Plant Biology*, 2003, 54: 519-546
- [15] Xia Y, Liu J, Wang Y, Zhang X, Shen Z, Hu Z. Ectopic expression of *Vicia sativa* Caffeoyl-CoA *O*-methyltransferase (*VsCCoAOMT*) increases the uptake and tolerance of cadmium in *Arabidopsis*. *Environmental and Experimental Botany*, 2018, 145: 47-53
- [16] Walker A M, Sattler S A, Regner M R, Jones J P, Ralph J, Vermerris W, Sattler S M, Kang C. Determination of the structure and catalytic mechanism of *Sorghum bicolor* caffeoyl-CoA *O*-methyltransferase. *Plant Physiology*, 2016, 172 (1): 78-92