

张晓东,李彩霞,王元忠. 滇龙胆草 7-脱氧番木鳖酸-7-羟化酶基因的克隆与表达分析[J]. 江苏农业科学,2019,47(17):76-81.  
doi:10.15889/j.issn.1002-1302.2019.17.018

# 滇龙胆草 7-脱氧番木鳖酸-7-羟化酶基因的克隆与表达分析

张晓东<sup>1</sup>, 李彩霞<sup>1</sup>, 王元忠<sup>2</sup>

(1. 玉溪师范学院化学生物与环境学院,云南玉溪 653100; 2. 云南省农业科学院药用植物研究所,云南昆明 650223)

**摘要:**克隆滇龙胆 7-脱氧番木鳖酸-7-羟化酶基因 *GrDL7H*,并进行表达分析,为龙胆苦苷生物合成途径的解析奠定基础。根据滇龙胆转录组 *GrDL7H* 序列,使用 RT-PCR(逆转录 PCR)和 3'RACE(cDNA 末端快速扩增)技术从滇龙胆幼叶中克隆该基因及其启动子,并进行序列分析和组织特异性表达分析。结果表明,*GrDL7H* 基因(登录号:KT306971)全长 2 154 bp,包含 5 个外显子和 4 个内含子,其中 *GrDL7H* 基因(登录号:KP340980)开放阅读框长 1 542 bp,编码 513 个氨基酸;*GrDL7H* 蛋白质相对分子质量为 59.08 ku,等电点(pI 值)为 8.88,属于 CYP450 蛋白超家族成员,可能定位于叶绿体;*GrDL7H* 蛋白质无信号肽,为亲水稳定蛋白质,主要由  $\alpha$ -螺旋和环构成;*GrDL7H* 蛋白质具有 CYP450 蛋白质保守结构域,功能为参与氧化还原反应;滇龙胆 *GrDL7H* 蛋白质与金鸡纳树 *CcDL7H* 蛋白质的亲缘关系最近;*GrDL7H* 基因启动子主要包含 6 个光应答元件、1 个赤霉素应答元件、6 个参与茉莉酸甲酯应答的顺式调控元件和 1 个热胁迫应答元件等;组织特异性表达分析结果表明,*GrDL7H* 基因主要在叶片中表达。

**关键词:**滇龙胆;7-脱氧番木鳖酸-7-羟化酶;基因克隆;表达分析

**中图分类号:** S188;Q78 **文献标志码:** A **文章编号:** 1002-1302(2019)17-0076-06

滇龙胆(*Gentiana rigescens* Franch. ex Hemsl.)为多年生草本植物,主要分布于云南、四川等地,其根部主要用来治疗肝炎和胆囊炎<sup>[1-2]</sup>。目前,由于滇龙胆的市场需求量逐年增加,野生滇龙胆遭到人为大肆破坏<sup>[3]</sup>。从合成生物学角度来看,滇龙胆的主要药效成分为龙胆苦苷,要大量合成龙胆苦苷,首先必须先弄清龙胆苦苷的生物合成途径及其调控机制。

萜类生物合成途径包括基本前体形成、碳骨架结构形成和萜类化合物的结构后修饰 3 个阶段<sup>[4]</sup>。7-脱氧番木鳖酸-7-羟化酶(*CrDL7H*)是萜类龙胆苦苷生物合成途径第二阶段重要的酶,能够催化 7-脱氧番木鳖酸生成番木鳖酸<sup>[5]</sup>。在长春花中,有研究者通过病毒诱导的基因沉默技术鉴定了 7-脱氧番木鳖酸-7-羟化酶,*CrDL7H* 基因沉默后,裂环番木鳖苷含量减少了 70%,7-脱氧番木鳖酸鲜质量含量增加到 4 mg/g,且在裂环番木鳖苷途径中羟基化先于羧基甲基化<sup>[5]</sup>。使用表达 *CrDL7H* 酶的酵母微体进行酶动力学分析,结果表明其米氏常数  $K_m$  和最大速率  $V_{max}$  分别为  $(111.07 \pm 14.80) \mu\text{mol}/\text{min}$  和  $(5.50 \pm 0.77) \mu\text{mol}/\text{min}$ <sup>[6]</sup>。目前,*CrDL7H* 酶基因已从长春花、小曼长春花、萝芙木、秀丽藤、金银花、金鸡纳树、水甘草等许多植物中分离<sup>[5]</sup>,但是目前仅在长春花中进行过详细的研究。*DL7H* 基因表达具有组

织特异性。在川西獐牙菜中,*SmDL7H* 基因在叶中的表达量最高,其次是茎和花,在根中的表达量最低;并且与 *G10H*、*IS*、*IO*、*SLS2* 等基因表达量相比,*DL7H* 基因的表达量相对较低<sup>[7]</sup>。在长春花中,*CrDL7H* 基因主要在叶中表达,在第 1 对叶中表达量最高,在第 2、3、4、5 对叶中的表达量逐渐降低,且对应的长春碱与裂环番木鳖苷的含量逐渐减少;其次是茎和花,表达量最低的为根<sup>[8]</sup>。

目前,由于龙胆科植物的基因组并未测序,导致龙胆苦苷生物合成途径并不完全清楚<sup>[9]</sup>。本研究根据笔者于 2013 年在滇龙胆转录组数据库中检索到的 *GrDL7H* 基因序列,设计特异性引物,通过逆转录 PCR(RT-PCR)技术成功从滇龙胆叶片中扩增到 *GrDL7H* 基因,然后根据所克隆基因序列设计引物,对该基因的启动子序列进行扩增和分析,最后对二年生滇龙胆根、茎、叶、花等不同组织中 *GrDL7H* 基因的表达情况进行分析,以期为滇龙胆龙胆苦苷生物合成途径的解析提供帮助。

## 1 材料与方法

### 1.1 试验材料

基因克隆和启动子扩增所用材料为滇龙胆无菌苗幼叶,采自玉溪师范学院分子生物学实验室,经云南省农业科学院药用植物研究所金航研究员鉴定为滇龙胆,采样日期为 2014 年 5 月 17 日。qPCR(实时定量基因扩增)分析使用的二年生滇龙胆植株采自临沧市云县临沧耀阳生物药业科技有限公司后箐基地,由该公司技术经理董顺福鉴定为滇龙胆,采样日期为 2018 年 4 月 30 日。

### 1.2 试验方法

1.2.1 *GrDL7H* 基因的扩增 采用多糖植物组织提取试剂盒

收稿日期:2018-06-03

基金项目:云南省科技计划(编号:2016FD113);云南省地方本科高校基础研究联合专项(编号:2017FH001-024);云南省自然科学基金重点项目(编号:2017FA049)。

作者简介:张晓东(1980—),男,河南新郑人,博士,副教授,主要从事植物代谢基因工程研究。E-mail:zxd95@126.com。

通信作者:王元忠,博士,副研究员,主要从事药用植物资源评价与利用研究。E-mail:boletus@126.com。

提取滇龙胆幼叶总 RNA,采用逆转录试剂盒合成 cDNA,采用植物基因组 DNA 提取试剂盒[宝生物工程(大连)有限公司]提取滇龙胆叶片 DNA。根据前期测序的滇龙胆转录组 *GrDL7H* 基因转录本序列,设计 1 对特异引物 *GrDL7H* EcoRⅠ-F、*GrDL7H*-R(表 1)。以 cDNA 为模板进行 PCR 扩增,反应体系如下:PrimeSTAR Max Premix(2×, TaKaRa)25 μL,cDNA 模板 2 μL,正反向引物(10 μmol/L)各 1 μL,加 ddH<sub>2</sub>O 补足 50 μL。PCR 反应条件如下:98 ℃ 10 s,55 ℃ 10 s,72 ℃ 10 s,30 个循环;72 ℃ 5 min。PCR 产物经 1.0% 琼脂糖凝胶电泳分离,然后割胶,使用胶回收试剂盒(德国 QIAGEN)对目的片段进行回收,将其连接到 pMD19-T 载体上。转化大肠杆菌 DH5α[宝生物工程(大连)有限公司]后进行蓝白斑筛选,挑取 12 个白斑摇菌;使用试剂盒(北京百泰克生物技术有限公司)提取质粒,经酶切检测正确后进行测序[生工生物工程(上海)股份有限公司],获得重组质粒 pMD19-*fGrDL7H*。根据测序结果,设计基因特异性引物 *GrDL7H* 3RACE-F,进行 3' cDNA 末端快速扩增(RACE)。使用 DNAMAN 软件比对后,进行基因片段拼接,获得 *GrDL7H* 基因 ORF(开放阅读框)全长。设计基因特异性引物,分别以 cDNA 和 gDNA 为模板进行 PCR 扩增,PCR 产物进行 TA 克隆后,分别进行 DNA 测序。

1.2.2 *GrDL7H* 基因启动子的克隆 根据“1.2.1”节中克隆到的滇龙胆 *GrDL7H* 基因序列,设计基因特异引物 *GrDL7H*-sp2 和 *GrDL7H*-sp3(表 1)。采用 Universal Genome Walker 2.0 试剂盒,分别使用 *Dra*Ⅰ、*EcoR*Ⅴ、*Pvu*Ⅱ 和 *Stu*Ⅰ 酶切滇龙胆基因组 DNA,加接头后,构建文库 DL1、DL2、DL3 和 DL4,然后进行 2 轮 PCR 扩增。第 1 轮 PCR 总体系为 25 μL,具体成分如下:DL1 gDNA 1 μL,50× Advantage 2 聚合酶混合物 0.5 μL,10× Advantage 2 PCR Buffer 1 μL,API(10 μmol/L)1 μL,dNTP(10 mmol/L)1 μL, *GrDL7H*-sp2(10 μmol/L)1 μL,ddH<sub>2</sub>O 19.5 μL,反应条件为 94 ℃ 25 s,72 ℃ 3 min,7 个循环;94 ℃ 25 s,67 ℃ 3 min,32 个循环;63 ℃ 7 min。第 2 轮 PCR 总体系为 25 μL,具体成分如下:第 1 轮 PCR 产物 1 μL,50× Advantage 2 聚合酶混合物 0.5 μL,10× Advantage 2 PCR 1 μL,API(10 μmol/L)1 μL 缓冲液,dNTP(10 mmol/L)1 μL, *GrDL7H*-sp3(10 μmol/L)1 μL,ddH<sub>2</sub>O 19.5 μL,反应

条件为 94 ℃ 25 s,72 ℃ 3 min,5 个循环;94 ℃ 25 s,67 ℃ 3 min,20 个循环;63 ℃ 7 min。第 2 轮 PCR 产物经 TA 克隆,获得重组质粒 pMD19-*pGrDL7H*,菌液送生工生物工程(上海)股份有限公司进行 DNA 测序。

1.2.3 *GrDL7H* 基因及启动子的生物信息学分析 用美国国立生物技术信息中心(NCBI)网站的 BLAST 程序进行序列比对,用 Genetyx 6.1.8 软件进行翻译,用 DNAMAN 7 进行多序列比对;用 Clustal X 2.1 进行比对,然后使用 MEGA 6.0 软件内置的 NJ(邻接)法构建系统进化树,设置 Bootstrap(自抽样)=1 000;利用在线数据库([http://molbiol.edu.ru/eng/scripts/01\\_11.html](http://molbiol.edu.ru/eng/scripts/01_11.html))进行稀有密码子分析。使用 ChloroP 服务器 v 1.1 进行叶绿体转运肽预测;使用 Interpro 软件进行保守结构域预测;使用 ProtScale 软件进行疏水性分析,方法为 Hphob./Kyte & Doolittle 量表法;使用 PredictProtein 对二级结构进行预测;使用 Phyre2(<http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index>)对三级结构进行预测;利用 Expasy 中的 TMHMM 工具预测蛋白质的跨膜螺旋区;利用在线工具 Wolf Psort 预测蛋白质的亚细胞定位情况。使用 GT-AG 原则,将 ORF 序列与 gDNA 序列进行比对,进行外显子与内含子的分析。使用启动子在线分析网站 PlantCARE(<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>)对克隆到的 *GrDL7H* 基因启动子序列进行分析。

1.2.4 *GrDL7H* 基因的荧光实时定量分析 分别取二年生滇龙胆的根、茎、叶、花,使用 MiniBEST Plant RNA Extraction Kit[宝生物工程(大连)有限公司]提取总 RNA,使用 PrimeScript RT reagent Kit with gDNA Eraser[宝生物工程(大连)有限公司]处理除去基因组 DNA,然后反转录合成第 1 链 cDNA。以转录组中 *GrACTIN* 基因(GenBank 登录号为 KM061807)作为内参和 *GrDL7H* 基因 ORF 序列设计特异性引物(表 1)。PCR 体系如下:TB GreenPremix Ex TaqⅡ(Tli RNaseH Plus)10 μL,ddH<sub>2</sub>O 7 μL,cDNA 1 μL,正反向引物各 1 μL。qPCR 条件:95 ℃ 5 min;95 ℃ 15 s,60 ℃ 15 s,72 ℃ 20 s,重复 4 次。在 LightCycler 480Ⅱ 荧光定量 PCR 仪(Roche,瑞士)上进行扩增反应,扩增曲线、溶解曲线、标准曲线由软件自动生成。使用内参基因 *GrACTIN* 表达校准后,计算根、茎、叶中 *GrDL7H* 基因的相对表达量。采用比较 C<sub>T</sub> 值的 2<sup>-ΔΔC<sub>T</sub></sup> 的方法进行定量

表 1 *GrDL7H* 基因克隆和荧光定量分析的相关引物序列

基因用途	引物名称	引物序列 (5'→3')	退火温度 T <sub>m</sub> (℃)
<i>GrDL7H</i> 转录本扩增	<i>GrDL7H</i> EcoRⅠ-F	GAATTCATGGGATTAGACTTGGTGCC	55
	<i>GrDL7H</i> -R	TGCATAGTGAAGTCTGTGAAGG	55
3'RACE 扩增	<i>GrDL7H</i> 3RACE-F	GCTGAACAAACGAAAGTGCGCG	63
	NUP	AAGCAGTGGTATCAACGCAGAGT	63
<i>GrDL7H</i> 基因克隆	<i>GrDL7H</i> EcoRⅠ-F	GAATTCATGGGATTAGACTTGGTGCC	55
	<i>GrDL7H</i> EcoRⅠ-R	CTCGAGTTAGACTTTGTGCAAGATCAAATGAG	55
<i>GrDL7H</i> 基因启动子克隆	<i>GrDL7H</i> -sp2	ACTCTCACTGGTCTCCCAAG	67
	<i>GrDL7H</i> -sp3	TCCAGTCCAAGAATCTGTGTACC	67
<i>GrDL7H</i> 基因 qPCR	<i>GrDL7H</i> -F	GCGATGGAGTTCAACCCAGA	60
	<i>GrDL7H</i> -R	GTGCATGAACATACGACGGC	60
<i>GrACTIN</i> 基因 qPCR	<i>GrACTIN</i> -F	GCGGATCGTATGAGCAAGGA	60
	<i>GrACTIN</i> -R	GGGCCCGGATTCTTCGTATT	60

数据的分析处理。

## 2 结果与分析

### 2.1 滇龙胆 *GrDL7H* 基因的克隆

以滇龙胆幼叶 cDNA 为模板,使用表 1 中的基因特异性引物 *GrDL7H* EcoR I - F 和 *GrDL7H* - R 扩增出约 1 500 bp 的片段。通过 TA 克隆获得重组质粒 pMD19 - *fGrDL7H*, 然后进行 DNA 测序。结果表明, *fGrDL7H* 长 1 502 bp, 不含 3' 末端。

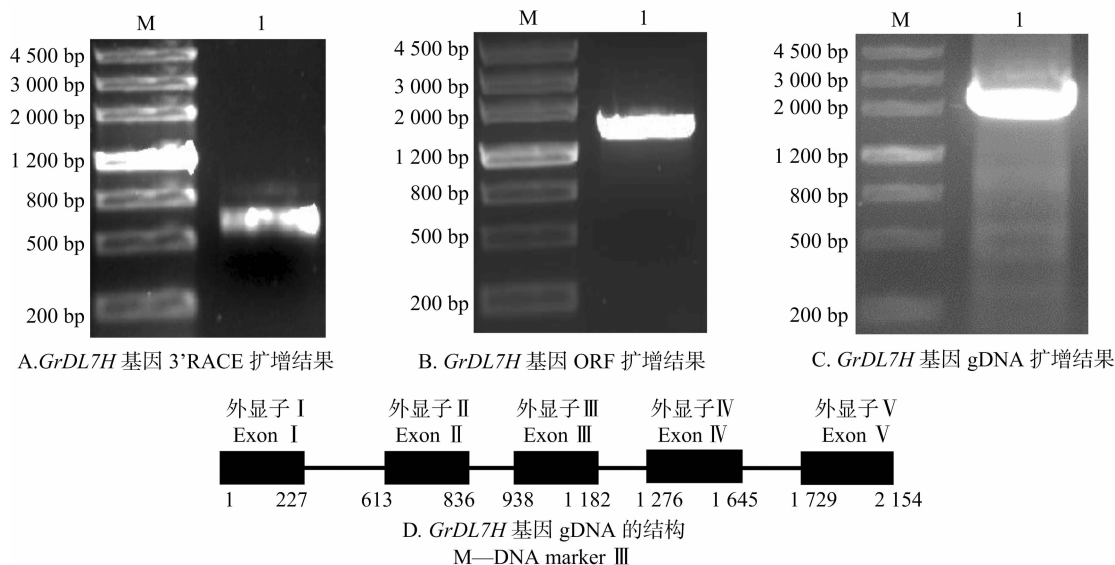


图1 滇龙胆 *GrDL7H* 基因的 PCR 扩增结果及其 gDNA 的结构

### 2.2 *GrDL7H* 基因的生物信息学分析

2.2.1 *GrDL7H* 基因的 ORF 分析和多序列比对分析 利用 Genetyx 6 软件对 *GrDL7H* 基因的 ORF 序列进行分析,该基因 (GenBank 登录号为 KP722033.1) 长 1 542 bp, 编码 513 个氨基酸。

利用 GenBank 数据库的 BLASTP 程序对 *GrDL7H* 蛋白质进行比对分析,结果表明,滇龙胆 *GrDL7H* 与萝卜木 *RsDL7HT* (78.17%)、金鸡纳树 *CcDL7H* (77.78%)、长春花 *CrDL7H* (77.73%) 的序列相似性均较高,与拟南芥 *AtDL7H* (51.47%) 的蛋白质相似性稍低。利用 DNAMAN 7 将 *GrDL7H* 蛋白质序列与 NCBI 中相似性较高的部分序列进行多序列比对分析,结果表明, *GrDL7H* 蛋白质与已知蛋白质序列的保守性较高(图 2)。利用 MEGA 7.0 软件将 *GrDL7H* 氨基酸序列与 NCBI 中相似性较高的部分序列进行系统发育分析,结果显示,滇龙胆 *GrDL7H* 蛋白质与金鸡纳树 *CcDL7H* 的亲缘关系较近(图 3)。

2.2.2 *GrDL7H* 蛋白质的理化特性分析 使用 ExPASy ProtParam tool 对 *GrDL7H* 蛋白质进行分析,结果表明, *GrDL7H* 蛋白质单体的相对分子质量为 59 077 u,理论等电点 (pI 值) 为 8.88,这与长春花 *CrDL7H* 蛋白质类似<sup>[8]</sup>;带正电氨基酸残基 (Arg + Lys) 为 70 个,带负电氨基酸残基 (Asp + Glu) 为 63 个,化学方程式为  $C_{2686}H_{4240}N_{708}O_{743}S_{23}$ ;不稳定指数为 31.37,属稳定蛋白质;脂肪指数为 94.81;总平均疏水性 (GRAVY) 为 -0.173,为亲水蛋白质(图 4)。 *GrDL7H* 蛋白质含有 20 种基本氨基酸,其中亮氨酸含量最高,为 11.9%;其次是赖氨酸、异亮氨酸、苏氨酸,含量分别为 7.60%、6.40%、

使用 SMART RACE 试剂盒进行扩增,获得 600 bp DNA 片段(图 1 - A)。使用 DNAMAN 软件对 *GrDL7H* 基因片段进行组装,结果发现 *GrDL7H* 基因 ORF 全长 1 542 bp(图 1 - B,登录号: KP340980)。使用 *GrDL7H* EcoR I - F 和 *GrDL7H* EcoR I - R 对 *GrDL7H* 基因的基因组 DNA (gDNA) 进行扩增,结果显示 *GrDL7H* 基因长 2 154 bp(图 1 - C,登录号: KT306971),包含 5 个外显子和 4 个内含子(图 1 - D)。

6.40%;半胱氨酸含量最低,为 1.00%。

2.2.3 *GrDL7H* 蛋白质的二级结构、三级结构分析 利用 PredictProtein 软件对 *GrDL7H* 蛋白质进行二级结构分析,结果表明:该蛋白质二级结构中  $\alpha$  - 螺旋 (H) 占 54.58%,环 (L) 占 37.43%,延伸带 (E) 占 7.99%。利用 Phyre 2 软件预测 *GrDL7H* 蛋白质的三级结构,结果见图 5,该模型以细胞色素 p450b1 酶 [c5t6qA] 为模板,在第 15 ~ 513 位氨基酸处建模,序列相似度为 24%。

2.2.4 *GrDL7H* 蛋白质序列的分析归类 使用 InterPro 6.0 在线工具对 *GrDL7H* 蛋白质进行分析和归类,结果表明, *GrDL7H* 蛋白质属于细胞色素 P450 超家族 E 类、类群 I (IPR002401) 的成员,包含细胞色素 P450 保守结构域 (IPR036396) (图 6)。GO 预测结果表明,在生物过程方面 *GrDL7H* 蛋白质参与氧化还原过程,在分子功能方面,该蛋白质参与铁离子结合、氧化还原酶活性、血红素结合,在细胞成分方面未预测到结果。

2.2.5 *GrDL7H* 蛋白质信号肽分析 利用 ExPASy SignalP 4.1 服务器分析 *GrDL7H* 蛋白质,未发现信号肽,表明该蛋白质为非分泌型蛋白质。利用 TMHMM 2.0 工具预测 *GrDL7H* 蛋白质的跨膜螺旋区,结果表明: *GrDL7H* 蛋白质在 5 ~ 27 位氨基酸含有跨膜区域,其余的 1 ~ 4 位氨基酸在细胞内,28 ~ 513 位氨基酸在细胞外(图 7)。

2.2.6 *GrDL7H* 蛋白质亚细胞定位分析 使用 WoLF PSORT 软件进行 *GrDL7H* 蛋白的亚细胞定位分析,结果显示, *GrDL7H* 蛋白质在叶绿体、细胞核、核骨架、细胞膜、细胞质、液泡的定位系数分别为 6.0、3.5、2.5、2 和 1.0。

滇龙胆( <i>Gentiana rigescens</i> )(AKI87771.1)	MGLDLVAFALTALIALSTLYVHRFLDWVWF1PKKLEKLRDQGLKGNRYRLDGDQYFSGKLIKESKPIGVEEDIKRRIVPHILK	87
水甘草( <i>Amsonia hubrichtii</i> )(AGX93058.1)	.MENFKSVAVLGFVFCVALYVYRLDWWFEPKKLEKLRDQGFEGKNGYRLDGDQYFSGKLIKESKPIGLEEFVKKRIIPHILK	86
罗芙木( <i>Rauvolfia serpentina</i> )(AGX93059.1)	MEVSFKSVTVLGFVGLALYVYRVLDWWFEPKKLEKLRDQGFEGKNGYRLDGDQYFSGKLIKESKPIGVEEDVKKRIIPHILK	87
长春花( <i>Catharanthus roseus</i> )(AGX93062.1)	MELNFKSIIFLVFSLTLYVYRLDWWFEPKKLEKLRDQGFEGKNGYRLDGDQYFSGKLIKESKPIGVEEDVKKRIIPHILK	87
秀丽藤( <i>Tabernaemontana elegans</i> )(AGX93060.1)	MEMNFKLVAVLGFIGLALYVYRVLDWWFEPKKLEKLRDQGFEGKNGYRLDGDQYFSGKLIKESKPIGVEEDVKKRIIPHILK	87
黄金鸡纳( <i>Cinchona calisaya</i> )(AGX93057.1)	MELNFTSIAILGFIGLSIYVYRVLDWWFEPKKLEKLRDQGFEGKNGYRLDGDQYFSGKLIKESKPIGVEEDVKKRIIPHILK	87
忍冬( <i>Lonicera japonica</i> )(AGX93056.1)	MMMSYNLIGGSLIFGVITYVYVSLNMFEPKKLEKLRDQGFEGKNGYRLDGDQYFSGKLIKESKPIGVEEDVKKRIIPHILK	87
一致序列; 相同序列	y l w w p k k l e k l r g g g n y r l l g d q s i a p i d r p h l k	
滇龙胆( <i>Gentiana rigescens</i> )(AKI87771.1)	AVQTHGKSFMMWGRIPRVFVITPELIREVLTLYKYKFNKHHCLDPITKILTGIGSLGEGDFAWRRRIINSAFHEFKLKLMLPAFY	174
水甘草( <i>Amsonia hubrichtii</i> )(AGX93058.1)	TVETHGKSNFMMWGRIPRVQVITPELIRKVLTKYKFKCNHHLDPITKILTGIGSLGEGDFAWRRRIINSAFHEFKLKLMLPAFY	173
罗芙木( <i>Rauvolfia serpentina</i> )(AGX93059.1)	TVETHGKSNFMMWGRIPRVQVITPELIRKVLTKYKFKCNHHLDPITKILTGIGSLGEGDFAWRRRIINSAFHEFKLKLMLPAFY	174
长春花( <i>Catharanthus roseus</i> )(AGX93062.1)	TVCTHGKSNFMMWGRIPRVNITPELIRKVLTKYKFKCNHHLDPITKILTGIGSLGEGDFAWRRRIINSAFHEFKLKLMLPAFY	174
秀丽藤( <i>Tabernaemontana elegans</i> )(AGX93060.1)	TVETHGKSNFMMWGRIPRVHVTPELIRKVLTKYKFKCNHHLDPITKILTGIGSLGEGDFAWRRRIINSAFHEFKLKLMLPAFY	174
黄金鸡纳( <i>Cinchona calisaya</i> )(AGX93057.1)	TVQTHGKSNFMMWGRIPRVNITPELIRKVLTKYKFKCNHHLDPITKILTGIGSLGEGDFAWRRRIINSAFHEFKLKLMLPAFY	174
忍冬( <i>Lonicera japonica</i> )(AGX93056.1)	TMNHHGKSNFMMWGRIPRLHITPELIRKVLTKYKFKCNHHLDPITKILTGIGSLGEGDFAWRRRIINSAFHEFKLKLMLPAFY	174
一致序列; 相同序列	h g k s f m w g r i p r v t p e l i r e v l t k y k f c n h h l d p i t k l t g i g s l e g d f a w r r i n s a f h e f k l k l m l p a f y	
滇龙胆( <i>Gentiana rigescens</i> )(AKI87771.1)	LSORDVSKWKNKVE. EGGSSEVVDVHDIETLTGDVLSRTLFSGNVEGRRIFELMKELTALTITIVIRSVYIPGQRFTPKRNNRMR	260
水甘草( <i>Amsonia hubrichtii</i> )(AGX93058.1)	LSORDVSKWKNKVE. EGGSSEVVDVHDIETLTGDVLSRTLFSGNVEGRRIFELMKELTALTITIVIRSVYIPGQRFTPKRNNRMR	259
罗芙木( <i>Rauvolfia serpentina</i> )(AGX93059.1)	LSORDVSKWKNKVE. EGGSSEVVDVHDIETLTGDVLSRTLFSGNVEGRRIFELMKELTALTITIVIRSVYIPGQRFTPKRNNRMR	260
长春花( <i>Catharanthus roseus</i> )(AGX93062.1)	LSORDVSKWKNKVE. EGGSSEVVDVHDIETLTGDVLSRTLFSGNVEGRRIFELMKELTALTITIVIRSVYIPGQRFTPKRNNRMR	260
秀丽藤( <i>Tabernaemontana elegans</i> )(AGX93060.1)	LSORDVSKWKNKVE. EGGSSEVVDVHDIETLTGDVLSRTLFSGNVEGRRIFELMKELTALTITIVIRSVYIPGQRFTPKRNNRMR	260
黄金鸡纳( <i>Cinchona calisaya</i> )(AGX93057.1)	LSORDVSKWKNKVE. EDSSSEVVDVHDIETLTGDVLSRTLFSGNVEGRRIFELMKELTALTITIVIRSVYIPGQRFTPKRNNRMR	260
忍冬( <i>Lonicera japonica</i> )(AGX93056.1)	LSORDVSKWKNKVE. SSKGSSEVVDVHDIETLTGDVLSRTLFSGNVEGRRIFELMKELTALTITIVIRSVYIPGQRFTPKRNNRMR	261
一致序列; 相同序列	s c d v k w v s e v h d e t l t g d v i r t l f g s n e g i f e l m e l t i t i v s v y i p g r f t p k r n	
滇龙胆( <i>Gentiana rigescens</i> )(AKI87771.1)	AIDKQVRVRIETIINNMKMLKAGTAA..ASDFGLILLECNLNEIREQGNKNAGMTIEQILGCKLFYFAGQDITSTLLWMTVIL	339
水甘草( <i>Amsonia hubrichtii</i> )(AGX93058.1)	TIDKQVRVRIETIINNMKMLKAGTAA..GDFGLILLECNLNEIREQGNKNAGMTIEQILGCKLFYFAGQDITSTLLWMTVIL	344
罗芙木( <i>Rauvolfia serpentina</i> )(AGX93059.1)	AIDKQVRVRIETIINNMKMLKAGTAA..AADDGLILLECNLNEIREQGNKNAGMTIEQILGCKLFYFAGQDITSTLLWMTVIL	346
长春花( <i>Catharanthus roseus</i> )(AGX93062.1)	AIDKQVRVRIETIINNMKMLKAGTAA..GDFGLILLECNLNEIREQGNKNAGMTIEQILGCKLFYFAGQDITSTLLWMTVIL	347
秀丽藤( <i>Tabernaemontana elegans</i> )(AGX93060.1)	AIDKQVRVRIETIINNMKMLKAGTAA..GDFGLILLECNLNEIREQGNKNAGMTIEQILGCKLFYFAGQDITSTLLWMTVIL	345
黄金鸡纳( <i>Cinchona calisaya</i> )(AGX93057.1)	AIDKQVRVRIETIINNMKMLKAGTAA..GDFGLILLECNLNEIREQGNKNAGMTIEQILGCKLFYFAGQDITSTLLWMTVIL	345
忍冬( <i>Lonicera japonica</i> )(AGX93056.1)	KIDKQVRVRIETIINNMKMLKAGTAA..SSDFGLILLECNLNEIREQGNKNAGMTIEQILGCKLFYFAGQDITSTLLWMTVIL	345
一致序列; 相同序列	i d v r i i i k k g d f i l l e c n l n e i r e q g n k n a g m t i e q i l g c k l f y f a g o d t t s t l l w m t v i l	
滇龙胆( <i>Gentiana rigescens</i> )(AKI87771.1)	SREPWQTRAREEVLOVFGDKPTDYDGLISHLKIVTMIILEVRLIYTPVAELTKVAHEDTOLCKYHIFAGVQIMMPQMLLHHPDITWC	426
水甘草( <i>Amsonia hubrichtii</i> )(AGX93058.1)	SREPWQTRAREEVLOVFGDKPTDYDGLISHLKIVTMIILEVRLIYTPVAELTKVAHEDTOLCKYHIFAGVQIMMPQMLLHHPDITWC	431
罗芙木( <i>Rauvolfia serpentina</i> )(AGX93059.1)	SREPWQTRAREEVLOVFGDKPTDYDGLISHLKIVTMIILEVRLIYTPVAELTKVAHEDTOLCKYHIFAGVQIMMPQMLLHHPDITWC	433
长春花( <i>Catharanthus roseus</i> )(AGX93062.1)	SREPWQTRAREEVLOVFGDKPTDYDGLISHLKIVTMIILEVRLIYTPVAELTKVAHEDTOLCKYHIFAGVQIMMPQMLLHHPDITWC	434
秀丽藤( <i>Tabernaemontana elegans</i> )(AGX93060.1)	SREPWQTRAREEVLOVFGDKPTDYDGLISHLKIVTMIILEVRLIYTPVAELTKVAHEDTOLCKYHIFAGVQIMMPQMLLHHPDITWC	432
黄金鸡纳( <i>Cinchona calisaya</i> )(AGX93057.1)	SREPWQTRAREEVLOVFGDKPTDYDGLISHLKIVTMIILEVRLIYTPVAELTKVAHEDTOLCKYHIFAGVQIMMPQMLLHHPDITWC	432
忍冬( <i>Lonicera japonica</i> )(AGX93056.1)	SREPWQTRAREEVLOVFGDKPTDYDGLISHLKIVTMIILEVRLIYTPVAELTKVAHEDTOLCKYHIFAGVQIMMPQMLLHHPDITWC	432
一致序列; 相同序列	s p w q a r e e v v f g p d y d g i l k t m i l e l r y t p v a e l t k v a h e d t o l c k y h i f a g v q m p m l l h h i w g	
滇龙胆( <i>Gentiana rigescens</i> )(AKI87771.1)	EDVMEKPEREAGVILKATKSQGSFPFSLGPRMCIGQNFALIAKAMALILIRFSFELSISYVHAPYTLITMQPQGAHLILHK	512
水甘草( <i>Amsonia hubrichtii</i> )(AGX93058.1)	EDVMEKPEREAGVILKATKSQGSFPFSLGPRMCIGQNFALIAKAMALILIRFSFELSISYVHAPYTLITMQPQGAHLILHK	517
罗芙木( <i>Rauvolfia serpentina</i> )(AGX93059.1)	EDVMEKPEREAGVILKATKSQGSFPFSLGPRMCIGQNFALIAKAMALILIRFSFELSISYVHAPYTLITMQPQGAHLILHK	519
长春花( <i>Catharanthus roseus</i> )(AGX93062.1)	EDVMEKPEREAGVILKATKSQGSFPFSLGPRMCIGQNFALIAKAMALILIRFSFELSISYVHAPYTLITMQPQGAHLILHK	520
秀丽藤( <i>Tabernaemontana elegans</i> )(AGX93060.1)	EDVMEKPEREAGVILKATKSQGSFPFSLGPRMCIGQNFALIAKAMALILIRFSFELSISYVHAPYTLITMQPQGAHLILHK	518
黄金鸡纳( <i>Cinchona calisaya</i> )(AGX93057.1)	EDVMEKPEREAGVILKATKSQGSFPFSLGPRMCIGQNFALIAKAMALILIRFSFELSISYVHAPYTLITMQPQGAHLILHK	518
忍冬( <i>Lonicera japonica</i> )(AGX93056.1)	EDVMEKPEREAGVILKATKSQGSFPFSLGPRMCIGQNFALIAKAMALILIRFSFELSISYVHAPYTLITMQPQGAHLILHK	518
一致序列; 相同序列	d e f p r f g v k a t k s q g s f p f s l g p r m c i g n f a l e a k m a l i l i r f s f e l s s y v h a p y t l i t m q p q g a h l i l h k	

黑色部分表示相似性=100%; 深灰色部分表示相似性的范围: 75%≤相似性<100%; 浅灰色部分表示相似性的范围: 50%≤相似性<75%

图2 滇龙胆 GrDL7H 蛋白质与其他植物 DL7H 蛋白质的多序列比对分析



图3 滇龙胆 GrDL7H 蛋白质与其他植物 GrDL7H 蛋白质的系统发育分析

2.2.7 GrDL7H 基因稀有密码子分析 使用在线软件对 GrDL7H 基因进行稀有密码子分析,结果表明,在大肠杆菌和酿酒酵母中,GrDL7H 基因中稀有密码子分别占 1.36%、0.19%,且均无二联或三联稀有密码子连续出现的情况,因此可选用大肠杆菌表达菌 BL21、Rosetta (DE3) 或酿酒酵母进行蛋白质表达。

### 2.3 GrDL7H 基因启动子的克隆

以滇龙胆叶片 DNA 为模板,使用基因特异性引物分别进行 2 轮 PCR 扩增,第 2 轮 PCR 扩增产物通过 TA 克隆获得重

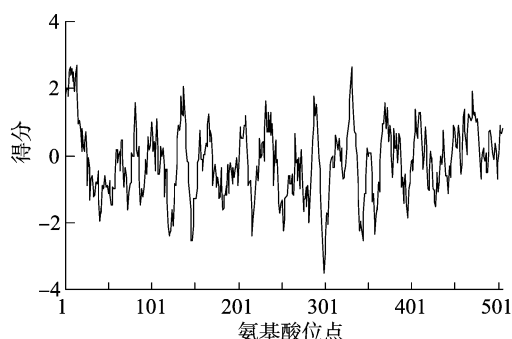


图4 滇龙胆 GrDL7H 蛋白质的疏水性分析

质粒 pMD19-pGrDL7H,测序后,上传至 NCBI 数据库,获得登录号 (GenBank 登录号: MG832585)。对获得的 GrDL7H 启动子进行生物信息学分析,结果表明,该启动子不但在 -30 区域附近含有转录起始的核心启动子元件 TATA 盒,而且含有启动子和增强子区域通用顺式作用元件 CAAT 盒;此外,还具有 6 个光应答元件、1 个赤霉素应答元件、6 个参与茉莉酸



图5 滇龙胆 GrDL7H 蛋白质的三维(3D) 结构预测结果



图6 滇龙胆 GrDL7H 蛋白质保守结构域的预测结果

甲酯应答的顺式调控元件、1 个热胁迫应答元件和 1 个参与生物钟控制的顺式调控元件(表 2)。

2.4 GrDL7H 基因的组织表达分析

取二年生滇龙胆的根、茎、花和叶,通过荧光定量 PCR (qRT-PCR)检测 GrDL7H 基因在根茎花叶等不同组织中的表达情况。结果显示,GrDL7H 基因表达量最高的组织是叶,

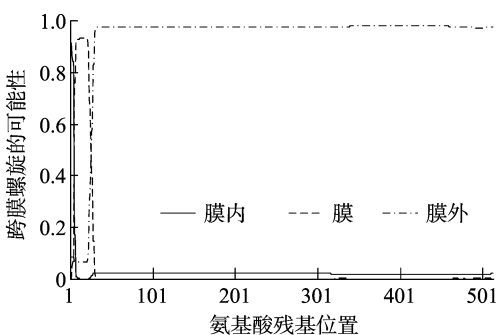


图7 滇龙胆 GrDL7H 蛋白质跨膜螺旋的可能性检测结果

其表达量大约分别是根、茎、花中的 50、32、6 倍;其次是花,表达量最低的是根(图 8)。这与川西獐牙菜 SmDL7H 基因在根茎花叶中的表达模式类似<sup>[7]</sup>。在长春花中,CrDL7H 基因主要在叶中表达,随着叶龄的增加,其表达量逐渐降低,相应的长春碱与裂环番木鳖苷的含量逐渐减少;其次是茎和花,表达量最低的是根<sup>[8]</sup>。这些结果表明,叶是滇龙胆龙胆苦苷生物合成的主要场所。

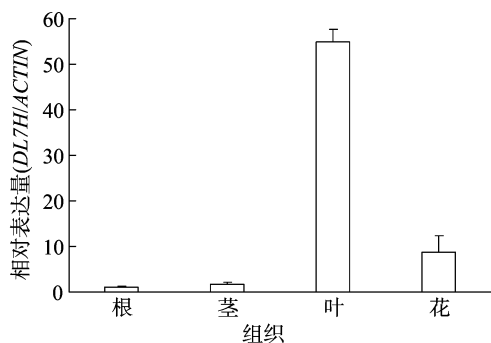
3 讨论

滇龙胆的主要药效成分为龙胆苦苷,属于环烯醚萜类物质。该类物质的合成主要包括 3 个阶段,第 1 阶段是通过甲

表 2 滇龙胆 GrDL7H 基因启动子功能预测(除 CAAT-box 和 TATA-box 外)

位点名称	物种	位点	链	得分	序列	功能
A-box	欧芹( <i>Petroselinum crispum</i> )	702	+	6.0	CCGTCC	顺式作用调控元件
AC-I	菜豆( <i>Phaseolus vulgaris</i> )	298	-	8.5	(T/C)C(T/C)(C/T)ACC(T/C)ACC	
Box 4	香芹( <i>Petroselinum crispum</i> )	292	+	6.0	ATTAAT	参与光响应的保守 DNA 模块的一部分
Box I	豌豆( <i>Pisum sativum</i> )	267	+	7.0	TTTCAAA	光响应元件
Box III	豌豆( <i>Pisum sativum</i> )	37	-	9.0	CATTTACACT	蛋白结合位点
CAT-box	拟南芥( <i>Arabidopsis thaliana</i> )	579	+	6.0	GCCACT	与分生组织表达相关的顺式作用调控元件
CAT-box	拟南芥( <i>Arabidopsis thaliana</i> )	733	-	6.0	GCCACT	与分生组织表达相关的顺式作用调控元件
CCGTCC-box	拟南芥( <i>Arabidopsis thaliana</i> )	702	+	6.0	CCGTCC	与分生组织特异性激活相关的顺式作用调控元件
CGTCA-motif	大麦( <i>Hordeum vulgare</i> )	306	-	5.0	CGTCA	参与茉莉酸甲酯(MeJA)响应的顺式调控元件
CGTCA-motif	大麦( <i>Hordeum vulgare</i> )	630	-	5.0	CGTCA	参与 MeJA 响应的顺式调控元件
CGTCA-motif	大麦( <i>Hordeum vulgare</i> )	618	+	5.0	CGTCA	参与 MeJA 响应的顺式调控元件
GAG-motif	大麦( <i>Hordeum vulgare</i> )	796	-	7.0	GGAGATG	光响应元件的一部分
GT1-motif	燕麦( <i>Avena sativa</i> )	139	-	7.0	GGTTAAT	光响应元件
GT1-motif	拟南芥( <i>Arabidopsis thaliana</i> )	140	-	6.0	GGTTAA	光响应元件
HSE	甘蓝( <i>Brassica oleracea</i> )	422	+	9.0	AAAAAATTTTC	参与热应激反应中的顺式作用元件
I-box	皱叶烟草( <i>Nicotiana plumbaginifolia</i> )	47	+	9.0	CTCTTATGCT	光响应元件的一部分
P-box	水稻( <i>Oryza sativa</i> )	165	+	7.0	CCTTTTTC	赤霉素应答元件
Skn-1_motif	水稻( <i>Oryza sativa</i> )	147	-	5.0	GTCAT	胚乳表达所需的顺式作用调控元件
Skn-1_motif	水稻( <i>Oryza sativa</i> )	741	-	5.0	GTCAT	胚乳表达所需的顺式作用调控元件
Skn-1_motif	水稻( <i>Oryza sativa</i> )	262	-	5.0	GTCAT	胚乳表达所需的顺式作用调控元件
Skn-1_motif	水稻( <i>Oryza sativa</i> )	259	+	5.0	GTCAT	胚乳表达所需的顺式作用调控元件
Skn-1_motif	水稻( <i>Oryza sativa</i> )	305	-	5.0	GTCAT	胚乳表达所需的顺式作用调控元件
TGACC-motif	大麦( <i>Hordeum vulgare</i> )	306	+	5.0	TGACC	参与 MeJA 响应的顺式调控元件
TGACC-motif	大麦( <i>Hordeum vulgare</i> )	630	+	5.0	TGACC	参与 MeJA 响应的顺式调控元件
TGACC-motif	大麦( <i>Hordeum vulgare</i> )	618	-	5.0	TGACC	参与 MeJA 响应的顺式调控元件
circadian	番茄( <i>Lycopersicon esculentum</i> )	270	+	6.0	CAANNNNATC	参与昼夜节律控制的顺式作用调控元件

注:“+”表示 DNA 正链;“-”表示 DNA 负链。



以根为参照, 设定其中的表达量为 1

图8 滇龙胆 *GrDL7H* 基因在不同组织中的相对表达

羟戊酸(MVA)和甲基赤藓糖-4-磷酸(MEP)途径合成 IPP (异戊烯焦磷酸)和 DMAPP(二甲基丙烯基二磷酸-三铵盐)的基本前体,第2阶段是 IPP 和 DMAPP 在异戊烯基转移酶和萜类合酶作用下形成萜类化合物和中间产物,第3阶段是萜类化合物的碳骨架结构在后修饰酶的催化下,分别进行羟基化、甲基化、糖基化等,形成结构稳定、具有活性的萜类化合物<sup>[4]</sup>。为阐明龙胆苦苷生物合成途径及其调控机制,本研究克隆了1个萜类合成第2阶段的 *GrDL7H* 基因,其表达情况可能影响龙胆苦苷的生物合成。由于许多萜类合成途径中的底物不能通过商业化途径获得,导致学界对 *DL7H* 基因的研究较少,目前研究比较清楚的仅有长春花中的 *CrDL7H* 基因。本研究从滇龙胆幼叶中克隆获得 *GrDL7H* 基因,通过多序列比对分析,表明滇龙胆 *GrDL7H* 与其他植物 *DL7H* 蛋白质具有较高的相似性,特别是与长春花 *CrDL7H* 蛋白质的序列相似性高达 77.73%,表明二者具有相似的结构与功能。在系统发育分析中,滇龙胆 *GrDL7H* 与黄金鸡纳树 *CeDL7H* 的亲缘关系较近,且 *GrDL7H* 蛋白质具有典型的 CYP450 蛋白质结构域,属于 E 类家族成员,这些结果表明,所克隆的基因为 *DL7H* 基因。

在长春花中,*CrDL7H* 基因在根、茎、叶、花芽和花中均有表达,但在第1对幼叶中的表达量远远高于其他组织<sup>[8]</sup>。组织表达特异性研究结果表明,滇龙胆 *GrDL7H* 基因在叶中的表达量最高,远远高于根、茎和花,因此推测滇龙胆叶片为龙胆苦苷生物合成的主要器官,这与朱宏涛等的研究结果<sup>[10]</sup>一致。目前研究表明,滇龙胆全株可以入药<sup>[11-12]</sup>,但其根中龙胆苦苷含量最高<sup>[13]</sup>。因此,根中的龙胆苦苷主要是由叶片和茎中合成的龙胆苦苷通过运输到达根中进行积累而成的。

在真核生物基因中,启动子由核心启动子和上游启动子2个部分组成,是位于转录起始位点及其5'端上游近端大约 100~200 bp 的1组具有独立功能的 DNA 序列,每个元件长度约为 7~20 bp,是决定 RNA 聚合酶转录起始和转录频率的关键元件<sup>[14]</sup>。本研究扩增到的滇龙胆 *GrDL7H* 基因启动子片段大小为 818 bp,不但具有核心启动子 TATA 盒上游启动子 CAAT 盒,而且具有 6 个光应答的元件、1 个参与赤霉素应答元件、6 个参与茉莉酸甲酯应答的顺式调控元件、1 个热胁迫应答元件、1 个参与生物钟控制的顺式调控元件,表明

*GrDL7H* 基因的表达受光、赤霉素、茉莉酸甲酯、热胁迫、生物钟等多种因素的调控,这同时也表明,可通过调控基因表达来提高龙胆苦苷的含量。

本研究为滇龙胆 *GrDL7H* 基因功能的解析奠定了基础。今后将对 *GrDL7H* 基因进行蛋白质表达、纯化和酶活分析、基因调控机制分析来研究其功能,从而为龙胆苦苷生物合成途径及其调控机制的阐明奠定基础。

#### 参考文献:

- [1] Suyama Y, Tanaka N, Kawazoe K A, et al. Rigenolides B and C, conjugates of norsecoiridoid and secoiridoid glucoside from *Gentiana rigescens* Franch [J]. *Tetrahedron Letters*, 2017, 58 (15): 1459 - 1461.
- [2] Suyama Y, Kurimoto S I, Kawazoe K, et al. Rigenolide A, a new secoiridoid glucoside with a cyclobutane skeleton, and three new acylated secoiridoid glucosides from *Gentiana rigescens* Franch [J]. *Fitoterapia*, 2013 (91): 166 - 172.
- [3] 金航, 张霁, 张金渝, 等. 滇龙胆 [M]. 昆明: 云南科技出版社, 2013.
- [4] 黄璐琦, 刘昌孝. 分子生药学 [M]. 3 版. 北京: 科学出版社, 2015.
- [5] Salim V, Yu F, Altarejos J, et al. Virus - induced gene silencing identifies *Catharanthus roseus* 7 - deoxyloganic acid - 7 - hydroxylase, a step in iridoid and monoterpene indole alkaloid biosynthesis [J]. *Plant Journal*, 2013, 76 (5): 754 - 765.
- [6] 王彩云, 张晓东, 沈涛, 等. 龙胆苦苷生物合成途径研究进展 [J]. *江苏农业科学*, 2014, 42 (3): 4 - 10.
- [7] Liu Y E, Wang Y, Guo F X, et al. Deep sequencing and transcriptome analyses to identify genes involved in secoiridoid biosynthesis in the Tibetan medicinal plant *Swertia mussotii* [J]. *Scientific Reports*, 2017 (7): 43108.
- [8] Salim V, Yu F, Altarejos J, et al. Virus - induced gene silencing identifies *Catharanthus roseus* 7 - deoxyloganic acid - 7 - hydroxylase, a step in iridoid and monoterpene indole alkaloid biosynthesis [J]. *Plant Journal*, 2013, 76 (5): 754 - 765.
- [9] Zhang X, Allan A C, Li C, et al. De Novo assembly and characterization of the transcriptome of the Chinese medicinal herb, *Gentiana rigescens* [J]. *International Journal of Molecular Sciences*, 2015, 16 (5): 11550 - 11573.
- [10] 朱宏涛, 郑传伟, 赵平, 等. 野生坚龙胆及其组培苗中龙胆苦苷的含量分析 [J]. *天然产物研究与开发*, 2011, 23 (3): 482 - 485.
- [11] 赵志莲, 张琳, 刘卫红, 等. 滇龙胆裂环烯醚萜类活性成分积累规律的研究 [J]. *时珍国医国药*, 2016, 27 (9): 2073 - 2075.
- [12] 游维丽. 龙胆地上、地下部分药理作用的比较 [J]. *内蒙古中医药*, 2017, 36 (8): 135 - 136.
- [13] 杨美权, 张金渝, 沈涛, 等. 不同栽培模式对滇龙胆中龙胆苦苷含量的影响 [J]. *江苏农业科学*, 2011 (1): 287 - 289.
- [14] 朱玉贤, 李毅, 郑晓峰, 等. 现代分子生物学 [M]. 4 版. 北京: 高等教育出版社, 2013.