

珍稀药用铁皮石斛 4 个 *CIPKs* 基因的鉴定与表达分析

李依民¹, 张 娜¹, 沈 霞¹, 李 欢¹, 黑小斌¹, 郭顺星^{2*}, 张 岗^{1,2*}

(1. 陕西中医药大学药学院/陕西省中药基础与新药研究重点实验室, 陕西 西安 712046;

2. 中国医学科学院、北京协和医学院药用植物研究所, 北京 100193)

摘要: 类钙调磷酸酶 B 蛋白 (calcineurin B-like protein, CBL) 互作蛋白激酶 (CBL-interacting protein kinase, CIPK) 通过感应钙信号, 在植物生长发育、逆境生理等生命活动中发挥重要作用。本研究利用 RACE 技术首次从珍稀药用铁皮石斛中克隆到 4 个 *CIPK* 基因 cDNA 全长, *DoCIPK1*、*DoCIPK2*、*DoCIPK3* 和 *DoCIPK4* (GenBank 注册号 KT957557、KT957558、KT957559 和 KT957560), 编码蛋白分别含 473、449、451、440 个氨基酸, 分子质量依次为 53.50、50.93、51.50、50.16 kDa, 等电点 7.99、9.25、8.81、9.11; 与多种植物 *CIPKs* 蛋白一致性在 70%~90%、69%~80%、78%~93%、66%~82% 之间; 各含 1 个蛋白激酶的保守结构域 (分别为第 21~275、14~268、16~271、12~266 位氨基酸)、1 个 *CIPK* 家族特有的 NAF/FISL 结构域 (335~391、313~370、310~369、305~362) 和多个功能基元; 4 个蛋白均无信号肽或跨膜域, 预测主要分布在质膜、内质网等亚细胞水平, 与拟南芥 *AtCIPK24* 三维结构类似; *DoCIPK1*、*DoCIPK3* 分别隶属于拟南芥和水稻 *CIPKs* 分子进化树的 E、A 类群, *DoCIPK2* 和 *DoCIPK4* 属于 C 类群; *DoCIPK1* 基因在石斛叶和茎中的相对表达量无显著差异, 根中表达量为叶中的 0.35 倍; *DoCIPK3* 转录本在茎和根中的丰度分别为叶中 3.36 倍和 3.47 倍; *DoCIPK2* 与 *DoCIPK4* 的表达模式相似, 二者在叶和茎中的相对表达量无显著差异, 根中表达量分别为叶中的 2.08 倍和 7.86 倍。*DoCIPK1*、*DoCIPK2*、*DoCIPK3* 和 *DoCIPK4* 基因克隆、生物信息特征及表达特性为下一步研究基因在铁皮石斛生理适应中的分子功能奠定基础。

关键词: 铁皮石斛; 激酶; 定量 PCR; 钙信号; 结构域

中图分类号: R931

文献标识码: A

文章编号: 0513-4870 (2018) 02-0304-09

Identification and expression analysis of four *CIPKs* genes from a rare and medicinal plant, *Dendrobium officinale*

LI Yi-min¹, ZHANG Na¹, SHEN Xia¹, LI Huan¹, HEI Xiao-bin¹, GUO Shun-xing^{2*}, ZHANG Gang^{1,2*}

(1. College of Pharmacy and Shaanxi Provincial Key Laboratory for Chinese Medicine Basis and New Drugs Research, Shaanxi University of Chinese Medicine, Xi'an 712046, China; 2. Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100193, China)

Abstract: The calcineurin B-like protein (CBL) – interacting protein kinase (CIPK) plays a vital role in the growth, development, and stresses adaptation in plants by interaction with the calcium signaling. In this study, four full length cDNAs of *CIPKs* genes, namely *DoCIPK1*, *DoCIPK2*, *DoCIPK3* and *DoCIPK4* (GenBank accession No. KT957557, KT957558, KT957559 and KT957560, respectively) were cloned from the rear and

收稿日期: 2017-10-25; 修回日期: 2017-11-15.

基金项目: 国家自然科学基金资助项目 (31101608, 31170314); 陕西省自然科学基金资助项目 (2017JM8030); 陕西省科技计划青年科技新星项目 (2012KJXX-44); 陕西省普通高校青年杰出人才计划项目; 咸阳市中青年科技领军人才项目.

*通讯作者 Tel / Fax: 86-29-38185165, E-mail: jay_gumling2003@aliyun.com;

Tel / Fax: 86-10-62829619, E-mail: sxguo1986@163.com

DOI: 10.16438/j.0513-4870.2017-1066

medicinal plant, *Dendrobium officinale*, by rapid amplification of cDNA ends (RACE) for the first time. The corresponding encoded proteins, consisting of 473, 449, 451 and 440 amino acids (aa), respectively, with a molecular weight of 53.50, 50.93, 51.50 and 50.16 kDa, and an isoelectric point (pI) of 7.99, 9.25, 8.81 and 9.11, respectively, shared 70%–90%, 69%–80%, 78%–93%, and 66%–82% identities CIPKs with various plants. Each deduced protein contained a conserved protein kinase domain (respectively at 21–275, 14–268, 16–271 and 12–266 aa position), a CIPK family characteristic NAF/FISL domain (respectively at 335–391, 313–370, 310–369 and 305–362 aa position) and some functional motifs. The four DoCIPK proteins, without signal peptide or transmembrane region, were located in the plasma membrane and endoplasmic reticulum at the subcellular level. The three dimensional structure of the proteins were similar to that of *Arabidopsis* AtCIPK24. DoCIPK1 and DoCIPK3 were respectively clustered in the group E and A of the *Arabidopsis* and rice CIPK evolutionary tree, while DoCIPK2 and DoCIPK4 belonged to group C. The relative expression of *DoCIPK1* showed no significant difference in the leaves and stems, and its transcripts in the roots was 0.35 fold over that in the leaves. The abundance of *DoCIPK3* transcripts in the stems and the roots were 3.36 fold and 3.47 fold higher, respectively, than those in the leaves. *DoCIPK2* exhibited similar expression pattern to *DoCIPK4*. Their relative expression in the leaves and the stems had no apparent difference, and the transcript levels were higher in the roots than that in the leaves, with 2.08 fold and 7.86 fold, respectively. Cloning, bioinformatics analyses, and expression patterns of the four *DoCIPK* genes provide a basis for functional elucidation of these genes further during the physiological responses in *D. officinale*.

Key words: *Dendrobium officinale*; kinase; quantitative PCR; calcium signal; domain

钙既是植物体生长发育过程中的一种必需矿质元素, 又是一种非常重要的第二信使, 主导胞内信号转导, 参与调节植物对光、病原菌、非生物胁迫等各种环境的生理适应^[1, 2]。钙信号一般由钙调素、类钙调磷酸酶 B 蛋白 (calcineurin B-like protein, CBL)、钙网蛋白或钙依赖蛋白激酶等钙结合蛋白感应, 传递信号并激活下游基因表达, 进而调节细胞代谢^[2]。CBL 互作蛋白激酶 (CBL-interacting protein kinases, CIPKs) 特异的结合 CBL 蛋白, 感应钙信号, 共同构成 Ca^{2+} 介导的 CBL-CIPK 调控网络, 介导植物应对各种环境的生理反应^[3, 4]。CIPKs 蛋白由一个 N-末端激酶结构域和一个 C-末端调控域构成, 后者包含一个保守的 NAF/FISL 基序, 负责与绑定 Ca^{2+} 的 CBL 互作, 靶向激活 CIPKs^[5]。此外, C-末端还有一个蛋白磷酸酶 2C 结合结构域^[6]。

植物 CIPKs 系较大的基因家族。截止目前, CIPKs 在拟南芥基因组中已鉴定 26 个^[3, 7]、水稻 34 个^[4]、玉米 43 个^[8]、杨树 27 个^[7]、油菜 23 个^[9]、木薯 27 个^[10]。大量研究证实, CIPKs 在介导植物矿质元素转运与吸收、激素信号转导、生物胁迫或非生物胁迫等生命活动中起重要调控作用。AtCIPK23 与 AtCBL1/CBL9 形成的复合物, 通过调节 K^+ 通道 AtAKT1 维系细胞 K^+ 稳态^[11], 番茄 SICIPK24 和野大麦 HbCIPK2 也有类似保持 K^+ 平衡的作用^[12, 13]。AtCIPK24 (AtSOS2) 与 AtCBL4 (AtSOS3) 的互作激活质膜 Na^+/H^+ 反向

转运体 AtNHX7 (SOS1) 和液泡 H^+ -ATPase, 进而增强植物耐盐性^[14, 15]。GhCIPK6 提高棉花对脱落酸 (abscisic acid, ABA)、盐和干旱等非生物胁迫^[16]。AtCIPK26 在种子萌发的 ABA 信号通路中与 2C 类蛋白磷酸酶 ABI1、ABI2、ABI5 相互作用中起重要作用^[17]。AtCIPK6 能运输生长素进而影响根系发育和响应盐胁迫^[18]。最新研究揭示, 番茄 CBL10-CIPK6 能够通过钙信号、活性氧信号转导机制, 介导植物对病菌的免疫反应^[19]。可见, CBL-CIPKs 网络在植物生长发育、抗逆胁迫中起着至关重要的作用。

铁皮石斛 *Dendrobium officinale* Kimura et Migo 为兰科石斛属最为珍稀名贵的种, 药用部位为新鲜或干燥茎, 具有益胃生津、滋阴清热、润肺止咳、明目强身等功效^[20]。石斛属植物普遍含有倍半萜类、菲酚类、双萜酚类、多糖以及生物碱等主要活性成分。石斛碱是药用石斛最重要的有效成分之一, 具有止痛解热、抗肿瘤的功效, 对心血管、胃肠道等疾病有抑制作用^[21]。铁皮石斛已在生物学、化学、药理学等方面取得一定的研究进展^[20–23]。但是基因功能研究相对匮乏, 限制了对石斛资源的可持续开发与利用。前期, 本课题组利用 SSH 富集菌根真菌侵染铁皮石斛的差异表达基因^[24], 分离得到长度分别为 557、530、450、555 bp 的 4 条差异基因, BLASTx 分析显示其与多种植物 CIPKs 蛋白高度一致, 推测参与调控铁皮石斛生长发育和菌根生理适应。本研究

以珍稀濒危药用植物铁皮石斛为对象, 采用 RACE 技术克隆获得 4 个 *CIPKs* 基因全长 cDNA (*DoCIPK1*、*DoCIPK2*、*DoCIPK3* 和 *DoCIPK4* 基因) 并进行生物信息学、系统进化及组织表达模式等特征分析, 旨在为深入研究基因分子作用提供基础。

材料与方法

材料 野生铁皮石斛植物材料于 2015 年 7 月采自云南西双版纳, 取其根、茎、叶组织样品, 液氮速冻后置 -80°C 保存备用。

核糖核酸 (RNA) 提取和互补脱氧核糖核酸 (cDNA) 合成 按照 EASYspin Plus 植物 RNA 快速提取试剂盒 (北京爱德莱) 操作说明提取纯化各样品总 RNA; 采用 NanoDropTM 2000 分光光度计 (Thermo Fisher, USA) 测定总 RNA 质量和纯度, 1.2% 琼脂糖凝胶电泳鉴定总 RNA 的完整性。使用 M-MLV Reverse Transcriptase kit (Promega, USA) 反转录合成 cDNA 第一链, -20°C 保存备用。

cDNA 末端快速扩增 (rapid amplification of cDNA ends, RACE) 克隆 以 4 条 EST 为模板, 设计 4 对特异引物 (表 1), 按照 SMARTTM RACE cDNA Amplification Kit (Clontech, Japan) 说明书分别进行 5'/3'-RACE。

分别以 4 个基因 5'/3'-RACE-R/F 引物与试剂盒提供 UPM 引物组合, 进行各基因 5'/3'-RACE 扩增。反

应体系为: 2.5 μL 10 \times Advantage[®] 2 PCR buffer, 0.5 μL 10 $\text{mmol}\cdot\text{L}^{-1}$ dNTPs, 0.5 μL 10 $\mu\text{mol}\cdot\text{L}^{-1}$ 5'/3'-RACE-R/F, 0.5 μL 10 $\mu\text{mol}\cdot\text{L}^{-1}$ UPM, 1.0 μL 5'/3'-RACE ready cDNA, 0.5 μL 5 U $\cdot\mu\text{L}^{-1}$ 50 \times Advatange[®] 2 Polymerase Mix, 补 ddH₂O 至 25 μL 。PCR 程序为: 94°C 3 min; 94°C 30 s, 70°C 30 s, 72°C 2 min, 5 个循环; 94°C 30 s, 68°C 30 s, 72°C 2 min 24 个循环; 72°C 10 min, 4°C 保温。PCR 产物经 1.8% 琼脂糖凝胶电泳分析, 回收目标条带、克隆, 随机选取 3 个重组质粒, 利用 ABI 3130xl 遗传分析仪 (Applied Biosystems, USA) 双向测序分析。

基因全长获取和验证 利用 CAP3 将 5'、3' 端 cDNA 序列与核心序列拼接后, 进行 BLASTX 和 ORF Finder 分析, 设计 4 对跨开放阅读框 (open reading frame, ORF) 引物 (表 1), 分别进行 4 个全长基因的 RT-PCR 验证。反应体系为: 2.0 μL 10 \times Ex TaqTM buffer, 0.4 μL 10 $\text{mmol}\cdot\text{L}^{-1}$ dNTPs, 10 $\mu\text{mol}\cdot\text{L}^{-1}$ ORF-F/R 各 0.4 μL , 1.0 μL cDNA, 0.2 μL 5 U $\cdot\mu\text{L}^{-1}$ Ex TaqTM Polymerase, 补 ddH₂O 至 20 μL 。PCR 程序: 94°C 3 min; 94°C 30 s, 58°C 30 s, 72°C 2 min, 32 个循环; 72°C 延伸 10 min; 4°C 保温。PCR 反应产物经 1.2% 琼脂糖凝胶电泳检测, 回收目的条带、克隆并进行测序分析。

序列分析 使用 BLAST (<http://www.ncbi.nlm.nih.gov/blast.cgi>)、CAP3 (<http://pbil.univ-lyon1.fr/cap3.php>)、

Table 1 RACE, RT-PCR and qPCR primer sequences of the four calcineurin B-like protein-interacting protein kinase (*CIPK*) genes in *D. officinale*

Gene symbol	Primer sequences (5'-3')	Amplicon size/bp
<i>DoCIPK1</i>	1-5'-RACE-R: GTCGTCCTGCCGATCTGTTCG	541
	1-3'-RACE-F: CGGTTTCCGAACAGATGCGGCAGG	1 086
	1-ORF-F: ATGGCGTCGGAGCCG	1 422
	1-ORF-R: TTA CTCTCTTT CAGAGTCATATTCAGAT	
	1-qPCR-F: GGTGAGTCCGTCGCCATTAAGG	396
	1-qPCR-R: TCCGTCCTGCCGATCTGT	
<i>DoCIPK2</i>	2-5'-RACE-R: AGCTGGAGTCCCACAAGTAGTATGGAGTAAG	840
	2-3'-RACE-F: CTA CTGTGGACTCCAGCTTATGTCCG	1 061
	2-ORF-F: TCTGTCTAAGTCCAATAGGGCT	1 636
	2-ORF-R: CCAAAGTCACAATTTTAGCTC	
	2-qPCR-F: TCCTTAAGCCAGCTTCCACCA	280
	2-qPCR-R: CACCTTGCCATGCCACACAA	
<i>DoCIPK3</i>	3-5'-RACE-R: GCGTGCTCATCTTCTCATCTTCCAT	681
	3-3'-RACE-F: GCATCCAAATGTTGTTCCGCTCCACG	1 461
	3-ORF-F: ATGCACAAGATGAGCACACCA	1 356
	3-ORF-R: CTAGTACTGCATCTCTGATCTGAAA	
	3-qPCR-F: GTGGGAGGAGGAGGAGAAAGCA	390
	3-qPCR-R: CGTGGAGGCGAACAACATTGG	
<i>DoCIPK4</i>	4-5'-RACE-R: ATCCTTCTCAAGCCTGCCCTTAGCAATC	824
	4-3'-RACE-F: TGTGGA ACTCCAGCTTATGTTGCTCCTGAG	994
	4-ORF-F: ATGGTGAATAAAGGGTGTATTG	1 323
	4-ORF-R: TCAAATCTGCAGCTGCTGTAAATA	
	4-qPCR-F: GCTAAGGGCAGGCTTGAGAAGG	134
	4-qPCR-R: CCAGCTTCATCCAGGAGGAGGT	

ORF Finder (<http://www.ncbi.nlm.nih.gov/gorf/gorf.html>) 等在线工具分析 *CIPKs* cDNA 序列。

采用 ExPASy Proteomics Server 在线软件 ProtParam (<http://www.expasy.ch/tools/protparam.html>) 进行 4 个 *CIPKs* 基因编码蛋白的理化性质分析; 使用 SOPMA (http://npsa-pbil.ibcp.fr/cgi-bin/secpred_sopma.pl) 预测蛋白质二级结构; 利用 InterProScan (<http://www.ebi.ac.uk/cgi-bin/iprscan/>) 和 PROSITE SCAN (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_proscan.html) 预测蛋白保守结构域和基元; 采用 SWISS-MODEL (<http://swissmodel.expasy.org/>) 进行 DoCIPKs 蛋白质三维建模分析; 用 SignalP 4.1 (<http://www.cbs.dtu.dk/services/SignalP/>) 进行分泌蛋白预测; 借助 PSORT (<http://psort.ims.u-tokyo.ac.jp/form.html>) 进行蛋白亚细胞定位预测; 利用 TMHMM 2.0 (<http://www.cbs.dtu.dk/services/TMHMM/>) 预测蛋白跨膜区; 用 DNASTAR 7.0 和 MEGA 6.0 软件分别进行氨基酸序列比对和进化树分析。

实时定量 PCR (real time quantitative PCR, qPCR) 分析 分别以 2 μg 根、茎、叶样品总 RNA 反转录合成 cDNA, 以 *EF1 α* 为内参基因^[20], 采用 qPCR 检测 4 个 *CIPKs* 基因的组织表达模式, 各基因引物见表 1。应用 ABI PRISM 7500 实时荧光定量 PCR 仪 (Applied Biosystems, USA) 进行 qPCR。反应体系为: 12.5 μL 2 \times SYBR[®] Premix *Ex Taq*[™] Master Mix, qPCR-F/R 各 0.5 μL (10 $\mu\text{mol}\cdot\text{L}^{-1}$), 0.5 μL Dye, 2 μL cDNA 模板, 9 μL ddH₂O。PCR 程序为: 95 $^{\circ}\text{C}$ 30 s, 95 $^{\circ}\text{C}$ 5 s, 60 $^{\circ}\text{C}$ 1 min, 40 个循环, 同时绘制熔解曲线。每个反应重复 3 次, 包括不加模板的对照 (NTC), 本实验重复 3 次。根据 ABI PRISM 7500 SDS 软件 (Applied Biosystems, USA) 生成的 CT (Cycle threshold) 值, $2^{-\Delta\Delta\text{CT}}$ 计算基因相对表达量^[25]。

结果

1 *DoCIPKs* 基因克隆和全长验证

利用 1-5'-RACE-R、2-5'-RACE-R、3-5'-RACE-R、4-5'-RACE-R 与 UPM 组成 4 对引物组合, 5'-RACE ready cDNA 为模板, 5'-RACE 结合克隆与测序分别获得 541、840、680 和 824 bp 的 cDNA 序列。同时, 利用 1-3'-RACE-R、2-3'-RACE-R、3-3'-RACE-R、4-3'-RACE-R 与 UPM 组成 4 对引物组合, 3'-RACE ready cDNA 为模板, 3'-RACE 分别获得 1086、1061、1461 和 994 bp 的 cDNA 序列。将 5'、3' 末端 cDNAs 序列,

分别与原 EST 拼接获得 4 个 cDNAs, 长度分别为 1 597、1 881、2 006、1 954 bp, 命名为 *DoCIPK1*、*DoCIPK2*、*DoCIPK3* 和 *DoCIPK4*, 提交 GenBank 获得注册号为 KT957557、KT957558、KT957559 和 KT957560。ORF Finder 分析显示 *DoCIPK1*、*DoCIPK2*、*DoCIPK3* 和 *DoCIPK4* 基因 ORF 分别为 1 422、1 350、1 356、1 323 bp, 5'-UTR 分别长 15、306、330、497 bp, 3'-UTR 分别长 160、225、320、134 bp, 含有真核生物 mRNA 末端 polyA 尾巴; 除 *DoCIPK3* 外, 其他 3 个基因起始密码子附近碱基序列符合 KOZAK 规则^[26]。

使用 1-ORF-F/R、2-ORF-F/R、3-ORF-F/R、4-ORF-F/R 进行 RT-PCR 扩增验证基因全长, 获得长度分别为 1 422、1 636、1 356、1 323 bp 的单一条带 (图 1), 克隆、测序分析显示其含有 4 个基因的完整 ORF, 说明已成功获得 4 个 *CIPKs* 基因全长 cDNAs。

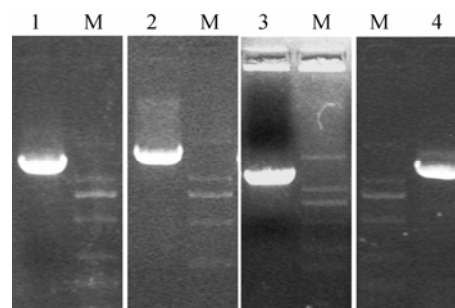


Figure 1 The agarose gel electrophoresis of RT-PCR verification of the four full length *CIPK* genes. No. 1 to 4: The four *DoCIPK1-4* respectively; M: DL 2000 standard molecular marker

2 *DoCIPKs* 基因编码蛋白的序列分析

利用生物信息学软件分别对 *DoCIPK1*、*DoCIPK2*、*DoCIPK3* 和 *DoCIPK4* 基因进行同源性、理化特性、二级结构、结构域、基元、亚细胞定位等预测分析, 结果见表 2。

2.1 *DoCIPKs* 基因编码蛋白同源性、理化特性分析

BLASTP 分析显示 *DoCIPK1*、*DoCIPK2*、*DoCIPK3* 和 *DoCIPK4* 基因编码蛋白与 GenBank 已注册植物 CIPKs 蛋白高度同源, 一致性分别在 70%~90%、69%~80%、78%~93%、66%~82% 之间, 与小兰屿蝴蝶兰 *Phalaenopsis equestris* XP_020574335、XP_020585983、XP_020585435、XP_020588196 的一致性分别为 90%、80%、93%、82%。ProtParam 分析获得 *DoCIPK1*、*DoCIPK2*、*DoCIPK3* 和 *DoCIPK4* 基因推定蛋白的分子式, 氨基酸数目分别为 473、449、451、440 个, 分子质量依次为 53.50、50.93、51.50、50.16 kDa, 等电点分别为 7.99、9.25、8.81、9.11。

Table 2 Bioinformatics analyses of the four deduced DoCIPKs proteins in *D. officinale*. CpT: Chloroplast thylakoid; ER: Endoplasmic reticulum; Glycos: Glycosylation; MB: Microbody; MW: Molecular weight; Myr: Myristoylation; PK: Protein kinase; PM: Plasma membrane; PS: Phosphorylation site

		DoCIPK1	DoCIPK2	DoCIPK3	DoCIPK4
Physiochemical	Formula	C ₂₃₉₆ H ₃₇₉₀ N ₆₅₆ O ₇₀₀ S ₁₆	C ₂₃₀₀ H ₃₆₅₀ N ₆₁₄ O ₆₅₇ S ₁₆	C ₂₃₁₆ H ₃₆₆₈ N ₆₃₂ O ₆₇₀ S ₁₃	C ₂₂₆₇ H ₃₆₀₁ N ₅₉₉ O ₆₃₈ S ₂₂
property	Amino acid	473	449	451	440
	MW (KD)	53.50	50.93	51.50	50.16
	pI	7.99	9.25	8.81	9.11
Secondary	α helix	34.04%	37.19%	41.69%	39.32%
structure	Random coil	34.25%	27.62%	27.05%	24.77%
	Extended strand	18.60%	23.39%	17.07%	22.73%
	β turn	13.11%	11.80%	14.19%	13.18%
Domain	PK domain	21–275	14–268	16–271	12–266
	NAF/FISL domain	333–357	310–334	308–332	302–326
Motif	N-glycos site	461–464	296–299	296–299	
	PK C PS	119–121, 193–195, 325–327, 388–390, 470–472	163–165, 346–348, 365–367	6–8, 10–12, 67–69, 344–346, 441–443	161–163, 252–254, 301–303, 355–357, 373–375, 427–429
	Casein kinase II PS	265–268, 293–296, 308–311, 318–321, 323–326, 376–379, 464–467, 466–469, 470–473	122–125, 301–304, 325–328, 410–413	21–24, 97–100, 286–289, 354–357, 425–428, 433–436	120–123, 293–296, 314–317, 347–450, 402–485
	Amidation site		185–188	381–384	183–186, 375–378
	Tyrosine kinase PS	430–437, 429–437		112–119	
	N-Myr site	175–180, 400–405	168–173, 331–36	171–176	166–171
	PKs ATP-binding signature	27–50, 27–59	20–43	22–45	18–41
	Ser/Thr PKs active-site signature	139–151	132–144	135–147	130–142
	cAMP- and cGMP-dependent PK PS			338–341, 420–423	331–334
Localization	Nucleus				70.0%
	PM	60.0%	70.0%	70.0%	44.0%
	ER (membrane)	30.0%	20.2%	20.0%	85.0%
	MB (peroxisome)		20.0%	11.5%	47.4%
	CpT membrane	30.4%	13.7%	16.2%	
	Golgi body	40.0%			

2.2 DoCIPKs 基因编码蛋白的二级结构、结构域和基元预测分析 SOPMA 分析显示 *DoCIPK1*、*DoCIPK2*、*DoCIPK3* 和 *DoCIPK4* 基因编码蛋白质的二级结构中 α -螺旋 (alpha helix) 和延伸链 (extended strand) 占主要比例, 其次为 β -折叠 (beta turn), 无规卷曲 (random coil) 数量较少。

InterProScan 分析显示, *DoCIPK1*、*DoCIPK2*、*DoCIPK3* 和 *DoCIPK4* 基因编码蛋白各含 1 个蛋白激酶的保守结构域 (IPR000719; 分别为第 21~275、14~268、16~271、12~266 位氨基酸) 和 1 个 CIPK 家族特有的 NAF/FISL 结构域 (IPR018451; 335~391、313~370、310~369、305~362), 符合已报道 CIPKs 蛋白的典型结构特征^[2]。PROSITE Scan 分析表明, 4 个 CIPKs 蛋白含有数目不等的保守基元 (motif), 共有蛋白激酶 C 磷酸化位点、酪蛋白激酶 II 磷酸化位点、丝氨酸/苏氨酸蛋白激酶活性基序、蛋白激酶 ATP

结合基序和 *N*-豆蔻酰化位点 5 种基元; 另有 4 种基元不一致, *DoCIPK4* 不含 *N*-糖基化位点, *DoCIPK1* 无酰胺化位点, *DoCIPK2*、*DoCIPK4* 不含酪氨酸激酶磷酸化位点, *DoCIPK1*、*DoCIPK2* 无 cAMP/cGMP 依赖的蛋白激酶磷酸化位点。

2.3 DoCIPKs 基因编码蛋白信号肽、跨膜和定位预测分析 SignalP 4.1 和 TMHMM 分析 4 个蛋白均无信号肽或跨膜结构。PSORT 预测显示 *DoCIPK1* 定位于质膜、内质网 (膜)、叶绿体内囊体膜、高尔基体; *DoCIPK2*、3 在质膜、内质网 (膜)、微体、叶绿体内囊体膜中均有一定分布; 然而, *DoCIPK4* 除了分布在质膜内质网 (膜) 和微体, 还以较高比率定位于细胞核内。

2.4 DoCIPKs 基因编码蛋白的三维建模分析 以拟南芥 AtCIPK24 (PDB No.: 4czt.1) B 链为模板^[27], 利用 SWISS-MODEL 进行 4 个 CIPKs 蛋白三维结构建

模。图 2 结果显示, DoCIPK1、DoCIPK2、DoCIPK3 和 DoCIPK4 与拟南芥 *AtCIPK24* 蛋白一致性分别为 52.78%、57.48%、66.59%、53.49%, 空间结构类似。

2.5 DoCIPKs 蛋白多序列和分子进化分析 运用 DNASTar 7.0 中的 MegAlign 程序对 DoCIPK1~4 蛋白进行多序列比对分析。图 3 结果表明, 4 个 DoCIPKs 蛋白具有一定的保守性, N 端氨基酸序列保守性很高,

C 端氨基酸组成存在较大差异; N 端蛋白激酶结构域内包含自 DFG 氨基酸残基起始到 APE 氨基酸残基结束、30 个氨基酸构成的激活环 (activation loop), 其中有磷酸化丝氨酸、苏氨酸、酪氨酸残基位点; C 端包含 CIPKs 蛋白特有的 NAF/FISL 结构域。

进一步用 MEGA 6.0 软件邻接法 (Neighbour-joining) 构建 DoCIPK1、DoCIPK2、DoCIPK3 和 DoCIPK4 与拟南芥和水稻 CIPK 蛋白家族成员的系统进化树。结果 (图 4) 表明, 所选 CIPKs 蛋白聚成 A、B、C、D、E 五大类群, DoCIPK1、DoCIPK3 分别属于 E 和 A 类群, DoCIPK2 和 DoCIPK4 同属于 C 类群。

3 DoCIPKs 基因表达特异性分析

分别提取铁皮石斛根、茎、叶的 RNA, 用 qPCR 技术检测 4 个基因的组织表达模式。图 5 结果表明, 4 个 *DoCIPKs* 基因在铁皮石斛 3 种器官中均为组成型表达, 但表达模式各有特点。以叶为校正样本, *DoCIPK1* 基因在茎中相对表达量无较大变化, 根中的相对表达量较低, 仅为叶中的 0.35 倍; *DoCIPK2* 在茎和叶中的相对表达量无显著差异, 根中表达量为叶中的 2.08 倍; *DoCIPK3* 的转录本丰度在茎和根中处于同一水平, 均显著高于叶, 分别为叶中 3.36 和 3.47 倍; *DoCIPK4* 在茎与叶中的相对表达量无显著差异, 根中表达量为叶中的 7.86 倍。此外, 通过各基因 CT 值比较分析发现, 4 个基因丰度依次为 *DoCIPK2* > *DoCIPK4* > *DoCIPK1* > *DoCIPK3*。

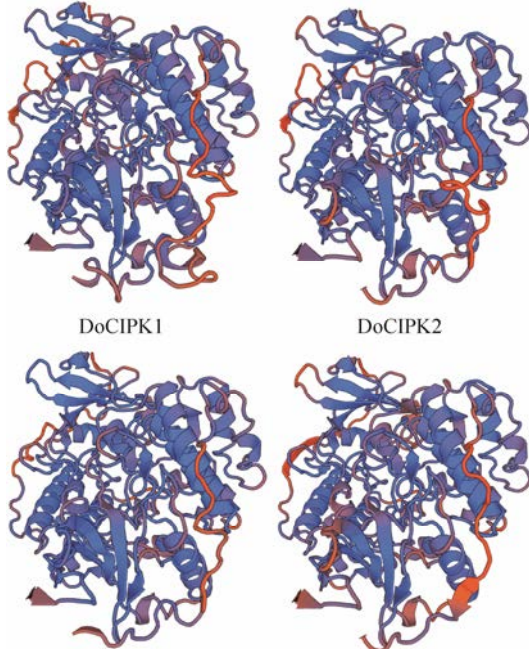


Figure 2 Three-dimensional structure of the four deduced DoCIPKs proteins using SWISS-MODEL

1	MASEPLAKTKDQASIIILGRYELGKILGHGTFFAKVYHARNVNTGESVAIKVLDKERILKSLGLVAHIKREIATLFRVRHPY	DoCIPK1
1	-----MGSEBNGKSVLMCKYELGRLLGQGTFAKVYHARNIRSSHVAIKMIDKERVFKVGLRDQIKREISVMRLVKHPN	DoCIPK2
1	-----MHRMSTFKVSRRVKCYELGRITIGGTFFAKVRHARNSETGDFVAIKLLDKEKLLKHKLVKIQKREIATMLIKHPN	DoCIPK3
1	-----MVNKGCLLHKYELGRLLGQGTFAKVYAKNVIITNQCVAIKMDKRNFRAGLMDQIKREISIMRLVRHPN	DoCIPK4
81	IVQLFEVMATKTKIYFVMEYVRGGELFGKVA--KGRLEKEDTARRYFQQLISAVGFCFARGVVFHRDLKPENLLDDKGLKLV	DoCIPK1
74	IVQLYEVMASKSKIYFVLEYYVKGGELEFNKVL--KGRITKEDVARKYFQQLISAVDFCHSRGVYHRDLKPENLLDDNGNLKV	DoCIPK2
76	VVRLHEVMGSKSKITFVLEYYVKGGELEFNKVA--KGRLEKEDTARRYFQQLINAVDFCHSRGVYHRDLKPENLLDVAGNLKV	DoCIPK3
72	IVHRYEVMATKTKIYFVMEYAKGGELFNKHA--KGRLEKEDVAGKYFQQLISAVGFCFNRGVYHRDLKPENLLDDAGNLKL	DoCIPK4
	Activation loop	
160	SDFGLSALVSEOMRQDGLLHTEFCGTPPAYVAPEVLSRKGYDGAQVILVWVSCGVILFVLMAGYLPFHDLNMMAMRYKIKGEFR	DoCIPK1
153	SDFGLSALAESKQDGLLHTTCGTPPAYVAPEVIGRKGYDGAQADIWSCGVILFVLMAGYLPFHDLNLMEMYYKIKGAEFK	DoCIPK2
156	SDFGLSALSQQVRDGLLHTTCGTPPAYVAPEVINDRGYDGSADLVWVSCGVILFVLLAGYLPFHDDNMMNYKIKIGTEGFT	DoCIPK3
151	SDFGLSALVESKRQDGLLHTTCGTPPAYVAPEVIGRKGYDGAQADIWSCGVILVLLAAYLPFHDTNLMEMRYKIKGAEFR	DoCIPK4
	* * *	
240	CPNWFSSDLTKLLSRLLDNPOTRITITPEIMENRWFKGRHRVRFYVENDEVHSLDEPQVPHDDASDSCGESDSSVASC	DoCIPK1
233	YPNWFSSDVRRLITRLLDNPATRIISIKIKENSWFRKGYNAN---QIKNOAEAKELVPE-----LDVNAAFNSS	DoCIPK2
236	FPSWLSYGAKRLLSRLLDNPMTFRITITPEILKDEWFRKGYKPP-----IFYEKMETN-----LDDIDAAFNS	DoCIPK3
231	CPNWFPTKARKLLSRLLDNPSTRIPIVKTMSHSWFRKGF-----EAESKEKIE-----TALISTESS	DoCIPK4
	NAF/FISL domain	
320	FEITFSEKRLGLPRPASLNAFDIIISFSQGFDSLGLFEE---RGEETRLSSEFVSNIIISKLEELAKRVVKTVRKRD--CR	DoCIPK1
299	-ETSFAEE--KQEIFSKPILNNAFDIISLSSGFDSLGLFEAS--FCKRMVRFSTSSKPASTIIVLKEEVAKQLSLKIRKKDGGV	DoCIPK2
299	-QVSHVVR--KE---EQPAALNAFELIHALSKGLNPNENLEIEQELKRETRFTSKCPAKELISKLEEAAPKLGFDIHKNN--YK	DoCIPK3
290	GSNSQVEQ--KEDSMKATSNNAFDIISLSDGFDSLGLFVEL--DRRRESRFTSNLPAEAIISKLEEVAKSLKFKVQKDDGI	DoCIPK4
396	VSLEGTREGEKGLPIIAVEIFELTPPIIVLVDIKKAGDRGEYEEFCNKELKPEGLQQLKFGTAALANITSESEY--SERE	DoCIPK1
376	LKMEGFKEGGKIVLAFDVEIFELTPSHMVEITQTNGDTEYQKILKQEMKPAKLDIIVWAWQGDRSLPQSQQL	DoCIPK2
374	IRLENIKAGRKNLNVATEIFQVAPSLHVMVEVRRKAGDITLFFQKFKK--LSTSLKDVLANSEEDAQVHTSKFS--QRMQY	DoCIPK3
368	MKMEVTKKGRKIVLADAIIFEITPSPHMVELKMTDGDITLEYEKLECKEDIRPALKDIIIVWRGEQGYLQQLQI	DoCIPK4

Figure 3 Multiple sequence alignments of four DoCIPKs proteins. The shaded indicate identical residues; the boxed are the activation loop and FISL motif, respectively; three highly conserved residues (Ser, Tyr and Thr) are marked with an asterisk in the activation loop

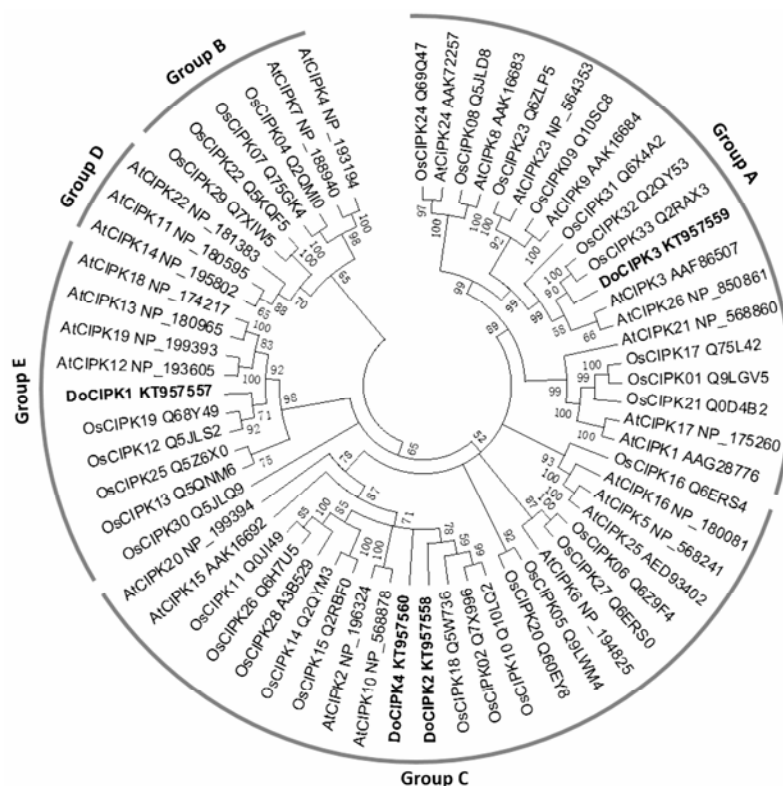


Figure 4 Phylogenetic tree of four DoCIPKs with CIPK family proteins from *Arabidopsis* and rice

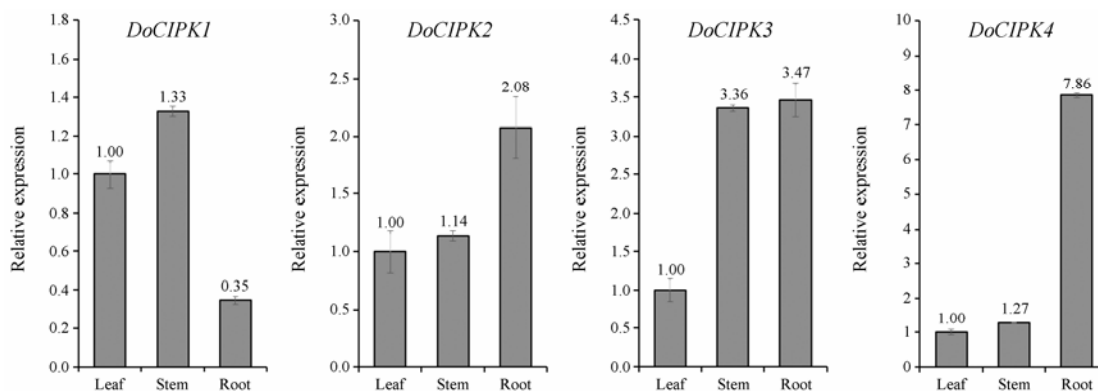


Figure 5 Organ-specific expression patterns of the four *DoCIPKs* genes using qPCR analyses

讨论

CIPKs 作为植物中特有的一类丝氨酸/苏氨酸蛋白激酶，通过与 CBL 蛋白互作来感应钙信号，参与信号网络传递，在植物生理代谢中起重要调控作用。对拟南芥、水稻、玉米、杨树、油菜、木薯等模式植物或重大作物的系统性分析初步揭示了植物 CIPKs 家族成员结构特点、蛋白分子进化、基因时空表达特性和部分基因的生物功能^[3,4,7-10]。CIPKs 家族基因结构与分子功能各有差异，协同表达参与调控植物体生命活动。本研究利用 RACE 技术首次从珍稀濒危兰科药用铁皮石斛中分离到 4 个 CIPKs 基因 cDNA 全长，DoCIPK1、DoCIPK2、DoCIPK3 和 DoCIPK4 与

多种植物 CIPKs 基因高度同源。4 个基因编码蛋白具有蛋白激酶保守结构域及激活环、NAF/FISL 结构域，为 CIPK 与 CBL 相互作用的位点^[6]，符合 CIPKs 蛋白的典型结构特征。进化分析将 4 个 DoCIPKs 蛋白分别聚为 E、C、A3 大类群，与已报道的拟南芥、水稻和葡萄等多种植物 CIPKs 蛋白分子进化一致^[3,8-11]。这些结果说明 4 个 CIPKs 基因是铁皮石斛中新鉴定的类钙调磷酸酶 B 蛋白互作蛋白激酶编码基因。

基因表达模式研究可为基因功能提供初步线索。CIPKs 基因表达分子特性与调控生理代谢过程密切相关。大量研究表明，植物 CIPKs 家族成员基因不仅具有组织表达特性，参与调控细胞的基本生理代谢，

而且能够响应外界环境胁迫表达, 参与调控植物细胞钙信号、激素信号转导。水稻 *CIPKs* 家族基因在水稻发育 3 个营养阶段和 11 个生殖阶段的组织中差异表达, 且响应盐、低温和干旱胁迫^[4]。油菜 12 个 *CIPKs* 基因受盐、干旱、冷、热、脱落酸、甲基紫精、低钾等外源环境诱导或抑制^[9]。木薯 21 个 *CIPKs* 基因在 3 个根发育阶段、茎和叶中都有表达, 有 4 个基因转录本未能检测, 9 个能响应盐、冷、脱落酸、 H_2O_2 等外源刺激^[10]。本研究 qPCR 分析揭示了 4 个 *CIPKs* 基因组织表达模式, *DoCIPK1* 基因在根中相对表达量很低, *DoCIPK2* 和 *DoCIPK4* 表达模式相似, 茎与叶中表达无差异, 后者根中表达量很高, *DoCIPK3* 在茎和根中表达水平一致且高于叶, 结合 4 个基因在器官水平上的整体丰度差异, 推测这 4 个基因通过不同的表达和作用方式参与石斛营养器官的生长发育过程。

铁皮石斛是珍稀濒危兰科药用植物的重要代表, 不仅能提取临床疗效较好的名贵药材, 而且是兰科菌根的理想研究材料。但是, 石斛属植物生长条件特殊, 种子野生条件萌发率低, 限制了其繁殖及市场供应。当前, 铁皮石斛基因组草图序列已经公布, 有助于下一步解析该植物及其近似种的系统繁育及进化的分子机制^[28]。铁皮石斛基因组中至少发现有 26 个 *CIPK* 编码基因, 本研究获得的 4 个 *CIPK* 基因全长, 与基因组预测 *CIPK* 基因编码蛋白 (XP_020702232、XP_020705240、XP_020701813、XP_020683888) 相似为 99%~100% (未发表数据), 证明了 *DoCIPK1*~*4* 的可靠性。然而, *CIPK* 家族成员众多, 分子功能、信号网络如何相互作用, 如何参与调控石斛菌根发育或种子萌发, 尚需更加系统的基因功能研究, 从而有助于揭示 *CIPK* 网络调控石斛生理适应的分子机制, 将为全面解析石斛类地道药材形成分子机制及良种培育提供科学依据。

References

- [1] Clapham DE. Calcium signaling [J]. *Cell*, 2007, 131: 1047–1058.
- [2] Kudla J, Batistic O, Hashimoto K. Calcium signals: the lead currency of plant information processing [J]. *Plant Cell*, 2010, 22: 541–563.
- [3] Kolukisaoglu U, Weini S, Blazevic D, et al. Calcium sensors and their interacting protein kinases: genomics of the *Arabidopsis* and rice CBL-CIPK signaling networks [J]. *Plant Physiol*, 2004, 134: 43–58.
- [4] Kanwar P, Sanyal SK, Tokas I, et al. Comprehensive structural, interaction and expression analysis of CBL and CIPK complement during abiotic stresses and development in rice [J]. *Cell Calcium*, 2014, 56: 81–95.
- [5] Ohta M, Guo Y, Halfter U, et al. A novel domain in the protein kinase SOS2 mediates interaction with the protein phosphatase 2C ABI2 [J]. *Proc Natl Acad Sci U S A*, 2003, 100: 11771–11776.
- [6] Albrecht V, Ritz O, Linder S, et al. The NAF domain defines a novel protein-protein interaction module conserved in Ca^{2+} -regulated kinases [J]. *EMBO J*, 2001, 20: 1051–1063.
- [7] Yu Y, Xia X, Yin W, et al. Comparative genomic analysis of *CIPK* gene family in *Arabidopsis* and *Populus* [J]. *Plant Growth Regul*, 2007, 52: 101–110.
- [8] Chen X, Gu Z, Xin D, et al. Identification and characterization of putative *CIPK* genes in maize [J]. *J Genet Genomics*, 2011, 38: 77–87.
- [9] Zhang H, Yang B, Liu WZ, et al. Identification and characterization of *CBL* and *CIPK* gene families in canola (*Brassica napus* L.) [J]. *BMC Plant Biol*, 2014, 14: 8.
- [10] Hu W, Xia Z, Yan Y, et al. Genome-wide gene phylogeny of *CIPK* family in cassava and expression analysis of partial drought-induced genes [J]. *Front Plant Sci*, 2015, 6: 914.
- [11] Xi Y, Liu J, Dong C, et al. The *CBL* and *CIPK* gene family in grapevine (*Vitis vinifera*): genome-wide analysis and expression profiles in response to various abiotic stresses [J]. *Front Plant Sci*, 2017, 8: 978.
- [12] Quan R, Lin H, Mendoza I, et al. SCABP8/CBL10, a putative calcium sensor, interacts with the protein kinase SOS2 to protect *Arabidopsis* shoots from salt stress [J]. *Plant Cell*, 2007, 19: 1415–1431.
- [13] Li R, Zhang J, Wu G, et al. HbCIPK2, a novel CBL-interacting protein kinase from halophyte *Hordeum brevisubulatum*, confers salt and osmotic stress tolerance [J]. *Plant Cell Environ*, 2012, 35: 1582–1600.
- [14] Qiu QS, Guo Y, Dietrich MA, et al. Regulation of SOS1, a plasma membrane Na^+/H^+ exchanger in *Arabidopsis thaliana*, by SOS2 and SOS3 [J]. *Proc Natl Acad Sci U S A*, 2002, 99: 8436–8441.
- [15] Batelli G, Verslues PE, Agius F, et al. SOS2 promotes salt tolerance in part by interacting with the vacuolar H^+ -ATPase and upregulating its transport activity [J]. *Mol Cell Biol*, 2007, 27: 7781–7790.
- [16] He L, Yang X, Wang L, et al. Molecular cloning and functional characterization of a novel cotton CBL-interacting protein kinase gene (*GhCIPK6*) reveals its involvement in multiple abiotic stress tolerance in transgenic plants [J]. *Biochem Biophys Res Commun*, 2013, 435: 209–215.

- [17] Lyzenga WJ, Liu H, Schofield A, et al. *Arabidopsis* CIPK26 interacts with KEG components of the ABA signalling network and is degraded by the ubiquitin-proteasome system [J]. *J Exp Bot*, 2013, 64: 2779–2791.
- [18] Tripathi V, Parasuraman B, Laxmi A, et al. CIPK6, a CBL-interacting protein kinase is required for development and salt tolerance in plants [J]. *Plant J*, 2009, 58: 778–790.
- [19] De la Torre F, Gutiérrez-Beltrán E, Pareja-Jaime Y, et al. The tomato calcium sensor Cbl10 and its interacting protein kinase CIPK6 define a signaling pathway in plant immunity [J]. *Plant Cell*, 2013, 25: 2748–2764.
- [20] Zhang G, Liu SS, Yang XJ, et al. Molecular cloning and characterization of a novel *DoSWEET1* gene from *Dendrobium officinale* [J]. *Acta Pharm Sin (药学报)*, 2016, 51: 991–997.
- [21] Chen XM, Wang CL, Yang JS, et al. Research progress on chemical composition and chemical analysis of *Dendrobium officinale* [J]. *Chin Pharm J (中国药学杂志)*, 2013, 48: 1634–1640.
- [22] Wang AH, Wu QQ, Yang L, et al. Genetic relationship of *Dendrobium* and genetic diversity of *Dendrobium nobile* in Qinling-Daba mountains revealed by ISSR [J]. *Acta Bot Boreal-Occid Sin (西北植物学报)*, 2016, 36: 1977–1983.
- [23] Zhao MM, Zhang G, Zhang DW, et al. ESTs analysis reveals putative genes involved in symbiotic seed germination in *Dendrobium officinale* [J]. *PLoS One*, 2013, 8: e72705.
- [24] Li B, Tang K, Zhang G, et al. Construction of differential cDNA library on the roots of *Dendrobium candidum* induced by a symbiotic mycorrhizal fungus [J]. *Chin Pharm J (中国药学杂志)*, 2012, 47: 1790–1795.
- [25] Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR [J]. *Nucleic Acids Res*, 2001, 29: e45.
- [26] Kozak M. An analysis of 50-noncoding sequences from 699 vertebrate messenger RNAs [J]. *Nucleic Acids Res*, 1987, 15: 8125–8132.
- [27] Chaves-Sanjuan A, Sanchez-Barrena MJ, Gonzalez-Rubio JM, et al. Structural basis of the regulatory mechanism of the plant CIPK family of protein kinases controlling ion homeostasis and abiotic stress [J]. *Proc Natl Acad Sci U S A*, 2014, 111: E4532–E4541.
- [28] Yan L, Wang X, Liu H, et al. The genome of *Dendrobium officinale* illuminates the biology of the important traditional Chinese orchid herb [J]. *Mol Plant*, 2015, 8: 922–934.