

## 基于全基因组鉴定华石斛优良内参基因

王 雨, 裴行杰, 傅钊钧, 霍恺森, 梁重钧, 刘丽岩, 王 佳, 钮 俊\*

海南大学林学院热带特色林木及观赏植物遗传与种植创新重点实验室, 海南海口 570228

**摘 要:** 华石斛是海南岛特有的热带兰科植物, 具有很高的观赏和药用价值。然而, 未确定的内参基因限制了该物种基因表达分析。本研究利用之前华石斛转录组数据, 筛选出 164 个变异系数低 ( $CV \leq 0.2$ ) 和表达量中等 (TPM 在 10~300 之间) 的候选内参基因。为了避免共表达现象对内参基因稳定性分析的影响, 根据 GO 功能聚类分析结果, 从不同的功能群中共筛选出 24 个候选基因。为避免模板中可能残留的基因组 DNA 在 RT-qPCR 扩增过程中产生假阳性, 本研究设计了跨越内含子引物, 并通过电泳筛选出符合要求的 15 个候选内参基因的扩增引物。为了确保 RT-qPCR 扩增的特异性, 使用 LightCycler 96 进行熔解曲线分析。依据 15 个候选基因的 RT-qPCR 数据, 同时综合多个算法分析, 在干旱胁迫下, 不同软件分别确定 *CLP1 & SEC23*、*SEC23 & CPY71*、*CLP1 & SEC23* 和 *RNGIL & PECT* 为最佳内参基因, 综合排序后选取 *SEC23* 和 *CLP1* 为华石斛干旱胁迫下最稳定内参基因。在华石斛不同组织中, 不同软件分别确定了 *ADF11 & IBR5*、*PRP19 & CLP1*、*ADF11 & CLP1*、*ACBP2 & IBR5* 为最佳内参基因, 综合分析后确定 *ADF11* 和 *ACBP2* 为不同组织中最稳定内参基因。结合上述干旱胁迫和不同组织分析数据, 最终确定 *IBR5* 和 *CPY71* 为综合条件下较稳定的内参基因。内参基因验证表明, 与传统的 *Actin* 内参相比, 本研究筛选出的内参基因可获得更为精确的 RT-qPCR 校准结果, 表明所筛选的内参基因具有极高的稳定性。综上所述, 本研究为利用转录组数据鉴定稳定内参基因提供思路, 首次在华石斛全基因组范围内鉴定稳定内参基因, 为进一步研究基因表达和功能奠定基础。

**关键词:** 华石斛; 内参基因; 转录组; 稳定性分析

中图分类号: S682.31 文献标识码: A

## Genome-wide Identification of Superior Reference Genes in *Dendrobium sinense*

WANG Yu, PEI Xingjie, FU Yijun, HUO Kaisen, LIANG Chongjun, LIU Liyan, WANG Jia, NIU Jun\*

Key Laboratory of Genetics and Germplasm Innovation of Tropical Special Forest Trees and Ornamental Plants, College of Forestry, Hainan University, Haikou, Hainan 570228, China

**Abstract:** *Dendrobium sinense* is a tropical plant endemic to Hainan Island with great ornamental and medicinal values. However, unascertained reference genes restrict the quantitative analysis of gene expression in the species. In this study, using our previous transcriptome data of *D. sinense*, the 164 candidate genes were identified with low variance ( $CV \leq 0.2$ ) and moderate expressions (TPM between 10–300). By GO clustering, 24 candidate genes were selected from different functional groups to avoid the effect of co-expression on further stability analysis. In order to avoid the false positive of genomic DNA that may remain in the template during RT-qPCR amplification, primers across introns were designed in this study, and amplification primers that met the requirements of 15 candidate reference genes were screened by electrophoresis. To ensure the specificity of RT-qPCR amplification, LightCycler 96 was used for melting curve analysis. According to the RT-qPCR data of 15 candidate genes, four tools were used to evaluate the expression stability. The results of multiple algorithm analysis showed that *CLP1 & SEC23*, *SEC23 & CPY71*, *CLP1 & SEC23* and *RNGIL & PECT* were the best reference genes under drought stress. After comprehensive sorting, *SEC23* and *CLP1*

收稿日期 2022-09-01; 修回日期 2022-09-29

基金项目 海南省自然科学基金项目 (No. 320RC469); 海南大学科研启动项目 (No. KYQD(ZR)-22056)。

作者简介 王 雨 (1998—), 男, 硕士研究生, 研究方向: 植物次生代谢。\*通信作者 (Corresponding author): 钮 俊 (Niu Jun), E-mail: niujun@hainanu.edu.cn。

were selected as the most stable reference genes under drought stress. In different tissues of *D.sinense*, *ADF11* & *IBR5*, *PRP19* & *CLP1*, *ADF11* & *CLP1*, *ACBP2* & *IBR5* were identified as the best reference genes, respectively. After comprehensive analysis, *ADF11* and *ACBP2* were identified as the most stable reference genes in different tissues. Combined with the above drought stress and different tissue analysis data, *IBR5* and *CPY71* were finally determined as stable reference genes under comprehensive conditions. The validation of internal reference genes showed that compared with the traditional *Actin* internal reference, the internal reference genes screened in this study could obtain more accurate RT-qPCR calibration results, indicating that the generated internal reference genes had extremely high stability. In summary, this study provides ideas for identifying stable reference genes using transcriptome data. For the first time, stable reference genes were identified in the whole genome of *D.sinense*, which laid a foundation for further study of gene expression and function.

**Keywords:** *Dendrobium sinense*; reference genes; transcriptome; stability analysis

**DOI:** 10.3969/j.issn.1000-2561.2023.07.006

华石斛 (*Dendrobium sinense*) 是海南岛特有的兰科植物, 具有很高的观赏价值和经济价值。前期的活性成分分析表明, 从华石斛中可以分离到萜类、双苄基类、黄酮类和木脂素类等化合物<sup>[1-5]</sup>。联苄是石斛属植物的主要酚类成分之一, 具有抗肿瘤、抗糖尿病、神经保护等作用<sup>[6]</sup>。近年来, 随着高通量测序技术的快速发展, 华石斛公布的基因数量也迅速增加<sup>[4]</sup>。对探索华石斛花叶发育、胁迫响应和生物活性物质合成途径有关的功能基因和转录因子研究提供了数据基础。然而, 华石斛稳定内参基因尚未鉴定, 这不可避免地会限制对华石斛基因表达及其功能的探究。

基因表达分析是生物学研究中重要且常用的方法, 研究基因表达有助于了解基因的功能<sup>[7]</sup>。实时定量聚合酶链反应 (RT-qPCR) 已经成为目前主流方法, 其具有灵敏度高、重复性好、高通量、通用性强等特点<sup>[8]</sup>。然而, RT-qPCR 产生的相对表达量受许多客观因素的影响, 如 RNA 的数量、质量和逆转录效率等<sup>[9]</sup>。RT-qPCR 相对定量时需要稳定内参基因为参考, 从而才能获得更准确可靠的表达结果<sup>[10]</sup>。

目前, 管家基因被广泛地用作内参基因进行 RT-qPCR 的定量分析<sup>[11]</sup>。管家基因在不同组织、细胞或实验处理中的表达水平理论上应该是稳定的<sup>[11-12]</sup>, 但是许多研究表明, 管家基因在不同的实验条件下仍会存在较大的表达波动<sup>[13-14]</sup>。在不同的物种中, 管家基因的稳定性必须经过严格的验证才能作为内参基因进行使用<sup>[15]</sup>。

随着测序技术的飞速发展, 高通量 RNA 测序 (RNA-seq) 已成为非模式物种基因挖掘及功能分析的有力工具。目前, 第二代测序 (NGS) 和单分子实时测序 (SMRT) 是 RNA-seq 最常用的

测序技术。2 种 RNA-seq 技术被广泛应用于植物的转录组文库构建及差异表达基因识别<sup>[4, 16]</sup>。通过高通量数据分析, 那些在各种组织类型或条件下表现稳定表达水平的基因, 可能是潜在的优良内参基因<sup>[9, 17]</sup>。

基于前期的华石斛转录组数据<sup>[4]</sup>, 本研究通过表达矩阵的归一化处理的 TPM (transcripts per kilobase of exon model per million mapped reads) 算法, 计算华石斛文库各个基因的数字表达水平, 从中筛选基因表达水平较稳定的潜在候选基因。通过基因 GO 聚类分析, 从不同的功能聚类中分别筛选内参基因。同时, 设计所筛选内参基因的跨内含子引物, 利用凝胶成像技术筛选符合条件的内参引物。根据 RT-qPCR 数据, 利用  $\Delta Ct$ 、BestKeeper、NormFinder 和 geNorm 四种不同算法软件, 评估候选内参基因表达的稳定性并从中确定最佳内参基因。本研究旨在确定华石斛的最佳内参基因, 为进一步研究基因表达和功能奠定基础。

## 1 材料与方法

### 1.1 材料

野生华石斛采自海南省昌江霸王岭国家级自然保护区。华石斛在 25 °C、相对湿度  $\geq 95\%$  的光照培养箱中培养 60 d。依据之前的干旱胁迫方法处理植物材料<sup>[4]</sup>, 随后收集华石斛根, 假鳞茎和叶组织, 经液氮速冻并 -80 °C 保存。

### 1.2 方法

1.2.1 转录数据和表达分析 从 NCBI Short Read Archive 数据库中下载转录组数据, 华石斛叶片转录组 SRR15112264、SRR15112265、SRR15112266, 假鳞茎转录组 SRR15112267、SRR15112268、SRR15112269, 根转录组 SRR15112270、SRR1511

2271、SRR15112272。另外 SRR14306704、SRR14306705 和 SRR14306706 是极度干旱胁迫的华石斛转录组数据 (DSC, 相对湿度 < 5%), SRR14306707、SRR14306708 和 SRR14306709 为一般干旱胁迫的华石斛转录组数据 (DSB,  $45\% \leq$  相对湿度  $\leq 50\%$ ) SRR14306710, SRR14306711 和 SRR14306712 代表未经干旱胁迫的华石斛转录组数据 (DSA, 相对湿度  $\geq 95\%$ )。根据全长转录组文库<sup>[4]</sup>, 用 TBtools 软件通过 TPM 算法计算基因表达量<sup>[18]</sup>。

1.2.2 基因筛选与基因聚类 根据 RNA-seq 数据的 TPM 结果, 如果满足变异系数 (CV)  $\leq 0.2$ , TPM 水平在 10~300 之间, 那么就被当作候选基因。为了避免基因共表达可能会影响内参稳定性分析结果, 使用 Blast2GO 软件对筛选出的候选内参基因进行 GO 功能聚类分析<sup>[19]</sup>。

1.2.3 引物设计 华石斛基因组 DNA (gDNA) 的提取采用植物基因组 DNA 提取试剂盒 (天根, 北京, 中国)。总 RNA 的提取采用 RNA 纯化试剂盒 (天根, 北京, 中国)。Micro-Drop (BIO-DL, 上海, 中国) 检测 RNA 纯度和浓度。用反转录试剂盒 (Monad, 苏州, 中国) 进行 RNA 反转录, RNA 用量 2  $\mu\text{g}$ 。根据与华石斛同源的铁皮石斛基因组信息 (GenBank 登录代码: GCA\_001605985.2), 利用 Primer Quest™ Tool 在不同外显子中设计 RT-qPCR 引物。gDNA 和 cDNA 的 PCR 产物用 1.5% 琼脂糖凝胶电泳验证。

1.2.4 RT-qPCR 以 1:5 稀释的 cDNA 作为模板进行 RT-qPCR 分析。RT-qPCR 结果使用 Light-Cycler 96 (Roche, Penzberg, Germany) 进行分析。根据 MonAmp™ SYBR®Green qPCR Mix (Monad, 广州, 中国) 制备 20  $\mu\text{L}$  RT-qPCR 反应体系。RT-qPCR 扩增从初始步骤 95  $^{\circ}\text{C}$  30 s 开始, 然后 95  $^{\circ}\text{C}$  变性 5 s, 55  $^{\circ}\text{C}$  退火 10 s, 72  $^{\circ}\text{C}$  延伸 30 s, 循环 40 次。为了确保扩增的特异性, 使用 Light-Cycler 96 进行熔解曲线分析。RCD1 (Radical-induced cell death 1) 是一种参与多种非生物胁迫反应的应激反应基因<sup>[20]</sup>。叶绿素合成酶 (CHLG) 参与叶绿素 a 合成的最后步骤之一<sup>[21]</sup>。本研究利用这 2 个基因对所选内参基因的可靠性进行验证。

1.2.5 基因表达稳定性分析 根据  $E=10^{\text{slope}}$  公式计算不同内参引物的扩增效率<sup>[22]</sup>。为了评估内参基因的稳定性, 使用 RefFinder (<http://blooge.cn/RefFinder/>) 筛选最佳内参基因<sup>[23]</sup>。这个网站收集

了目前主要算法 (geNorm<sup>[8]</sup>、Normfinder<sup>[24]</sup>、BestKeeper<sup>[25]</sup>和比较  $\Delta\text{Ct}$  方法<sup>[26]</sup>) 来对测试的候选内参基因进行稳定性比较和排序。

## 2 结果与分析

### 2.1 基于华石斛转录组数据筛选候选内参基因

根据之前的华石斛转录组数据, 利用 TPM 计算基因表达。最终, 共鉴定出 164 个符合变异系数 (CV  $\leq 0.2$ ) 和中等表达 (TPM 在 10~300 之间) 的候选基因。GO 聚类分析显示, 这些候选基因归属于 36 个 GO 类别 (图 1)。为了避免基因共表达对内参基因稳定性分析的影响, 因此从不同 GO 聚类分支中分别挑选候选基因, 共筛选出 24 个候选内参基因 (表 1)。例如, *RHN1* 参与囊泡运输, *ADF11* 参与肌动蛋白丝结合, *MGE2* 参与 ATP 依赖的蛋白易位, *SGS3* 参与转录后基因调控, *RNGIL* 参与蛋白泛素化。

### 2.2 内参基因引物设计

根据铁皮石斛的基因组数据, 对 24 个候选基因进行同源性鉴定, 提取基因的内含子/外显子结构, 设计的 RT-qPCR 引物至少跨越一个内含子 (表 1)。为验证 RT-qPCR 引物是否跨内含子, 分别以 cDNA 和 gDNA 为模板进行 PCR。cDNA 模板中 *RNG1*、*MGE2*、*RPS7A*、*IPUT1*、*TFIIA1*、*TF2H2*、*OOPDA*、*ANM1*、*SLX1* 基因的扩增片段与 gDNA 模板相似 (图 2A), 说明扩增产物之间不存在内含子, 这些候选基因可能会在 RT-qPCR 扩增过程中产生假阳性。*ADF11* 和 *PECT* 的扩增产物在 cDNA 模板中达到了预期的大小, 而在 gDNA 模板中未观察到扩增条带 (图 2A)。这说明在 gDNA 模板中, *ADF11* 和 *PECT* 的内含子过长, 无法完成 PCR 扩增, 因此可以当作候选基因。此外, *SGS3*、*RNGIL*、*ACBP2*、*RPAC1*、*IBR5*、*CLP1*、*VDAC1*、*SR542*、*UBP25*、*MED17*、*PRP19*、*CPY71*、*SEC23* 以 gDNA 为模板的扩增条带较 cDNA 模板中扩增条带长 (图 2A)。这说明在正向引物和反向引物的扩增区域中至少包含一个内含子。综上所述一共鉴定出 15 个候选内参基因扩增引物具有跨内含子结构。RT-qPCR 扩增后, 这 15 个内参基因进行了熔解曲线分析, 均呈现单峰 (图 2B), 表明引物特异扩增。因此, 选择这 15 个基因进行进一步的稳定性分析。

### 2.3 表达稳定性分析

根据 15 个候选基因的 RT-qPCR 数据, 采用 4

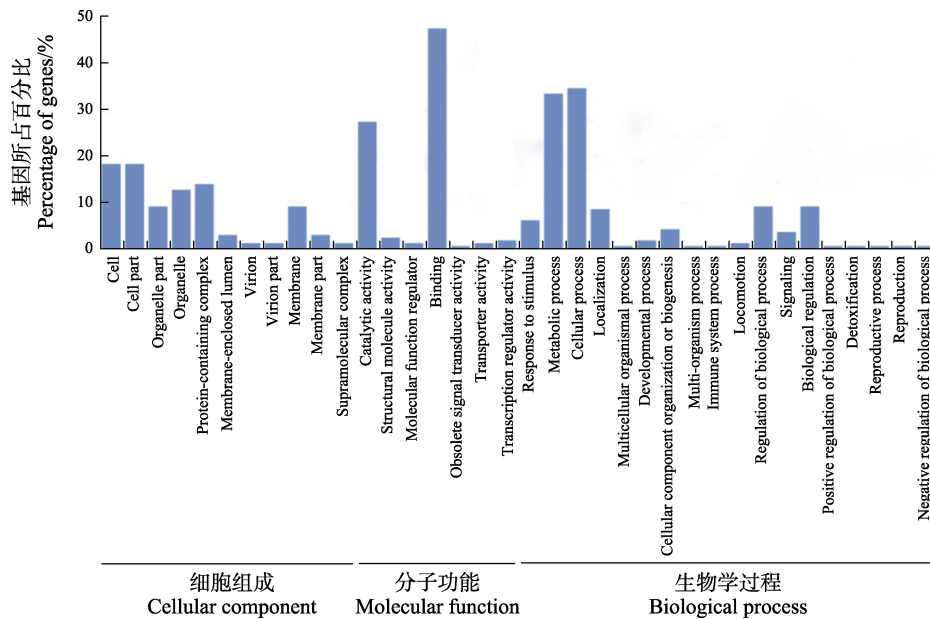


图 1 164 个候选内参基因的 GO 聚类分析

Fig. 1 Functional clustering of 164 candidate reference genes

表 1 候选内参基因及其引物序列

Tab. 1 Candidates reference genes and their primer sequences

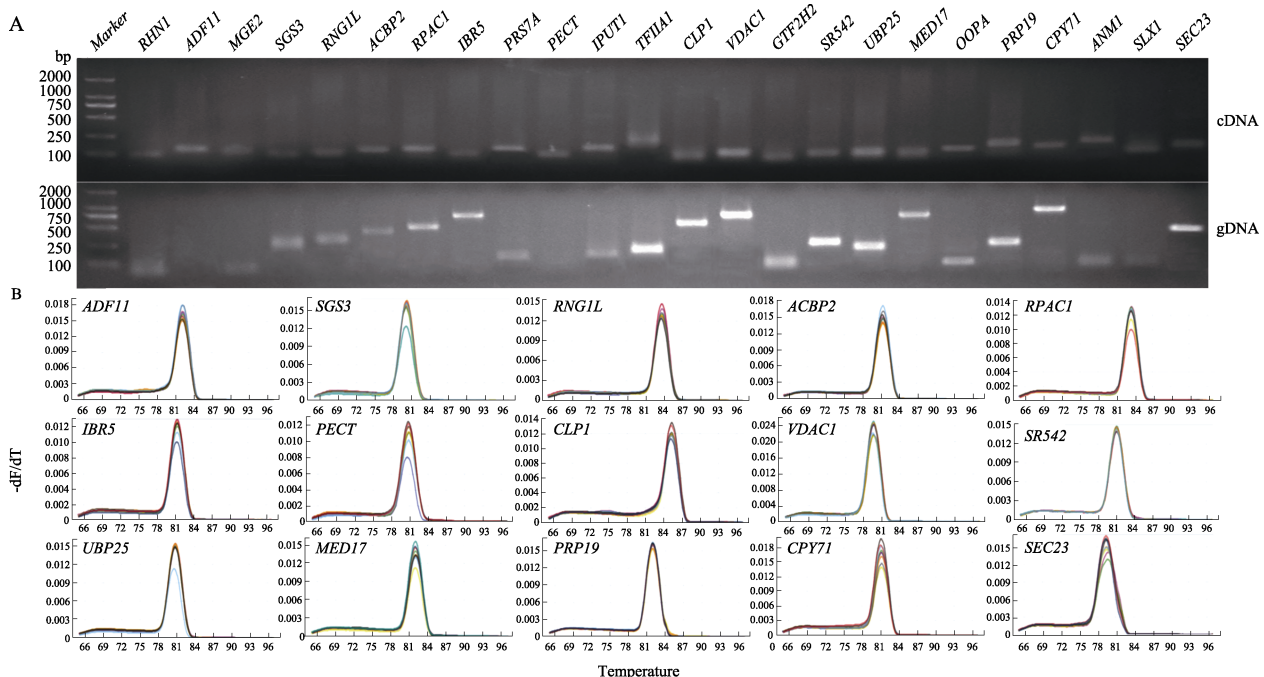
基因名称 Gene name	缩写 Abbr.	正向引物(5'-3') Forward primer (5'-3')	反向引物(5'-3') Reverse primer (5'-3')	长度 Amplicon size/bp	同源基因编号 Homologous locus <sup>a</sup>	E 值 E-value	位置 Loc. <sup>b</sup>
Ras-related protein RHN1	<i>RHN1</i>	GTGCTAGAGCTG CAGAAACA	CCTCCAATGTCAC CTTCCTTT	97	NW_021320048.1	5.0E-88	4/5-6
Actin-depolymerizing factor 11	<i>ADF11</i>	TTGACTTCGTCAC CCAAGAAA	CCATCTAAGTCTC GCCTGAATC	133	NW_021318832.1	2.0E-131	2/3
Mitochondrial GrpE homolog 2	<i>MGE2</i>	GGACAAAAGCGTG AGGTAGAA	AAGTTCCTTTGA CTGCACTAGA	115	NW_021319432.1	0	2/3
Suppressor of gene silencing 3	<i>SGS3</i>	GAGGGTTTGGGT AACAAAGAGA	AATGGCCTTCCTT CCCATATAC	94	NW_021319151.1	0	2/3
RING finger protein 1-like	<i>RNG1L</i>	AGTGGGAATGAT GAAACTTCGG	GTACAAAGAGTTG GGCTCGTT	106	NW_021318693.1	0	1/2
Acyl-CoA binding protein 2	<i>ACBP2</i>	GTAGATCGTGGCC ACCTAAAG	CGAAGGCTTCTCT CTCACAAA	124	NW_021318608.1	0	5/6
RNA polymerases I and III subunit AC1	<i>RPAC1</i>	GGATATCCACAA CGTCAGCTC	CAAACCCTATCGC CGCATAA	128	NW_021318864.1	6.0E-104	1/2
Indole-3 butyric acid response 5	<i>IBR5</i>	ATCTAGAGGTTGG CGGTTTC	AATCAACTGCTCA CGGACAA	94	NW_021319178.1	0	4/5
26S protease regulatory subunit 7A	<i>PRS7A</i>	GGAAGTGAGCTC GTTCAGAAATA	CACCTATCGCATC CACTTCATC	121	NW_021319370.1	2.0E-178	8/8-9
Phosphate-ethanolamine cytidyltransferase	<i>PECT</i>	CTGATCCGAATGC CCGAATAA	GCTTTAGCAAGCC GCAATATC	88	NW_021319309.1	0	4/5
Inositol phosphorylceramide glucuronosyltransferase 1	<i>IPUT1</i>	CTGCAGATGTGA GGAAGAGTAG	GGCCCAAGTGAT AGTGGATAA	145	NW_021318711.1	2.0E-165	4/5
Transcription initiation factor IIA subunit 1	<i>TFIIA1</i>	CCAGAATCCAGA AGGGTCTTATC	CAACCACAGACAT CCGACTG	139	NW_021319178.1	7.0E-139	7/8
Cleavage factor polyribonucleotide kinase subunit 1	<i>CLP1</i>	GTGAACGTACAC GCGATACT	GCATCCTGCACAG ACTACTT	102	NW_021319342.1	4.0E-135	3/4
Voltage-dependent anion channel protein 1	<i>VDAC1</i>	GGTTACAGTCAC AGCTGAAGAG	CAGGAGTCCCAA GTTTCATCAAT	105	NW_021318595.1	1.0E-164	1/2
General transcription factor IIIH subunit 2	<i>GTF2H2</i>	CCAGAATCCAGA AGGGTCTTATC	CAACCACAGACAT CCGACTG	102	NW_021318614.1	0	1/2

续表 1 候选内参基因及其引物序列  
Tab. 1 Candidates reference genes and their primer sequences (continued)

基因名称 Gene name	缩写 Abbr.	正向引物(5'-3') Forward primer (5'-3')	反向引物(5'-3') Reverse primer (5'-3')	长度 Amplicon size/bp	同源基因编号 Homologous locus <sup>a</sup>	E 值 E-value	位置 Loc. <sup>b</sup>
Signal recognition particle 54 kDa protein 2	SR542	GAACTGGAGAGCA TATGGATGAA	TCCATGAAGCCAG ACCAATC	94	NW_021319083.1	0	4/5
Ubiquitin carboxyl-terminal hydrolase 25	UBP25	CTCCAGTGCTTAA CCTACACT	GAAGGGGACACTC CTTCTTCTTATT	108	NW_021319800.1	0	1/2
Mediator of RNA polymerase II transcription subunit 17	MED17	GGTTGATCTTCCT GTGGCTATT	AGTCCCTGCTCAC CTTCATA	101	NW_021394673.1	0	8/9
Organelar oligopeptidase A	OOPA	ACTTCGATTTCCC GCCTTT	TCCACACTCCTCT CCAGTT	112	NW_021318492.1	7.0E-166	1/2
Pre-mRNA-processing factor 19	PRP19	AGGAAAGGGATG AAGCAAGG	TTCCCGTCAGGAC CAAATTC	139	NW_021319460.1	4.0E-142	1/2
Cyclophilin 71	CPY71	CCTAAGCGCCCTC TCAAAT	GACACTGCAACAT GTGTAACC	110	NW_021553810.1	2.0E-121	1/2
Arginine N-methyltransferase 1	ANM1	AGCATGTGTATGC GGTTGAG	TCCACATGAGTGA CTGGTAGAT	135	NW_021318883.1	0	2/4
Syndetin-like isoform X1	SLX1	GGTGAAGCATTCT GTGGATTAC	TCCAATTCTCCCT CTCCAATAC	133	NW_021319127.1	0	12/13
Protein transport protein SEC23	SEC23	CGACCGCTGTCAT TAGGTTT	GAAGCATGTGCCT CAGATGATA	122	NW_021319945.1	0	1/2

注：<sup>a</sup> 基于华石斛基因组数据的同源基因；<sup>b</sup> 对角线前后的数字分别表示正向引物和反向引物在外显子中的预测位置。

Note: <sup>a</sup> The homologues based on genomic data of *D. catenatum*; <sup>b</sup> The number before and after the diagonal bar indicated the predicted location of forward and reverse primer in exon, respectively.



A: RT-qPCR 获得扩增产物; B: 15 个候选内参基因的溶解曲线分析。

A: Amplicons obtained by RT-qPCR; B: The melting curve analysis of 15 candidate reference genes.

图 2 内参基因引物筛选

Fig. 2 Performance of the amplification primers

种工具对其表达稳定性进行评估。基因稳定性分析结果用热图呈现 (图 3), 稳定值越低, 颜色越红, 说明该基因表达越稳定。在干旱胁迫下,

通过比较  $\Delta Ct$  法、BestKeeper、Normfinder 和 geNorm 分析分别确定 *CLP1* & *SEC23*、*SEC23* & *CPY71*、*CLP1* & *SEC23* 和 *RNGIL* & *PECT* 为最

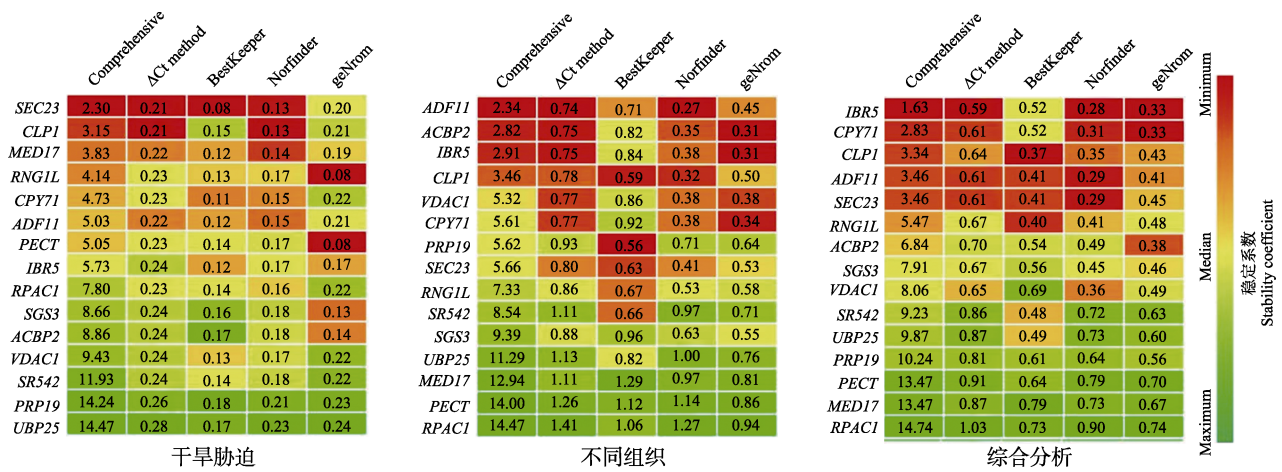


图 3 华石斛候选内参基因的稳定性分析

Fig. 3 Stability analysis of candidate reference genes in *D. sinense*

佳内参基因，综合排序后选取 *SEC23* 和 *CLP1* 作为干旱胁迫下最稳定内参基因。在华石斛不同组织中，4 种算法分别确定了 *ADF11* & *IBR5*、*PRP19* & *CLP1*、*ADF11* & *CLP1*、*ACBP2* & *IBR5* 为最佳内参基因，综合分析后确定 *ADF11* 和 *ACBP2* 为不同组织中最稳定内参基因。结合上述干旱胁迫和不同组织分析数据，最终确定 *IBR5* 和 *CPY71* 为综合条件下较稳定的内参基因。

2.4 内参基因验证

为验证本研究筛选出内参基因的准确性，以最常见的管家基因 *Actin* 作为内参基因。分别验证了 *CHLG* 在不同组织和 *RCD1* 在干旱胁迫下的表达水平（图 4）。与预期的一样，不同内参基因归一化后 *CHLG* 和 *RCD1* 的表达量与转录组测序结果呈现出相似的模式，但存在一些细微的差异。例如，在不同组织分析中，以 *Actin* 作为内参基因时，*CHLG* 在根和假鳞茎中的表达量没有显著差异，而以 *ADF11* 和 *ACBP2* 作为内参基因时，*CHLG* 在假鳞茎中的表达量显著上调。后者与转录组数

据结果更为一致，这说明本研究筛选出的内参基因可以提供更准确的定量结果。

3 讨论

华石斛是海南岛特有的热带兰科植物，具有很高的观赏和药用价值<sup>[1]</sup>。该植物附生于海拔 1000 m 的山地疏林中树干上，干旱胁迫是其生存中的最主要危害。研究表明华石斛具有丰富的活性药物成分，其生物合成机理已成为研究热点<sup>[1-2, 5-6]</sup>。然而，华石斛的稳定内参基因尚未鉴定出来，这限制了相关基因的表达和功能研究。因此，为了推进华石斛基因功能、调控机制、遗传机理等研究，稳定的内参基因鉴定至关重要。

高通量测序技术的快速发展使得非模式物种的基因资源挖掘和表达研究成为可能<sup>[4, 9]</sup>。此前的华石斛转录组测序为本研究提供了大量的基因数据<sup>[4]</sup>。将 clean reads 映射到基因文库，通过 TPM 对基因表达水平进行初步筛选。研究表明内参基因的表达水平过低或过高都会影响表达校准结果<sup>[13]</sup>，

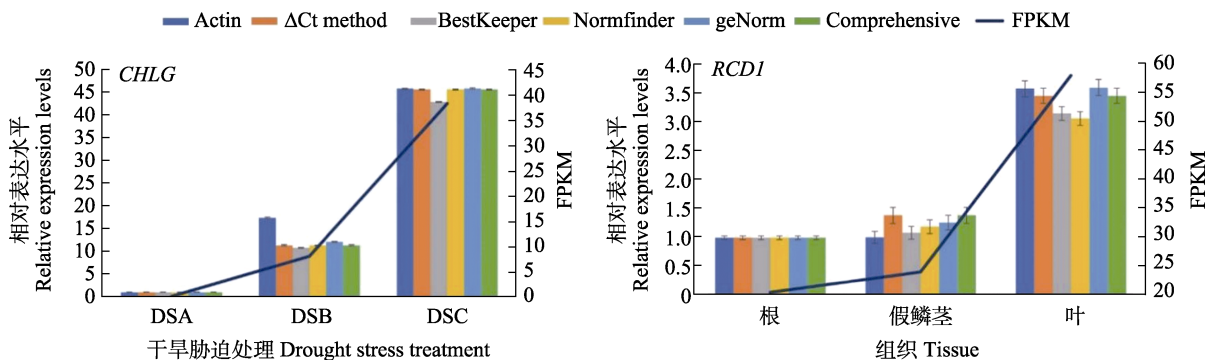


图 4 内参基因的表达校准验证

Fig. 4 Expression calibration validation of internal reference genes

本研究共鉴定出 164 个低变异系数 ( $CV \leq 0.2$ ) 和中等表达 (TPM 在 10~300 之间) 的候选内参基因。此外, 如果不同内参基因同属一个功能类群, 那么很可能这些基因存在共表达现象<sup>[27]</sup>。为避免共表达对内参基因稳定性分析的潜在影响, 候选基因从不同功能簇中挑选, 包括一些常规内参基因 (如 *ADF11*、*PRS7A*、*CPY71* 和 *UBP25*) 和不常见基因 (如 *SGS3*、*IPUT1*、*VDAC1* 和 *MED17*)。

虽然在提取 RNA 的过程中, 样品经过 DNAase 处理, 去除 gDNA 污染, 但并不清楚样品中是否存在残留的 gDNA。为避免残留 gDNA 对 RT-qPCR 分析的影响, 本研究按照前人的方法设计了跨越内含子的 RT-qPCR 引物<sup>[27]</sup>。根据凝胶电泳结果显示, 9 个候选基因 (*RNG1*、*MGE2*、*RPS7A*、*IPUT1*、*TFIIA1*、*TF2H2*、*OOPDA*、*ANMI*、*SLX1*) 的扩增引物不符合要求, 在后续分析中被剔除。溶解曲线分析表明, 剩余的 15 个候选基因引物进行特异扩增, 可用于稳定性分析。

为了评估这些候选内参基因的表达稳定性, 本研究使用了 4 种主流算法, 包括 *geNorm*<sup>[8]</sup>、*Normfinder*<sup>[24]</sup>、*BestKeeper*<sup>[25]</sup> 和比较  $\Delta Ct$  方法<sup>[26]</sup>。稳定值最低的候选基因被认为具有最高的稳定性。在 *geNorm* 分析中, 所有内参基因的稳定性值均低于阈值 1.5<sup>[8]</sup>, 说明本研究产生的内参基因的稳定性极高。这一结果可能是由于本研究基于转录组数据对内参基因进行了初步筛选, 已将不稳定的基因剔除, 与前人对苹果内参基因研究一致<sup>[9]</sup>。内参基因验证表明, 与传统的 *Actin* 内参相比, 本研究筛选出的内参基因可获得更为精确的 RT-qPCR 校准结果, 表明所产生的内参基因具有极高的稳定性。

#### 4 结论

基于华石斛前期的转录组数据, 通过广泛的筛选获得理想的 RT-qPCR 内参基因。本研究将 164 个低变异系数及中等表达的候选基因划分为 36 个 GO 分类, 从中筛选了 24 个候选基因。以 cDNA 和 gDNA 为模板, 筛选出了 15 个跨越内含子的候选基因引物。综合多种算法分别选取了 *SEC23* 与 *CLP1*、*ADF11* 与 *ACBP2*、*IBR5* 与 *CPY71* 分别为干旱胁迫、不同组织和综合情况下最稳定的内参基因。本研究首次在华石斛全基因组范围内鉴定了稳定内参基因, 为进一步研究基因表达和功能奠定了基础。

#### 参考文献

- [1] 蔡彩虹, 谭彩银, 陈惠琴, 王昊, 梅文莉, 宋希强, 戴好富. 华石斛化学成分研究(II)[J]. 广西植物, 2020, 40(9): 1368-1374.  
CAI C H, TAN C Y, CHEN H Q, WANG H, MEI W L, SONG X Q, DAI H F. Chemical constituents from *Dendrobium sinense*(II)[J]. Guihaia, 2020, 40(9): 1368-1374. (in Chinese)
- [2] CHEN X J, MEI W L, CAI C H, GUO Z K, SONG X Q, DAI H F. Four new bibenzyl derivatives from *Dendrobium sinense*[J]. Phytochemistry Letters, 2014, 9: 107-112.
- [3] CHEN X J, MEI W L, ZUO W J, ZENG Y B, GUO Z K, SONG X Q, DAI H F. A new antibacterial phenanthrenequinone from *Dendrobium sinense*[J]. Journal of Asian Natural Products Research, 2013, 15(1): 67-70.
- [4] ZHANG C, CHEN J, HUANG W, SONG X, NIU J. Transcriptomics and metabolomics reveal purine and phenylpropanoid metabolism response to drought stress in *Dendrobium sinense*, an endemic orchid species in Hainan Island[J]. Front Genet, 2021, 12: 1039.
- [5] 谭彩银, 梅文莉, 赵友兴, 黄圣卓, 孔凡栋, 杨宁宁, 宋希强, 戴好富. 华石斛化学成分研究[J]. 热带亚热带植物学报, 2017, 25(2): 189-194.  
TAN C Y, MEI W L, ZHAO Y X, HUANG S Z, KONG F D, YANG N N, SONG X Q, DAI H F. Chemical constituents from *Dendrobium sinense*[J]. Journal of Tropical and Subtropical Botany, 2017, 25: 189-194. (in Chinese)
- [6] HE L, SU Q, BAI L, LI M, LIU J, LIU X, ZHANG C, JIANG Z, HE J, SHI J. Recent research progress on natural small molecule bibenzyls and its derivatives in *Dendrobium* species[J]. European Journal of Medicinal Chemistry, 2020, 204: 112530.
- [7] VANGUILDER H D, VRANA K E, FREEMAN W M. Twenty-five years of quantitative PCR for gene expression analysis[J]. Biotechniques, 2008, 44(5): 619-626.
- [8] VANDESOMPELE J, DE PRETER K, PATTYN F, POPPE B, VAN ROY N, DE PAEPE A, SPELEMAN F. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes[J]. Genome Biology, 2002, 3(7): 1-12.
- [9] ZHOU Z, CONG P, TIAN Y, ZHU Y. Using RNA-seq data to select reference genes for normalizing gene expression in apple roots[J]. PLoS One, 2017, 12(9): e0185288.
- [10] DERVEAUX S, VANDESOMPELE J, HELLEMANS J. How to do successful gene expression analysis using real-time PCR[J]. Methods, 2010, 50(4): 227-230.
- [11] DHEDA K, HUGGETT J F, BUSTIN S A, JOHNSON M A, ROOK G, ZUMLA A. Validation of housekeeping genes for

- normalizing RNA expression in real-time PCR[J]. *Biotechniques*, 2004, 37(1): 112-119.
- [12] CHEN C M, LU Y L, SIO C P, WU G C, TZOU W S, PAI T W. Gene ontology based housekeeping gene selection for RNA-seq normalization[J]. *Methods*, 2014, 67(3): 354-363.
- [13] THELLIN O, ZORZI W, LAKAYE B, DE BORMAN B, COUMANS B, HENNEN G, GRISAR T, IGOUT A, HEINEN E. Housekeeping genes as internal standards: use and limits[J]. *Journal of Biotechnology*, 1999, 75(2/3): 291-295.
- [14] SUZUKI T, HIGGINS P J, CRAWFORD D R. Control selection for RNA quantitation[J]. *Biotechniques*, 2000, 29(2): 332-337.
- [15] LI F, CHENG Y, MA L, LI S, WANG J. Identification of reference genes provides functional insights into meiotic recombination suppressors in *Gerbera hybrida*[J]. *Horticultural Plant Journal*, 2022, 8(1): 123-132.
- [16] YE J, CHENG S, ZHOU X, CHEN Z, KIM S U, TAN J, ZHENG J, XU F, ZHANG W, LIAO Y. A global survey of full-length transcriptome of *Ginkgo biloba* reveals transcript variants involved in flavonoid biosynthesis[J]. *Industrial Crops and Products*, 2019, 139: 111547.
- [17] CZECHOWSKI T, STITT M, ALTMANN T, UDVARDI M K, SCHEIBLE W R. Genome-wide identification and testing of superior reference genes for transcript normalization in *Arabidopsis*[J]. *Plant Physiol*, 2005, 139(1): 5-17.
- [18] CHEN C, CHEN H, ZHANG Y, THOMAS H R, FRANK M H, HE Y, XIA R. TBtools: an integrative toolkit developed for interactive analyses of big biological data[J]. *Molecular Plant*, 2020, 13(8): 1194-1202.
- [19] GÖTZ S, GARCÍAGÓMEZ J M, TEROL J, WILLIAMS T D, NAGARAJ S H, NUEDA M J, ROBLES M, TALÓN M, DOPAZO J, CONESA A. High-throughput functional annotation and data mining with the Blast2GO suite[J]. *Nucleic Acids Research*, 2008, 36(10): 3420-3435.
- [20] FUJIBE T, SAJI H, WATAHIKI M K, YAMAMOTO K T. Overexpression of the RADICAL-INDUCED CELL DEATH 1 (*RCD1*) gene of *Arabidopsis* causes weak *rcd1* phenotype with compromised oxidative-stress responses[J]. *Bioscience Biotechnology and Biochemistry*, 2006, 70(8): 1827-1831.
- [21] SCHMID H C, OSTER U, KÖGEL J, LENZ S, RÜDIGER W. Cloning and characterisation of chlorophyll synthase from *Avena sativa*[J]. *Biological Chemistry*, 2001, 382: 903-911.
- [22] RAMAKERS C, RUIJTER J M, DEPREEZ R H L, MOORMAN A F. Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data[J]. *Neuroscience Letters*, 2003, 339(1): 62-66.
- [23] XIE F, XIAO P, CHEN D, XU L, ZHANG B. miRDeepFinder: a miRNA analysis tool for deep sequencing of plant small RNAs[J]. *Plant Molecular Biology*, 2012, 80(1): 75-84.
- [24] ANDERSEN C L, JENSEN J L, ØRNTOFT T F. Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets[J]. *Cancer Research*, 2004, 64(15): 5245-5250.
- [25] ZHANG K, FAN W, CHEN D, JIANG L, LI Y, YAO Z, YANG Y, QIU D. Selection and validation of reference genes for quantitative gene expression normalization in *Taxus* spp[J]. *Scientific Reports*, 2020, 10(1): 1-10.
- [26] HOJA-LUKOWICZ D, MACIĄŻEK D, KOŚCIELNIAK P, JANIK M E. Innovative GenExpA software for selecting suitable reference genes for reliable normalization of gene expression in melanoma[J]. *Scientific Reports*, 2022, 12(1): 3331.
- [27] NIU J, ZHU B, CAI J, LI P, WANG L, DAI H, QIU L, YU H, HA D, ZHAO H. Selection of reference genes for gene expression studies in Siberian Apricot (*Prunus sibirica* L.) germplasm using Quantitative Real-Time PCR[J]. *PLoS One*, 2014, 9(8): e103900.