

巴斯德毕赤酵母不同生长阶段内参基因的筛选与验证

吕梦娇^{1,2}, 战春君^{1,2}, 白仲虎^{1,2}, 杨艳坤^{*1,2}

1 江南大学粮食发酵与食品生物制造国家工程研究中心, 江苏 无锡 214122

2 江南大学江苏省生物活性制品加工工程技术研究中心, 江苏 无锡 214122

吕梦娇, 战春君, 白仲虎, 杨艳坤. 巴斯德毕赤酵母不同生长阶段内参基因的筛选与验证[J]. 微生物学通报, 2022, 49(11): 4586-4597

Lü Mengjiao, Zhan Chunjun, Bai Zhonghu, Yang Yankun. Identification and validation of reference genes for RT-qPCR normalization in *Komagataella phaffii* at different growth stages[J]. Microbiology China, 2022, 49(11): 4586-4597

摘要:【背景】巴斯德毕赤酵母(*Komagataella phaffii*)是一种甲基营养型酵母, 近年来作为生产重组蛋白和构建生物合成途径的细胞工厂受到广泛关注。实时荧光定量 PCR (real-time quantitative PCR, RT-qPCR)是巴斯德毕赤酵母表达系统研究中一种快速、高效的基因表达水平检测技术, 但需要进行归一化处理才能保证所得结果的可靠性。【目的】筛选并验证巴斯德毕赤酵母在不同生长阶段最稳定的内参基因用于精准归一化 RT-qPCR 的结果。【方法】通过转录组数据分析初步筛选出 16 个候选内参基因(*rps8b*、*rpl35a*、*rpl10*、*EIF5A*、*rpl19a*、*por1*、*rpl23b*、*0887*、*tif1*、*ole1*、*rpl14b*、*gs*、*sun*、*sdh2*、*trx1* 和 *ccp1*)。通过 RT-qPCR 技术得到候选内参基因的 C_t 值, 利用 qBASE 软件中的 geNorm 程序综合 NormFinder 算法评估内参基因的表达稳定性。【结果】通过 geNorm 分析得出精准归一化所需的最佳内参基因个数为 2, 最稳定的基因是 *rpl19a* 和 *tif1*, NormFinder 分析得到稳定性最高的内参基因为 *tif1*。此外, 利用甲酸脱氢酶编码基因 *fdh* 和乙醇脱氢酶/甲醛脱氢酶双功能酶的编码基因 *afdh* 对候选内参基因进行验证。【结论】巴斯德毕赤酵母不同生长阶段的 RT-qPCR 进行精准归一化需要 *tif1* 和 *rpl19a* 这 2 个内参基因, 为相关功能基因的表达定量提供了可靠的分析依据, 补充了 RT-qPCR 分析中的内参基因, 为巴斯德毕赤酵母不同生长阶段的基因表达调控及其应用研究提供了新的参考。

关键词: 巴斯德毕赤酵母; 内参基因; 实时荧光定量 PCR; geNorm; NormFinder

基金项目: 国家重点研发计划(2021YFC2100203); 国家自然科学基金(21908077)

Supported by: National Key Research and Development Program of China (2021YFC2100203); National Natural Science Foundation of China (21908077)

*Corresponding author: E-mail: yangyankun@jiangnan.edu.cn

Received: 2022-03-15; Accepted: 2022-04-21; Published online: 2022-05-09

Identification and validation of reference genes for RT-qPCR normalization in *Komagataella phaffii* at different growth stages

LÜ Mengjiao^{1,2}, ZHAN Chunjun^{1,2}, BAI Zhonghu^{1,2}, YANG Yankun^{*1,2}

1 National Engineering Research Center of Cereal Fermentation and Food Biomanufacturing, Jiangnan University, Wuxi 214122, Jiangsu, China

2 Jiangsu Provincial Engineering Research Center for Bioactive Product Processing, Jiangnan University, Wuxi 214122, Jiangsu, China

Abstract: [Background] The methylotrophic yeast *Komagataella phaffii* (*K. phaffii*) or *Pichia pastoris* has received extensive attention as a cell factory to produce recombinant proteins and build biosynthetic pathways. Real-time quantitative PCR (RT-qPCR) is a rapid and efficient technique for detecting gene expression levels in the study of the *K. phaffii* expression system. However, normalization is required to ensure the reliability of the results obtained. **[Objective]** To screen and verify the most stable reference genes in *K. phaffii* at different growth stages for accurate normalization of RT-qPCR. **[Methods]** Sixteen candidate reference genes (*rps8b*, *rpl35a*, *rpl10*, *EIF5A*, *rpl19a*, *por1*, *rpl23b*, *0887*, *tif1*, *ole1*, *rpl14b*, *gs*, *sun*, *sdh2*, *trx1*, and *ccp1*) were preliminarily screened by transcriptome data analysis. After the C_t values of the candidate reference genes were detected by RT-qPCR, their expression stability was analyzed by applying qBASE geNorm combined with NormFinder. **[Results]** The optimal number of reference genes by geNorm analysis was 2, and the most stable genes were *rpl19a* and *tif1*. The NormFinder analysis demonstrated that the most stable gene was *tif1*. In addition, we validated the candidate reference genes using *fdh* encoding formate dehydrogenase and *afdh* encoding the bifunctional alcohol/formaldehyde dehydrogenase. **[Conclusion]** Two reference genes, *tif1* and *rpl19a*, were required for accurate normalization of RT-qPCR at different growth stages of *K. phaffii*. This study provided a reliable basis for the analysis of expression quantification of related functional genes and enriched the reference genes in RT-qPCR analysis, facilitating the study of gene expression regulation and application at different growth stages of *K. phaffii*.

Keywords: *Komagataella phaffii*; reference gene; RT-qPCR; geNorm; NormFinder

巴斯德毕赤酵母(*Komagataella phaffii*)也称为 *Pichia pastoris*, 是一种单细胞真核微生物, 同时也是一种甲基营养型酵母^[1]。*K. phaffii* 具有调控严格的表达系统和翻译后修饰机制, 与传统的模式微生物酿酒酵母(*Saccharomyces cerevisiae*)相比, *K. phaffii* 糖基化少、生长速度快, 在基础培养基中也可以实现高密度发酵, 可以正确且高效地表达和分泌外源蛋白, 被广泛用作生产重组蛋白的宿主^[2]。此外, 近年来 *K. phaffii* 也同样被用于各种化合物的生物合成,

如诺卡酮^[3]、类胡萝卜素^[4-5]和透明质酸^[6]。*K. phaffii* 基因组序列已经公布, 再加上已经实现商品化的表达载体、工程菌株及操作手册, 人们在基因水平上对 *K. phaffii* 有了更清晰、更全面的认知^[7]。

目前, 实时荧光定量 PCR (real-time quantitative PCR, RT-qPCR)是最常用于检测基因转录水平的方法, 具有高灵敏度、高特异性、高准确度和高效率的优点^[8-9]。然而, RT-qPCR 结果的可靠性会受到 RNA 和 cDNA

的质量、引物的特异性、模板的数量和 PCR 扩增效率等因素的影响^[10-11]。为了减少这些因素的影响,可以通过使用内参基因来避免 RT-qPCR 分析存在的潜在偏倚^[12]。内参基因必须在特定的实验条件下满足稳定表达的要求。

用 *K. phaffii* 构建细胞工厂时,需要检测不同生长阶段目的基因的转录情况,目前在 *K. phaffii* 中尚无不同生长阶段表达量稳定的内参基因的研究报道,这不利于使用 RT-qPCR 方法研究重组 *K. phaffii* 中功能基因的表达情况。

在本研究中,通过分析转录组数据筛选出了 16 个稳定表达的基因,利用 RT-qPCR 技术评估了它们在不同生长阶段的表达情况,通过 geNorm^[13]和 NormFinder^[14]算法分析了它们的表达稳定性,筛选出了在 *K. phaffii* 中不同生长阶段的最佳内参基因,有利于菌株构建工程菌和重组蛋白生产,为今后开展基因表达研究提供了新的参考。

1 材料与方法

1.1 材料

供试菌株 *Komagataella phaffii* GS115 购于 Thermo Fisher Scientific 公司。

酵母浸出粉胨葡萄糖培养基(yeast extract peptone dextrose medium, YPD): 酵母浸出物 0.5 g/L, 蛋白胨 1.0 g/L, NaCl 0.5 g/L。

超纯 RNA 提取试剂盒,北京康为世纪科技有限公司; HiScript III RT SuperMix for qPCR (+gDNA wiper), 南京诺唯赞生物科技股份有限公司; ChamQ Universal SYBR qPCR Master Mix, 南京诺唯赞生物科技股份有限公司。

759S 紫外可见分光光度计,上海棱光技术有限公司; 多功能样品均质器, 欧姆尼(中

国)有限公司; Q5000 NanoDrop 紫外可见分光光度计, Quawell 公司; T series Multi-Block Thermal Cycler, 杭州朗基科学仪器有限公司; StepOnePlus RealTime PCR System 荧光定量 PCR 仪, Applied Biosystems 公司。

1.2 方法

1.2.1 菌株培养条件

用 YPD 培养基在 30 °C、200 r/min 的条件下培养野生型 *K. phaffii* GS115。每隔 2 h 取发酵液 2 mL, 利用紫外可见分光光度计测定 OD_{600} 并进行镜检, 绘制生长曲线。8、16、24、32 h 取 OD_{600} 为 10 的菌液, 12 000 r/min 离心 2 min 并用液氮速冻, 保存在 -80 °C 冰箱中。

1.2.2 总 RNA 提取和 cDNA 合成

将保存在 -80 °C 的菌体用超纯 RNA 提取试剂盒提取纯净的总 RNA。使用 NanoDrop 紫外可见分光光度计检测总 RNA 的浓度和质量。通过 2% 的琼脂糖凝胶电泳检测总 RNA 的完整性, 然后用 HiScript III RT SuperMix for qPCR (+gDNA wiper) 试剂盒去除基因组 DNA 并且反转录(表 1)。首先加入 4×gDNA wiper mix 4 μL, 加入 800 ng RNA 作为模板, 补足体系至 16 μL, 42 °C 持续 2 min, 加入 4 μL HiScript III RT SuperMix, 50 °C 持续 15 min, 85 °C 失活 5 s, cDNA 产物保存在 -20 °C 备用。

1.2.3 选择候选内参基因和引物设计

在 *K. phaffii* 已知的基因组注释数据库中寻找并鉴定各候选内参基因及对照基因 *act1*、*gap* 的开放阅读框, 利用 NCBI 网站(ncbi.nlm.nih.gov)设计引物, T_m 值为 58–62 °C, 引物长度为 18–23 bp, 扩增片段长度为 80–120 bp。引物均在苏州金唯智生物科技有限公司合成, 通过 RT-qPCR 检测引物的特异性与扩增效率, 选择熔解曲线单一峰并且扩增效率高的引物作为最终实验引物。

表 1 RNA 的浓度及纯度

Table 1 The concentration and purity of RNA

样品 Sample	$A_{260/280}$	$A_{260/230}$	浓度 Concentration (ng/ μ L)
8 h-1	2.124	2.002	153.08
16 h-1	2.182	2.029	222.48
24 h-1	2.168	2.055	210.56
32 h-1	1.963	2.144	178.04
8 h-2	2.176	1.848	185.60
16 h-2	2.187	1.998	283.12
24 h-2	2.096	1.380	156.48
32 h-2	2.215	2.242	169.52
8 h-3	1.963	1.844	161.04
16 h-3	1.633	2.440	226.36
24 h-3	2.176	1.848	195.60
32 h-3	2.214	2.094	173.12

1.2.4 RT-qPCR 条件

qPCR 反应液由 2×ChamQ Universal SYBR qPCR Master Mix 5 μ L、10 μ mol/L 的正向及反向引物各 0.1 μ L、cDNA 10 ng 和 ddH₂O 组成, 总体积为 10 μ L, 在 StepOnePlus RT-qPCR 系统上进行反应。反应条件: 95 $^{\circ}$ C 10 min; 95 $^{\circ}$ C 15 s 和 60 $^{\circ}$ C 1 min 交替进行 40 个循环。常规循环 40 次后进行熔解曲线分析(60–95 $^{\circ}$ C)以验证引物特异性。为计算每个基因的扩增效率 (E) 和相关系数 (R^2), 将反转录的 cDNA 混合后, 以 10 倍倍比稀释成 5 个浓度梯度(1 000、100、10、1 和 0.1 ng/ μ L), 并用其生成标准曲线。每个 RT-qPCR 有 3 个技术和生物学重复。

1.2.5 基因表达稳定性评估

利用 qBASE 软件中的 geNorm 程序^[13]和 NormFinder 算法^[14]对 RT-qPCR 的结果进行分析来评估基因表达的稳定性。geNorm 程序用 geNorm M 值来表征内参基因表达的稳定性, geNorm M 值越小则稳定性越高^[13]。同时 geNorm 中 $V_{n/n+1}$ 的 n 值代表了精确归一化所需的最小的内参基因数^[13,15]。NormFinder 是通过

方差分析评估候选内参基因表达的稳定性, 其结果表示为稳定性值, 稳定性值越低则基因越稳定^[16]。综合两者结果确定适用于 *K. phaffii* 不同生长阶段的最佳内参基因。

2 结果与分析

2.1 确定筛选内参基因取样节点

为确定筛选内参基因所需要的取样节点, 绘制了野生型巴斯德毕赤酵母 GS115 菌株在 YPD 培养基中的生长曲线(图 1)。菌株生长的调整期为 0–9 h, 对数期为 9.0–25.5 h, 稳定期为 25.5–37.0 h。根据绘得的生长曲线, 确定取样节点为处于调整期的 8 h 和处于对数期的 16、24 h 及处于稳定期的 32 h。

2.2 提取的总 RNA 质量检测

提取的总 RNA 的浓度及纯度检测结果如表 2 所示。4 个时间点的样品浓度均大于 150 ng/ μ L, $A_{260/280}$ 和 $A_{260/230}$ 均在 2.0 左右(表 1), 说明 RNA 纯度较高、污染较少。

通过 2% 的琼脂糖凝胶电泳检测 RNA, 可以观察到明显的 28S 和 18S 亚基条带(图 2), 表

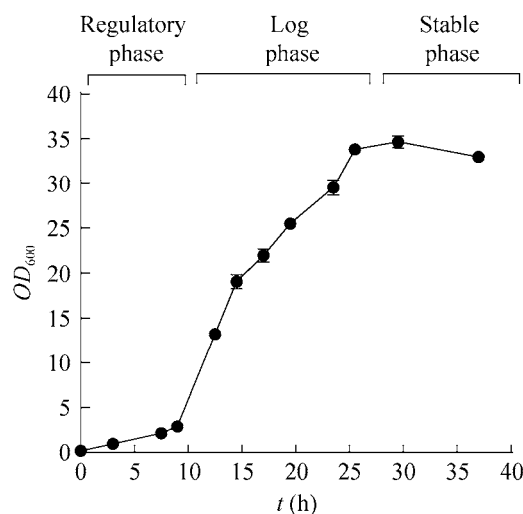


图 1 野生型 *Komagataella phaffii* GS115 生长曲线
Figure 1 Growth curve of wild type *Komagataella phaffii* GS115.

表 2 基于转录组数据筛选出的候选内参基因

Table 2 Candidate reference genes were screened based on transcriptome data

基因 Gene	产物 Production	位点 Locus
<i>rps8b</i>	40S ribosomal protein S8	PAS_chr1-1_0439
<i>rpl35a</i>	60S ribosomal protein L35	PAS_chr1-1_0302
<i>rpl10</i>	60S ribosomal protein L10	PAS_chr2-2_0054
<i>eif5a</i>	Translation initiation factor eIF-5A	PAS_chr3_1048
<i>rpl19a</i>	60S ribosomal protein L19	PAS_chr3_0091
<i>por1</i>	Mitochondrial porin	PAS_chr2-2_0392
<i>rpl23b</i>	60S ribosomal protein L23	PAS_chr1-3_0300
<i>0887</i>	Hypothetical protein	PAS_chr2-1_0887
<i>tif1</i>	Translation initiation factor eIF4A	PAS_chr3_0595
<i>ole1</i>	Delta(9) fatty acid desaturase	PAS_chr2-1_0072
<i>rpl14b</i>	60S ribosomal protein L14	PAS_c131_0011
<i>gs</i>	Glutamine synthetase	PAS_chr4_0785
<i>sun</i>	Mitochondrial outer membrane and cell wall localized SUN family member	PAS_chr4_0046
<i>sdh2</i>	Iron-sulfur protein subunit of succinate dehydrogenase	PAS_chr3_1111
<i>trx1</i>	Cytoplasmic thioredoxin isoenzyme of the thioredoxin system	PAS_chr4_0284
<i>ccp1</i>	Mitochondrial cytochrome-c peroxidase	PAS_chr2-2_0127

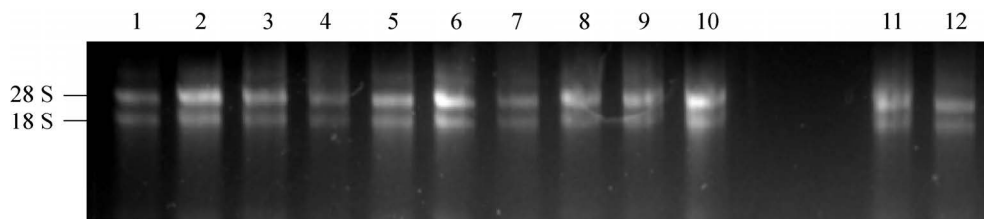


图 2 RNA 琼脂糖凝胶电泳检测结果 泳道 1、5、9 分别为 8 h 样品的 3 个生物学重复 8 h-1、8 h-2 和 8 h-3；泳道 2、6、10 分别为 16 h 样品的 3 个生物学重复 16 h-1、16 h-2 和 16 h-3；泳道 3、7、11 分别为 24 h 样品的 3 个生物学重复 24 h-1、24 h-2 和 24 h-3；泳道 4、8、12 分别为 32 h 样品的 3 个生物学重复 32 h-1、32 h-2 和 32 h-3

Figure 2 Electrophoresis of total RNA. Lanes 1, 5, and 9 are the three biological replicates 8 h-1, 8 h-2, and 8 h-3 of 8 h sample; lanes 2, 6, and 10 are the three biological replicates 16 h-1, 16 h-2, and 16 h-3 of 16 h; lanes 3, 7, and 11 are the three biological replicates 24 h-1, 24 h-2, and 24 h-3 of 24 h; lanes 4, 8, and 12 are the three biological replicates 32 h-1, 32 h-2, and 32 h-3 of 32 h.

示提取的总 RNA 完整性较高, 无降解现象, 可以用于后续 cDNA 合成。

2.3 基于转录组数据的候选内参基因筛选结果

通过分析实验室前期得到的不同培养条件下的转录组数据^[17-18]筛选出稳定表达的梯度强度基因作为候选内参基因。为确定这些基因表

达的稳定性, 计算出了它们的变异系数(CV)值, CV 值越小代表基因越稳定。以 *K. phaffii* 中常用的内参基因 *gap* 和 *act1* 作为对照, 筛选出了 16 个在不同生长阶段 CV 值小于对照基因的候选基因(表 3)。

2.4 内参基因引物扩增效率和特异性的评估

将反转录的 cDNA 模板梯度稀释作为模板

建立 RT-qPCR 引物的标准曲线, 并根据标准曲线的斜率(slope)计算出每一个基因所对应引物的扩增效率 $E=(10^{-1/\text{slope}}-1)$, 所得结果如表 3 所

示。扩增效率在 95.05%–101.19%之间, 线性相关系数(R^2)大于 0.988 9。表明所设计的引物可以进行有效扩增。

表 3 内参基因和目的基因的 RT-qPCR 引物及其扩增参数

Table 3 RT-qPCR primers for reference genes and target genes and their amplification parameters

基因 Gene	引物序列 Primer sequence (5'→3')	扩增片段长度 Amplified fragment length (bp)	扩增效率 E (%)	相关系数 R^2
<i>rps8b</i>	F CCTTGAGCCACCAGTTCCTT	105	98.91	0.989 4
	R CTAGTTCCTCGCCCTCCAAA			
<i>rpl35a</i>	F CAAGTCGGAGCAGGAGTTGG	112	95.64	0.990 1
	R ACGAACAGCCTTAATTTCTGGG			
<i>rpl10</i>	F CAAGCCGTTCCCAAAGTCCA	100	101.19	0.997 4
	R TGGGAAGTCATCGACAGCAG			
<i>eil5a</i>	F CCATGTTGTGGGTGGATGGA	100	100.08	0.996 3
	R AAGACTGGTAAGCACGGTCA			
<i>rpl19a</i>	F AGACCCAAACGAGACTGAGC	109	96.36	0.989 9
	R TCTGGAGTGTCCACAGGAG			
<i>por1</i>	F AGCGTCCTTCTTCTGTCTCGG	105	100.13	0.995 6
	R ACTGCCAATGCCCTAGACAC			
<i>rpl23b</i>	F CATGTCTCCAACAGAGGCGG	106	99.73	0.994 9
	R CATTAACTGCGCCGACAACA			
0887	F TCCAACCTCGACTGGTTCACA	105	96.26	0.989 3
	R GCTGGCCCTTGTAACAGTA			
<i>tif1</i>	F CCAAGACATCACGACCAGCA	100	96.30	0.991 4
	R GAGCAAATCCTGAGAGCCGT			
<i>ole1</i>	F TCAACCGTTCGATGACAGGA	101	100.01	0.999 4
	R CTGATGGGAACCTCGTGGTGG			
<i>rpl14b</i>	F TCAGCTCACTTCTTCTCTGCT	107	95.05	0.992 1
	R CTCCGATGCTGTCAAGGAGG			
<i>gs</i>	F CCTACTACGCCGACCCTTTC	112	100.02	0.988 9
	R AGCTTGGCACACTCATGTCT			
<i>sun</i>	F CCTGGAACACAGAGGGTGTCT	104	100.11	0.996 3
	R CCTCCAATCAGCCATCCGA			
<i>sdh2</i>	F CGTAGATCGTGTCTGAGGG	104	97.10	0.997 8
	R TCCGTGCTCTCATTCTGGTC			
<i>trx1</i>	F TGAGTGAGGTTGCTGCTCGT	111	95.91	0.991 5
	R GCTTGATAGCAGCAGCATTGG			
<i>ccp1</i>	F GTGACGGTGCCAATAACGGA	102	99.76	0.998 2
	R CAGCCAAGGTCCACAAGTCA			
<i>act1</i>	F TTTGGACTCTGGTGACGGTG	105	98.37	0.989 3
	R TCGGTCAAATCTCTACCGGC			
<i>gap</i>	F ATGGAGCAGTGATGACGACC	106	99.54	0.993 7
	R GCTGGTGTCTGACTACGTCAT			
<i>afd1</i>	F CCAGGTCGTTCTCCTTGACG	91	97.63	0.989 6
	R AAGGCAGTCAACCCAACCAA			
<i>fdh</i>	F AAGCACATTCTGACGCTGA	112	98.71	0.990 6
	R CACCAGCAACGACCAACAAC			

汇总 RT-qPCR 得到的熔解曲线如图 3 所示, 所有引物均只有单峰。综上所述, 所设计的引物扩增效率较高、特异性高, 满足 RT-qPCR 对引物的要求。

2.5 候选内参基因表达水平的评估

为评估转录组筛选出来的候选内参基因

的表达稳定性, 首先利用 RT-qPCR 检测这些基因在不同生长阶段的样品中的表达水平。 C_t 值表示转录水平, C_t 值越高则转录水平越低。如图 4 所示, 17 个内参基因 C_t 值的平均值在 15.15–22.11 之间, 说明各个基因之间的表达水平差异很大。其中 *rps8b* 的表达水平最高, 为

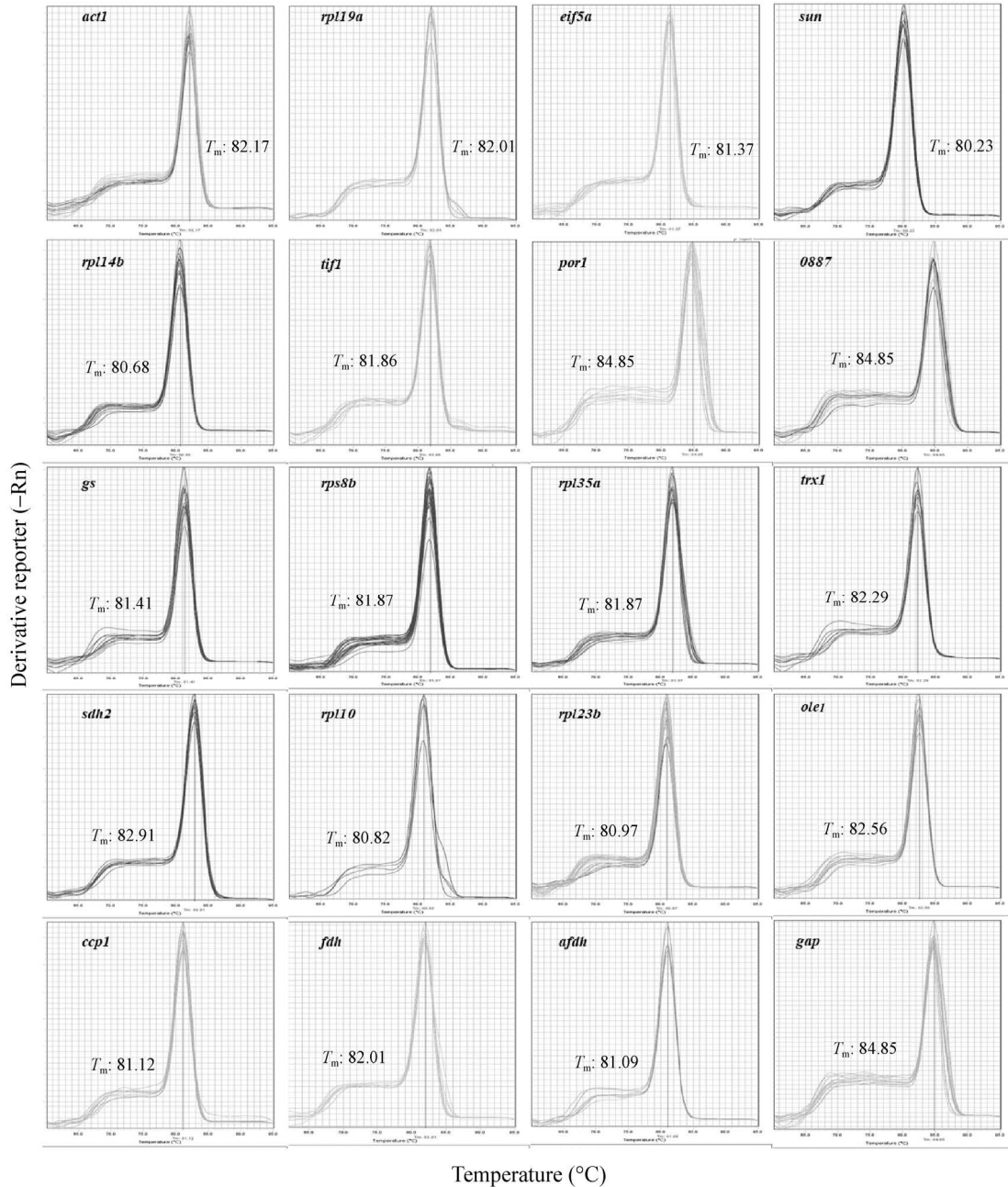


图 3 候选内参基因和目的基因的熔解曲线

Figure 3 Fusion curves of candidate reference genes and target genes.

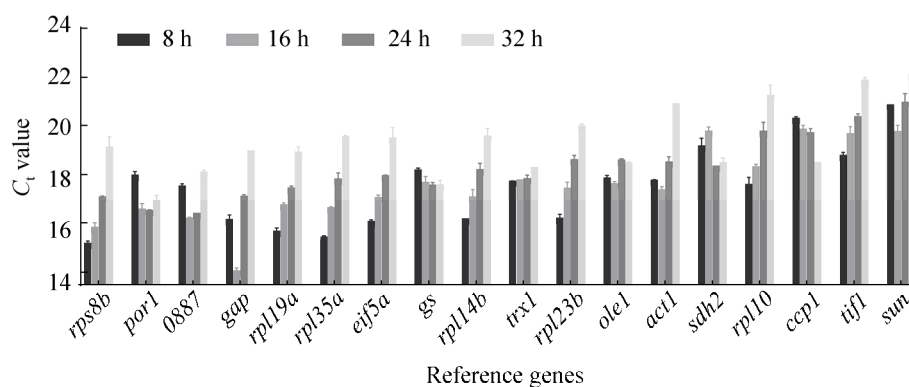


图 4 候选内参基因的 C_t 值分布

Figure 4 C_t values for candidate reference genes.

16.77±1.74, *sun* 的表达水平最低(20.92±0.97)。根据各基因在不同生长阶段中的 C_t 值比较可以发现, 有 12 个基因在 32 h 的表达水平最低, 然而在没有统一的归一化程序处理 RT-qPCR 数据的情况下, 这种基因的表达差异也有可能是由于样品的异质性或模板浓度的差异引起的, 因此, 筛选内参基因对于分析基因的水平极其重要。

2.6 候选内参基因表达稳定性的评估

利用 qBASE 软件的 geNorm 程序来分析 RT-qPCR 的结果从而评估候选基因的稳定性。geNorm 分析所得到的 M 值可以用于表征基因的稳定性, M 值最高的基因稳定性最低。 M 值小于 0.7 的基因可以用作内参基因。在 16 个候选基因中有 10 个基因在不同培养条件下的 geNorm M 值均小于 0.7, 说明这些基因是较稳定的, 足以用作 *K. phaffii* RT-qPCR 时的内参基因。其中 *rps8b*、*rpl35a*、*rpl10*、*rpl23b*、*rpl14b*、*ef5a*、*tif1* 和 *rpl19a* 的 geNorm M 值小于 *act1* 和 *gap*, 说明这 8 个基因比常用的两个内参基因 *act1* 和 *gap* 更稳定, 更适合用于 RT-qPCR 的归一化(图 5A)。其中, 最稳定的基因是 *rpl19a* (0.041)。

此外, 在 RT-qPCR 实验中应用多个内参基因进行归一化可以使结果更可信。geNorm 的成对变异分析 $V_{n/n+1}$ 显示了精确归一化所需的

最小内参基因数。当 $V_{n/n+1}$ 的结果小于 0.15 时, n 就是可靠归一化所需的最少的内参基因数。在不同生长阶段 $V_{2/3}$ 的值为 0.042, 小于 0.15, 说明最佳的内参基因数为 2 (图 5B)。因此, 在不同的生长阶段, 最佳的内参基因的组合为 *rpl19a* (0.098) 和 *tif1* (0.112)。

NormFinder 可以进一步验证 geNorm 结果的可靠性。基于方差的算法通过计算组间和组内变异来表示稳定性^[16]。最稳定的基因将得到最低的稳定性值。这个程序可以减少候选基因相关表达的影响, 但是只能筛选出一个最稳定的内参基因。计算出的 16 个候选内参基因表达稳定性的排序如图 6 所示, 最稳定的内参基因为 *tif1* (0.45)。

综上所述, 在不同的生长阶段, 适用于巴斯德毕赤酵母 RT-qPCR 结果精准归一化的最佳内参基因是 *rpl19a* 和 *tif1*。

2.7 内参基因在 RT-qPCR 中的稳定性验证

分别选用巴斯德毕赤酵母甲酸脱氢酶编码基因 *fdh* 和乙醇脱氢酶/甲醛脱氢酶双功能酶的编码基因 *afdh* 作为目的基因, 验证已筛选的稳定内参基因并与最常用的内参基因 *act1* 和 *gap* 进行比较。

发现用不同的内参基因对 *fdh* 基因在不同生长阶段进行相对定量时, *gap*、*act1*、*tif1*、

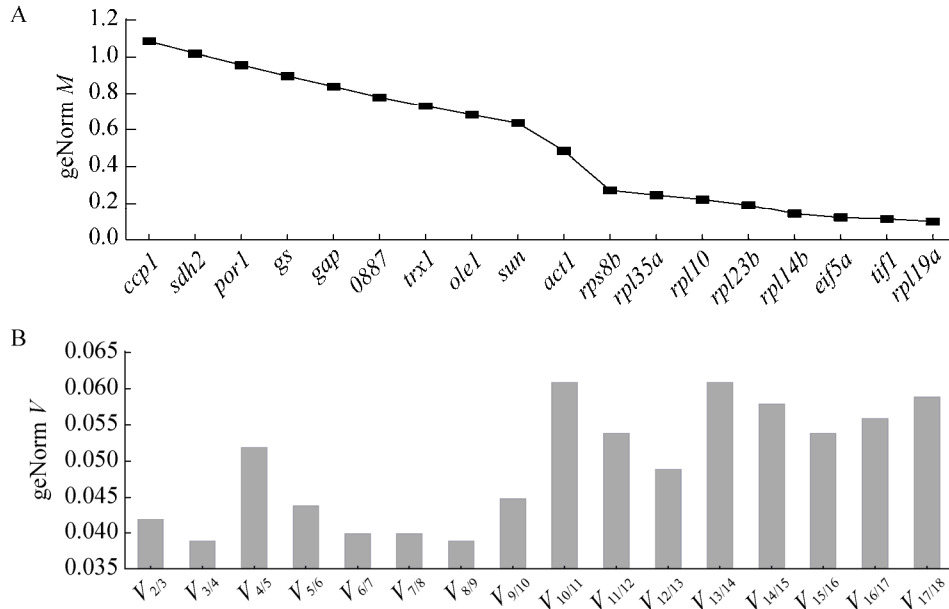


图5 geNorm 程序分析结果 geNorm 程序分析 *K. phaffii* 不同生长阶段内参基因的表达稳定性排序(A)和精确归一化所需的最小内参基因数(B)

Figure 5 Results of the geNorm program analysis. Expression stability ranking (A) and the minimum number (B) of reference genes required for accurate normalization of reference genes at different growth stages for *K. phaffii* based on geNorm program analysis.

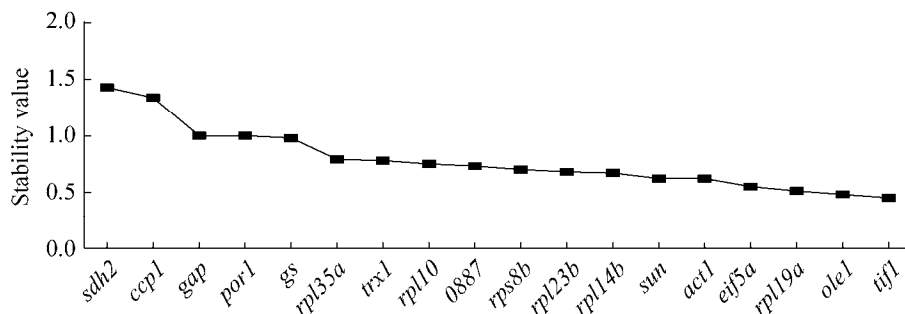


图6 NormFinder 算法分析 *Komagataella phaffii* 不同生长阶段内参基因的表达稳定性排序

Figure 6 Expression stability ranking of reference genes at different growth stages *Komagataella phaffii* based on NormFinder algorithm analysis.

rpl19a 和 *tif1* & *rpl19a* 存在相似的变化趋势，但是内参基因 *tif1* & *rpl19a* 比 *act1* 和 *gap* 倍数变化差异较大，说明所筛选出的内参基因 *tif1* & *rpl19a* 对目的基因变化的表征更加明显(图 7A)。

用不同的内参基因对 *afdH* 基因在不同生长阶段进行相对定量时，*afdH* 在响应时间上略有

差异，在用 *act1* 和 *gap* 进行归一化时，*afdH* 的表达量在 24 h 达到了最高点，但是利用 *tif1*、*rpl19a* 及 *tif1* 和 *rpl19a* 进行归一化时，该基因的表达量在 16 h 达到最高，而且 *tif1* & *rpl19a* 对于目的基因的变化表征同样比 *act1* 更加明显(图 7B)。

以上结果说明选择不同的内参基因进行归

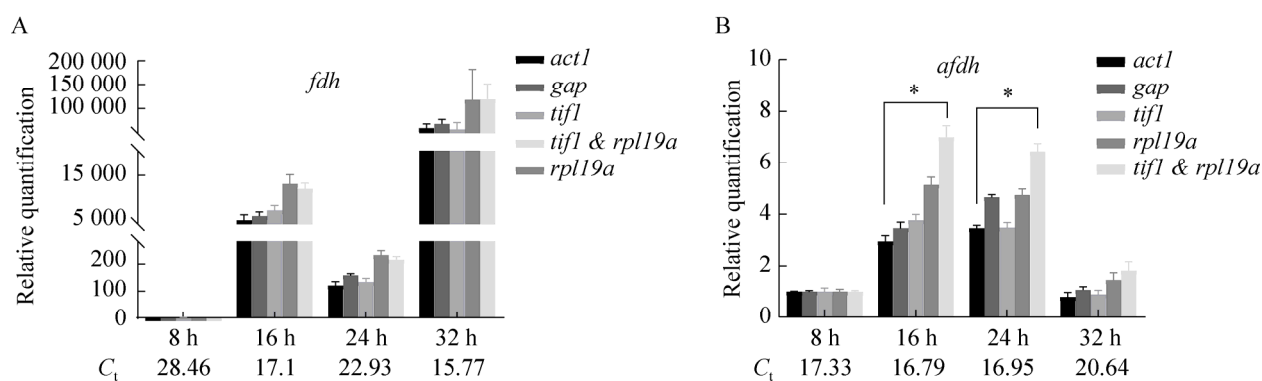


图 7 在不同生长阶段 *fdh* (A)和 *afd*h (B)的相对表达量 用 *gap*、*act1*、*tif1*、*rpl19a* 和 *tif1 & rpl19a* 作为数据归一化的内参基因，*代表具有显著性差异

Figure 7 Relative expression of *fdh* (A) and *afd*h (B) at different growth stages. *Gap*, *act1*, *tif1*, *rpl19a*, and *tif1 & rpl19a* were used as reference genes for data normalization. * represents significant difference.

一化会影响 RT-qPCR 结果分析的准确性，所筛选获得的内参基因 *tif1 & rpl19a* 可适用于巴斯德毕赤酵母不同生长阶段功能基因的实时荧光定量分析。

3 讨论与结论

RT-qPCR 是目前比较流行的评估基因表达水平的技术方法，但是为获得准确的 RT-qPCR 结果，必须使用一个稳定的内参基因进行归一化，以尽量减少样品之间由于非生物差异造成的误差。以往的研究表明，并无一种单一的内参基因可以用于归一化所有的实验条件或者物种^[19-20]。

gap 广泛存在于多种生物，常被选作内参基因，但是 *K. phaffii* 在不同碳源中培养时 *gap* 基因的表达差异巨大^[21]。此外，*act1* 在 *K. phaffii* 中广泛应用于 RT-qPCR 的内参基因。尽管在以往的研究中，*gap* 和 *act1* 常被用于 RT-qPCR 分析的内参基因，在 *Cryptococcus punctulatus* 的不同发育阶段和组织中，*act1* 为最稳定的内参基因^[22]，在不同季节的蜜蜂组织，*gap* 为最适内参基因^[23]。在 *Yarrowia lipolytica* 的研究中发现，*act1* 稳定性受生长阶

段和酵母菌株类型的影响，更适合作为不同培养基成分下的内参基因^[24]。但在本研究中，*gap* 和 *act1* 在不同的生长阶段并未表现出明显的稳定性。此外，*Miscanthus sacchariflorus* 在非生物胁迫的条件下，*gap* 和 *act1* 是最不稳定的基因^[25]。事实上，Northern blotting^[26] 和 RT-qPCR^[27] 证明，从指数期到稳定期，*act1* 的表达受到严重抑制。在 *S. cerevisiae* 中，如果不仔细考虑生长阶段相关的变异性，利用 *act1* 作为内参基因可能会导致对稳定生长阶段目标转录水平的高估超过 10 倍^[27-28]。以往的研究表明，*gap* 在哺乳动物系统中存在不稳定性^[29]，本研究将这一现象的范围扩大到毕赤酵母。上述结果说明，即使是同源基因，在不同宿主和不同实验条件下也会表现出不同的稳定性。因此，在每种实验环境中都有必要对内参基因进行评估和筛选。

在本研究中，从 *K. phaffii* 的转录组测序数据库初步筛选到较稳定的 16 个候选内参基因，经过 geNorm 和 NormFinder 计算，筛选到不同生长阶段的最佳内参基因 *tif1* 和 *rpl19a*。用这些内参基因及 *gap*、*act1* 进行归一化时，靶向基因 *afd*h 和 *fd*h 的表达水平发生了明显的

变化,说明内参基因的选择会影响 RT-qPCR 分析的结果。同样地, Liu 等^[19]在 *Paspalum vaginatum* 中对 *mt2a*、*pip1*、*vp1* 和 *cor413* 表达模式用 RT-qPCR 进行检测的过程中也发现,使用稳定和不稳定的内参基因进行归一化会导致目的基因的表达水平存在显著差异,这会引致实验结果的误差。

本研究的结果表明,与 *act1* 和其他内参基因的表达相比, rRNA (*rps8b*、*rpl35a*、*rpl10*、*rpl23b*、*rpl14b* 和 *rpl19a*) 的表达具有更好的一致性。在多种表达系统及各种实验条件下,已有研究比较了 18S rRNA 基因和 28S rRNA 基因相较于其他内参基因的稳定表达水平,包括哺乳动物和酵母细胞^[30-33]。因为 rRNA 是由一种独特的 RNA 聚合酶转录的,在影响 mRNA 表达的条件下, rRNA (占总 RNA 的 80%) 的水平不太可能发生变化^[32-33]。

在 *K. phaffii* 发酵重组蛋白的过程中,某些产物会与生长关联,如用于动物饲料添加剂的单细胞蛋白^[34]。因此,本研究筛选的稳定的内参基因为 *K. phaffii* 不同生长阶段的相关功能基因的表达定量提供了可靠的分析依据,同时也提高了功能基因研究的重复性、稳定性和可靠性,有助于研究 *K. phaffii* 在发酵过程中关键基因的表达,从而使生产效率最大化。

REFERENCES

- [1] Kurtzman CP. Biotechnological strains of *Komagataella (Pichia) pastoris* are *Komagataella phaffii* as determined from multigene sequence analysis[J]. *Journal of Industrial Microbiology and Biotechnology*, 2009, 36(11): 1435
- [2] Parashar D, Satyanarayana T. Enhancing the production of recombinant acidic α -amylase and phytase in *Pichia pastoris* under dual promoters [constitutive (GAP) and inducible (AOX)] in mixed fed batch high cell density cultivation[J]. *Process Biochemistry*, 2016, 51(10): 1315-1322
- [3] Wriessnegger T, Augustin P, Engleder M, Leitner E, Müller M, Kaluzna I, Schürmann M, Mink D, Zellnig G, Schwab H, et al. Production of the sesquiterpenoid (+)-nootkatone by metabolic engineering of *Pichia pastoris*[J]. *Metabolic Engineering*, 2014, 24: 18-29
- [4] Araya-Garay JM, Feijoo-Siota L, Rosa-dos-Santos F, Veiga-Crespo P, Villa TG. Construction of new *Pichia pastoris* X-33 strains for production of lycopene and β -carotene[J]. *Applied Microbiology and Biotechnology*, 2012, 93(6): 2483-2492
- [5] Bhataya A, Schmidt-Dannert C, Lee PC. Metabolic engineering of *Pichia pastoris* X-33 for lycopene production[J]. *Process Biochemistry*, 2009, 44(10): 1095-1102
- [6] Jeong E, Shim WY, Kim JH. Metabolic engineering of *Pichia pastoris* for production of hyaluronic acid with high molecular weight[J]. *Journal of Biotechnology*, 2014, 185: 28-36
- [7] Sturmberger L, Chappell T, Geier M, Krainer F, Day KJ, Vide U, Trstenjak S, Schiefer A, Richardson T, Soriaga L, et al. Refined *Pichia pastoris* reference genome sequence[J]. *Journal of Biotechnology*, 2016, 235: 121-131
- [8] Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, et al. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments[J]. *Clinical Chemistry*, 2009, 55(4): 611-622
- [9] Zhang C, Li T, Hou CL, Shen XY. Selection of reference genes from *Shiraia bambusicola* for RT-qPCR analysis under different culturing conditions[J]. *AMB Express*, 2017, 7(1): 14
- [10] Hao XY, Horvath DP, Chao WS, Yang YJ, Wang XC, Xiao B. Identification and evaluation of reliable reference genes for quantitative real-time PCR analysis in tea plant (*Camellia sinensis* (L.) O. Kuntze)[J]. *International Journal of Molecular Sciences*, 2014, 15(12): 22155-22172
- [11] Huang LK, Yan HD, Jiang XM, Zhang XQ, Zhang YW, Huang X, Zhang Y, Miao JM, Xu B, Frazier T, et al. Evaluation of candidate reference genes for normalization of quantitative RT-PCR in switchgrass under various abiotic stress conditions[J]. *BioEnergy Research*, 2014, 7(4): 1201-1211
- [12] Mascia T, Santovito E, Gallitelli D, Cillo F. Evaluation of reference genes for quantitative reverse-transcription polymerase chain reaction normalization in infected tomato plants[J]. *Molecular Plant Pathology*, 2010, 11(6): 805-816
- [13] 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): RESEARCH0034
- [14] Andersen CL, Jensen JL, Ørntoft TF. 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
- [15] Tong ZG, Gao ZH, Wang F, Zhou J, Zhang Z. Selection of reliable reference genes for gene expression studies in peach using real-time PCR[J]. *BMC Molecular Biology*, 2009, 10: 71
- [16] Løvdal T, Saha A. Reference gene selection in *Carnobacterium maltaromaticum*, *Lactobacillus curvatus*, and *Listeria innocua* subjected to temperature and salt stress[J]. *Molecular Biotechnology*, 2014, 56(3): 210-222
- [17] Li X, Yang YK, Zhan CJ, Zhang ZY, Liu XX, Liu HB, Bai ZH. Transcriptional analysis of impacts of glycerol transporter 1 on methanol and glycerol metabolism in *Pichia pastoris*[J]. *FEMS Yeast Research*, 2017, 18(1): fox081
- [18] Wang RB, Zhao TY, Zhuo JL, Zhan CJ, Zhang FM, Linhardt RJ, Bai ZH, Yang YK. MAPK/HOG signaling pathway induced stress-responsive damage repair is a mechanism for *Pichia pastoris* to survive from hyperosmotic stress[J]. *Journal of Chemical Technology & Biotechnology*, 2021, 96(2): 412-422
- [19] Liu Y, Liu J, Xu L, Lai H, Chen Y, Yang ZM, Huang BR. Identification and validation of reference genes for seashore *Paspalum* response to abiotic stresses[J]. *International Journal of Molecular Sciences*, 2017, 18(6): 1322
- [20] Wang XY, Peng F, Dong GB, Sun Y, Dai XF, Yang YK, Liu XX, Bai ZH. Identification and validation of appropriate reference genes for qRT-PCR analysis in *Corynebacterium glutamicum*[J]. *FEMS Microbiology Letters*, 2018, 365(8): fny030
- [21] Waterham HR, Digan ME, Koutz PJ, Lair SV, Cregg JM. Isolation of the *Pichia pastoris* glyceraldehyde-3-phosphate dehydrogenase gene and regulation and use of its promoter[J]. *Gene*, 1997, 186(1): 37-44
- [22] Li Z, Li XR, Zhang QW, Yuan L, Zhou XG. Reference gene selection for transcriptional profiling in *Cryptocercus punctulatus*, an evolutionary link between *Isoptera* and *Blattodea*[J]. *Scientific Reports*, 2020, 10: 22169
- [23] Jeon JH, Moon K, Kim Y, Kim YH. Reference gene selection for qRT-PCR analysis of season- and tissue-specific gene expression profiles in the honey bee *Apis mellifera*[J]. *Scientific Reports*, 2020, 10: 13935
- [24] Borkowska M, Białas W, Celińska E. A new set of reference genes for comparative gene expression analyses in *Yarrowia lipolytica*[J]. *FEMS Yeast Research*, 2020, 20(7): foaa059
- [25] Zong JQ, Chen JB, Li L, Li JJ, Li DD, Wang JJ, Liu J, Liu JX. Reference gene selection for quantitative RT-PCR in *Miscanthus sacchariflorus* under abiotic stress conditions[J]. *Molecular Biology Reports*, 2022, 49(2): 907-915
- [26] Wenzel TJ, Teunissen AWRH, Steensma HY. PDA1 mRNA: a standard for quantitation of mRNA in *Saccharomyces cerevisiae* superior to ACT1 mRNA[J]. *Nucleic Acids Research*, 1995, 23(5): 883-884
- [27] Teste MA, Duquenne M, François JM, Parrou JL. Validation of reference genes for quantitative expression analysis by real-time RT-PCR in *Saccharomyces cerevisiae*[J]. *BMC Molecular Biology*, 2009, 10: 99
- [28] Monje-Casas F, Michán C, Pueyo C. Absolute transcript levels of thioredoxin- and glutathione-dependent redox systems in *Saccharomyces cerevisiae*: response to stress and modulation with growth[J]. *The Biochemical Journal*, 2004, 383(Pt 1): 139-147
- [29] Brunner AM, Yakovlev IA, Strauss SH. Validating internal controls for quantitative plant gene expression studies[J]. *BMC Plant Biology*, 2004, 4: 14
- [30] Blanquicett C, Johnson MR, Heslin M, Diasio RB. Housekeeping gene variability in normal and carcinomatous colorectal and liver tissues: applications in pharmacogenomic gene expression studies[J]. *Analytical Biochemistry*, 2002, 303(2): 209-214
- [31] Bond BC, Virley DJ, Cairns NJ, Hunter AJ, Moore GBT, Moss SJ, Mudge AW, Walsh FS, Jazin E, Preece P. The quantification of gene expression in an animal model of brain ischaemia using TaqMan™ real-time RT-PCR[J]. *Molecular Brain Research*, 2002, 106(1/2): 101-116
- [32] Zhong H, Simons JW. Direct comparison of GAPDH, beta-actin, cyclophilin, and 28S rRNA as internal standards for quantifying RNA levels under hypoxia[J]. *Biochemical and Biophysical Research Communications*, 1999, 259(3): 523-526
- [33] 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
- [34] Ahmad M, Hirz M, Pichler H, Schwab H. Protein expression in *Pichia pastoris*: recent achievements and perspectives for heterologous protein production[J]. *Applied Microbiology and Biotechnology*, 2014, 98(12): 5301-5317