

基因表达转录分析中内参基因的选择*

张艳君^{1,2)} 朱志峰¹⁾ 陆融^{1,3)} 徐琼¹⁾

石琳熙¹⁾ 简序¹⁾ 刘俊燕¹⁾ 姚智^{1)**}

(¹⁾天津医科大学免疫学教研室, 天津 300070; ²⁾天津医科大学生物化学教研室, 天津 300070;

³⁾深圳康哲药业有限公司, 深圳 518029)

摘要 目前基因表达的转录分析多采用单一看家基因作为内参来校正目标基因的表达量. 实验中以人肝癌 BEL-7402 细胞为研究对象, 应用实时荧光定量 PCR 技术, 观察了新型三肽化合物酪丝缬肽作用后 RPL13A、UBC、EIF4A、B2M、GAPDH 和 ACTB 共 6 个看家基因 mRNA 水平的表达情况. 经过 geNorm 程序统计学分析处理, 结果表明, 这 6 个看家基因的表达存在差异, 确定了 RPL13A、UBC 2 个看家基因用于校正目标基因的表达量. 基因表达转录分析中内参基因选择的必要性在实验中得以证明, 更重要的是为各种实验因素影响下(尤其是新物质作用下)内参基因的选择介绍和提供了一种行之有效的方法.

关键词 内参基因, geNorm 程序, 基因表达, 实时定量 PCR

学科分类号 R 331

基因表达分析在生命科学的众多研究领域变得日趋重要, 其研究的深入将为探索疾病相关基因、了解基因表达调控的复杂网络、解析生命奥秘、最终为人类服务大有裨益. 在转录水平进行基因表达定量的方法, 如实时定量 PCR、RNA 印迹、核糖核酸酶保护分析、基因芯片等, 在蛋白质水平进行定量的方法如蛋白质印迹等, 都需要应用内参基因对目标基因表达量进行校正, 以期获得真实可靠的结果. 看家基因有几百种, 最常用的为 ACTB、GAPDH、18SrRNA、28SrRNA, 理想的看家基因应在各种实验因素条件下, 各种类型的组织或细胞中均恒定表达^[1]. 然而, 大量的研究表明, 任何一种看家基因的所谓恒定表达都只是在一定类型的细胞或实验因素作用下“有范围”的恒定. 在其他类型的细胞中或实验因素作用下则是变化的, 有时可以是十几倍、几十倍甚至是上百倍的差异^[2,3]. 盲目地使用一种看家基因作为内参, 一方面可能使基因表达的微小差异难以发现, 另一方面可能引致错误甚至相反的结论^[4,5]. 本研究中的酪丝缬肽(tyroservatide, YSV)是一种新三肽化合物, 化学结构组成为 L-酪氨酰-L-丝氨酰-L-缬氨酸, 分子结构式为 C₁₇H₂₅N₃O₆, 分子质量为 367.40, 先前研究表明 YSV 对人肝癌 BEL-7402 裸小鼠移植瘤

具有显著的抑制作用^[6]. 为了深入研究 YSV 的作用机理, 我们首先需要明确 YSV 对看家基因的影响情况. 本研究选择 RPL13A、UBC、EIF4A、B2M、GAPDH 和 ACTB 共 6 个看家基因, 通过实时定量 PCR 技术对 YSV 作用下 BEL-7402 细胞中各看家基因的转录水平进行检测, 同时采用 Vandesompele 等^[7]编写的 geNorm 程序对检测结果进行统计学处理, 分析哪些看家基因适于 YSV 对肿瘤基因表达水平影响的研究.

1 材料和方法

1.1 人肝癌 BEL-7402 细胞培养和总 RNA 提取

将人肝癌 BEL-7402 细胞经 0.25%胰酶-0.02% EDTA 消化后, 用含 10% FBS 的 RPMI-1640 完全培养基调整成 5×10⁴ 个/ml 的单细胞悬液, 每瓶 4 ml 接种于 25 cm² 细胞培养瓶中, 37℃、5% CO₂ 的细胞培养箱中培养 24 h 后, 实验组分别加入含

*国家高技术研究发展计划(863)(2004AA2Z3170, 2005AA2Z3D40), 国家重点基础研究发展计划(2003CCA04300)和教育部重点项目(03007)资助.

**通讯联系人. Tel: 022-23542817, E-mail: yaozhi@tmu.cn

收稿日期: 2006-09-08, 接受日期: 2006-11-18

不同浓度 YSV 的完全培养基(深圳康哲药业有限公司生产, 批号: 200504), 使 YSV 的终浓度分别为 100 mg/L、200 mg/L, 阴性对照组加入等体积的完全培养基, 每个样本做 3 个平行对照. 培养 48 h 后终止细胞培养, 使用 TRI 试剂(MRC 公司, USA)按说明书提取总 RNA, 采用无 RNA 酶的 DNA 酶(Promega 公司)按照说明书处理总 RNA 以去除基因组 DNA.

1.2 实时定量反转录 PCR

取 2 μg 总 RNA 按照 M-MLV 逆转录酶(Promega 公司)说明书逆转为 cDNA, 并对 cDNA 原液进行 10 倍稀释作为实时定量 PCR 的 cDNA 工作液. PCR 通过 ABI 7500 Real-Time PCR System 和

7500 System Software(Applied Biosystems 公司)得以运行. 引物序列见表 1, 反应体系参照 Real time PCR Master Mix(TOYOBO, JAPAN)说明书, 简言之, 每个 PCR 反应管内加入 Syber Green mix 12.5 μl、上下游引物(2.5 μmol/L)各 2 μl、cDNA 工作液 5 μl 和高压灭菌去离子水 5.5 μl, 每个样本的 mRNA 检测均重复 4 次 PCR. 经摸索和优化 PCR 的反应条件最终设置如下: 95℃ 1 min (预变性)后进入循环; 95℃ 15 s(变性), 56℃ 15 s(退火), 72℃ 45 s(延伸), 重复 40 个循环. 设置在延伸阶段收集荧光信号. 经 7500 System Software 分析可得到样本各看家基因的循环阈值(*Ct* 值).

Table 1 Primer sequences for internal control genes

Gene	Forward primer (5'~3')	Reverse primer (5'~3')
RPL13A	CCTGGAGGAGAAGAGGAAAGAGA	TTGAGGACCTCTGTGTATTTGTCAA
UBC	ATTGCGGTCGCGAGTTCTTG	TGCCTTGACATTCTCGATGGT
EIF4A	TGGAGATTGAGTTCAAGGAGACC	TGACAAGTGGCTCCCATATAGT
B2M	TATCCAGCGTACTCCAAAGA	GACAAGTCTGAATGCTCCAC
GAPDH	CTCTCTGCTCCTCTGTTCG	TGGCAACAATATCCACTTTACC
ACTB	TTGCCGACAGGATGCAGAAGGA	AGGTGGACAGCGAGGCCAGGAT

1.3 geNorm 程序统计学分析

从 <http://medgen.ugent.be/~jvdesomp/genorm/> 网址下载 geNorm 程序, 在 Excel 中设定不同样本中某一 看家基因 *Ct* 值最小者的表达量为 1, 其他样本此看家基因的相对表达量则为 $2^{-\Delta C_t}$ (ΔC_t = 各样本 *Ct* 值-最小 *Ct* 值), 将这些数据导入 geNorm 程序. 利用该程序计算基因表达稳定度的平均值 *M*, 对看家基因的表达稳定度进行排序(*M* 值越小, 表达越稳定), 并通过看家基因标准化因子的配对差异分析($V_{n/n+1}$)来判定看家基因的最适数目.

2 结 果

基因表达能否准确定量与实验体系是否选用了合适的看家基因高度相关. 为此, Vandesompele 和 Katleen 等于 2002 年编写了 geNorm 这一用于微软 Excel 平台的 VBA 宏程序, 用于候选看家基因的稳定性排序及找出合适的看家基因. 在本研究中, RPL13A、UBC、EIF4A、B2M、GAPDH、ACTB 表达稳定度的平均值 *M* 依次为 0.168、0.153、0.219、0.304、0.233、0.169, 表达稳定度由高到低

排序依次为 UBC > RPL13A > ACTB > EIF4A > GAPDH > B2M (图 1).

geNorm 软件以 0.15 为默认取舍值, 即当 $V_{n/n+1} < 0.15$ 时, 说明没必要使用数量 $\geq n+1$ 的看家基因作为内参, 本实验 $V_{2/3} = 0.028$, 看家基因的最适数目为 2 个(图 2), 分别是表达相对最稳定的 RPL13A 和 UBC.

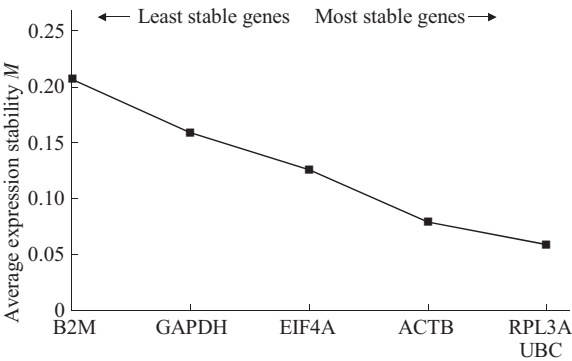


Fig. 1 Average expression stability values of remaining control genes

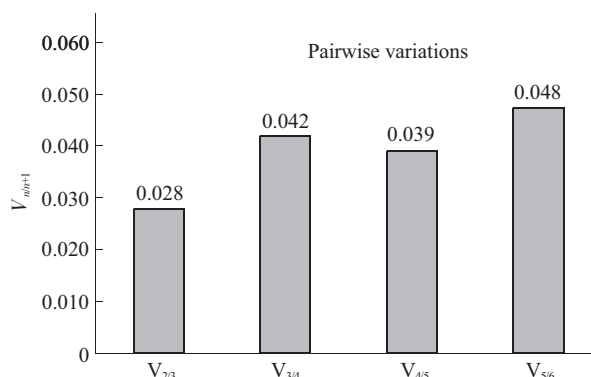


Fig. 2 Determination of the optimal number of control genes for normalization

3 讨 论

本研究选用 RPL13A、UBC、EIF4A、B2M、GAPDH 和 ACTB 6 个不同功能的常用看家基因作为候选内参. 对于基因表达 mRNA 水平的标准化策略, 除了最常用的看家基因, 很多研究者们提出可以用总 RNA 来进行标准化, 然而, 这种方法除了不能标准化逆转录及 PCR 反应等步骤引起的实验误差, 还有一个重要原因是该方法依赖于准确而可靠的 RNA 定量方法. 通常的 RNA 定量仪器分光光度计都存在灵敏度较低, 易受残存 DNA 影响, 以及低于一定样品浓度定量不准确的问题. 生化分析仪虽优于分光光度计, 但其价格十分昂贵, 很多实验室都没有配备, 并且总 RNA 作内参不能校正逆转录效率差异等引起的实验误差. 另外, 一些实验人员基于 rRNA 占总 RNA 量 80% 这一事实, 认为 rRNA 是 RNA 定量标准化的最适选择. 但是, rRNA 的调节独立于 mRNA, 前者由 RNA 聚合酶 I 合成, 后者由 RNA 聚合酶 II 合成^[7]. 越来越多的研究表明, rRNA 的转录易受到各种生物因素和药物的影响, 如在有丝分裂期间, 28 S、18 S rRNA 可明显减少或停止表达^[8~10]. 除此之外, rRNA 相对于目标基因常呈高丰度表达, 将使研究者在实时定量 PCR 数据分析时难以准确减去基线值. 选用 rRNA 以外的 RPL13A 等看家基因作为内参, 不仅简便, 而且因为目标基因与看家基因都经过逆转录、PCR 等相同步骤, 所以能对获得 PCR 结果所需的所有步骤都进行标准化.

放之四海而皆准的看家基因其实是不存在的, 找出合适的看家基因作为内参对于基因表达的准确定量至关重要, 然而, 没有标准和方法对内参基因进行挑选却一直是困扰广大科研工作者的一大

难题. geNorm 作为看家基因选择的良好工具, 其核心原理在于: 不同实验条件或不同细胞中, 2 个理想看家基因表达水平的比值在所有样本中应一致, 表达水平比值变异度的增加意味着待测单一或 2 个看家基因表达稳定性的下降. 该程序将某一看家基因与其他看家基因表达水平的两两比值经对数变换后, 计算其平均标准差作为基因表达稳定度的平均值 M , 对所有候选看家基因的表达稳定度进行排序, 并对标准化因子进行配对差异分析来判定所需看家基因的最适数目^[9]. 利用该程序, 研究者可以对任何组织或细胞的任意数量看家基因进行筛选, 择出 2 个以上而不是传统地使用单一看家基因作为内参, 将有助于校正系统偏差以得到更可靠的结果, 这对于基因的准确定量, 尤其是细微表达差异的生物学研究具有极其重要的意义.

在本研究中, geNorm 首先通过计算表达稳定度平均值(M)实现了 6 个看家基因表达稳定度的排序, 并通过标准化因子的配对差异分析($V_{n/n+1}$)挑选出最适的看家基因. 具体过程为: geNorm 程序先计算出 V_{56} 为 0.048, 由于其值小于 geNorm 的默认取舍值 0.15, 而使 B2M 作为 6 个看家基因中稳定性最差者(M 最大)被第一个排除. 接着 geNorm 程序对剩余的 5 个看家基因重新进行配对差异分析, 得出 $V_{45}=0.039(<0.15)$, 结果 GAPDH 作为 5 个看家基因中稳定性最差者继 B2M 第二个被排除. 基于相同原理, EIF4A 和 ACTB 亦被 geNorm 程序相继排除, 最终挑选出 RPL13A 和 UBC 可用于 YSV 对人肝癌 BEL-7402 细胞的研究.

本研究无论是 100 mg/L YSV 组还是 200 mg/L YSV 组, 其 B2M 或 GAPDH 的 mRNA 水平均低于阴性对照组, 我们对此现象进行了分析. B2M 全名为 Beta-2 微球蛋白, 存在于除红细胞和胎盘滋养层细胞以外的所有有核细胞, 在免疫细胞中起重要作用. 血清 B2M 浓度在原发性肝癌和肝硬化病人均高于正常人, 在原发性肝癌病人又高于肝硬化或肝脓肿病人, 因此血清 B2M 可作为甲胎蛋白以外肝癌的又一肿瘤标记物^[11]. GAPDH (3-磷酸甘油醛脱氢酶)是参与糖酵解的一种关键酶, Robert 等^[12]对 72 种人体组织的研究表明, GAPDH 的表达可能在组织或细胞的能量需求增高时上调. YSV 作用下这两者的波动是否因为其免疫抗肿瘤作用和抑制肿瘤细胞株导致后者能量需求下降, 仍需要通过实验进一步证实.

在运用 geNorm 的同时, 我们 also 需注意其应用

的基础或前提: 提取高质量的总 RNA; 对相同量的总 RNA 在相同条件下进行逆转录; 尽可能减少逆转录和 PCR 步骤中的实验误差等等. 因为毕竟 geNorm 分析的循环域值(C_t 值)直接受以上因素的影响, 控制这些因素将有助于使基因表达结果最大程度地反映实验因素对看家基因的调节情况, 从而尽可能地减少这些实验处理因素以外的误差对正确挑选看家基因的干扰.

无论是使用 RNA 印迹、核糖核酸酶保护分析、基因芯片等其他在 mRNA 水平对基因表达进行定量的技术, 还是使用蛋白质印迹等在蛋白质水平进行定量的技术, 都需要应用内参基因对目标基因表达量进行校正, 因此选择合适看家基因对于这些技术来说也十分重要. 从程序原理上分析, geNorm 不仅可以用于实时定量 PCR 转录分析中内参基因的选择, 也可用于以上基因表达定量方法内参基因的选择, 只是基因表达的相对量不再是用前面提到的 $2^{-\Delta\Delta C_t}$, 而是用荧光信号强度差或者光密度值差等, 但均可使用 geNorm 程序.

总之, 看家基因的选择对于基因表达的准确定量(尤其是新物质基因表达研究)至关重要. 研究者可通过使用 geNorm 程序, 从一系列不同功能的候选看家基因中选择出最适数目的看家基因用于目标基因的标准化, 从而达到准确定量的实验目的.

参 考 文 献

- 1 Suzuki T, Higgins P J, Crawford D R. Control selection for RNA quantitation. *Biotechniques*, 2000, **29** (2): 332~337
- 2 Liu D W, Chen S T, Liu H P. Choice of endogenous control for gene expression in nonsmall cell lung cancer. *Eur Respir J*, 2005, **26** (6): 1002~1008
- 3 Selvey S, Thompson E W, Matthaei K, *et al.* Beta-actin—an unsuitable internal control for RT-PCR. *Mol Cell Probes*, 2001, **15** (5): 307~311
- 4 Vandesompele J, De Preter K, Pattyn F, *et al.* Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol*, 2002, **3** (7): RESEARCH0034
- 5 Huggett J, Dheda K, Bustin S, *et al.* Real-time RT-PCR normalisation; strategies and considerations. *Genes Immun*, 2005, **6** (4): 279~284
- 6 Jia J, Lu R, Qiu S, *et al.* Preliminary investigation of the inhibitory effects of the tyroservaltide (YSV) tripeptide on human hepatocarcinoma BEL-7402. *Cancer Biol Ther*, 2005, **4** (9): 993~997
- 7 Tricarico C, Pinzani P, Bianchi S, *et al.* Quantitative real-time reverse transcription polymerase chain reaction: normalization to rRNA or single housekeeping genes is inappropriate for human tissue biopsies. *Anal Biochem*, 2002, **309** (2): 293~300
- 8 Spanakis E. Problems related to the interpretation of autoradiographic data on gene expression using common constitutive transcripts as controls. *Nucleic Acids Res*, 1993, **21** (16): 3809~3819
- 9 Johnson M L, Redmer D A, Reynolds L P. Quantification of lane-to-lane loading of poly(A) RNA using a biotinylated oligo(dT) probe and chemiluminescent detection. *Biotechniques*, 1995, **19** (5): 712~715
- 10 Warner J R. The economics of ribosome biosynthesis in yeast. *Trends Biochem Sci*, 1999, **24** (11): 437~440
- 11 王 凯, 袁孟彪, 王东盛, 等. β_2 -微球蛋白对原发性肝癌鉴别诊断的价值. *临床肝胆病杂志*, 2002, **18** (5): 286~287
Wang K, Yuan M B, Wang D S, *et al.* *Chinese Journal of Clinical Hepatology*, 2002, **18** (5): 286~287
- 12 Robert D. Barber, Dan W. Harmer, Robert A. Coleman, and Brian J. Clark. GAPDH as a housekeeping gene: analysis of GAPDH mRNA expression in a panel of 72 human tissues. *Physiol Genomics*, 2005, **21**: 389~395

Selection of Control Genes in Transcription Analysis of Gene Expression*

ZHANG Yan-Jun^{1,2}, ZHU Zhi-Feng¹, LU Rong^{1,3},
XU Qiong¹, SHI Lin-Xi¹, JIAN Xu¹, LIU Jun-Yan¹, YAO Zhi¹**

¹Department of Immunology, Tianjin Medical University, Tianjin 300070;

²Department of Biochemistry, Tianjin Medical University, Tianjin 300070;

³Shenzhen Kangzhe Pharmaceutical co., Shenzhen 518029)

Abstract At present, transcription analysis of gene expression commonly uses a single housekeeping gene as control for normalization. Through quantitative real-time PCR expression the levels of 6 housekeeping genes (including RPL13A, UBC, EIF4A, B2M, GAPDH, and ACTB) in BEL-7402 cell line were estimated under the role of a new tripeptide compound-YSV. Differences in expression levels were observed by analysis of geNorm program, and RPL13A, UBC were finally determined as suitable internal control genes. In conclusion, the necessity of choosing control genes was proved, and a good way was introduced to select control genes when experiments were handled by different empirical factors (especially under the effect of new materials).

Key words internal control genes, geNorm program, gene expression, quantitative real-time PCR

*This work was supported by grants from Hi-Tech Research and Development Program of China (2004AA2Z3170, 2005AA2Z3D40), National Basic Research Program of China (2003CCA04300) and Important Project of Ministry of Education of China (03007).

**Corresponding author. Tel: 86-22-23542817, E-mail: yaozhi@tmu.cn

Received: September 8, 2006 Accepted: November 18, 2006