Selection of suitable housekeeping genes for expression analysis in glioblastoma using quantitative RT-PCR

Valeria Valente*[1,2], Silvia A Teixeira[1], Luciano Neder[3], Oswaldo K Okamoto[4], Sueli M Oba-Shinjo[5], Suely KN Marie[5], Carlos A Scrideli[6], Maria L Paçó-Larson[2] and Carlos G Carlotti Jr[1]

[1] Department of Surgery and Anatomy, Faculty of Medicine, University of São Paulo, Av. dos Bandeirantes 3900, 140490-900, Ribeirão Preto, SP, Brazil; [2] Department of Cellular and Molecular Biology, Faculty of Medicine, University of São Paulo, Av. dos Bandeirantes 3900,140490-900, Ribeirão Preto, SP, Brazil; [3] Department of Pathology, Faculty of Medicine, University of São Paulo, Av. dos Bandeirantes 3900, 140490-900, Ribeirão Preto, SP, Brazil; [4] Department of Neurology and Neurosurgery, Federal University of São Paulo, R. Botucatu 740, 04023- 900, São Paulo, SP, Brazil; [5] Department of Neurology, School of Medicine, University of São Paulo, Av. Dr. Arnaldo 455, 01246903, São Paulo, SP, Brazil; [6] Department of Pediatrics, Faculty of Medicine, University of São Paulo, Av. dos Bandeirantes 3900, 140490-900, Ribeirão Preto, SP, Brazil

Deepali Mathur

Department of Functional Biology, University of Valencia, Valencia, Spain

Background

Until recently expression of housekeeping genes is presumed to remain constant in different cell types, across developmental stages and various experimental conditions hence they are employed as internal controls to normalize gene expression and protein content data. Many studies conducted recently on plant and animal samples have however, reported variable expression of constitutive genes during cellular development[1] and experimental conditions[2-5]. The reason for such variation may be that they are not exclusively involved in performing the basic function of the cell but they also contribute in other cellular processes[6,7]. The variation in their expression is therefore likely and it is advised to consider prior validation of the expression stability of these reference genes, a routine step in all laboratories to accurately normalize the target gene(s) expression. Normalization, a fundamental step is used to compare the target gene expression with the constitutively expressed housekeeping genes. Through this process the technical and biological variations that are introduced during sample preparation or any other step during the experiment are controlled.

In this report, the authors evaluated a panel of candidate reference genes for their expression stability in glioblastoma tissues using statistical tools namely Genorm[8], Normfinder[9] and equivalence test[10]. The hypothesis was to identify a set of housekeeping genes with minimal variability in glioblastoma tumors compared to non-neoplastic white matter. In addition, the significance of normalization by measuring the transcription levels of twelve target genes using different reference genes is demonstrated.

Study design

The study was intended to validate a panel of reference genes for their expression stability in glioblastoma (GBM) samples, the most common type of tumor in the central nervous system. These malignant gliomas are the most common and most aggressive type of glial tumors. Due to several complicating factors treatment of these patients is difficult. Thus, there is an urgent requirement to identify novel therapeutic targets for its therapy. In this context, many recent studies are conducted to explore high-throughput gene expression profiles using real time PCR in order to identify novel genes differentially expressed in GBM.

The authors first extracted mRNA from thirty glioblastoma tissues and nine non-neoplastic white matter tissue samples and subjected them to quantitative PCR. The $2-\Delta\Delta CT$ equation was applied to calculate the relative expression of tumor samples and non-neoplastic brain tissues. To evaluate stably expressed reference genes in this experimental set up, the authors chose seven frequently used housekeeping genes (HKGs) from literature. Candidate HKGs assessed were beta-actin (ACTB), glyceraldeyde-3-phosphate dehydrogenase (GAPDH), beta-glucuronidase (GUSB), hydroxymethylbilane synthase (HMBS), hypoxanthine guanine phosphoribosyl transferase 1 (HPRT1), TATA-box binding protein (TBP) and 18S ribosomal RNA (18S rRNA). The gene expression stability was assessed using publicly available softwares namely Genorm, Normfinder and equivalence test. Genorm program defines the gene stability as the average pairwise variation of a particular gene with all other control genes and ranks the genes according to their average expression stability denoted by M. The gene with minimum M value is considered to be highly stable whereas the gene with highest M value is least stable and can be excluded. An alternative program, Normfinder, ranks the candidate reference genes based on the combined estimates of both intra- and intergroup variations. The lower stability values indicate the more stable expressed candidate genes. The equivalence test estimates the significance of differences in the median expression values between tumor and normal tissues of each individual gene.

For comparison of housekeeping transcription levels, the cycle threshold (CT) values were plotted directly, assuming similar threshold for all genes assessed. The seven housekeeping genes analyzed showed a wide expression range, with CT values between 14 and 32. The CT is defined as the number of cycles needed for fluorescence to reach a specific threshold level of detection and is inversely correlated with the amount of RNA template present in the reaction. The CT values for all these genes showed normal distribution according to the Kolmogorov and Smirnov method, in both tumor and non-neoplastic samples. These genes were clearly distributed into different expression level categories. The extremely abundant 18S rRNA, which represents the bulk of total RNA in the cell, presented CT values below 18 cycles; genes coding for highly expressed mRNAs, such as ACTB and GAPDH, with majority of CT values between 18 and 22 cycles; and the moderately expressed genes, GUSB,

HMBS, HPRT1 and TBP, showing CT values between 26 and 30 cycles. Additionally, the authors found slightly higher levels in the abundance of GUSB and HPRT1 mRNAs than in HMBS and TBP mRNAs.

The authors reported that amongst seven frequently used housekeeping genes TBP and HPRT1 were adequate references for glioblastoma gene expression analysis and that expression levels of ACTB, GAPDH, GUSB and 18S rRNA, differed significantly between tumor and normal tissues on the basis of the examination of raw CT values. This indicates that these genes are regulated and cannot be utilized for target gene normalizations. It is recommended that this initial analysis is a crucial step in order to exclude highly unstable genes from further calculations. Based on data from a previously analyzed microarray, the authors selected 12 target genes signified as TG 1-12, candidates of being over-expressed in GBM, for quantitative RT-PCR experiments to investigate the correctness of overexpression suggestion. The authors found five genes whose expression levels significantly differ between tumor and normal samples. Surprisingly, two of them were decreased (TG4 and TG6 and three (TG1, TG8 and TG12) were increased in tumor tissue. Furthermore, Holliday Junction Recognizing Protein, a novel DNA repair protein over expressed in lung cancer, is found to extremely over-expressed in GBM samples. Evaluation of the expression levels of these genes utilizing different endogenous controls revealed that the normalization method applied might introduce errors in the estimation of relative quantities. Genes presenting expression levels which do not significantly differ between tumor and normal tissues can be considered either increased or decreased if unsuitable reference genes are applied. Most importantly, genes showing significant differences in expression levels between tumor and normal tissues can be missed indicating the significance of prior validation of candidate housekeeping genes for any experimental set up in a laboratory, especially when small differences are intended to be detected.

Implications

The authors found a fluctuated expression of *Gapdh*, *beta-actin and 18S rRNA* in glioblastoma samples which are frequently used as internal controls in majority of the laboratories. It is now demonstrated that expression of housekeeping genes, which are used as endogenous controls for normalizing mRNA transcripts, may not be necessarily stable in all cells/tissues under all conditions. A gene may show stable expression in one condition whereas unstable in another. Therefore it is indispensible to pre-validate the expression stability of reference genes to accurately normalize the gene expression data.

References

1. Al-BaderT MD, Al-Sarraf HA. Housekeeping gene expression during fetal brain development in the rat—validation by semi-quantitative RT-PCR. Developmental Brain Research 2005; 156: 38–45.
2. Deindl E, Boengler K, van RN, Schaper W. Differential expression of GAPDH and beta3-actin in growing collateral arteries. Mol Cell Biochem 2002; 236: 139–146.
3. Glare EM, Divjak M, Bailey MJ, Walters EH. beta-Actin and GAPDH housekeeping gene expression in asthmatic airways is variable and not suitable for normalising mRNA levels. Thorax 2002; 57: 765–770.
4. Hamalainen HK, Tubman JC, Vikman S, Kyrola T, Ylikoski E, et al. Identification and validation of endogenous reference genes for expression profiling of T helper cell differentiation by quantitative real-time RT-PCR. Anal Biochem 2001; 299: 63–70.
5. Radonic A, Thulke S, Mackay IM, Landt O, Siegert W, et al. Guideline to reference gene selection for quantitative realtime PCR. Biochem Biophys Res Commun 2004; 313: 856–862.
6. Ishitani R, Sunaga K, Hirano A, Saunders P, Katsube N, Chuang DM. Evidence that glyceraldehyde-3-phosphate dehydrogenase is involved in age-induced apoptosis in mature cerebellar neurons in culture. J Neurochem 1996; 66(3): 928-935.
7. Singh R, Green M. Sequence-specific binding of transfer RNA by glyceraldehyde-3-phosphate dehydrogenase. Science 1993; 259(5093): 365–368.
8. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, 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.
9. Andersen CL, Jensen JL, Orntoft 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. Cancer Res 2004; 64(15): 5245–5250.
10. Haller F, Kulle B, Schwager S, Gunawan B, von Heydebreck A, et al. Equivalence test in quantitative reverse transcription polymerase chain reaction: confirmation of reference genes suitable for normalization. Anal Biochem 2004; 335(1):1–9.