

Identification of Reference Genes in Chordoma Cells Allows Cross-Comparison of Expression Studies Across Subtypes

Didem SEVEN¹, Nehir KIZILILSOLEY², Emrah NIKEREL², Omer Faruk BAYRAK¹, Ugur TURE³

¹Yeditepe University, Faculty of Medicine, Department of Medical Genetics, Istanbul, Turkey

²Yeditepe University, Faculty of Engineering, Department of Genetics of Bioengineering, Istanbul, Turkey

³Yeditepe University, Faculty of Medicine, Department of Neurosurgery, Istanbul, Turkey

Corresponding author: Omer Faruk BAYRAK ✉ ofbayrak@yeditepe.edu.tr

ABSTRACT

AIM: To present the best housekeeping genes including clival/sacral based chordoma, and the nucleus pulposus cells.

MATERIAL and METHODS: We investigated 13 candidate reference genes in public chordoma array transcriptome datasets, validated these genes by using RT-PCR, and evaluated their stability with NormFinder, geNorm, and Bestkeeper.

RESULTS: YWHAZ, TBP and PGK1 genes were identified as the most stable reference genes as confirmed with three different approaches. Conversely, KRT8, KRT19 and GAPDH genes are less stable and not appropriate for use in chordoma research.

CONCLUSION: For normalization of RT-PCR experiments in gene profiling of chordoma, we recommend the use of the stable genes YWHAZ, TBP and PGK1.

KEYWORDS: Chordoma cell lines, Reference genes, YWHAZ, TBP, PGK1

INTRODUCTION

Chordoma is a rare, locally invasive and slow-growing tumor that arises from the remnants of the notochord (4), with the incidence rate being 0.08 per 100000. Slow growth of chordoma causes the resistance to radiotherapy and chemotherapy (24). Since the pathology of chordoma is not entirely defined, the elucidation of its molecular mechanism is crucial to e.g., point out the new targets for therapy. Such an approach calls typically for expression studies, for which selection of a reference gene is of key importance. Reference genes allow for normalization and comparison of expression datasets obtained from different subtypes, different techniques, as well as from different labs as these datasets may span several orders of magnitude.

A reference gene is typically defined as an internal control or housekeeping gene (HKG). Housekeeping genes are reported

(i) to be required for the maintenance of cellular functions and (ii) to be expressed in all cells of an organism (25). These genes are expected to have stable expression, e.g., not to alter significantly among tissues and experimental conditions (1,18). There are several, commonly employed reference genes used in RT-PCR studies such as Beta actin (ACTB), glyceraldehyde-3-phosphate dehydrogenase (GAPDH), hypoxanthine guanine phosphoribosyl transferase (HPRT1), some ribosomal proteins (e.g., RPS13, RPS29, RPL30) or the members of keratin family (e.g., KRT8 and KRT19). However, the expression of many of these genes do vary unacceptably (9,14), which is not surprising *per se*, as some HKGs may have functions in various cellular process, in addition to cell maintenance (21). All in all, the selection of HKGs for RT-PCR studies should be experimentally justified; however, to the best of our knowledge, no such justification is available for chordoma cell lines used in chordoma research.

Didem SEVEN  : 0000-0003-3406-5905
Nehir KIZILILSOLEY  : 0000-0002-8798-4561
Emrah NIKEREL  : 0000-0002-9157-8662

Omer Faruk BAYRAK  : 0000-0001-7562-6604
Ugur TURE  : 0000-0002-2499-7525

In this study, we evaluated candidate reference genes to evaluate their stability in chordoma cell lines. For this, we analyzed publicly available datasets GSE95084, GSE48779, GSE101866 and GSE86036 belonging to chordoma cell lines, listed stably expressed genes and selected 13 frequently used reference genes. Once the candidate reference genes were selected, we carried out an RT-PCR study in 7 chordoma cell lines and a healthy control, nucleus pulposus cells. The candidate reference genes are selected as: YWHAZ, TBP, PGK1, HPRT1, KRT19, KRT8, ACTB, GAPDH, B2M, RPS29, RPS13, RPL30 and their stabilities, e.g. their performances as reference genes, and chordoma biomarker TBXT gene were assessed by three different known statistical approaches: Normfinder, geNorm, and BestKeeper.

■ MATERIAL and METHODS

High Throughput Expression Data

Ethical approval is not required for this study as the data was obtained from a publicly available database (<https://www.ncbi.nlm.nih.gov/gds>).

Available data from the literature were explored using GENT2 toolbox, both for differential expression and for the discovery of stable genes, among normal and tumor tissues (16). The tools allowed searching for gene expression patterns across different normal and tumor tissues, obtained compiled from public gene expression datasets. The following datasets from different subtypes were selected: GSE95084 (clival), GSE48779 (sacral), GSE101866 (sacral) and GSE86036 (sacral), in such a way to take into account different chordoma subtypes.

Software for Analysis of Stability

Gene stability analyses were assessed based on RT-PCR measurements of candidate genes, via three known reference gene finding algorithms: NormFinder, geNorm and BestKeeper (2,17,22). Raw Ct values were used as input for each algorithm and the stability metrics for each gene for each algorithm were reported. While NormFinder and geNorm measure gene stability based on relative variance in Ct values for each gene (log transformed for geNorm), BestKeeper uses the significance of correlation between Ct values for the selected gene and geometric mean of all genes for each cell line. Additionally for geNorm, the so-called "V values" are also reported, pointing to the provisional number of selected candidate reference genes.

Cell Culture

Um-Chor, CH-22, JHC7, MugCC1, UCH1, UCH2 cell lines were provided from Chordoma Foundation (Durham, NC, USA), MugChor1 cell line was purchased from American Type Culture Collection (ATCC, Manassas, VA, USA) and Nucleus Pulposus cell line was acquired from AbmGood (AbmGood, Canada). The flasks were covered with % 1 of Gelatin before use. All cell lines have been checked and shown to be free of mycoplasma contamination. JHC7 cell line was cultured with DMEMF12 medium, Nucleus Pulposus cell is with PRigrow III medium (AbmGood, Canada) and the rest of chordoma

cell lines maintained with in a special medium of IMDM/RPMI (4:1). All cells' media contains %10 of FBS and %1 of Penicillin-Streptomycin. Cells were incubated at 37 °C in 5% CO₂. Trypsinization was applied when they reached %80 confluence and washed with PBS and collected for RNA isolation.

RNA isolation and cDNA synthesis

Total RNA was isolated from each cell line pellet by using TRIzol Reagent (Invitrogen Life Technologies, Germany) according to the manufacturer's protocol. RNA purity and yield were measured using a NanoDrop ND-1000 spectrophotometer. Total RNA (1 µg) was reverse-transcribed into cDNA in 20µL final volume, using the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Rotkreuz, Switzerland) with random primers and RNase inhibitor, according to the instructions.

RT-PCR Experiments

RT-PCR was performed by using a Step One Plus instrument (Applied Biosystem, USA) in triplicates. TaqMan PCR probes were used to detect genes of interest. The selected candidate genes were reported in Table I. Equal amount of cDNA (50ng) was added to the mixture of RT-PCR tube Amplification and real-time data acquisition were run using the following cycle conditions: 10 min at 95 °C, followed by 40 cycles of 15 s at 95 °C and 1 min at 60 °C.

■ RESULTS

Expression Profiles of the Candidate Reference Genes

To gain a better insight into the gene expression in chordoma cell lines with different subtypes, we first inspected the expression levels of different types of chordoma cell lines using publicly available transcriptome data (GSE95084, GSE48779, GSE101866 and GSE86036) in GEO database (7). In doing so, we focused on 13 candidate reference genes and considered (i) overall expression level for each candidate reference gene and (ii) their differential expression with respect to the different cell lines (Figure 1). Considering the stably expressed genes and literature information, we finalized the candidate gene list (Table I).

Genes with absolute log of fold change higher than 1 and Benjamini Hochberg adjusted p-value lower than 0.05 are considered to be differentially expressed. The ultimately stable genes then should have log fold change of 0 (that is, stably expressed or not changed) and a low p-value that would indicate statistical significance. The selected 13 genes were non-differentially expressed almost in each dataset, with the exceptions of KRT8 and TBXT genes in GSE95084, which is the dataset with the most similar experimental sample group to our study. YWHAZ, TBP, PGK1 and HPRT1 genes were stable in all datasets. Volcano plot of the datasets are shown in Figure 1.

Determination of Reference Gene Expression Stability

To detect the stable genes in chordoma cell lines UmChor, CH22, JHC7, MugCC1, UCH1, UCH2, MugChor1 and JHC7

Table I: Initially Selected Pool of Genes, Candidates for Reference

Gene Symbol	EnsemblID	Gene Name	TaqMan ID
T	ENSG00000164458	T-box transcription factor T	Hs00610080_m1
GAPDH	ENSG00000111640	Glyceraldehyde-3-phosphate dehydrogenase	Hs02786624_g1
ACTB	ENSG00000075624	Actin beta	Hs03023943_g1
TBP	ENSG00000112592	TATA-box binding protein	Hs00427620_m1
PGK1	ENSG00000102144	Phosphoglycerate kinase 1	Hs00943178_g1
KRT19	ENSG00000171345	Keratin 19	Hs00761767_s1
KRT8	ENSG00000170421	Keratin 8	Hs01595539_g1
B2M	ENSG00000166710	Beta-2-microglobulin	Hs00187842_m1
HPRT1	ENSG00000165704	Hypoxanthine phosphoribosyltransferase 1	Hs02800695_m1
YWHAZ	ENSG00000164924	Tyrosine3-monooxygenase/tryptophan 5-monooxygenase activation protein zeta	Hs01122445_g1
RPL30	ENSG00000156482	Ribosomal protein L30	Hs00265497_m1
RPS13	ENSG00000110700	Ribosomal protein S13	Hs06633284_g1
RPS29	ENSG00000213741	Ribosomal protein S29	Hs03004310_g1

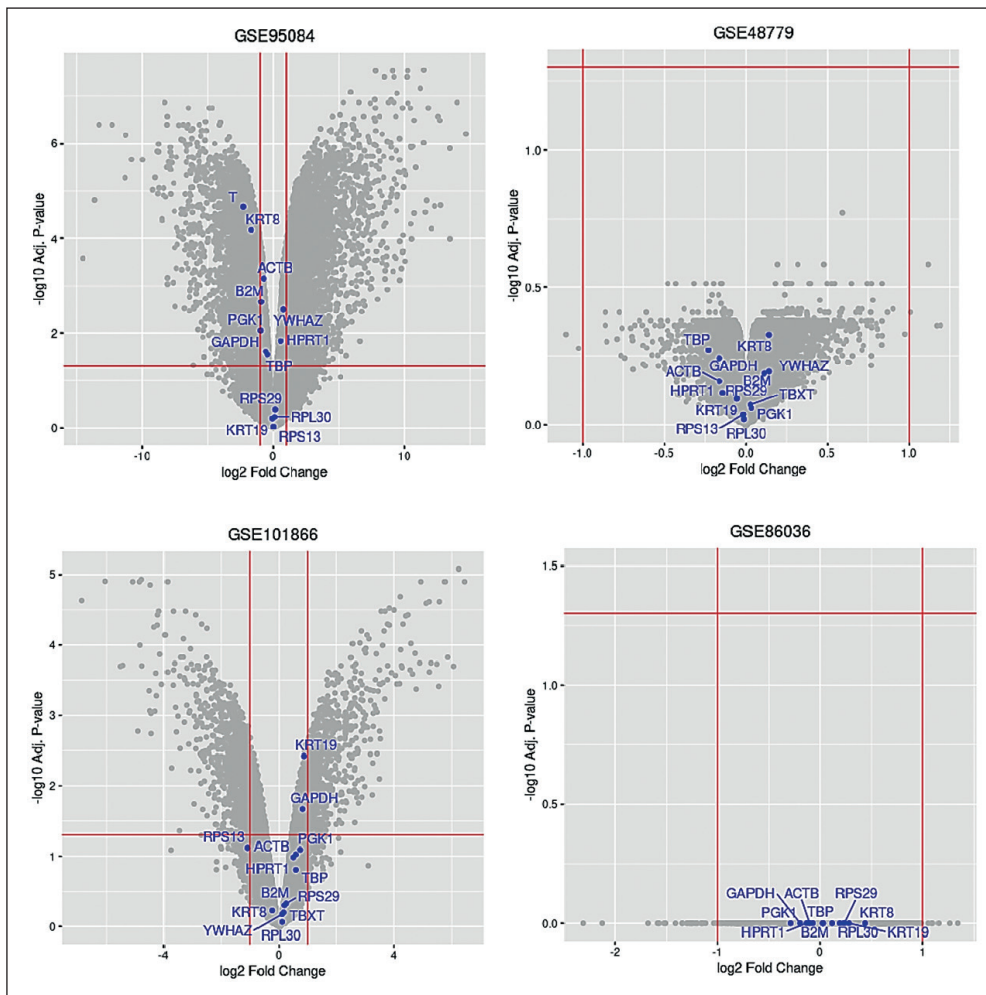


Figure 1: Volcano plot for the selected GEO Datasets, provided at the title of each plot. The selected candidate reference genes (Table I) are also shown in each subfigure. In each plot, the horizontal red line is the adjusted pValue threshold (pVal = 0.05) and vertical red lines are the thresholds for fold-changes ($\log_2|FC|=1$).

were used. Candidate reference genes YWHAZ, TBP, PGK1, HPRT1, KRT19, KRT8, ACTB, GAPDH, B2M, RPS29, RPS13, RPL30 and chordoma biomarker gene TBXT level were detected by using RT-PCR. The comparison for RT-PCR Ct values was performed between sacral versus clival chordoma cell lines. The specifications of the cell lines were given in Table II.

Table II: Specifications of the Chordoma Cell Lines

Cell lines	Disease Origin	Disease Status	Age (years)
JHC7	Sacral	Primary	61
MugChor	Sacral	Recurrent	57
UCH1	Sacral	Recurrent	56
CH22	Sacral	Recurrent	56
UmChor	Clival	Primary	66
UCH2	Sacral	Recurrent	72
MUGCC1	Clival	Primary	72

The stability measures were calculated using NormFinder, geNorm and BestKeeper algorithms and reported in Figure 2A, together with the V-values from the geNorm algorithm (Figure 2B). Overall, the stability values from different algorithms are in qualitative agreement among each other. The V-values additionally indicate that, from the initially selected gene pool, 5 genes can be further identified as reference genes for the current RT-PCR data from chordoma cell lines. We note that each algorithm points to a slightly different set of 5 candidate genes, so the final list is constructed by considering the consensus, i.e. genes selected by all three algorithms constitute the final reference gene list. NormFinder identifies YWHAZ, PGK1, TBP, RPL30 and HPRT1 as candidate genes, while geNorm identifies B2M instead of HPRT1 and BestKeeper GAPDH instead of RPL30. Further observation of Figure 2a notes a conflicting result among the algorithms for KRT19, GAPDH and RPS29. While geNorm and NormFinder identify them as unsuitable to be reference genes, BestKeeper points to significant correlation and renders them comparably stable. Finally, YWHAZ, PGK1 and TBP genes are found to be the most stable by all three algorithms.

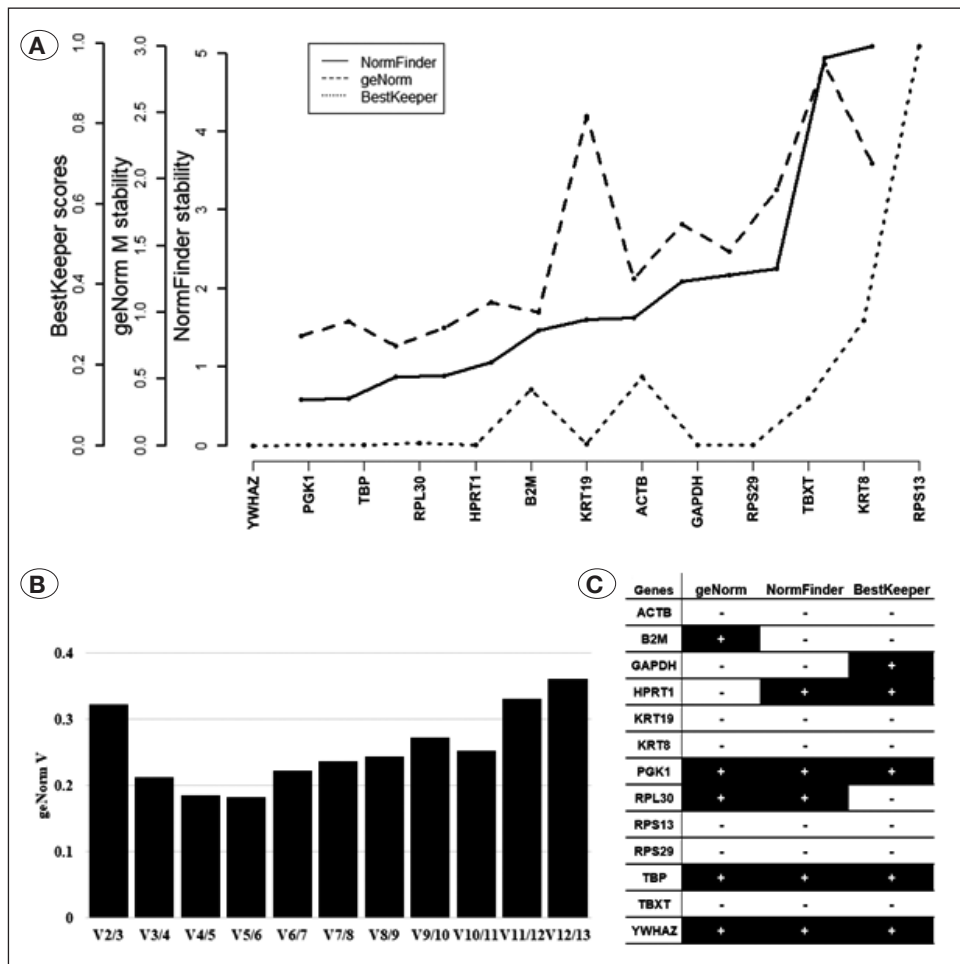


Figure 2: Stability analysis for determining the reference genes. **A)** Stability measures calculated by each algorithm. The x-axis represents the genes, sorted with respect to the NormFinder stability values. Leftmost genes are candidates as reference genes. **B)** geNorm V-values as indicator for the final number of reference genes from the selected gene pool. The final list is set to contain 5 most stable genes. **C)** Consensus table for each gene with each algorithm. The final list of candidate genes are formed by taking the genes identified as reference by three algorithms.

■ DISCUSSION

Chordoma is a rare bone tumor and its treatment is currently surgery and radiotherapy, yet chemotherapy is not an option (3). To create new chemotherapeutics for chordoma, it is crucial to understand the mechanism of both pathogenesis and therapeutics. In vitro tumor models are essential not only for searching for new substances with antitumor activity, but also for evaluating their efficacy (13). All in all, it is vital to have more information on chordoma disease models to understand the disease pathology (24). Gene expression analysis is key to investigate the underlying mechanisms of chordoma. Due to gene expressions differing among tissue types and experimental conditions, selecting stable reference genes is critical, allowing not only comparison within single studies, but also comparison across subtypes, platforms or labs.

The rare occurrence of chordoma patients and the difficulties in searching for new drugs often lead researchers to use cell lines to investigate chordoma biology. High throughput gene expression data from microarray or RNASeq based studies is useful for this purpose yet such data on chordoma is scarce. In addition, any large or moderate size study calls for data aggregating approaches e.g., across subtypes, which in turn requires stable reference genes. Similarly, gene expression analysis by low-throughput RT-PCR is yet another important detection quantification method, commonly used in cancer research as it is a gold standard technique allowing observation of relative changes in gene expression via quantification of mRNA (10, 23). However, there are several factors affecting the sensitivity and accuracy of the gene expression, like enzyme activity (e.g. Reverse transcriptase, DNA Polymerase), RNA integrity and quality, proper handling, designing appropriate primers- probes and variations between cells or tissue transcriptional activity. To take account of these variables, normalization is needed by using a reference gene to compare with the target gene (8). Taken collectively, the selection of the suitable reference gene set is vital as variation in HKG may cause misinterpretations in gene expression studies. The typical approach is to compile a list of reference genes, as opposed to relying on a single HKG (12). Several algorithms are available for validation and selection of the right reference gene sets taking into account variations across experiments as well as across replicates. The available ones are geNorm (22), NormFinder (2), which focus on within-sample variation, and Bestkeeper (17) which focuses on correlation with the average expression.

In this study, we determined the reference genes of both sacral and clival originated chordoma cell lines, and the nucleus pulposus cell line which is a remnant of the notochord and an appropriate healthy control for chordoma cell lines, by evaluating transcriptomic data and RT-PCR analysis. The putative reference genes that we investigated are commonly employed as internal controls, owing to their relatively steady expression and also approved by the results of high throughput data analysis. We focused particular attention to genes from various functional families. Through analysing 13 candidate reference genes by Normfinder, BestKeeper and GeNorm analysis, we detected the most suitable reference

genes that should be used in chordoma research.

In the present study, the initially selected 13 potential reference genes were investigated in four different public chordoma array datasets. It was shown that the genes are stable in four of the datasets with one exception that belonged to GSE95084 dataset that comprises three clival chordoma cell lines with differentially expressed brachyury (TBXT) and KRT8 genes. KRT8, KRT19 and TBXT genes are expressed in developing notochord (6,11), and different expression levels of these genes in notochord and chordoma cell lines were also previously shown (18). This was also in agreement with our RT-PCR results that showed that KRT8, and TBXT genes levels are not stable. Cytokeratins are frequently used as reference genes in most cancer research (19). However, our results showed that KRT8 and KRT19 genes are not suitable in chordoma studies as reference genes.

It has become clear that no single mRNA has a fixed expression level, and mRNA levels in reference genes can change quantitatively in response to a range of events, including developmental stage, cell cycle, and experimental circumstances, calling in turn for a set of reference genes. The number of reference genes identified by each of the three methods that we employed was 5. However, selection of the last two genes showed differences among the algorithms. The three algorithms produced contradictory results for KRT19, GAPDH, and RPS29. While geNorm and NormFinder exclude them out as reference genes, BestKeeper finds a substantial association and declares them comparable stable. This difference is due to the underlying idea of each algorithm. While geNorm and NormFinder deal with variance of Ct values over samples, BestKeeper deals with correlation (2,17,22). In our dataset, this might point to linear background bias: the expressions of these two genes change, yet this is measured to be in concert among cell lines.

In our study, the most stable genes were identified as YWHAZ, TBP and PGK1. YWHAZ, PGK1 genes and the more frequently used ACTB gene were previously stated as stable genes in chordoma fresh tissues. However, in our samples, the stability of ACTB is not among the top five genes. One of the reasons for the incompatibility can be the sample variation since the previous study includes only clival samples (20). Our work considers both subtypes and healthy cells covering a broader range in a biological context. Thus, we do not recommend the ACTB gene to other chordoma researchers when performing their gene expression analyses.

Another widely used reference gene glyceraldehyde-3-phosphate dehydrogenase (GAPDH) is an enzyme that catalyzes the sixth step of glycolysis. Variation in GAPDH levels in different tissue types were shown before (5). In our study GAPDH mRNA levels in chordoma and healthy cell line showed large variation and as a consequence of this, it was detected as a less stable gene for chordoma research.

YWHAZ (Tyrosine 3-Monooxygenase/Tryptophan 5-Monooxygenase Activation Protein Zeta) gene product is highly conserved and involved in many vital cellular processes (1). The most stable 3 genes in chordoma and healthy control nucleus

pulposus cell lines were YWHAZ, TBP, PGK1 in our study in compliance with BestKeeper, GeNorm and NormFinder. Our results are in accordance with research by Lopa et al. (2016) that pointed out that YWHAZ is a stable gene in intervertebral disc, endplate and articular cartilage cell lines (15), and skull base chordoma tissues. Taken together with sacral and clival chordoma cell lines and non-tumoral nucleus pulposus cells, the recommended reference genes for chordoma research are YWHAZ, TBP, PGK1 genes.

CONCLUSION

This work consists of a detailed analysis of gene expression data for focused list of selected reference gene set. For normalization in gene profiling studies of chordoma, we recommend the use of the stable genes YWHAZ, TBP and PGK1 genes, for both clival and sacral subtypes, and different experimental techniques.

Availability of data and material: The datasets created and analysed during the current study are available from the corresponding author upon reasonable request.

ACKNOWLEDGEMENT

We thank Dr. Andrew Harvey for the final proofreading of the manuscript.

The work is funded by Yeditepe University Research Fund.

AUTHORSHIP CONTRIBUTION

Study conception and design: DS

Data collection: DS

Analysis and interpretation of results: DS, NK, EN

Draft manuscript preparation: DS, EN

Other (study supervision, fundings, materials, etc...): OFB, UT

All authors (DS, NK, EN, OFB, UT) reviewed the results and approved the final version of the manuscript.

REFERENCES

- Aitken A, Howell S, Jones D, Madrazo J, Patel Y: 14-3-3 alpha and delta are the phosphorylated forms of raf-activating 14-3-3 beta and zeta. In vivo stoichiometric phosphorylation in brain at a Ser-Pro-Glu-Lys MOTIF. *J Biol Chem* 270(11):5706-5709, 1995
- 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. *Cancer Res* 64(15):5245-5250, 2004
- Bai J, Shi J, Li C, Wang S, Zhang T, Hua X, Zhu B, Koka H, Wu HH, Song L, Wang D, Wang M, Zhou W, Ballew BJ, Zhu B, Hicks B, Mirabello L, Parry DM, Zhai Y, Li M, Du J, Wang J, Zhang S, Liu Q, Zhao P, Gui S, Goldstein AM, Zhang Y, Yang XR: Whole genome sequencing of skull-base chordoma reveals genomic alterations associated with recurrence and chordoma-specific survival. *Nat Commun* 12(1):757, 2021
- Bakker SH, Jacobs WCH, Pondaag W, Gelderblom H, Nout RA, Dijkstra PDS, Peul WC, Vleggeert-Lankamp CLA: Chordoma: A systematic review of the epidemiology and clinical prognostic factors predicting progression-free and overall survival. *Eur Spine J* 27(12):3043-3058, 2018
- Barber RD, Harmer DW, Coleman RA, Clark BJ: GAPDH as a housekeeping gene: Analysis of GAPDH mRNA expression in a panel of 72 human tissues. *Physiol Genomics* 21(3):389-395, 2005
- Barresi V, Ieni A, Branca G, Tuccari G: Brachyury: A diagnostic marker for the differential diagnosis of chordoma and hemangioblastoma versus neoplastic histological mimickers. *Dis Markers* 2014:514753, 2014
- Barrett T, Suzek TO, Troup DB, Wilhite SE, Ngau WC, Ledoux P, Rudnev D, Lash AE, Fujibuchi W, Edgar R: NCBI GEO: Mining millions of expression profiles--database and tools. *Nucleic Acids Res* 33: D562-6, 2005
- Bustin SA, Benes V, Garson JA, Hellems J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, Vandesompele J, Wittwer CT: The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin Chem* 55(4):611-622, 2009
- de Kok JB, Roelofs RW, Giesendorf BA, Pennings JL, Waas ET, Feuth T, Swinkels DW, Span PN: Normalization of gene expression measurements in tumor tissues: Comparison of 13 endogenous control genes. *Lab Invest* 85(1):154-159, 2005
- Dheda K, Huggett JF, Bustin SA, Johnson MA, Rook G, Zumla A: Validation of housekeeping genes for normalizing RNA expression in real-time PCR. *Biotechniques* 37(1):112-114, 116, 118-119, 2004
- Götz W, Kasper M, Fischer G, Herken R: Intermediate filament typing of the human embryonic and fetal notochord. *Cell Tissue Res* 280(2):455-462, 1995
- Huggett J, Dheda K, Bustin S, Zumla: Real-time RT-PCR normalisation; strategies and considerations. *Genes Immun* 6(4):279-284, 2005
- Kitaeva KV, Rutland CS, Rizvanov AA and Solovyeva VV: Cell culture based in vitro test systems for anticancer drug screening. *Front Bioeng Biotechnol* 8:322, 2020
- Kozera B, Rapacz M: Reference genes in real-time PCR. *J Appl Genet* 54(4):391-406, 2013
- Lopa S, Ceriani C, Cecchinato R, Zagra L, Moretti M, Colombini A: Stability of housekeeping genes in human intervertebral disc, endplate and articular cartilage cells in multiple conditions for reliable transcriptional analysis. *Eur Cell Mater* 31:395-406, 2016
- Park SJ, Yoon BH, Kim SK, Kim SY: GENT2: An updated gene expression database for normal and tumor tissues. *BMC Med Genomics* 12(Suppl 5):101, 2019
- Pfaffl MW, Horgan GW, Dempfle L: Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. *Nucleic Acids Res* 30(9):e36, 2002
- Radonić A, Thulke S, Mackay IM, Landt O, Siegert W, Nitsche A: Guideline to reference gene selection for quantitative real-time PCR. *Biochem Biophys Res Commun* 313(4):856-862, 2004

19. Resutek L, Hsieh AH: The vacuolated morphology of chordoma cells is dependent on cytokeratin intermediate filaments. *J Cell Physiol* 234(4):3458-3468, 2019
20. Santegoeds RGC, Yakkioi Y, Jahanshahi A, Hoogland G, Temel Y, van Overbeeke JJ: Validation of reference genes in human chordoma. *Surg Neurol Int* 8:100, 2017
21. 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 Biotechnol* 75(2-3):291-295, 1999
22. 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. *Genome Biol* 3(7):00341-003411, 2002
23. VanGuilder HD, Vrana KE, Freeman WM: Twenty-five years of quantitative PCR for gene expression analysis. *Biotechniques* 44(5):619-626, 2008
24. Walcott BP, Nahed BV, Mohyeldin A, Coumans JV, Kahle KT, Ferreira MJ: Chordoma: Current concepts, management, and future directions. *Lancet Oncol* 13(2):e69-76, 2012
25. Zhu J, He F, Hu S, Yu J: On the nature of human housekeeping genes. *Trends Genet* 24(10):481-484, 2008