GAPDH as a housekeeping gene: analysis of GAPDH mRNA expression in a panel of 72 human tissues

Robert D. Barber, Dan W. Harmer, Robert A. Coleman, and Brian J. Clark

Pharmagene Laboratories Limited, Royston, Hertfordshire, United Kingdom

Submitted 15 January 2005; accepted in final form 8 March 2005

GAPDH as a housekeeping gene: analysis of GAPDH mRNA expression in a panel of 72 human tissues. Physiol Genomics 21: 389–395, 2005. First published March 15, 2005; doi: 10.1152/physiolgenomics.00025.2005.—Quantitative gene expression data are often normalized to the expression levels of control or so-called “housekeeping” genes. An inherent assumption in the use of housekeeping genes is that expression of the genes remains constant in the cells or tissues under investigation. Although exceptions to this assumption are well documented, housekeeping genes are of value in fully characterized systems. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) is one of the most commonly used housekeeping genes used in comparisons of gene expression data. To investigate the value of GAPDH as a housekeeping gene in human tissues, the expression of GAPDH mRNA was measured in a panel of 72 different pathologically normal human tissue types. Measurements were obtained from 371,088 multiplexed, quantitative real-time RT-PCRs with specific target genes. Significant differences in the expression levels of GAPDH mRNA were observed between tissue types and between donors of the same tissue. A 15-fold difference in GAPDH mRNA copy numbers was observed between the highest and lowest expressing tissue types, skeletal muscle and breast, respectively. No specific effect of either age or gender was observed on GAPDH mRNA expression. These data provide an extensive analysis of GAPDH mRNA expression in human tissues and confirm previous reports of the marked variability of GAPDH expression between tissue types. These data establish comparative levels of expression and provide standard values for levels of GAPDH mRNA expression in the tissues studied. These data confirm the variability of GAPDH mRNA expression previously indicated in published data and, as a result, may be used to define the variability of GAPDH expression between tissues. Furthermore, by providing normal copy numbers of GAPDH mRNA expression, these data can be used as a factor for the normalization of gene expression between tissues in experiments where GAPDH is being used as the experimental housekeeping control gene.

Materials and Methods

Rationale of the study. As part of an extensive analysis of genespecific RNA transcripts, the expression of GAPDH was measured as one of the reaction controls. The data provided in this study encompass the GAPDH “control” measurements from triplicate analyses of 1,718 genes in a panel of 72 human tissues.

Tissue acquisition. All tissues used in this study were obtained from ethically approved intermediaries. Tissue supply was governed by legal agreements and by stringent ethical review from local institutions.
research ethics committees. In addition, and in all cases, the informed consent of the donor or the donor’s next of kin was obtained for the use of the donated tissue for research.

Tissue was obtained either from living donors as a by-product/ consequence of surgical intervention or from postmortem donors. As soon as was practical, excised tissue was placed in ice-cold HEPS buffer and shipped to the laboratory on ice. Upon arrival (24 h/day, 7 days/wk) macroscopic dissection was performed, and the tissue was either snap frozen in liquid nitrogen or a small section of tissue was fixed in formalin. The snap-frozen tissue was stored at −80°C until RNA was harvested. The formalin-fixed tissue was paraffin wax embedded and, after being sliced on a microtome, was stained with hematoxylin and eosin. A certified pathologist examined the section and confirmed the origin and (if any) disease state of the tissue. Only tissue that was unaffected by a recognizable disease based on a combination of clinical history and pathological assessment was used in this study. Because it is known that GAPDH expression can be affected by disease or experimental manipulation, to our knowledge, none of the donors had any systemic illness that might have resulted in altered GAPDH expression in the donated tissues. For example, no tissue from a diabetic patient was included in these studies. In all cases, tissue used for RNA extraction was snap frozen no more than 24 h postexcision.

Purification of RNA and quality control. Total RNA was isolated from tissues using TRIzol (Invitrogen) according to the manufacturer’s instructions. Absorbance measurements at 260 nm in water were used to adjust the stock concentration of all RNA samples to 1 µg/µl, assuming an absorbance of 1 is equivalent to RNA at 40 µg/ml. Quality control standards were applied to all RNA samples in this study. Briefly, these were that the purity (A260:A280) was >1.7 and that an 18S rRNA band was present when samples were analyzed by agarose gel electrophoresis. This was a minimum first-pass analysis before more stringent criteria were applied. These were that samples must contain at least 10,000, 2,600, and 726 copies of GAPDH, the transferrin receptor, and β-actin RNA, respectively, per 100 ng total RNA when analyzed using Q-RT-PCR. Note that the primer set for GAPDH was designed to span the exon 1-exon 2 boundary, which restricted PCR amplification to cDNA templates only.

Reverse transcription. Before reverse transcription, all RNA samples (5 µg) were treated with DNase I (Invitrogen) according to the manufacturer’s instructions. Absorbance measurements at 260 nm in water were used to adjust the stock concentration of all RNA samples to 1 µg/µl, assuming an absorbance of 1 is equivalent to RNA at 40 µg/ml. Quality control standards were applied to all RNA samples in this study. Briefly, these were that the purity (A260:A280) was >1.7 and that an 18S rRNA band was present when samples were analyzed by agarose gel electrophoresis. This was a minimum first-pass analysis before more stringent criteria were applied. These were that samples must contain at least 10,000, 2,600, and 726 copies of GAPDH, the transferrin receptor, and β-actin RNA, respectively, per 100 ng total RNA when analyzed using Q-RT-PCR. Note that the primer set for GAPDH was designed to span the exon 1-exon 2 boundary, which restricted PCR amplification to cDNA templates only.

Q-RT-PCR. Q-RT-PCR was carried out using AmpliTaq Gold polymerase (ABI) in a standard PCR buffer containing cDNA prepared from 100 ng total RNA. Uracil N-glycosylase was included in all reactions to prevent cross-contamination and amplification of previous PCR products. Thermocycling conditions were 50°C × 2 min and 95°C × 2 min, followed by 40 cycles of 95°C for 15 s and 60°C for 30 s. Sequences of the primer-probe sets for GAPDH, β-actin, and the transferrin receptor are given in Supplemental Table S1 (the Supplemental Material for this article is available at the Physiological Genomics web site).1

Data acquisition. These data were obtained as part of a defined program of work to provide an expression map for 1,718 genes in the 72 tissues studied. Each Q-RT PCR experiment consisted of a multiplex reaction in which GAPDH was used as an internal control. Thus significant data concerning GAPDH expression can be extracted from these experiments. Data were accessed from the ABI software, and, in all experiments, the normalized reporter signal for GAPDH was <2. Threshold levels were established at either 0.04 or 0.03, and the cycle time (Ct) read from the interpolation of the threshold level and amplification curve. However, a number of reactions have been removed from the analysis: specifically, those that have Cts of <13, which occurred as a result of a noisy baseline, and those corresponding to Cts of >32.761. This is equivalent to 100 copies of GAPDH mRNA per 100 ng RNA but realistically corresponds to an absence of RNA and is a result of spurious amplification. With the use of these cutoffs to ensure data integrity, a total of 802 samples was not used in this analysis (17 samples with Cts <13 and 785 samples with Cts >32.761). This is equal to 0.22% of the total 371,088 data points.

RESULTS

Analysis of tissues and donors. The 72 human tissue types used in this study are listed in Supplemental Table S2. Within this study, 371,088 individual data points (72 tissues × 5,154 measurements) were analyzed. Data were derived from 1,595 individual samples (referred to by tissue type and donor), which were obtained from 618 donors. That there are fewer donors than samples indicates that individual donors have provided more than one tissue type. For example, from a whole heart, the left atria, left ventricle, and coronary artery may be obtained. Because of the differences in sizes of tissues received and of the level of difficulty of isolation of total RNA from different tissues, different numbers of donors were required for each tissue. Numbers of donors for each tissue type ranged from 10 to 66. The median number of donors per tissue type was 20, with 25.75 percentiles at 16 and 24 donors, respectively. These data are shown in Supplemental Fig. S1.

Generation of the GSC. Individual standard curves for 81 different genes were created from 150 different Q-RT PCR experiments in which genomic DNA was used as the template. All experiments were carried out on ABI Taqman 7700 machines, and thresholds of amplification were calculated from the linear portion of a plot of emitted fluorescence and Ct (Fig.

1 The Supplemental Material (Supplemental Figs. S1–S6 and Supplemental Tables S1 and S2) for this article is available online at http://physiolgenomics.physiology.org/cgi/content/full/00025.2005/DC1.
A. The data were pooled, and a linear relationship between log copy number (Cn) and the threshold cycle was derived (Fig. 1B). Linear regression analysis of the data yielded a slope of $-3.623$ and an ordinate intercept of $40.007$. By following strict rules of primer-probe design (see MATERIALS AND METHODS), the equation can be used to estimate the number of copies of a target transcript within a sample of total RNA. Specifically, the global standard curve (GSC) was used in these experiments to determine the number of copies of GAPDH mRNA that were present in a standard amount (100 ng) of total RNA in each reaction.

Measurement of GAPDH expression. Multiplex Q-RTPCR reactions were used to measure gene expression in the panel of 72 human tissues (Fig. 2). The Cts at which the amplification threshold was reached were calculated, and, with the use of the GSC, mean Cts were used to calculate the estimated copy numbers for GAPDH in each of the 72 tissues. Mean amplification Cts were calculated from the mean data for each tissue type per donor (Supplemental Fig. S2 and Supplemental Table S2), and the calculated copy numbers are given in Fig. 2. The spread of the data is narrow, and the SD for GAPDH mRNA expression in most tissue types is low (the mean SD for each of the 72 tissue types is $1.31 \pm 0.03$ Cts; Supplemental Fig. S2). This indicates that the overall variability between experiments is low. The range of expression is from $\sim 70,000$ copies (breast) to over $1,000,000$ copies (skeletal muscle) of GAPDH mRNA per 100 ng of total RNA. This represents an overall 15-fold variation between the tissue types studied and supports previously published variability in GAPDH mRNA expression between tissues.

Description of the data. Analysis of frequency-distribution histograms was carried out with the data from the Q-RTPCR experiments. Individual histograms with each data point for each tissue type were found to be significantly non-Gaussian and were skewed toward higher Ct values [significant values were obtained using the KS test, see Supplemental Fig. S3]. However, when the mean data for each sample from different donors were examined, normal distributions were observed for each tissue except for temporal cortex and mesenteric blood vessel (Supplemental Fig. S4).

Intertissue comparisons of GAPDH mRNA expression. The expression of GAPDH for each individual tissue was compared with that for every other tissue. It was found that the number of copies of GAPDH mRNA in each individual tissue type was significantly different from the other individual tissue types ($P < 0.05$ by ANOVA; Supplemental Fig. S3). Significant differences between tissue types were particularly evident in tissues with either very high or low Cts (see also Fig. 2).

Interdonor comparisons of GAPDH mRNA expression. It is known that there is significant biological variation between samples from different individuals. With the use of two tissues as examples, biological variation between donors can be illustrated. Mean amplification Cts derived from the 14 donors of the pyloric canal and from the 24 donors of the temporal cortex were analyzed (see Supplemental Fig. S4). For each tissue, the mean amplification Cts for samples from most donors are significantly different from each other ($P < 0.05$ by ANOVA).

Effects of age, gender, and postmortem delay. Factors that might have an effect on the expression of GAPDH in the tissues that have been analyzed are the age and sex of the donor and the time taken for the tissue to be processed. To investigate these factors, seven tissues were analyzed in further detail.
These are the mesenteric vessels, medulla oblongata, temporal cortex, umbilical cord, pyloric canal liver, and vas deferens. The levels of GAPDH mRNA expression (Ct) were assessed in relation to the age of the donor with the exception of the umbilical cord, as all donors are newborn. The data for each tissue were fitted with a linear regression (Fig. 3, A and B), for which the slope was calculated and compared with a flat line (i.e., where the slope = 0). For each of these tissues, with the exception of the vas deferens, none of the slopes was significantly different from zero (Fig. 3C), indicating that the age of the donors in this study had no effect on the expression of GAPDH mRNA in these tissues.

The second factor that may have affected the detection of GAPDH mRNA is the time between excision/cross-clamp and the snap freezing of the tissue. To evaluate this as a possibility, a comparison was made for the mesenteric vessels, liver, and vas deferens (Fig. 3D). These tissues are relatively readily available after surgical intervention (in contrast to tissues such as the medulla oblongata and temporal cortex, for which all donors in this study were postmortem). Statistical comparison of GAPDH mRNA expression in surgical and postmortem donors revealed no significant differences for mesenteric vessels and liver (P > 0.3 and P > 0.8 by ANOVA respectively). However, with the vas deferens, a significant difference was

---

**Fig. 3.** Effects of age, postmortem delay, and gender on GAPDH mRNA expression. **A:** linear regressions of the mean amplification Cts obtained for each donor against donor age. For each tissue, none of the slopes of the regressions was significantly different from zero. For the liver, dotted lines represent the 95% confidence interval of the regression. These data are not significantly different from zero but may be limited by the number of donors. **B:** linear regressions of the mean amplification Cts obtained for each donor against donor age in the vas deferens. The dotted line indicates the data obtained from all donors and the slope is significantly different from zero. When the postmortem donors are removed from the population, the slope is not significantly different from zero (solid line). **C:** table showing the equations of each of the linear regressions shown in A and B. The table also shows the 95% confidence interval of the slope, the error in the y-intercept, and the P value that each regression is not significantly different from zero. **D:** analysis of the effects of postmortem delay (P) on GAPDH mRNA expression compared with postsurgical delay (S). No difference is found except for the vas deferens. *P < 0.05. **E:** analysis of the effects of gender on GAPDH mRNA expression. No difference in expression was observed between the two genders.
found in the levels of GAPDH mRNA expression in surgical and postmortem donors (P < 0.01 by ANOVA). Consequently, in the vas deferens, together with the data above, it appears that donor age and/or postmortem delay may have affected the measurement of GAPDH mRNA expression.

There were four postmortem donors of the vas deferens with the remainder of the samples being donated after surgical contraception. Of these four postmortem donors, three were over 50 yr of age. Because tissues taken postmortem tended to be from older donors, it is difficult to distinguish the contributions of these two factors. To separate these two determinants, two approaches were taken. First, donors over 50 yr old (n = 7) were separated into two groups: surgical (n = 4) and postmortem (n = 3). When GAPDH mRNA expression was compared in these two groups, significantly lower expression (higher Ct) was observed in the postmortem group (P < 0.05 by ANOVA; Fig. 3D). Second, the surgical donors of all ages were separated from the postmortem donors. When the effects of age on the GAPDH mRNA expression in the vas deferens from surgical donors only was analyzed, the slope of the linear regression through the data was not significantly different from zero (Fig. 3B). These data imply specifically that postmortem delay affects the GAPDH mRNA expression in the vas deferens and that there is no evidence that the age of the donor has any influence on the levels of GAPDH mRNA expression.

The third factor that might have influenced the expression of GAPDH mRNA in each of these tissues was the gender of the donor. However, there was no significant difference in the levels of GAPDH mRNA between tissues from male and female donors in any tissue (P > 0.7 by ANOVA for all tissues except the medulla oblongata, for which P > 0.25; Fig. 3E). For obvious reasons, expression in the vas deferens could not be included in this analysis nor was the expression in the umbilical cord, as the gender of the infant had not been recorded.

**DISCUSSION**

This study presents data describing the expression of mRNA for GAPDH in an extensive panel of human tissues. GAPDH is one of the most common housekeeping genes and is often used to normalize gene expression data. The data presented here show that there are significant differences in the levels of GAPDH mRNA expression across the comprehensive panel of tissues studied, with a maximum difference between the highest and lowest expression of 15-fold. This is consistent with the data presented by Vandesompele et al. (13). The data provide no evidence that there is any effect of the age or gender of the donor and that, for most tissues, the influence of delay in processing surgical and postmortem tissues is negligible.

A total of 1,595 different tissue samples was used in this study. These were obtained from 618 donors, with some donors providing multiple tissue samples. The number of donors required for each tissue type varied due to the size of the sample donated and the ease with which RNA could be extracted from that sample. For example, tissues that are generally small, such as mesenteric vessels and the vas deferens, provide only small yields of RNA. For this reason, samples from 66 donors of mesenteric vessels were required, whereas only 10 donors were required for the stomach antrum and ileum. Other tissues from which it was more difficult to obtain high quality RNA were the adipose, breast, and umbilical cord, whereas the liver and some of the cardiovascular or respiratory tissues yielded substantial quantities of RNA, and thus fewer donors were required. It should be noted that each of the tissues that have been used in this study are made up of multiple cell types. Because both the availability of human tissue for research and the yields of RNA from microscopic dissections are low, this study only addressed each tissue as a whole and no attempt was made to study individual cell populations within any of the tissues.

To remove any variability caused by RNA isolation techniques and to ensure a consistent RNA quality, each RNA sample used in this study was subjected to and passed stringent quality assurance steps consistent with those of previous authors (10, 13). A review of the pitfalls of working with RNA is beyond the scope of the present report, but Thellin et al. (10) and manufacturers of RNA isolation kits and reagents provide adequate discussion.

One of the key quality control criteria set at the outset of the study was that tissues should yield at least 10,000 copies of GAPDH mRNA per 100 ng total RNA. Thus no data from tissues yielding <10,000 copies of GAPDH are included in the data presented in this study. With the benefit of hindsight, it may be that for tissues such as the breast, this number (10,000 copies) is too high, being only about sevenfold lower than the mean levels measured in this tissue. For others, such as skeletal muscle, it was probably too low, being some 100-fold lower than the mean copy number.

The data presented in this present study were derived from multiplexed Q-RTPCRs, in which GAPDH was included as a reaction control for the PCR stage. Transcript copy numbers were interpolated from a GSC, which was constructed with pooled data from 150 standard curves (for 81 different genes) prepared using genomic DNA as a PCR template. The slope of the GSC is −3.623, which is slightly different from the theoretical value of −3.322 for PCRs that are 100% efficient. This difference in slope indicates that the efficiency of the reaction as a whole is 89%, possibly due to a degree of inefficiency in primer annealing within the secondary structure of the genomic DNA.

The data demonstrate that there are significant differences in the levels of expression of GAPDH mRNA between many tissues. The copy numbers reported here are consistent with the relative intertissue expression levels described by Warrington et al. (14). Furthermore, the tissues appear in similar order when they are ranked by their levels of GAPDH mRNA expression. For example, in both studies, high levels of GAPDH mRNA have been found in the heart and brain and a relatively low level has been in the pancreas. These data are also consistent with the microarray data of Hsiao et al. (4), who demonstrated variability of GAPDH mRNA expression in 19 tissue types. While there was a considerable variation in GAPDH mRNA expression levels between different tissues, it was noted that there were similarities within groups of related tissues, such as the 1) stomach antrum, body, and fundus; 2) left atria and left ventricle; 3) kidney cortex, medulla, and pelvis; 4) duodenum, jejunum, ileum, and colon; 5) central nervous system tissues; or 6) blood vessels (Fig. 2).

Within the neurological tissues, there is a higher proportion of similar GAPDH copy numbers than in any other tissue grouping. The most obvious exceptions to this are the spinal...
cord and dorsal root ganglia. One explanation is that there is a difference between the cell bodies of the neurons in the brain compared with the peripheral neurons of the dorsal root ganglia and of the axons and interneurons and Schwann cells that make up the spinal cord.

These data can be compared with the data that have been published for GAPDH by the Cancer Genome Anatomy Project (CGAP) of the National Institutes of Health (8). From examination of the virtual Northern data of the expression pattern for cluster Hs.544577 (tag = TACCATCAAT), it is possible to see that both the expressed sequence tag and serial analysis of gene expression analyses indicate that there is substantial expression of GAPDH in a broad spectrum of tissues. The only major difference with the data in this study is that no expression is reported in the cervix by the CGAP. It is of note that high expression is reported in cervical cancers by the CGAP, and this is more akin with the data described here. Similarly, the tissues (muscle and brain regions) that were shown in this study to have high levels of GAPDH mRNA expression are the tissues with the highest representation in the data published by the CGAP.

The effects of age, gender, and possible differences in tissue processing were also investigated in this study. No differences in GAPDH mRNA expression could be determined between males and females. Similarly, no differences in GAPDH mRNA expression could be detected with donor age. Initially, it appeared that a difference in GAPDH mRNA expression with age may have been evident in the vas deferens. In reality, this difference was due to postmortem delay, and it was the age of the postmortem donors that distorted the conclusions in the vas deferens. By contrast, no other differences could be detected between surgical and postmortem donors in the other tissues analyzed.

Some of the variation in expression of GAPDH in different tissues may reflect the role of GAPDH in the cell. GAPDH is one of 10 enzymes that catalyze reactions in the glycolytic pathway. Specifically, GAPDH catalyzes the reversible oxidative phosphorylation of glyceraldehyde-3-phosphate via a thioester intermediate. The reaction catalyzed by GAPDH is allosterically regulated by nicotinamide adenine dinucleotide (NAD). Cellular levels of NAD+ must be replenished for glycolysis to occur either from the reduction of pyruvate to lactate in anaerobic conditions or by electron transfer from NADH to O2 through the electron transport chain. Consequently, it can be envisaged that rather than GAPDH expression being the same in all tissues, GAPDH expression may be higher in tissues with high energy demands. Thus it is not surprising that expression was found to be highest in the skeletal muscle, left ventricle, and many of the brain regions. This hypothesis is supported by evidence that in chickens at least, GAPDH expression levels are affected by nutrition status (7).

It has been demonstrated that not all of the frequency-distribution plots of the data are Gaussian. This includes the plot for the full data set, most of the plots describing the complete data set for each individual tissue type, and specifically those for the temporal cortex and mesenteric blood vessels. In neither of these latter cases is it likely that this is due to the use of insufficient donors, but it is possible that the source of this variation is in RNA quality or loading efficiencies, with RNA being readily biodegradable. However, there appears to be sufficient RNA to generate reproducible data with tight SDs (Supplemental Fig. S2; the mean SD for each of the 72 tissues is 1.31 ± 0.03 Cts), and the data and method of acquisition would appear to be valid, and, therefore, the most likely cause of the non-Gaussian distribution is simply an over- or underrepresentation of some samples from individual donors. Thus the data presented in this study describe the normal levels of GAPDH mRNA expression in 72 human tissues, and, in future experiments, measurement of GAPDH mRNA expression and comparison with these data will allow accurate normalizations to be carried out. The fact that different tissue types clearly exhibit different levels of GAPDH mRNA indicates that when GAPDH mRNA copy number is used as a quality control criterion, the cutoff level should be tailored to the tissue type.

Finally, it is apparent that the measurements of GAPDH levels within tissue types are reproducible and that determinations within donor replicates are robust. Therefore, these data suggest that in tissue samples from healthy donors at least, normalizations to GAPDH levels should at least be considered as a means of minimising intratissue sample variability in test mRNA expression.

In conclusion, the data presented here provide the most extensive collection of human tissue mRNA expression data for GAPDH, one of the most commonly used housekeeping genes. We have shown that within-tissue variation of GAPDH mRNA expression levels is generally small, whereas between-tissue variation can be substantial, according to tissue type. This provides further evidence that normalizations to a single gene across different tissue types are unwise. Furthermore, these data were obtained from tissues that were defined as pathologically normal. Because the variation observed between normal tissues of different types may in part be due to the different glycolytic demands of those tissues, these data suggest that comparisons within a tissue type between normal and diseased states are similarly unwise. The data provide information on normal levels of GAPDH mRNA in human tissues that can be regarded as standards for those tissue types, and, consequently, estimations of RNA quality can be made by comparison with these data. Finally, these standards may also be used to normalise the expression levels of a target gene in that specific tissue.

ACKNOWLEDGMENTS

The authors thank the members of Pharmagene’s Biomaterials Department past and present for RNA isolations and technical assistance. We are indebted to the members of the Molecular Expression Profiling-Genomics Department at Pharmagene for the technical expertise in generating the original data described in this study. We are grateful for the help of Pharmagene’s Informatics Group for extracting the data from the databases and to Dr G. A. Place and J. Carey for comments on the manuscript.

DISCLOSURES

R. D. Barber, D. W. Harmer, and B. J. Clark are salaried employees of Pharmagene Laboratories Limited and own shares in the company. R. A. Coleman was a joint founder of Pharmagene and remains a full-time employee; his full financial interests are reported in the company’s Annual Report.

REFERENCES


