Acta Biochim Biophys Sin 2014, 46: 330–337 | © The Author 2014. Published by ABBS Editorial Office in association with Oxford University Press on behalf of the Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences. DOI: 10.1093/abbs/gmt153.
Advance Access Publication 23 January 2014

Original Article

Identification of reference genes for qRT-PCR in human lung squamous-cell carcinoma by RNA-Seq

Cheng Zhan†, Yongxing Zhang†, Jun Ma, Lin Wang, Wei Jiang, Yu Shi*, and Qun Wang*

Department of Thoracic Surgery, Zhongshan Hospital, Fudan University, Shanghai 200031, China
†These authors contributed equally to this work.
*Correspondence address. Tel/Fax: +86-21-64041990-2915; E-mail: wang.qun@zs-hospital.sh.cn (Q.W.); shi.yu@zs-hospital.sh.cn (Y.S.)

Although the accuracy of quantitative real-time polymerase chain reaction (qRT-PCR) is highly dependent on the reliable reference genes, many commonly used reference genes are not stably expressed and as such are not suitable for quantification and normalization of qRT-PCR data. The aim of this study was to identify novel reliable reference genes in lung squamous-cell carcinoma. We used RNA sequencing (RNA-Seq) to survey the whole genome expression in 5 lung normal samples and 44 lung squamous-cell carcinoma samples. We evaluated the expression profiles of 15 commonly used reference genes and identified five additional candidate reference genes. To validate the RNA-Seq dataset, we used qRT-PCR to verify the expression levels of these 20 genes in a separate set of 100 pairs of normal lung tissue and lung squamous-cell carcinoma samples, and then analyzed these results using geNorm and NormFinder. With respect to 14 of the 15 common reference genes (*B2M*, *GAPDH*, *GUSB*, *HMBS*, *HPRT1*, *IPO8*, *PGK1*, *POLR2A*, *PPIA*, *RPLP0*, *TBP*, *TFRC*, *UBC*, and *YWHAZ*), the expression levels were either too low to be easily detected, or exhibited a high degree of variability either between lung normal and squamous-cell carcinoma samples, or even among samples of the same tissue type. In contrast, 1 of the 15 common reference genes (*ACTB*) and the 5 additional candidate reference genes (*EEF1A1*, *FAU*, *RPS9*, *RPS11*, and *RPS14*) were stably and constitutively expressed at high levels in all the samples tested. *ACTB*, *EEF1A1*, *FAU*, *RPS9*, *RPS11*, and *RPS14* are ideal reference genes for qRT-PCR analysis of lung squamous-cell carcinoma, while 14 commonly used qRT-PCR reference genes are less appropriate in this context.

Keywords reference gene; lung cancer; squamous-cell carcinoma; RNA sequencing

Received: November 7, 2013 Accepted: December 13, 2013

Introduction

Quantitative real-time polymerase chain reaction (qRT-PCR) is still the most common method to accurately quantify the mRNA transcription levels. qRT-PCR is fast, economical, and easy to use, and is especially suitable for the measurement of the expression of a limited number of genes. Given the highly quantitative nature of this technique, however, the use of a reliable normalization control is essential for meaningful comparison of the results between different samples. Accordingly, reference genes, usually housekeeping genes, are frequently used to normalize experimental deviation in the qRT-PCR datasets arising from differences in RNA quantity and quality, mRNA content, transcriptional activity, and operational deviation, among others. Since the accuracy of qRT-PCR is highly dependent on reliable reference genes, the expression levels of such genes should ideally be stably and constitutively elevated, and should not vary in different samples, nor be affected by experimental treatments. Since the expression patterns of many genes, including housekeeping genes, varies from tissue to tissue, proper validation of candidate reference genes is required prior to their use in qRT-PCR studies [1–3]. RNA sequencing (RNA-Seq) has in recent years emerged as a popular high-throughput technology for global gene expression analysis. In this technique, transcript levels are quantified in reads per kilobase of exon model per million mapped reads (RPKM), which reflects the molar concentration of a transcript normalized by the total read number in the measurement [4]. This normalization ideally avoids the experimental deviation in the qRT-PCR referred to above and facilitates transparent comparison of the expression levels of different genes between samples. As such, RNA-Seq is an ideal method to identify stably expressed reference genes across the entire genome.

Lung cancer is the most commonly diagnosed cancer worldwide, as well as the leading cause of cancer death [5]. Worldwide in 2008, it accounted for 13% (1.6 million) of the total cases and 18% (1.4 million) of the deaths [5]. Squamous-cell carcinoma is one of the major histological subtypes of lung cancer, and although a considerable number of qRT-PCR studies have focused on this subtype, they contain relatively few reliable reference genes. Although reference genes such as beta-actin (*ACTB*),

glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) are commonly used in this context, we found that the expression of certain reference genes in lung squamous-cell carcinoma is sufficiently variable to invalidate their use as internal controls. [6–10]. To solve this problem, here, we performed global gene expression profiling measurement in lung normal and squamous-cell carcinoma samples by using RNA-Seq, with the aim of identifying the most appropriate reference genes for the qRT-PCR analysis of lung squamous-cell carcinoma.

Materials and Methods

Ethics statement

This study was approved by Ethics Committee of Zhongshan Hospital, Fudan University (Approval No. 2011–219(2)) (Shanghai, China). All work conforms to the provisions of the Declaration of Helsinki. Written informed consent was obtained from all patients participating in this research at the time of hospitalization.

Tissue samples

Samples were obtained from patients with lung squamous-cell carcinoma who underwent surgical resection between July 2012 and December 2012 at Zhongshan Hospital, Fudan University. Normal lung specimens were resected at least 3 cm away from tumor margin, while tumor samples were carefully removed from the center of squamous-cell carcinoma. All samples were quickly frozen in liquid nitrogen after removal and then stored at −80°C. A portion of each sample was paraffin embedded, HE stained, and then checked by an experienced pathologist to ensure that no cancer cell existed in the normal tissues, and that >80% cells in every tumor sample were squamous carcinoma cells. Finally, 5 normal tissue samples and 44 lung squamous-cell carcinoma tissues were obtained for use in the RNA-Seq analysis. One hundred additional pairs of normal tissue samples and lung squamous-cell carcinoma tissues were used for qRT-PCR validation.

RNA preparation

Total RNA was extracted from each sample with Trizol (Invitrogen, Carlsbad, USA), then re-dissolved in diethypyrocarbonate-treated water, quantified using NanoVue Plus spectrophotometry (GE Healthcare, Fairfield, USA), and integrity-evaluated using agarose gel electrophoresis. DNA contamination was eliminated using gDNA Eraser (TaKaRa, Tokyo, Japan) according to the manufacturer's guidelines.

RNA-Seq

mRNA in total RNA was converted into a library of template molecules suitable for subsequent sequencing using a TruSeq®RNA Sample Preparation Kit v2 (Illumina, SanDiego, USA) according to the manufacturer's guidelines. These steps involved mRNA purification and fragmentation, first and second strand cDNA synthesis, end repair, 3′-end adenylation, ligation of adapters, and PCR amplification of cDNA libraries procedure, among others. Sequencing was then performed using Genome Analyzer II (Illumina) according to the manufacturer's recommendation. Sequence analysis was performed using the software Galaxy (http://galaxyproject.org) to calculate the RPKM of every transcript. RPKM values for all the transcripts of a given gene were summed to generate a measure of the expression of that gene. Each sample was sequenced twice and the average of the RPKM value of each gene was taken to reflect its actual expression level.

qRT-PCR analysis

cDNA synthesis was performed using the PrimeScript™ RT Master Mix (Perfect Real Time) (TaKaRa) according to the manufacturer's guidelines. qPCR reactions for each treatment were carried out in triplicate by the Mastercycler epgradient S realplex (Eppendorf, Hamburg, Germany) and SYBR® Premix Ex Taq™ (Tli RNaseH Plus II) (TaKaRa) using the following PCR procedure, 1 cycle of 5 s at 95°C, 40 cycles of 5 s at 95°C, 30 s at 55°C, 30 s at 72°C. Primers were designed according to sequences common to all transcripts of a given gene, avoided all single nucleotide polymorphisms (SNPs) or mutations reported in NCBI dbSNP database.

Statistical analysis

The RPKM data from RNA-Seq were analyzed using IBM SPSS for windows, version 20 (Armonk, USA). The mean RPKM value was used to evaluate the expression level of a gene while the CV [coefficient of variance, equal to standard deviation (SD) divided by mean] of RPKM was used to assess the stability of their expression.

Two different statistical algorithms, geNorm and NormFinder, both of which provide a stability value for each gene in a pool of genes were applied to analyze the qRT-PCR data [11,12]. The stability value was then used to validate the outcome of the RPKM data analysis.

Results

Expression profiles of commonly used reference genes in RNA-Seq data

Our criteria for appropriate reference genes were: constitutively elevated expression for easy detection; and stable, comparable expression levels in both normal lung tissues and lung squamous-cell carcinomas. A total of 15 commonly used reference genes were evaluated against these criteria, namely: *ACTB*, beta-2-microglobulin (*B2M*), *GAPDH*, beta-glucuronidase (*GUSB*), hydroxymethylbilane synthase (*HMBS*), hypoxanthine phosphoribosyltransferase 1 (*HPRT1*), importin

8 (*IPO8*), phosphoglycerate kinase 1 (*PGK1*), polymerase RNA II DNA directed polypeptide A 220 kDa (*POLR2A*), peptidylprolyl isomerase A (*PPIA*), ribosomal protein large P0 (*RPLP0*), TATA box binding protein (*TBP*), transferrin receptor (*TFRC*), ubiquitin C (*UBC*), and tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein zeta polypeptide (*YWHAZ*) (**Fig. 1** and **Table 1**) [13–21]. The full name, NCBI Gene ID, function, and location of these genes can be found in Supplementary Table S1.

As shown in **Fig. 1** and **Table 1**, several genes (*GUSB*, *HMBS*, *HPRT1*, *IPO8*, *POLR2A*, *TBP*, and *TFRC*) had very low RPKM values, and were expressed at such low levels that they would likely not be routinely detectable. Some common reference genes had widely varying expression levels between lung normal and squamous-cell carcinoma samples. These genes included *B2M*, *GAPDH*, *HMBS*, *HPRT1*, *PGK1*, *POLR2A*, *PPIA*, *TBP*, *TFRC*, *UBC*, and *YWHAZ*. *GAPDH*, for example, one of the most commonly used reference genes, had an average expression in lung squamous-cell carcinoma that was five folds more than that in normal lung tissues. And finally, the expression levels of several common reference genes (*B2M*, *GAPDH*, *HMBS*, *HPRT1*, *PGK1*, *TFRC*, and *YWHAZ*) varied considerably among samples in the same tissue type (normal or tumor). Of the 14 common reference genes tested, only *ACTB* met the criteria for elevated constitutive expression levels that were comparable between normal lung and squamous-cell carcinoma samples.

Identification of other candidate reference genes in RNA-Seq data

Currently, there is no existing method available, so we took the following steps to identify genes with consistent expression between normal lung and squamous-cell carcinoma tissues.

First, the genes with a <1 Mean(A) were excluded due to the possibility that RNA-Seq might be less robust at such low RPKM values [4]. Secondly, genes were excluded if Mean(N)/Mean(T) were <0.8 or >1.2, or if their expression levels were significantly different between normal and squamous-cell carcinoma samples ($P < 0.05$). Of the initial $>20{,}000$ genes, 3276 genes which met the criteria above were drawn in a scatter plot, with CV(A) as the *X*-axis and Mean(A) as the *Y*-axis (**Fig. 2**).

Here, Mean(A) denotes the mean of RPKM value of the gene in all samples. Mean(N) and Mean(T) denote the mean of RPKM value of the gene in lung normal samples and squamous-cell carcinoma samples, respectively. CV(A) denotes the CV of RPKM value of the gene in all the samples.

Finally, we selected six candidate reference genes from the top-left of **Fig. 2**: *ACTB*, eukaryotic translation elongation factor 1 alpha 1 (*EEF1A1*), Finkel-Biskis-Reilly murine sarcoma virus ubiquitously expressed (*FAU*), ribosomal protein S9 (*RPS9*), ribosomal protein S11 (*RPS11*), and ribosomal protein S14 (*RPS14*). The expression profiles of these six genes were displayed in **Fig. 3** and **Table 2**. The full name, NCBI Gene ID, function, and location of these genes can also be found in **Supplementary Table S1**. Our RNA-Seq data indicated that these six genes were consistently and robustly expressed, with the average RPKM values in excess of 1000, and the average CV values of no more than 0.3. These data indicated that these genes were the most highly and stably expressed genes in both lung normal and squamous-cell carcinoma samples. Interestingly, it is noteworthy that all six of these genes are considered 'housekeeping' genes [22].

Validation of candidate reference genes by qRT-PCR

Next, we used qRT-PCR to measure the expression levels of all 20 genes (the 15 common reference genes and the five novel candidate reference genes that we identified) in 100

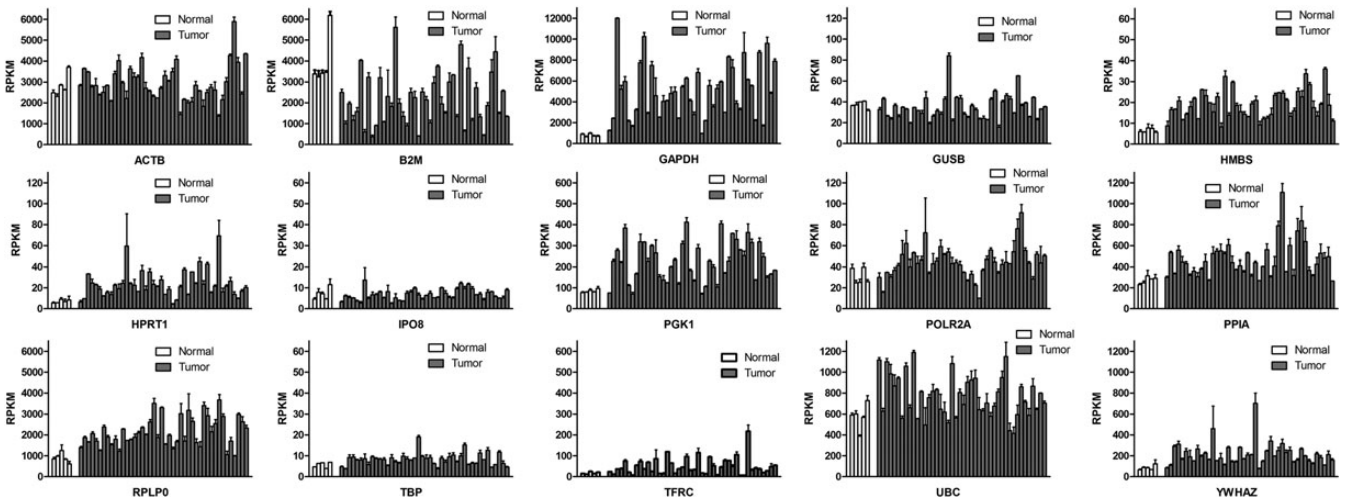

Figure 1. Expression profiling of 15 common reference genes in RNA-Seq data The expression of each gene in 5 lung normal samples and 44 lung squamous-cell carcinoma samples.

Table 1. The expression profile of 15 commonly used reference genes

Gene	Mean(N)	CV(N)	Mean(T)	CV(T)	Mean(N)/Mean(T)	Mean(A)	CV(A)
ACTB	2795	0.1951	2928	0.2948	0.9546	2915	0.2858
B2M	3941	0.3195	2140	0.5821	1.8416	2324	0.5815
GAPDH	808.4	0.1687	4988	0.5240	0.1621	4562	0.6105
GUSB	37.13	0.0951	34.46	0.3585	1.0775	34.73	0.3387
HMBS	6.582	0.1525	18.89	0.3541	0.3484	17.64	0.4180
HPRT1	7.062	0.2128	23.31	0.5536	0.3030	21.66	0.6093
IPO8	7.186	0.3933	6.719	0.3738	1.0695	6.766	0.3720
PGK1	84.79	0.0934	223.8	0.4212	0.3789	209.6	0.4716
POLR2A	30.94	0.2457	44.18	0.3338	0.7003	42.83	0.3432
PPIA	274.1	0.1238	466.0	0.3621	0.5882	446.4	0.3818
RPLP0	904.1	0.2579	2128	0.3241	0.4249	2003	0.3772
TBP	5.721	0.2285	8.262	0.3469	0.6925	8.003	0.3557
TFRC	17.65	0.2915	49.34	0.8142	0.3577	46.10	0.8517
UBC	575.7	0.2117	769.1	0.2591	0.7485	749.4	0.2680
YWHAZ	85.91	0.2699	213.8	0.4971	0.4018	200.7	0.5386

(N), (T), and (A) denote lung normal samples, squamous-cell carcinoma sample, and all samples, respectively. *ACTB* expressed highly and stably with high Mean(A), close to 1 Mean(N)/Mean(T), and low CV(A). Other 14 genes have either low Mean(A), Mean(N)/Mean(T) deviated to 1, or large CV(A) values, which means they were not highly and stably expressed in these samples.

pairs of normal lung and squamous-cell carcinoma samples. The primer sequences and melting temperature (T_m) used in our research and the product length, PCR efficiency, and correlation with dilution series (R^2) can be found in **Supplementary Table S2**. qRT-PCR data were then analyzed by both geNorm and NormFinder to evaluate the stability of these genes [11,12].

As shown in **Fig. 4**, the raw cycle threshold (CT) values of *B2M*, *GAPDH*, *HMBS*, *HPRT1*, *PGK1*, *RPLP0*, *TFRC*, and *YWHAZ* varied considerably between lung normal tissues and squamous-cell carcinoma samples, to an extent greater than that which could be explained by experimental errors. In contrast, the CT values of *ACTB* and the five novel candidate reference genes were sufficiently low to allow for routine detection. Moreover, the CT values of these six genes were comparable between normal lung and squamous-cell carcinoma tissues. Collectively, the results of qRT-PCR analysis were in broad agreement with those in the RNA-Seq dataset.

As shown in **Table 3**, analysis of the raw CT data using geNorm and NormFinder, indicated that *ACTB* and the five novel candidate reference genes were stable with very high rank of stability values. Of these six genes, *EEF1A1* was ranked by both analyses as the most stably expressed gene. In summary, the results of these analyses broadly validated the results of RNA-Seq analysis.

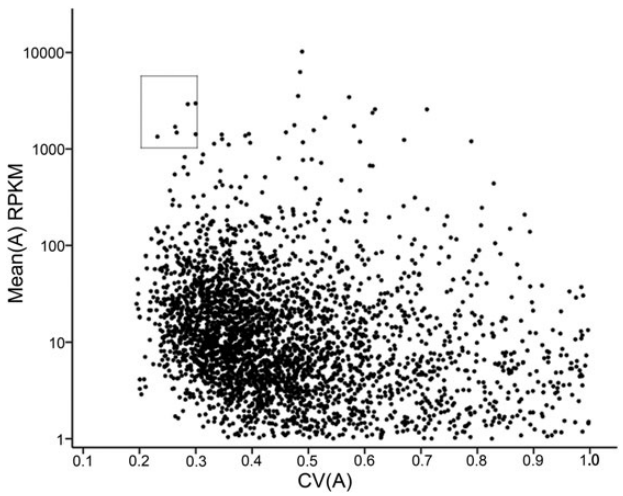

Figure 2. The distribution of Mean(A) and CV(A) of genes in RNA-Seq data Genes with CV(A) > 1.0 were not displayed. These six genes at the frame at the top-left of the figure (*ACTB*, *EEF1A1*, *FAU*, *RPS9*, *RPS11*, and *RPS14*) were considered the most highly and stably expressed genes.

Discussion

In this study, we carried out RNA-Seq global expression profiling of normal lung and squamous-cell carcinoma to identify candidate reference genes for qRT-PCR. To our knowledge, this is the first description of the use of RNA-Seq to identify candidate reference genes for qRT-PCR. We found that *ACTB*, *EEF1A1*, *FAU*, *RPS9*, *RPS11*, and *RPS14*, were highly and stably expressed in both tissue types and, as such, are suitable reference genes for qRT-PCR-based studies of lung squamous-cell carcinoma. In contrast, 14 commonly used reference genes (*B2M*, *GAPDH*, *GUSB*, *HMBS*, *HPRT1*, *IPO8*, *PGK1*, *POLR2A*, *PPIA*, *RPLP0*, *TBP*, *TFRC*, *UBC*, and *YWHAZ*) failed to meet our selection criteria and should

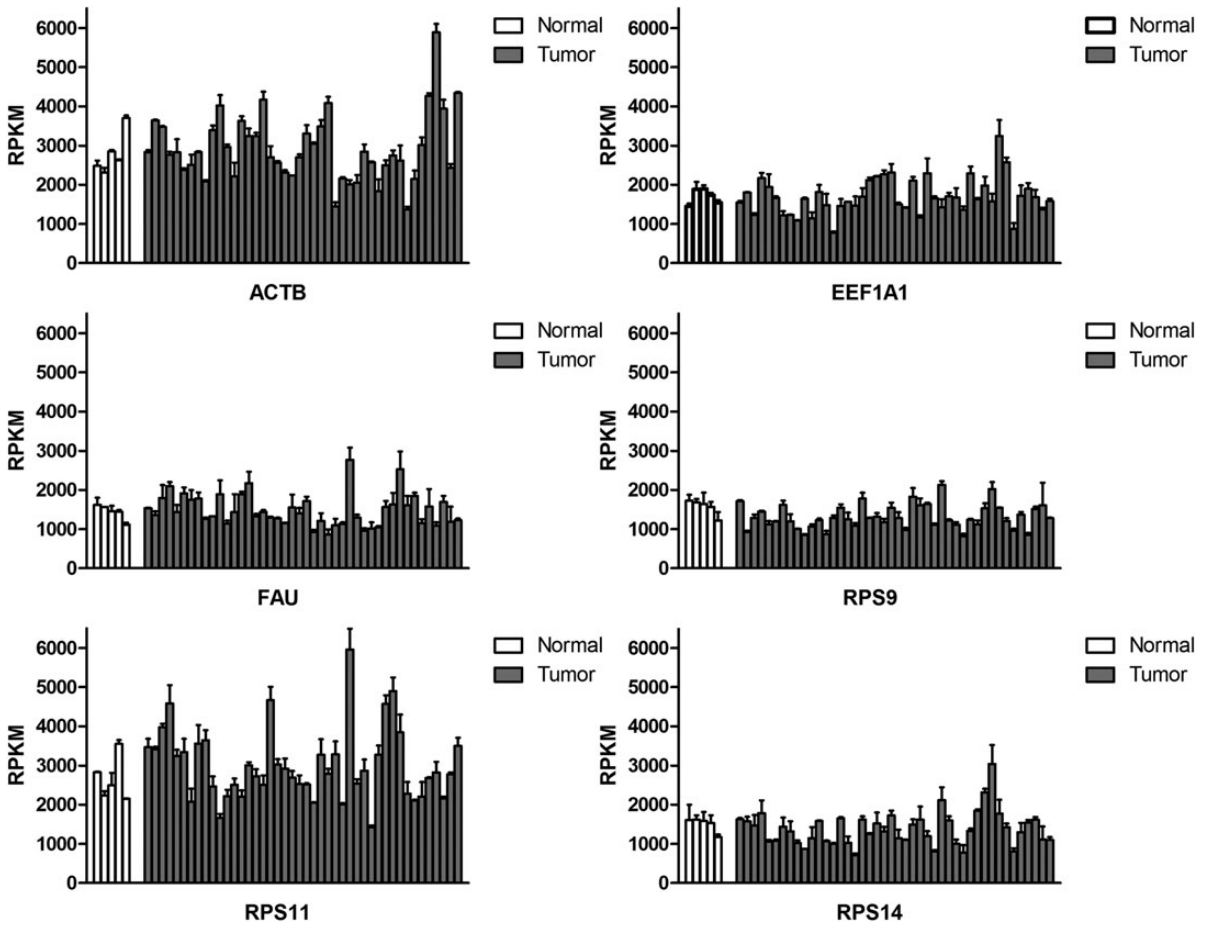

Figure 3. The expression profiling of six candidate reference genes in RNA-Seq data The expression of all 6 candidate genes was elevated and did not change significantly between all 5 normal and 44 squamous-cell carcinoma samples.

not be considered as suitable reference genes. We considered that our results were vitally important for those researchers investigating gene expression in lung squamous-cell carcinoma.

Another commonly used reference gene, *18S rRNA*, its expression data were not included in the results of the RNA-Seq since it could not be enriched at the mRNA purification step. So whether or not 1*8S rRNA* is suitable as endogenous control for qRT-PCR of lung squamous-cell carcinoma needs to be further verified.

Genes encoding ribosome proteins are universally conserved, and most of these genes were stably expressed in our samples. Three of our six successfully screened genes belong to this family, and many genes in this set have been verified in other studies to be suitable reference genes [23–25]. We also found that the expression levels of certain ribosome protein-encoding genes, such as ribosomal protein S7 (*RPS7*), ribosomal protein S10 (*RPS10*), *RPLP0*, and ribosomal protein L21 (*RPL21*) varied widely, ranging from single digits to tens of thousands of RPKM value. *GAPDH*, one of the most frequently used reference genes, has been widely used as a qRT-PCR reference gene in the expression analysis of lung squamous-cell carcinoma and other tumors [6,26–28]. However, tumor cells show various metabolic anomalies, in which the best known are the typical high rate of glycolysis and lactate production termed 'Warburg effect' [29]. How tumor cells establish this altered metabolic phenotype is not entirely clear, but here our RNA-Seq results showed that the expression levels of many enzymes involving in glucose metabolism's expression elevated many folds, such as *GAPDH*, *PGK1*, hexokinase 2 (*HK2*), the M2 type of pyruvate kinase (*PKM2*), and lactate dehydrogenase (*LDH*), echoed lots of previous research [29–35]. So our results indicated that *GAPDH* is not entirely suitable as the reference gene of qRT-PCR in lung squamous-cell carcinoma or even other kinds of cancers. Moreover, the expression of *GAPDH* has been reported to be upregulated in many other disorders such as inflammation, diabetes, hypoxia, and some respiratory diseases, suggesting that it should only be cautiously used as a reference gene in studies of these diseases [36–40]. Indeed, in a large-scale microarray meta-analysis, *GAPDH* was not recommended as a reference gene, except in heart or muscle [23].

geNorm and NormFinder are two robust strategies to identify the most suitable reference gene from a given set of genes, and have been frequently used in reference gene-finding analyses [11,12,17,41–44]. geNorm calculates a

Table 2. The expression profile of six candidate reference genes

Gene	Mean(N)	CV(N)	Mean(T)	CV(T)	Mean(N)/Mean(T)	Mean(A)	CV(A)
ACTB	2795	0.1951	2928	0.2948	0.9546	2915	0.2858
EEF1A1	1701	0.1180	1695	0.2759	1.0035	1695	0.2633
FAU	1440	0.1375	1485	0.2771	0.9697	1481	0.2660
RPS9	1567	0.1286	1313	0.2371	1.1935	1339	0.2317
RPS11	2653	0.2141	3009	0.3050	0.8817	2972	0.2996
RPS14	1503	0.1244	1414	0.3129	1.0629	1423	0.2998

The expression of *EEF1A1*, *FAU*, *RPS9*, *RPS11*, and *RPS14* were both highly and stably like *ACTB*.

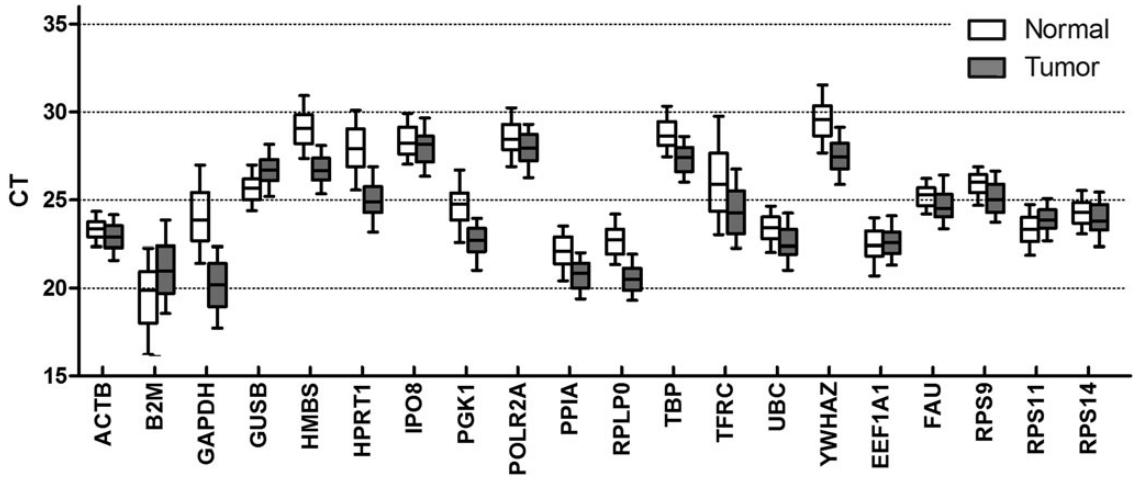

Figure 4. Raw qRT-PCR CT data for 20 genes *ACTB*, *EEF1A1*, *FAU*, *RPS9*, *RPS11*, and *RPS14* were highly and stably expressed with low and narrow range of raw CT values.

gene's stability measured as the SD of the log2-transformed expression ratios of each gene compared with all others tested throughout the samples [12]. NormFinder uses a model-based approach to estimate expression stability based on intra- and inter-group variations for candidate genes [11]. In both the strategies, the stability value of one gene is largely determined by its correlation with other candidate genes, such that a low stability value will be obtained if the expression of the given gene is highly correlated with most other genes. And when most of the genes in the given set express stably in all sample, both methods will efficiently select the most stably expressed gene. In the event that most genes are not stably expressed, however, then neither geNorm nor NormFinder would be in a position to evaluate stability. This is a possible explanation of why certain genes such as *PPIA*, *RPLP0*, and *TBP*, which were not stable in the RNA-Seq and raw CT datasets, ranked high with these programs, while *ACTB* was ranked in the middle position. Moreover, it might also explain the differences between RNA-Seq and qRT-PCR in identifying the best reference genes in lung squamous-cell carcinoma [20,21]. Based on this point, research based primarily on geNorm or NormFinder may need further verification.

Squamous-cell carcinoma and adenocarcinoma are the most two common histological subtypes of lung non-small-cell lung cancer. However, these two subtypes are quite different in host susceptibility, clonal evolution, molecular evolution, and molecular profiling [45]. Based on these caveats, in the absence of validation, reference genes which are suitable in lung squamous-cell carcinoma may not necessarily be directly applicable to other types of lung cancer.

Like RNA-Seq, microarray is capable of assaying genome-wide transcript expression levels in a single sample, and is frequently used in the screening of reference genes [23,46,47]. However, since microarray-based gene expression measurements are derived from nucleotide hybridization, which is prone to false signals arising from cross-hybridization, such data often require validation [48,49]. Another inherent limitation of microarray is that they are highly reliant upon existing knowledge about genome sequence, such that newly discovered genes or gene variants are often not available on a given platform [50].

Table 3. The geNorm and NormFinder analyses of the qRT-PCR results

Gene	geNorm value	NormFinder value	geNorm rank	NormFinder rank
ACTB	0.761	0.340	10	10
B2M	1.231	0.767	19	19
GAPDH	1.09	0.650	18	18
GUSB	0.767	0.348	11	12
HMBS	0.805	0.375	13	13
HPRT1	0.923	0.491	17	17
IPO8	0.827	0.397	14	14
PGK1	0.844	0.415	16	16
POLR2A	0.778	0.346	12	11
PPIA	0.722	0.277	3	2
RPLP0	0.743	0.307	9	6
TBP	0.743	0.302	8	4
TFRC	1.264	0.785	20	20
UBC	0.741	0.303	7	5
YWHAZ	0.831	0.397	15	15
EEF1A1	0.702	0.265	1	1
FAU	0.715	0.294	2	3
RPS9	0.727	0.321	4	8
RPS11	0.734	0.317	5	7
RPS14	0.738	0.334	6	9

ACTB, *EEF1A1*, *FAU*, *RPS9*, *RPS11*, and *RPS14* were considered stable with very low stability values. *EEF1A1* was ranked as the first most stable gene in both analyses.

RNA-Seq has recently emerged as a complementary technique to microarray, bringing advantages such as low background noise, capable of detecting novel transcripts and alternative splicing forms, exon and nucleotide resolution, high dynamic range in detection, high precision, and high reproducibility [51–54].

In conclusion, *ACTB*, *EEF1A1*, *FAU*, *RPS9*, *RPS11*, and *RPS14* are ideal reference genes in qRT-PCR research of lung squamous-cell carcinoma, while 14 commonly used reference genes, *B2M*, *GAPDH*, *GUSB*, *HMBS*, *HPRT1*, *IPO8*, *PGK1*, *POLR2A*, *PPIA*, *RPLP0*, *TBP*, *TFRC*, *UBC*, and *YWHAZ*, are less appropriate as reference genes in this context.

Supplementary Data

Supplementary data are available at *ABBS* online.

Funding

This work was supported by a grant from the Doctoral Program Foundation of Institutions of Higher Education of China (20110071120066).

References

1. Lin J and Redies C. Histological evidence: housekeeping genes beta-actin and GAPDH are of limited value for normalization of gene expression. Dev Genes Evol 2012, 222: 369–376.
2. Bustin SA. Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems. J Mol Endocrinol 2002, 29: 23–39.
3. Bustin SA and Nolan T. Pitfalls of quantitative real-time reverse-transcription polymerase chain reaction. J Biomol Tech 2004, 15: 155–166.
4. Mortazavi A, Williams BA, McCue K, Schaeffer L and Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods 2008, 5: 621–628.
5. Jemal A, Bray F, Center MM, Ferlay J, Ward E and Forman D. Global cancer statistics. CA Cancer J Clin 2011, 61: 69–90.
6. Ito S, Kawano Y, Katakura H, Takenaka K, Adachi M, Sasaki M and Shimizu K, *et al*. Expression of MAGE-D4, a novel MAGE family antigen, is correlated with tumor-cell proliferation of non-small cell lung cancer. Lung Cancer-J IASLC 2006, 51: 79–88.
7. Senchenko VN, Anedchenko EA, Kondratieva TT, Krasnov GS, Dmitriev AA, Zabarovska VI and Pavlova TV, *et al*. Simultaneous down-regulation of tumor suppressor genes RBSP3/CTDSPL, NPRL2/G21 and RASSF1A in primary non-small cell lung cancer. BMC Cancer 2010, 10: 75.
8. Skrzypski M, Dziadziuszko R, Jassem E, Szymanowska-Narloch A, Gulida G, Rzepko R and Biernat W, *et al*. Main histologic types of non-small-cell lung cancer differ in expression of prognosis-related genes. Clin Lung Cancer 2013, 14: 666–673.
9. Wright CM, Savarimuthu FS, Tan ME, Martins MU, Winterford C, Davidson MR and Duhig EE, *et al*. MS4A1 dysregulation in asbestos-related lung squamous cell carcinoma is due to CD20 stromal lymphocyte expression. PLoS ONE 2012, 7: e34943.
10. Sasaki H, Shitara M, Yokota K, Hikosaka Y, Moriyama S, Yano M and Fujii Y. RagD gene expression and NRF2 mutations in lung squamous cell carcinomas. Oncol Lett 2012, 4: 1167–1170.
11. de Kok JB, Roelofs RW, Giesendorf BA, Pennings JL, Waas ET, Feuth T and Swinkels DW, *et al*. Normalization of gene expression measurements in tumor tissues: comparison of 13 endogenous control genes. Lab Invest 2005, 85: 154–159.
12. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A and Speleman F. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol 2002, 3: H34.
13. Walter RF, Mairinger FD, Wohlschlaeger J, Worm K, Ting S, Vollbrecht C and Kurt WS, *et al*. FFPE tissue as a feasible source for gene expression analysis. A comparison of three reference genes and one tumor marker. Pathol Res Pract 2013, 209: 784–789.
14. van Rijn SJ, Riemers FM, van den Heuvel D, Wolfswinkel J, Hofland L, Meij BP and Penning LC. Expression stability of reference genes for quantitative RT-PCR of healthy and diseased pituitary tissue samples varies between humans, mice, and dogs. Mol Neurobiol 2013, doi:10.1007/s12035-013-8567-7.
15. Ceelen L, De Craene J and De Spiegelaere W. Evaluation of normalization strategies used in real-time quantitative PCR experiments in HepaRG cell line studies. Clin Chem 2013, doi:10.1373/clinchem.2013.209478.
16. Kolkova Z, Arakelyan A, Casslen B, Hansson S and Kriegova E. Normalizing to GADPH jeopardises correct quantification of gene expression in ovarian tumors – IPO8 and RPL4 are reliable reference genes. J Ovarian Res 2013, 6: 60.
17. Wu C, Wang X, Zhong M, Liu H, He Q, Yang X and Wen J, *et al*. Evaluation of potential reference genes for qRT-PCR studies in human hepatoma cell lines treated with TNF-α. Acta Biochim Biophys Sin 2013, 45: 780–786.

18. Shimamoto Y, Kitamura H, Niimi K, Yoshikawa Y, Hoshi F, Ishizuka M and Takahashi E. Selection of suitable reference genes for mRNA quantification studies using common marmoset tissues. Mol Biol Rep 2013, 40: 6747–6755.
19. Rienzo M, Schiano C, Casamassimi A, Grimaldi V, Infante T and Napoli C. Identification of valid reference housekeeping genes for gene expression analysis in tumor neovascularization studies. Clin Transl Oncol 2013, 15: 211–218.
20. Soes S, Sorensen BS, Alsner J, Overgaard J, Hager H, Hansen LL and Kristensen LS. Identification of accurate reference genes for RT-qPCR analysis of formalin-fixed paraffin-embedded tissue from primary non-small cell lung cancers and brain and lymph node metastases. Lung Cancer-J IASLC 2013, 81: 180–186.
21. Nguewa PA, Agorreta J, Blanco D, Lozano MD, Gomez-Roman J, Sanchez BA and Valles I, *et al*. Identification of importin 8 (IPO8) as the most accurate reference gene for the clinicopathological analysis of lung specimens. BMC Mol Biol 2008, 9: 103.
22. Eisenberg E and Levanon EY. Human housekeeping genes are compact. Trends Genet 2003, 19: 362–365.
23. Cheng WC, Chang CW, Chen CR, Tsai ML, Shu WY, Li CY and Hsu IC. Identification of reference genes across physiological states for qRT-PCR through microarray meta-analysis. PLoS ONE 2011, 6: e17347.
24. Shah KN and Faridi JS. Estrogen, tamoxifen, and Akt modulate expression of putative housekeeping genes in breast cancer cells. J Steroid Biochem Mol Biol 2011, 125: 219–225.
25. Shen Y, Li Y, Ye F, Wang F, Lu W and Xie X. Identification of suitable reference genes for measurement of gene expression in human cervical tissues. Anal Biochem 2010, 405: 224–229.
26. Inazu M, Yamada T, Kubota N and Yamanaka T. Functional expression of choline transporter-like protein 1 (CTL1) in small cell lung carcinoma cells: a target molecule for lung cancer therapy. Pharmacol Res 2013, 76: 119–131.
27. Barathidasan R, Pawaiya RS, Rai RB and Dhama K. Upregulated Myc expression in N-methyl nitrosourea (MNU)-induced rat mammary tumours. Asian Pac J Cancer Prev 2013, 14: 4883–4889.
28. Tabatabaeian H and Hojati Z. Assessment of HER-2 gene overexpression in Isfahan province breast cancer patients using real time RT-PCR and immunohistochemistry. Gene 2013, 531: 39–43.
29. Upadhyay M, Samal J, Kandpal M, Singh OV and Vivekanandan P. The Warburg effect: insights from the past decade. Pharmacol Ther 2012, 24: 62–67.
30. Guo C, Liu S and Sun MZ. Novel insight into the role of GAPDH playing in tumor. Clin Trans Oncol 2013, 15: 167–172.
31. Ahmad SS, Glatzle J, Bajaeifer K, Buhler S, Lehmann T, Konigsrainer I and Vollmer JP, *et al*. Phosphoglycerate kinase 1 as a promoter of metastasis in colon cancer. Int J Oncol 2013, 43: 586–590.
32. Sato-Tadano A, Suzuki T, Amari M, Takagi K, Miki Y, Tamaki K and Watanabe M, *et al*. Hexokinase II in breast carcinoma: a potent prognostic factor associated with hypoxia-inducible factor-1alpha and Ki-67. Cancer Sci 2013, 104: 1380–1388.
33. Zhan C, Shi Y, Lu C and Wang Q. Pyruvate kinase M2 is highly correlated with the differentiation and the prognosis of esophageal squamous cell cancer. Dis Esophagus 2013, 26: 746–753.
34. Haas M, Heinemann V, Kullmann F, Laubender RP, Klose C, Bruns CJ and Holdenrieder S, *et al*. Prognostic value of CA 19-9, CEA, CRP, LDH and bilirubin levels in locally advanced and metastatic pancreatic cancer: results from a multicenter, pooled analysis of patients receiving palliative chemotherapy. J Cancer Res Clin Oncol 2013, 139: 681–689.
35. Wang D, Moothart DR, Lowy DR and Qian X. The expression of glyceraldehyde-3-phosphate dehydrogenase associated cell cycle (GACC) genes correlates with cancer stage and poor survival in patients with solid tumors. PLoS ONE 2013, 8: e61262.
36. Della BC, Klawonn F, Menetski JP, Schumacher HJ and Pessler F. Evaluation of glyceraldehyde-3-phosphate, prolylpeptidyl isomerase A, and a set of stably expressed genes as reference mRNAs in urate crystal inflammation. BMC Res Notes 2011, 4: 443.
37. Higashimura Y, Nakajima Y, Yamaji R, Harada N, Shibasaki F, Nakano Y and Inui H. Up-regulation of glyceraldehyde-3-phosphate dehydrogenase gene expression by HIF-1 activity depending on Sp1 in hypoxic breast cancer cells. Arch Biochem Biophys 2011, 509: 1–8.
38. Wang T, Liang ZA, Sandford AJ, Xiong XY, Yang YY, Ji YL and He JQ. Selection of suitable housekeeping genes for real-time quantitative PCR in $CD4^+$ lymphocytes from asthmatics with or without depression. PLoS ONE 2012, 7: e48367.
39. Glare EM, Divjak M, Bailey MJ and Walters EH. β-Actin and GAPDH housekeeping gene expression in asthmatic airways is variable and not suitable for normalizing mRNA levels. Thorax 2002, 57: 765–770.
40. Wentzel P, Ejdesjo A and Eriksson UJ. Maternal diabetes *in vivo* and high glucose *in vitro* diminish GAPDH activity in rat embryos. Diabetes 2003, 52: 1222–1228.
41. Peng R, Zhai Y, Ding H, Di T, Zhang T, Li B and Shen W, *et al*. Analysis of reference gene expression for real-time PCR based on relative quantitation and dual spike-in strategy in the silkworm *Bombyx mori*. Acta Biochim Biophys Sin 2012, 44: 614–622.
42. Zhong M, Wang X, Wen J, Cai J, Wu C and Aly SM. Selection of reference genes for quantitative gene expression studies in the house fly (*Musca domestica* L.) using reverse transcription quantitative real-time PCR. Acta Biochim Biophys Sin 2013, 45: 1069–1073.
43. Zhai Z, Yao Y and Wang Y. Importance of suitable reference gene selection for quantitative RT-PCR during ATDC5 cells chondrocyte differentiation. PLoS ONE 2013, 8: e64786.
44. Wei R, Stewart EA and Amoaku WM. Suitability of endogenous reference genes for gene expression studies with human intraocular endothelial cells. BMC Res Notes 2013, 6: 46.
45. Herbst RS, Heymach JV and Lippman SM. Lung cancer. N Engl J Med 2008, 359: 1367–1380.
46. Maltseva DV, Khaustova NA, Fedotov NN, Matveeva EO, Lebedev AE, Shkurnikov MU and Galatenko VV, *et al*. High-throughput identification of reference genes for research and clinical RT-qPCR analysis of breast cancer samples. J Clin Bioinform 2013, 3: 13.
47. Meyer FR, Grausgruber H, Binter C, Mair GE, Guelly C, Vogl C and Steinborn R. Cross-platform microarray meta-analysis for the mouse jejunum selects novel reference genes with highly uniform levels of expression. PLoS ONE 2013, 8: e63125.
48. Okoniewski MJ and Miller CJ. Hybridization interactions between probesets in short oligo Microarrays lead to spurious correlations. BMC Bioinformatics 2006, 7: 276.
49. Royce TE, Rozowsky JS and Gerstein MB. Toward a universal microarray: prediction of gene expression through nearest-neighbor probe sequence identification. Nucleic Acids Res 2007, 35: e99.
50. Chen H, Liu Z, Gong S, Wu X, Taylor WL, Williams RW and Matta SG, *et al*. Genome-wide gene expression profiling of nucleus accumbens neurons projecting to ventral pallidum using both microarray and transcriptome sequencing. Front Neurosci 2011, 5: 98.
51. Wang Z, Gerstein M and Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet 2009, 10: 57–63.
52. Wilhelm BT and Landry JR. RNA-Seq-quantitative measurement of expression through massively parallel RNA-sequencing. Methods 2009, 48: 249–257.
53. Ozsolak F and Milos PM. RNA sequencing: advances, challenges and opportunities. Nat Rev Genet 2011, 12: 87–98.
54. Costa V, Angelini C, De Feis I and Ciccodicola A. Uncovering the complexity of transcriptomes with RNA-Seq. J Biomed Biotechnol 2010, 2010: 853916.