Malaysian J Pathol 2019; 41(3) : 313 – 326

ORIGINAL ARTICLE

Selection of reference genes for quantitative studies in acute myeloid leukaemia

Cindee LEE[1], Kang-Xian-Stephnie YIAU[2], Le-Jie LEE[2], Pei-Pei CHONG[3], Kian-Meng CHANG[4], Maha ABDULLAH[1,2]

[1]UPM-MAKNA Cancer Research Laboratory, Institute of Bioscience, Universiti Putra Malaysia, 43400 Serdang, Selangor, [2]Department of Pathology, Faculty of Medicine and Health Sciences, Universiti Putra Malaysia, 43400 Serdang, Selangor, [3]Department of Biomedical Sciences, Faculty of Medicine and Health Sciences, Universiti Putra Malaysia, 43400 Serdang, Selangor, and [4]Hospital Ampang, Jalan Mewah Utara, Pandan Mewah, 68000 Ampang, Selangor.

Abstract

Introduction: Quantitative polymerase chain reaction (qPCR) is commonly used in the investigation of acute myeloid leukaemias (AML). Stable reference genes (RG) are essential for accurate and reliable reporting but no standard method for selection has been endorsed. *Materials and Methods*: We evaluated simple statistics and published model-based approaches. Multiplex-qPCR was conducted to determine the expression of 24 candidate RG in AMLs (N=9). Singleplex-qPCR was carried out on selected RG (*SRP14, B2M* and *ATP5B)* and genes of interest in AML (N=15) and healthy controls, HC (N=12). *Results:* RG expression levels in AML samples were highly variable and coefficient of variance (CV) ranged from 0.37% to 10.17%. Analysis using GeNorm and Normfinder listed different orders of most stable genes but the top seven (*ACTB, UBE2D2, B2M, NF45, RPL37A, GK, QARS*) were the same. In singleplex-qPCR, *SRP14* maintained the lowest CV in AML samples. *B2M*, one of most stable reference genes in AML, was expressed near significantly different in AML and HC. GeNorm selected *ATP5B+SRP14* while Normfinder chose *SRP14+B2M* as the best two RG in combination. The median expressions of combined RG genes in AML compared to HC were less significantly different than individually implying smaller expression variation after combination. Genes of interest normalised with RG in combination or individually, displayed significantly different expression patterns. *Conclusions*: The selection of best reference gene in qPCR must consider all sample sets. Model-based approaches are important in large candidate gene analysis. This study showed combination of RG SRP14+B2M was the most suitable normalisation factor for qPCR analysis of AML and healthy individuals.

Keywords: reference gene, acute myeloid leukaemia, quantitative gene expression

INTRODUCTION

Quantitative analysis of nucleic acids and their derivatives provides essential information to understand and define states, conditions and changes in cells and tissues for research and diagnosis. Most common is the quantification of gene expression but single nucleotide polymorphism (SNP) genotyping, miRNA analysis, methylated DNA and copy number variation (CNV) are increasingly important. Various methods are applied but the most popular is real-time quantitative polymerase chain reaction (qPCR) because of its sensitivity, relatively rapid and convenience.

qPCR is routinely performed in laboratory diagnosis of acute myeloid leukaemia (AML). WHO (World Health Organisation) recommends testing for specific chromosomal translocations such as *AML1*/*ETO* t(8;21), *CBFB*/*MYH11* inv(16), *PML*/*RARA* t(15;17) and gene mutations in *NPM1, CEBPA FLT3* or *c-KIT* for prognostic evaluation. Treatment response is then monitored using the same abnormal markers. PCR-based assays may provide more sensitive

Address for correspondence: Assoc. Prof. Dr. Maha Abdullah, [1]UPM-MAKNA Cancer Research Laboratory, Institute of Bioscience, Universiti Putra Malaysia, 43400 Serdang, Selangor. [2]Department of Pathology, Faculty of Medicine and Health Sciences, Universiti Putra Malaysia, 43400 Serdang, Selangor. Tel: +603-89472375. Fax: +603-89412787. Email: maha@upm.edu.my

detection than FISH or routine chromosome analysis. Monitoring these markers during and after therapy for minimal residual disease (MRD) helps evaluate the risk of relapse.[1]

Earlier use of conventional reverse transcription (RT)-PCR generated only positive/negative results which did not allow timely assessment of therapeutic response as many patients remain positive for a long period even after achieving a cytogenetic response. Quantitative assessment, of BCR-ABL transcripts, for example, is clinically proven more useful because patients with high or increasing levels of BCR-ABL over the course of the disease have greater probability of relapse than those at steady state or decreasing levels of BCR-ABL.[2] Comparative studies on tissue sets from disease or healthy controls, however, are beset with difficulties in uneven cell numbers, poor RNA quality as well as variations in reverse transcription efficiencies in individual samples.

Reference genes, RGs (also known as housekeeping genes or internal control genes) are endogenous constitutively expressed genes and are used to compensate these variations. RGs were first used in qualitative studies to verify the quality of both tissue and technique. The commonly used reference genes were brought forward into quantitative studies also to standardise research reporting. At one time, *GAPDH* and *ACTB* were single control genes used in more than 90% of studies in high impact journals. These genes which are involved in important fundamental activities such as cell metabolism or structure were assumed to be most stably expressed genes.[3] Recent reports, however, demonstrated that the expression of commonly used housekeeping genes varied considerably under experimental circumstances. DeJonge *et al*.[4] evaluated the expression of 13,037 unique genes in 13,629 diverse samples, publicly available in microarray expression data from healthy and diseased tissues under various experimental conditions. None of the commonly used housekeeping genes (e.g. *ACTB, GAPDH, HPRT1* and *B2M*) were among the top 50 most stably expressed genes. These genes do not provide the best results as expressions are species and tissue-specific and may be altered by physiological conditions, diseased states as well as experimental conditions.[5] A universal internal control gene with expression levels that is constant across all thinkable tissue samples and cells as well as experimental treatments, and design samples is unlikely to exist.[6] Therefore, the best reference genes should be experimentally validated for specific experimental designs as well as types of samples.

The appropriate control gene for quantitative assessment should introduce the least systematic error when used as normalisation genes.[6] It must meet certain criteria including constitutive expression in all cell types studied at relatively stable levels with stability and expression levels comparable to the target gene.

Weisser *et al*.[1] considered several housekeeping genes (*B2M, PBGD, G6PDH* and *ABL*) to quantify fusion gene expression in AML and evaluated expression variation by comparing standard deviations. Correlation tests were also performed to compare expression levels in normal samples and ratio of target fusion gene to reference gene in AML samples. *ABL*, expressed at levels in the range of the fusion genes, was recommended as the most stable RG in normal donors and at diagnosis of *AML1ETO*- and *CBFB-MYH11*-positive AML.

Beillard *et al*.[7] evaluated several RG (*ABL, ACTB, B2M, GAPDH, PBGD, TBP* and *18S sRNA*) for fusion gene detection of MRD in leukaemic patients. Major exclusion criteria were the presence of pseudogenes, very high or low expression levels, significantly different expression in normal peripheral blood (PB) samples and leukaemic samples, and between PB and bone marrow (BM) samples. Although *ABL, B2M* and *GUS* were stably expressed in the samples, only *ABL* was proposed, as gene transcript expression did not differ significantly between normal and leukaemic samples at diagnosis.

More recently, model-based approaches using stability ranking algorithms such as geNorm,[8] NormFinder[6] and BestKeeper[9] were applied on haematopoietic neoplasia to identify the most stable RG. Potashnilova *et al*.[10] tested the systems on expressions of candidate reference genes on fresh clinical lymphoid samples in spleens, lymph nodes, and peripheral blood mononuclear cells from patients with different types of non-Hodgkin lymphomas as well as non-neoplastic lymphoid specimens. They concluded a normalisation-based approach using three reference genes (*YWHAZ, UBC* and *ACTB*) allows for robust reduction of variance in real-time PCR results.

These model-based approaches are widely used and tested on many other organisms to select the best reference genes (PUBMED search). They, however, are based on diverging

mathematical approaches which may not produce complementary results. In all cases, the selection is dependent on the candidate list and the number of potential genes included in the analysis.

So far, the best method for the selection of reference genes has yet to be decided. Researchers have used simple statistics while others depended on model-based approaches. Evaluation of these methods for the selection of best reference genes in acute myeloid leukaemia may help future decisions.

We evaluated various methods on 24 reference genes identified in a commercial kit on AML samples using a multiplex PCR-based method and evaluated the suitability of three potential reference genes expressed in AML and healthy control samples, using singleplex qPCR. We concluded simple statistics were just as important as model-based approaches for the selection of best reference genes.

METHODS AND MATERIALS

AML samples

Peripheral blood (PB) or bone marrow (BM) samples were collected from AML patients admitted into haematology wards in Hospital Ampang, Selangor. The collection was performed at diagnosis before chemotherapy, following informed consent. Healthy controls were apparently healthy individuals with no diagnosis of chronic diseases, at least two weeks free from acute infection or inflammatory inducing procedures. Approval to conduct the study was obtained from relevant CRC and the Medical Research and Ethics Committee, Ministry of Health, Malaysia. All procedures were in compliance with the Helsinki Declaration of 1975, revised in 2008. Samples were immediately isolated for mononuclear cell layer according to the protocol for density gradient centrifugation using Ficoll-Paque PLUS (GE Healthcare, Sweden). Cells were cryopreserved in liquid nitrogen until use.

Total RNA isolation

Total RNA was extracted with Tri-Reagent (Molecular Research, USA) containing phenol and guanidine thiocyanate, according to the manufacturer's instructions. Briefly, Tri-Reagent was added to homogenate cells. 1-bromo-3-chloropropane (1-BCP; Sigma, USA) was added for phase separation. RNA was then precipitated from the aqueous phase with isopropanol. After washing with 75% ethanol, RNA was dissolved in ultrapure water treated with 0.1% diethylpyrocarbonate (*DEPC*-treated *water)* containing RNase inhibitor (0.16 U/μl). Subsequently, RNA was subjected to DNase I (New England Biolabs, UK) treatment. The RNA quality and quantity were determined on a Nanodrop 1000 spectrophotometer (Thermo Scientific, Wilmington, USA) while integrity was confirmed by electrophoresis on 1% agarose gel stained with ethidium bromide (EtBr). RNA was subsequently reverse transcribed or stored at -80°C until further use.

Reference Genes

The GenomeLab GeXP Human Reference*Plex* Kit (Beckman Coulter, USA) which utilises the GenomeLab GeXP Analysis System, included 24 candidate reference genes represented in human RNA samples (Table 1). Reagents included a Control RNA Templates Human Reference*Plex*. Reverse transcription and polymerase chain reaction on AML samples were performed according to the product's user guide (A21780AG, December 2009) and supplemented with reagents from GenomeLab GeXP Start Kit (A85019AA, September 2009).

GenomeLab GeXP Analysis System is a multiplex quantitative PCR-based method to determine gene expression levels. It allows multiple reference genes, genes of interest and internal control to be analysed in a single well for improved accuracy. All reagents and materials used were purchased from Beckman Coulter, USA. PCR reactions were optimised for linear range detection.

Reverse Transcription (RT) for multiplex PCR

Total RNA was reverse transcribed using gene-specific reverse primers which add a flanking universal reverse sequence to the cDNAs. RT reaction mixture consisted of RT Buffer, RT Rev Primer Plex, Reverse Transcriptase, pre-diluted KAN^r RNA with RI, sample RNA (25-100 ng total) and DNase/RNase free water to a total of 20 µl. Reaction consisted of a series of incubation at 48°C for 1 min, 42°C for 60 min, 95°C for 5 min and held at 4°C in a Mastercycler® ep realplex (Eppendorf, USA).

Multiplex Polymerase Chain Reaction (PCR)

GeXP multiplex PCR reaction contained PCR Buffer, PCR Fwd Primer Plex, $MgCl_2$, Thermo-Start DNA polymerase and the cDNA sample and mixed according to the manufacturer's protocol. Incubation was conducted on a Mastercycler® ep realplex (Eppendorf, USA) based on the

TABLE 1: Reference Genes from the GenomeLab GeXP kit

	Gene name	ID[1]	Chr[2]	Pseudo[2]	Gene Symbol[2]
1	Histone deacetylase HD1	U50079	1	2	*HDAC1*
2	Ezrin	X51521	6	1	*EZRIN*
3	Ribosomal protein L37a	L06499	2	9	*RPL37A*
4	Transferrin Receptor	BC001188	3	NR	*TFR*
5	18kDa Alu RNA binding protein	NM_003134	15	4	*SRP14*
6	QRSHs glutaminyl-tRNA synthetase	X76013	3	NR	*QARS*
7	E2 Ubiquition conjugating enzyme UbcH5B	U39317	5	1	*UBE2D2*
8	Beta 2 microglubulin	NM_004048	15	NR	*B2M*
9	Nuclear factor NF45	U10323	1	2	*NF45*
10	Beta-actin	NM_001101	7	18	*ACTB*
11	Cyclophilin A	BC000689	7	79	*PPIA*
12	Glycerol kinase	NM_203391	X	3	*GK*
13	Acidic Ribosomal Protein	NM_001002	12	12	*RPLP0*
14	Hypoxanthine ribosyl tranferase	M31642.1	X	3	*HPRT1*
15	Elongation factor Ef-1alpha	NM_001402	6	42	*EEF1A1*
16	Proteasome subunit Y	D29012	17	NR	*PSMB6*
17	MLN51	X80199	17	NR	*MLN51*
18	Glutaraldehyde dehydrogenase	NM_002046	12	64	*GAPDH*
19	ATP synthase	X83218	21	1	*ATP5PO*
20	Beta-glucuronidase	NM_000181	7	18	*GUSB*
21	18s-rRNA	M10098			*18S rRNA*
22	Transcription factor IID	X97999	6	NR	*TBP*
23	Ca2-activted neutral protease large subunit	M23254	1	NR	*CAPN2*
24	Lysosomal hyaluronidase	AJ000099	3	NR	*HYAL2*

Chr=chromosome, Pseudo=pseudogene; [1]GenomeLab GeXP Human Reference*Plex* Kit manual; 2[11]

following profile: One cycle of 95°C for 10 min, 94°C for 0.5 min, 55°C for 0.5 min and 70°C for 1 min followed by 34 cycles of 94°C for 0.5 min, 55°C for 0.5 min and 70°C for 1 min and finally held at 4°C.

The resulting fragments were analysed on eXpress Analysis module where each PCR product is associated with the corresponding gene based on fragment sizes which eventually reports peak height and area. Each amplicon is designed to have a distinct length. "DefaultGeXPAnalysisParameters" was chosen when analysing the data obtained. After manually excluding unwanted fragments from the sample results, plate fragments were exported in .csv file for eXpress Profiler analysis.

A sample layout was performed to associate sample wells with multiplex. Peak binning was done to align the designed fragment size with GeXP estimated size. Normalisation was performed with internal control Kan^r provided with the kit, where necessary.

Model-based software

Several algorithms have been created and published to automate the selection of best reference genes from among a set of candidates for gene expression studies. We tested Normfinder[6] and GeNorm[8] on our list of candidates. Both function as Add-Ins on Microsoft Excel.

Normfinder website: https://moma.dk/Normfinder-software downloaded 20/2/2018.

Free online software. geNorm website: https://genorm.cmgg.be/

Genes of interest (GOI) and reference genes for singleplex qPCR

Thirteen AML associated genes, potentially involved in early relapse (<one year; poor prognosis, PP) and late relapse (>one year; good prognosis, GP) were selected from an earlier study[12] as genes of interest. They were *CALM2, CSTB, H2AFZ, EIF3M, TMSB4X* (GP) and *PBX3, SON, DDB2, PDCD61P, PGK1, SELL* and *SF3B1* (PP).

Two reference genes (*SRP14* and *B2M*) from the multiplex-PCR experiment were selected in addition to another commonly used gene, *ATP5B*. A total of 12 healthy controls and 15 acute myeloid leukaemia samples were examined using singleplex qPCR.

Conventional reverse transcription

Conventional reverse transcription was initiated with 4 μg total RNA and oligo(dT) 15 primer (Promega Corporation, WI, USA) which were heated and then chilled to prevent reforming of secondary structure. Subsequently, reagents added were M-MLV RT buffer, recombinant RNasin® ribonuclease inhibitor (Promega Corporation, WI, USA), dNTP mix (Fermentas, Lithuania, MBI) according to the manufacturer's manual. The product was stored at -20 °C until use.

Real-time Polymerase Chain Reaction (qPCR)

qPCR amplification was performed on the Mastercycler® ep realplex (Eppendorf, USA). Forward and reverse primers were synthesised by Next Gene Scientific, Malaysia. Taq polymerase buffer, $MgCl_2$, dNTP Mix (MBI Fermentas, Lithuania), Maxima SYBR Green qPCR Master Mix (Thermo Scientific, USA) were added according to the manufacturer's procedure. Primer sequences for genes examined are shown in Table 2. To perform standard curves for quantitation purposes, a serial dilution of cDNA from the leukaemia cell line Reh (CRL 8286) was conducted for each gene of interest including reference genes. Each sample was run in triplicate.

PCR profile was as follows: initial denaturation at 95°C for 10 minutes, then 34 cycles of 1 minute at 94°C, 30 s at 55°C, 30 s at 72°C, followed by melting curve from 50°C to 95°C with ramping time 15 minutes.

Analysis of Real-time PCR Data

Standard curves for all genes studied achieved R^2>0.97. Efficiency was >0.81. A melting curve analysis was performed to confirm the specificity of amplification and the appearance of primer dimers. Only one distinct peak should be seen in the melting curve. The relative expression of each gene was determined by using a standard curve method.

Statistical Analysis

Data were expressed as mean/median ± standard deviation (SD). Gene expressions were evaluated using non-parametric Spearman correlation test and Mann-Whitney T-test performed on IBM SPSS statistics version 25. The significant level was fixed at $p<0.05$.

RESULTS

Samples

Characteristics of AML samples and controls are shown in Table 3. FAB classification was not determined in all samples. M6 and M7 samples were excluded.

High variability in expressions of reference genes in AML samples in multiplex-PCR method

The reference genes tested on nine randomly selected AML samples showed very different levels of expressions (Supplementary S1). *SRP14* was among genes expressed at the highest level.

CV values and model-based approaches inconsistent in the selection of best reference genes in AML samples

Coefficient of variation (CV) and maximum fold change <2 (MFC, a ratio of maximum and minimum values within the dataset),[4] among simple statistics used to determine variation of the 24 reference genes in AML samples (N=9) are listed from lowest to highest CV values in Table 4. Gene with the lowest CV was *HD1*. *SRP14* was among the top five genes with the lowest CV. *B2M* was among the common reference genes with the lowest CV.

Two model-based approaches were chosen for analysis on quantitative output. GeNorm algorithm calculates gene stability M which is the average pairwise variation of a particular gene with all other control genes. Increasing variation corresponds to decreasing expression stability.[8] M values of the 24 reference genes are listed in Table 4. All values were below arbitrary number 1.5 and thus acceptable as reference genes. The

TABLE 2: Primer sequences of reference and genes of interest used in quantitative PCR

No.	Primer	Sequence	Amplicon
1	*SRP14*	5'-ACCAATTGACCACTGAATTGCTA-3' 3'-TGATAGCTTGCTCTTCACAGAGA-5'	250
2	*B2M*	5'-AAGATGAGTATGCCTGCCGTG-3' 3'-CTCCAAACTTCTACGGCGTA-5'	260
3	*ATP5B1*	F: 5'- CAGAGGTGTCTGCATTATTGG-3' R: 5'- CACATAGATAGCCTGTACAGAG-3'	140
4	*CALM-2*	F: 5'- AGAATCCCACAGAAGCAGAG-3' R: 5'- CATTGCCATCCTTATCAAACAC-3'	171
5	*CSTB*	F: 5'- CTTCATCAAGGTGCACGTC-3' R: 5'- GATGACTTTGTCAGTCTTCTGG-3'	187
6	*H2AFZ*	F: 5'- GCAGAGGTACTTGAACTGG-3' R: 5'- TTTCACAGAGATACAGTCCAC-3'	151
7	*EIF3M*	F: 5'- AAGTAGTTGTCAGTCATAGCAC-3' R: 5'- AACTCAGGTATCAGAAAGACTC-3'	137
8	*TMSB4X*	F: 5'- TCCAAAGAAACGATTGAACAGG-3' R: 5'- TGCCAGCCAGATAGATAGAC-3'	244
9	*PBX3*	F: 5'- AATCACAGGTGGATACCCTC-3' R: 5'- TAGGAGAAGTCACAGAAGATGG-3'	150
10	*SON*	F: 5'- CATTCCCTTCTCCTTCC-3' R: 5'- TTTGACACTTGGCATTA-3'	111
11	*DDB2*	F: 5'- CCCTTATGAATTGAGGACGA-3' R: 5'- AATGTGGTAACCCATTGCAG-3'	151
12	*PDCD61P*	F: 5'-GTTCATCCAGCAGACTTACC -3' R: 5'-GATCATAATATCTCAGGAGCGT -3'	152
13	*PGK1*	F: 5'- AACAACATGGAGATTGGCAC -3' R: 5'- GGCATTCTCATCAAACTTGTC -3'	144
14	*SELL*	F: 5'- TCTCAATGATTAAGGAGGGTG-3' R: 5'- GGGTCATTCATACTTCTCTTGG-3'	145
15	*SF3B1*	F: 5'- AATGGATAGAGACCTTGTACACAG-3' R: 5'- AACTGCCTGAATTACATGAGGA-3'	160

most stable gene was *UBE2D2*. The top seven most stable genes included common reference genes, *ACTB* and *B2M*. Neither *HD1* nor *SRP14* was included.

Multiple reference genes can also be combined to obtain a normalisation factor. The best number of reference gene combinations is a trade-off between practical considerations and accuracy.[8] GeNorm selected the best combination by eliminating worst-scoring gene first and recalculate for new M values until the final two most stable genes remains.[8] These were identified as *QARS* and *RPL37A* (Fig. 1A) which were not the original two most stable genes (Table 4). During recalculation, top stable genes such as *UBE2D2* and *ACTB* were eliminated early compared to *B2M* or *SRP14* (Fig. 1A).

Pairwise variation for N (minimum three) and N+1 genes are calculated to choose the best number of genes to combine. An arbitrary cut off point of 0.15 is used. A higher variation suggests the additional gene has a significant effect and should be included for a reliable normalisation factor. Fig. 1B shows very little change until the addition of the eight genes suggesting that the best was three, as eight genes were impractical.

Normfinder considers intragroup and intergroup variation for the estimation of expression variation.[6] As AML samples here were not assigned into groups, a complete analysis was

TABLE 3: Characteristics of AML patients and healthy controls included in the study

Baseline characteristics	AML patients		Healthy controls
	Reference gene – Multiplex-PCR (GeXP method)	Reference gene – Singleplex-qPCR method	Reference gene – Singleplex-qPCR method
Subjects (N)	9	15	12
Site	7-PB, 2-BM	12-BM, 3-PB, 1-NA	All-PB
FAB	3-M2, 1-M5b, 5-NA	1-M2, 2-M5, 13-NA	-
Blast cells Morphology IP Cytochemistry	>20% CD13/CD33+ or CD11c+ Auer rods+, cyMPO+ or cyMPO-/cyCD3-/CD19- and cyCD79a-, SBB+ or NSE+		None
Sex, N (%) Female Male	 3 (33.3%) 6 (66.7%)	 5 (33.3%) 10 (66.7%)	 7 (58.3%) 5 (41.6%)
Age range (years)	11-63	1-75	17-49

Diagnosis was determined by haematologists based on morphology, immunophenotyping and cytochemistry. IP=immunophenotyping, MPO=myeloperoxidase, SBB=Sudan Black, NSE=nonspecific esterase, NA=not available

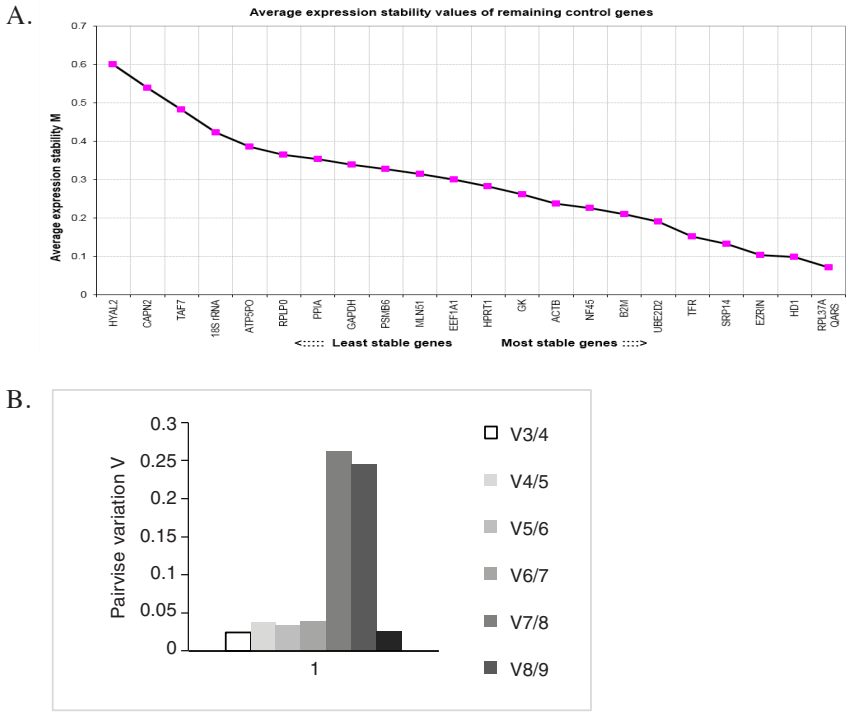


FIG. 1: GeNorm analysis to determine A) average expression stability values (M) of remaining control genes during stepwise exclusion of the least stable control genes in AML samples. B) Pairwise variation analysis of the 24 references in AML samples to determine the optimal number of control genes for accurate normalisation.

TABLE 4: Mean, standard deviation (SD), coefficient of variance (CV), maximum fold change (MFC) and stability determined by geNorm and Normfinder algorithms of reference genes expressed in AML samples (N=9) using multiplex-PCR

		LOG2(INTENSITY)				GENORM		NORMFINDER	
No	Gene	Ave	Std Dev	CV(%)	MFC	Gene	M	Gene	Stability value
1	*HDAC1*	17.28	0.06	0.37	1.01	*UBE2D2*	0.421	*ACTB*	0.051
2	*EZRIN*	17.25	0.11	0.62	1.02	*ACTB*	0.429	*UBE2D2*	0.065
3	*RPL37A*	17.28	0.12	0.70	1.02	*RPL37A*	0.437	***B2M***	0.100
4	*TFR*	17.24	0.14	0.81	1.02	***B2M***	0.438	*NF45*	0.135
5	***SRP14***	17.28	0.14	0.82	1.03	*QARS*	0.449	*RPL37A*	0.141
6	*QARS*	17.27	0.14	0.84	1.03	*NF45*	0.459	*GK*	0.157
7	*UBE2D2*	17.21	0.31	1.79	1.06	*GK*	0.484	*QARS*	0.163
8	***B2M***	17.20	0.32	1.84	1.06	*HDAC1*	0.485	*HPRT1*	0.164
9	*NF45*	17.08	0.34	1.98	1.06	*HPRT1*	0.490	*EEF1A1*	0.174
10	*ACTB*	17.17	0.35	2.03	1.07	*EZRIN*	0.503	*PSMB6*	0.178
11	*PPIA*	16.45	0.40	2.42	1.08	*EEF1A1*	0.505	*MLN51*	0.182
12	*GK*	17.00	0.44	2.58	1.08	*MLN51*	0.513	*GAPDH*	0.191
13	*RPLP0*	16.00	0.43	2.69	1.08	*PSMB6*	0.514	*HDAC1*	0.214
14	*HPRT1*	16.93	0.46	2.71	1.08	*GAPDH*	0.520	*EZRIN*	0.239
15	*EEF1A1*	16.87	0.48	2.87	1.10	*TFR*	0.522	*TFR*	0.249
16	*PSMB6*	16.61	0.48	2.87	1.10	***SRP14***	0.555	*RPLP0*	0.260
17	*MLN51*	16.72	0.52	3.09	1.11	*RPLP0*	0.569	*PPIA*	0.266
18	*GAPDH*	16.44	0.56	3.39	1.12	*PPIA*	0.573	***SRP14***	0.286
19	*ATP5PO*	16.59	0.67	4.06	1.12	*ATP5PO*	0.638	*ATP5PO*	0.317
20	*GUSB*	16.70	0.68	4.34	1.16	*18S rRNA*	0.843	*18S rRNA*	0.511
21	*18S rRNA*	16.03	0.77	4.83	1.14	*CAPN2*	1.085	*CAPN2*	0.703
22	*TBP*	16.98	0.94	5.51	1.20	*TBP*	1.118	*TBP*	0.726
23	*CAPN2*	16.33	1.23	7.54	1.26	*HYAL2*	1.246	*HYAL2*	0.825
24	*HYAL2*	13.47	1.37	10.17	1.39				

not performed. The stability value of reference genes is listed in Table IV. Unlike geNorm, *ACTB* was selected as the most stable reference gene. Nevertheless, the top seven most stable genes were the same in both models though not in the same order (Table 4). Only three of these were of genes with the lowest CV.

Supportive information from simple statistics and model-based approaches to select best reference genes in AML and HC samples

SRP14, highly expressed with small CV, and *B2M*, among most stable genes by model-based approaches, were reference genes selected for further assessment in AML samples and healthy controls (HC) using singleplex qPCR. Another common reference gene, *ATP5B*, was also included. The selection of these genes was also based on the absence of pseudogene.

Expression variations of the three genes, in AML and HC samples, tested using Mann-Whitney is shown in Fig 2A. *SRP14* and *ATP5B* were expressed higher in AML samples while *B2M* was expressed higher in HC samples. The smallest variation between AML and HC samples was observed for *ATP5B*. None was statistically significant.

Co-expression analysis with Spearman test showed *ATP5B* in HC was strongly correlated with *SRP14* (R2=0.797, p=0.002) but not *B2M* (R2=0.084, p=0.795). In AMLs, *ATP5B* was strongly correlated with both *SRP14* (R2=0.975, p=0.000) and *B2M* (R2=0.896, p=0.000). *SRP14* was poorly correlated with *B2M*

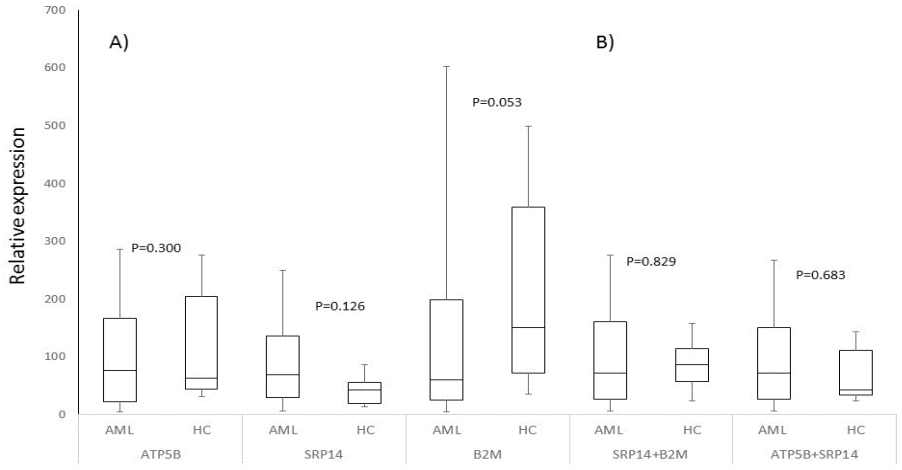


FIG. 2: Median expression levels of reference genes in AML and HC A) as individual and B) in gene combination using singleplex qPCR. Mann-Whitney U test analysis showed no significant difference (p>0.05) in the expression of reference genes, individually or in combination when compared between AML and HC samples.

((R2=0.273, p=0.391) in HC but well correlated in AML samples (R2=0.818, p=0.000). These observations suggest the genes were stably expressed in AML but not healthy controls.

The CV of log2 transformed expressions of *SRP14, B2M* and *ATP5B* in AML samples were 28.8%, 36.3% and 33.6%, respectively, which was similar in order to the multiplex-qPCR experiment but increased in magnitude. By combining AML and HC tissues, CV was reduced to 25.5%, 29.8% and 27.6%, respectively. This suggested the multiplex-qPCR method may be more robust than the singleplex-qPCR method.

GeNorm analysis of the three RGs in AML samples (N=15), showed *ATP5B* was most stable followed by *SRP14* and *B2M* with M values of 0.760, 0.811 and 1.148. Combining AML (N=15) and HC (N=12) samples also maintained the same order, M values of 1.188, 1.285 and 1.558, respectively. GeNorm selected the best two genes by default when the least stable gene was eliminated leaving *ATP5B* with *SRP14* (M=0.915).

Normfinder also selected *ATP5B* as best gene in AML samples, followed by *SRP14* and *B2M* (stability value = 0.146, 0.316 and 0.769, respectively). Normfinder also identified *ATP5B* as the best gene in combined AML and HC samples, followed by *SRP14* and *B2M* (stability values = 0.234, 0.502 and 0.588, respectively). The best gene combination (in HC and AML) was *SRP14* with *B2M,* with a stability value of 0.300. (Normfinder is able to calculate best two gene combination when at least two groups of study samples are available).

Variation in geomean expression when the best two genes were combined was analysed by Mann-Whitney statistics and presented in Fig. 2B. Larger p values were achieved indicating smaller variation in expression, in particular for combined *SRP14* with *B2M*.

Variation in p-values and expression patterns in genes of interest (GOI) normalised to individual or combined reference genes

Comparison of median fold-change of 12 AML-associated GOI in AML and HC samples after normalisation with reference genes individually or in combination are shown in Table 5. Different patterns of GOI expression in AML and significance were reached. The number of GOI having the same pattern of expression and significance was compared with the best RG combination (*SRP14+B2M*; p=0.829). Unsurprisingly, the number decreased from least variable to most variable: 8 (*ATP5B+SRP14*; p=0.683), 6 *(ATP5B*; p=0.300), 4 (*SRP14*; p=0.126) to 2 (*B2M*; p=0.053).

DISCUSSION

AML is a heterogeneous disease with high variability in genetic makeup and treatment outcomes. Diversity in morphology, immunophenotype and stages in progenitor cell immaturity[13] are challenges in classification and diagnosis of the disease. Within this disease, many studies attempt to relate its cancer biology to drug resistance and disease relapse through

TABLE 5: Median fold-change in genes of interest after normalization with individual and combined reference genes in healthy controls, HC (N=12) and acute myeloid leukaemia, AML (N=15) samples. Results are expressed as median ± STDEV. Statistical analysis using Mann-Whitney T-test is accepted as significantly different when p<0.05.

			Normalized with (fold-change)																							
			ATP5B					*SRP14*					*B2M*					*SRP14+B2M*					*ATP5B+SRP14*				
	HC	AML	HC		AML			HC		AML			HC		AML			HC		AML			HC		AML		
Gene	N	N	Median	SD	Median	SD	*p*	Median	SD	Median	SD	*p*	Median	SD	Median	SD	*p*	Median	SD	Median	SD	*p*	Median	SD	Median	SD	*p*
CALM2	12	15	0.82	0.45	0.33	0.17	**0.028**	2.38	1.50	0.36	0.14	**0.002**	0.92	0.71	0.32	0.25	0.231	1.45	1.03	0.27	0.15	0.103	1.32	0.79	0.34	0.14	**0.009**
CSTB	12	15	3.15	2.40	2.09	3.12	0.200	8.92	7.76	1.89	2.45	**0.001**	2.15	1.85	2.05	3.23	0.899	3.93	2.26	2.34	2.28	**0.014**	5.14	3.95	2.21	2.66	**0.001**
H2AFZ	12	15	2.39	0.85	2.86	1.10	0.373	6.13	3.74	2.49	1.23	**0.002**	2.59	2.39	3.06	2.55	0.631	4.23	3.10	2.36	1.68	0.456	4.46	1.73	2.49	1.11	0.670
EIF3M	12	15	2.13	2.44	4.06	1.55	0.183	5.35	5.50	3.42	1.15	**0.004**	1.98	1.01	4.13	2.90	**0.017**	3.32	1.19	3.89	1.60	0.256	3.54	3.29	3.91	1.22	0.792
TMSB4X	12	15	2.29	8.72	3.90	2.47	0.792	6.82	24.34	4.26	2.55	0.631	1.19	5.87	3.92	2.55	**0.036**	2.88	10.71	3.52	2.06	0.683	4.20	14.00	4.08	2.44	0.867
PBX3	10	15	0.71	0.29	0.66	0.89	0.807	1.68	38.63	0.48	1.02	**0.007**	0.63	20.77	0.81	0.70	0.687	0.98	0.57	0.69	0.74	0.531	1.17	0.23	0.55	0.95	0.103
SON	12	15	5.81	7.99	18.25	13.97	**0.004**	14.49	27.22	19.64	12.64	0.347	2.18	2.70	17.70	17.59	**0.000**	4.43	6.03	16.30	11.21	**0.002**	11.21	14.05	19.17	12.78	0.053
DDB2	12	15	0.76	0.40	0.42	0.37	**0.001**	2.03	1.48	0.36	0.29	**0.000**	0.89	0.59	0.49	0.30	0.193	1.33	0.92	0.40	0.21	**0.019**	1.08	0.74	0.39	0.32	**0.000**
PDCD61P	12	15	0.89	6.14	2.13	1.64	0.581	2.13	12.90	2.16	1.56	0.742	0.73	0.54	1.56	1.88	**0.011**	2.55	1.35	1.07	0.64	**0.021**	1.37	8.33	2.16	1.56	0.905
PGK1	12	15	1.64	0.53	1.24	0.48	**0.005**	4.43	1.51	1.03	0.59	**0.000**	1.54	1.15	1.02	0.89	0.781	2.55	1.35	1.07	0.64	**0.021**	2.89	0.62	1.09	0.51	**0.000**
SELL	12	15	52.62	274.85	46.33	67.15	0.217	156.67	487.27	39.26	50.93	**0.000**	51.44	22.10	41.12	60.82	0.193	107.21	90.18	47.50	49.12	**0.001**	89.08	342.95	42.65	57.43	**0.012**
SF3B1	12	15	3.89	5.37	5.20	1.02	0.256	11.12	7.29	5.61	1.74	**0.000**	3.15	1.62	5.05	3.56	**0.027**	5.18	1.49	4.51	1.98	0.486	6.32	5.78	5.43	1.29	0.235

gene expression.

To select the best reference genes for AML studies, we evaluated a set of 24 reference genes in human. These are relevant to leukaemia studies as 10 of 14 candidate control genes were evaluated in a multicentre study for normalisation of normal, ALL, CML and AML tissues in diagnosis and residual disease detection in leukaemia patients.[7]

Early common reference genes are highly expressed in most tissues. Very highly expressed genes are, however, not recommended in quantitative studies as it may increase the risk of false-negative results, particularly when target genes are expressed at low levels or if the quality of the sample is low. Best control genes are expressed within the expression range of genes of interest.[1] In the study by Beillard *et al*.[7], reference genes, *18S rRNA* and *TBP* were excluded for very high and low expressions, respectively. Selection for *B2M* appears to be ambivalent as one study excluded it for high expression[1] while another maintained.[7] In the latter study *B2M* was eventually excluded as a comparison between normal and leukaemic samples showed a significant difference. Exclusion based on high expression levels cannot be done when studies do not include the gene of interest among its candidates.

In our study, we found mean levels of *18S rRNA* among one of the lowest with *TBP* being slightly higher. This may be due to the different primers used in the studies. In the multicentre study, *B2M* expressions differed between a commercial kit (ABI) and in-house primers, which resulted in one being preferred over the other.[7] In multigene studies, it may be difficult to find an internal control within the range of each genes of interest. It is not practical to have separate reference genes for each GOI incurring extra difficulty for GOIs that are less investigated. The recommendation to select RGs with high median expression and at least one with a medium median expression may help.[7]

Simple statistics including standard deviation (SD), coefficient of variance (CV) and maximum fold change (MFC) are used to measure expression variation in candidate reference genes. DeJonge *et al*.[4] compared CV for each gene and observed the top 15 (CV <4%) did not include commonly used housekeeping genes. The majority were genes that encode ribosomal proteins involved in protein biosynthesis. One of these was *SRP14*. In our 24 candidates, *RPL37A* and *SRP14* were also among the top five with a lowest CV which did not include any of the common reference genes in the kit (i.e. *B2M, ACTB, HPRT1, GLUB, GAPDH* and *18S rRNA*). On the other hand, two commonly used reference genes (*ACTB* and *B2M*) were among the top five best genes. Only one was a ribosomal protein gene, *RPL37a*.

A useful method to compare expression variation among different sample sets is to conduct statistical tests. Similar to Beillard *et al*.[7], the nonparametric Mann-Whitney test was performed to compare median expression levels of healthy controls and AML samples. P-value is used to indicate the degree of variation for the higher-lower expression patterns seen in two samples. Using a global linear model, Beillard *et al*.[7] excluded *B2M* and *GUSB* as unsuitable reference genes since both were significantly different in normal and leukaemic samples. In this study, expressions of *B2M* were near significantly different in AML and healthy controls and had the highest variation among the three RGs examined. *B2M* was expressed higher in healthy controls as was also demonstrated in the study by Beillard *et al*.[7]

ABL was selected as the best reference gene for fusion gene analysis in AML[1,7] and also recommended for quantification of *BCR-ABL* transcripts in CML samples.[1] *ABL* is a partner fusion gene in the *BCR-ABL* chromosome translocation, highly associated with chronic myeloid leukaemia (95%) and B-cell ALL.[14] Wang *et al*.[2] expressed concern whether a translocation partner is qualified to serve as an internal control gene for normalisation as the probe detects both fusion and wild type transcripts. It is likely, this will affect the accuracy of *BCR-ABL* measured particularly the *BCR-ABL* to control ratio. The International Standardization Group now recommends three genes *BCR, ABL* and *GUSB* as internal controls.[15] Two percent of AML (M1/M2) with adverse prognosis carries the *BCR-ABL* translocation,[16] suggesting a need for the careful use of *ABL* as an internal control for AML studies.

In CML, genes located on chromosomes 8, 17, 19, or 22, are not recommended as they are frequently subjected to rearrangements.[2] Common chromosomes translocated in AML include chromosomes 3, 6, 8, 9, 11, 21, 22, 16, 15, 17 as well as gene mutations in chromosome 4 (*KIT*), 13 (*FLT1*), 1 (*NRAS*) and 5 (*NPM1*).[13,17] X chromosome genes are preferably excluded[1]. *SRP14* and *B2M* are located on chromosome 15. *ATP5B* is located on chromosome 12.[11]

Weisser *et al*.[1] recommended excluding genes with pseudogenes to avoid amplification of genomic DNA. Pseudogenes are common, nonfunctional sequences with close similarities to the genes from which they are derived. In RT-PCR studies it can lead to false-positive results and underestimate quantification.[18]

Commonly used reference genes, except *B2M*, have abundant pseudogenes: *ACTB* (18), *HPRT1* (3), *GLUB* (18), *GAPDH* (64).[11] Pseudogenes are also common to ribosomal proteins with 2000 being reported.[19] The number of pseudogene in *RPL37a* (9) exceeded that in *SRP14*.[4] *ATB5B* has one pseudogene.

The model-based approaches used here both selected *ATP5B* as the most suitable reference gene among the three RG selected. On the other hand, both systems chose different best gene combinations. GeNorm uses the ratio of two internal control genes to indicate normalisation error. Gene stability increases with a decrease in the average pairwise variation of a particular gene with all other control genes.[8] Andersen *et al*.[6] argues that this will tend to select for genes with the highest degree of similarity in expression patterns, thus candidates with minimal expression variation do not necessarily become top-ranked. Normfinder evaluates every gene independently and selects for candidates with minimal estimated intra- and intergroup variation.[6] Interestingly, Normfinder selected the best RG combination (*ATP5B and B2M*) based on the least variation of expression in AML and HC samples.

The most suitable gene selected is obviously limited to the list of candidate genes that are analysed. Therefore, the best gene is only as good as what is on the list. GeNorm provided an arbitrary cut off point of M<1.5 which was fulfilled by *ATP5B* and *SRP14* but not *B2M*. Normfinder did not provide a cut-off point but with the lowest stability value of 0.234, *ATP5B* was also selected as the most suitable gene.

Multiple control genes are an option when an optimal reference gene is not available. The assumption is the average variation in multiple genes is smaller than variation in individual genes. However, the set of genes has to be cautiously selected to consider intergroup variation. An average intergroup variation of almost zero is most desirous.[6] GeNorm selected the final best combination to be *ATP5B* with *SRP14* which reduced the average M value, suggesting increased stability. Normfinder selected *SRP14* with *B2M* as the best combination but stability decreased. These conflicting results did not allow us to conclude on the best reference gene(s) based on model-based approaches.

Mann-Whitney statistical testing compared median expression and provided a p-value. Increasing p-value suggests less variability in gene expression in samples. The highest p-value was observed for geomean of *SRP14+B2M*. Normalisation of the gene of interests (GOI) with reference genes separately and in combinations revealed different patterns of expression and significance. Using combined *SRP14+B2M* as a benchmark (highest p-value), the number of GOI sharing the same pattern and significance decreased with decreasing p-value. This clearly showed that increasing expression variability between sample sets will shift and change the patterns of gene expression, again emphasising the importance of selecting the most suitable reference gene. Combining two genes (*SRP14 and B2M*) with opposing patterns of intergroup variations balanced the differences and reduced variability. This method was suitable in selection when shortlist of candidate genes was considered.

It is practical to deliberate on the number of reference genes to be included when there is a limitation in the amount of RNA sample, as multiple reference genes have to be measured. This also applies when only a small number of target genes are considered.[6] Both geNorm and Normfinder provide means to select an optimal number of normalisation gene.

However, as individual labs proceed in this manner to select for own reference genes, data will become less comparable in literature. External controls with predefined transcript levels may in future be considered as reference materials to unify laboratory specific measurements particularly in diagnosis to achieve a common standardized scale which is independent of any testing laboratory and its specific RT-qPCR method.[20]

CONCLUSION

Choosing the most suitable reference gene for quantitative PCR analysis remains a challenge. Model-based approaches provide a systematic method for selection from a large candidate pool. Genorm and Normfinder software identified *ACTB, UBE2D2, B2M* and *RPL37A* as the top five most stable reference genes in AML. We found statistical testing was suitable for the selection of best gene(s) in small number candidates. A combination of *SRP14+B2M* reference genes was potentially the best normalisation factor

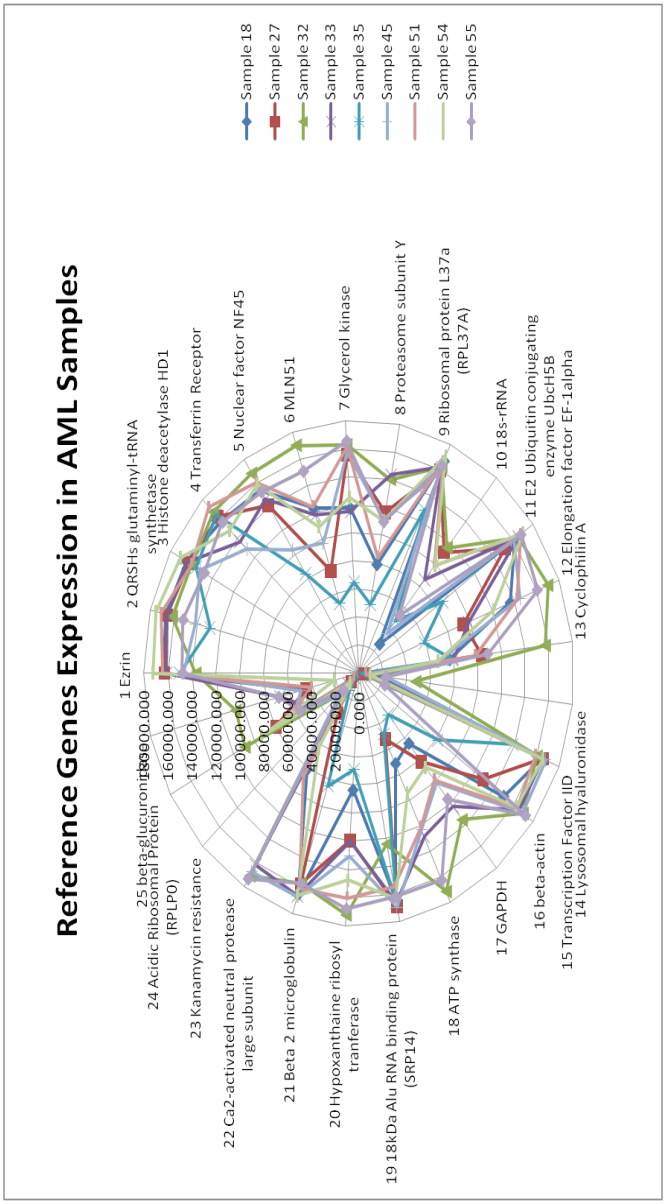
in studies in sample sets involving AML and healthy controls.

ACKNOWLEDGEMENT

This study was supported by research grants from the Ministry of Science, Technology and Innovation (Project No: 02-01-04-SF008) and Ministry of Higher Education, Malaysia (Project No: 04-11-08-624R).

REFERENCES

1. Weisser M, Haferlach T, Schoch C, Hiddemann W, Schnittger S. The use of housekeeping genes for real-time PCR-based quantification of fusion gene transcripts in acute myeloid leukemia. Leukemia. 2004; 18(9): 1551-3.
2. Wang YL, Lee JW, Cesarman E, Jin DK, Csernus B. Molecular monitoring of chronic myelogenous leukemia: identification of the most suitable internal control gene for real-time quantification of BCR-ABL transcripts. J Mol Diagn. 2006; 8(2): 231-9.
3. Suzuki T, Higgins PJ, Crawford DR. Control selection for RNA quantitation. Biotechniques. 2000; 29(2): 332-7.
4. de Jonge HJ, Fehrmann RS, de Bont ES, *et al*. Evidence based selection of housekeeping genes. PLoS One. 2007; 2(9): e898.
5. Schmittgen TD, Zakrajsek BA. Effect of experimental treatment on housekeeping gene expression: validation by real-time, quantitative RT-PCR. Biochem Biophys Methods 2000; 46: 69-81.
6. 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. 2004; 64(15): 5245-50.
7. Beillard E, Pallisgaard N, van der Velden VH, *et al*. Evaluation of candidate control genes for diagnosis and residual disease detectionin leukemic patients using 'real-time' quantitative reverse-transcriptasepolymerase chain reaction (RQ-PCR) - a Europe against cancer program. Leukemia. 2003; 17(12): 2474-86.
8. Vandesompele J, De Preter K, Pattyn F, *et al*. Accurate normalization of real-time quantitative RT PCR data by geometricaveraging of multiple internal control genes. Genome Biol. 2002; 3(7): RESEARCH0034.
9. Pfaffl M, Tichopad A, Prgomet C, Neuvians TP. Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper – Excel-based tool using pair-wise correlations. Biotechnology Letters. 2004; l26 (6): 509-15.
10. Potashnikova D, Gladkikh A, Vorobjev IA. Selection of superior reference genes' combination for quantitative real-time PCRin B-cell lymphomas. Ann Clin Lab Sci. 2015 Winter; 45(1): 64-72.
11. Gene. Bethesda (MD): National Library of Medicine (US), National Center for Biotechnology Information; 2004. Available from: https://www.ncbi.nlm.nih.gov/gene/
12. Ngiow SF, Abdullah M, Lim J, Cheong SK, Seow HF. Subtractive Hybridization Identifies Stem Cell-Associated Genes in an Acute Myeloid Leukemia with Poor Prognosis. Mal J Med Health Sciences 2016; 12 (1): 19-21.
13. Arber DA, Orazi A, Hasserjian R, *et al*. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. Blood. 2016; 127(20): 2391-405.
14. De Braekeleer E, Douet-Guilbert N, Rowe D, *et al*. ABL1 fusion genes in hematological malignancies: a review. Eur J Haematol. 2011; 86(5): 361-71.
15. Hughes T, Deininger M, Hochhaus A, *et al*. Monitoring CML patients responding to treatment with tyrosine kinase inhibitors: review and recommendations for harmonizing current methodology for detecting *BCR-ABL* transcripts and kinase domain mutations and for expressing results. Blood. 2006; 108: 28-37.
16. Martens JH, Stunnenberg HG. The molecular signature of oncofusion proteins in acute myeloid leukemia. FEBS Lett. 2010; 584(12):2662-9.
17. Kumar CC. Genetic Abnormalities and Challenges in the Treatment of Acute Myeloid Leukemia. Genes Cancer. 2011; 2(2): 95-107.
18. Smith RD, Ogden CW, Penny MA. Exclusive amplification of cDNA template (EXACT) RT-PCR to avoid amplifying contaminating genomic pseudogenes. Biotechniques. 2001; 31(4):776-8.
19. Tonner P, Srinivasasainagendra V, Zhang S, Zhi D. Detecting transcription of ribosomal protein pseudogenes in diverse human tissues from RNA-seq data. BMC Genomics. 2012; 13: 412.
20. Zhen C, Wang YL. Molecular monitoring of chronic myeloid leukemia: international standardization of BCR-ABL1 quantitation. J Mol Diagn. 2013; 15(5): 556-64.



Supplementary figure 1. Expression of 23 reference genes in individual AML samples (N=9)