

Identification and Validation of Reference Genes for qPCR Detection of Serum microRNAs in Colorectal Adenocarcinoma Patients

Guixi Zheng, Haiyan Wang, Xin Zhang, Yongmei Yang, Lili Wang, Lutao Du, Wei Li, Juan Li, Ailin Qu, Yimin Liu, Chuanxin Wang*

Department of Clinical Laboratory, Qilu Hospital of Shandong University, Jinan, Shandong Province, China

Abstract

Serum microRNAs (miRNAs) have become a highlighted research hotspot, especially for their great potential as a novel promising non-invasive biomarker in cancer diagnosis. The most frequently used approach for serum miRNAs detection is quantitative real time polymerase chain reaction (qPCR). In order to obtain reliable qPCR data of miRNAs expression, the use of reference genes as endogenous control is undoubtedly necessary. However, no systematic evaluation and validation of reference genes for normalizing qPCR analysis of serum miRNAs has been reported in colorectal adenocarcinoma. We firstly profiled pooled serum of colorectal adenocarcinoma, colorectal adenoma and healthy controls and selected a list of 13 miRNAs as candidate reference genes. U6 snRNA (U6) and above-mentioned 13 miRNAs were included in further confirmation by qPCR. As a result, 5 miRNAs (miR-151a-3p, miR-4446-3p, miR-221-3p, miR-93-5p and miR-3184-3p) were not detected in all samples and 2 miRNAs (miR-197-3p and miR-26a-5p) were relatively low with median Cq more than 35, and were excluded from further stability analysis. Then variable stability of other 6 miRNAs (miR-103b, miR-484, miR-16-5p, miR-3615, miR-18a-3p and miR-191-5p) and U6 were evaluated using two algorithms: geNorm and NormFinder which both identified miR-191-5p as the most stably expressed reference gene and selected miR-191-5p and U6 as the most stable pair of reference genes. After validating in an independent large cohorts and selecting miR-92a-3p as target miRNA to evaluate the effect of reference gene, we propose that combination of miR-191-5p and U6 could be used as reference genes for serum microRNAs qPCR data in colorectal adenocarcinoma, colorectal adenoma and healthy controls.

Citation: Zheng G, Wang H, Zhang X, Yang Y, Wang L, et al. (2013) Identification and Validation of Reference Genes for qPCR Detection of Serum microRNAs in Colorectal Adenocarcinoma Patients. PLoS ONE 8(12): e83025. doi:10.1371/journal.pone.0083025

Editor: Alfons Navarro, University of Barcelona, Spain

Received: September 3, 2013; **Accepted:** November 6, 2013; **Published:** December 11, 2013

Copyright: © 2013 Zheng et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This study was financially supported by National Science Foundation for Yong Scholars of China (Grant No. 81301506). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

* E-mail: cxwang@sdu.edu.cn

Introduction

Colorectal cancer (CRC) is one of the most leading causes of cancer-related death with increased incidence and mortality in the past several decades. The progression to CRC is considered as a stepwise process with the accumulation of different genetic and epigenetic alterations resulting in a transformation from normal mucosa to precancerous lesion and finally to malignant tumor that is a well known normal mucosa-adenoma-adenocarcinoma (NM-A-AC) sequence [1]. Currently, several CRC screening studies pay more attention to the detection rate of precancerous adenoma which is associated with a high risk of progression to invasive lesion and represents the optimal target lesion for preventing CRC [2-4]. There is an urgent need to seek new biomarkers which should have high

sensitivity and specificity for early-stage CRC and precancerous lesions.

MicroRNAs(miRNAs) are single-stranded RNA molecules, approximately 19-24 nucleotides in length, that regulate gene expression at the post-transcriptional level. Recently, serum miRNAs have been shown to be stable and reproducible which open a new opportunity of non-invasive test for the early diagnosis of cancer [5-7]. Quantitative real time polymerase chain reaction (qPCR) is the most frequently used approach for measurement of serum miRNAs due to its accuracy, sensitivity, specificity, reproducibility and robustness [8,9]. The accuracy of miRNAs expression analysis depends largely on a proper normalization. The use of reference genes as endogenous control is the most common method for normalizing qPCR data of miRNAs expression. The identification of suitable reference

genes play a crucial role in miRNAs research because normalization to unreliable reference genes may lead to incorrect determination of miRNAs of interests [10,11]. The chosen of reference genes for detection of serum miRNAs is a major issue to be solved.

It has become clear that no single reference gene is constitutively expressed in all sample types, different kinds of disease and under all experimental designs which indicates that the expression stability of reference genes has to be verified before each experiment [12,13]. As far as aware, no systematic evaluation and validation of reference genes for normalizing qPCR analysis of serum miRNAs in colorectal adenocarcinoma has been published. In this study, we firstly made an effort to screen candidate reference genes using Miseq sequencing. Secondly, the reliability of candidate reference genes for normalization was evaluated by qPCR assays to select the most suitable reference genes. Then the selected reference genes were validated in an independent cohort study. To test for the effect of reference genes, we selected serum miR-92a-3p as target miRNA.

Materials and Methods

Ethics statement

Written informed consent was obtained from every participants for the use of the venous blood samples in this study. This project was approved by the Clinical Research Ethics Committee of Qilu Hospital of Shandong University.

Study design and subjects

All the colorectal adenocarcinoma and adenoma patients were recruited from Department of General surgery and Gastroenterology, Qilu Hospital of Shandong University between 2010 and 2013. And the healthy controls were recruited from the Department of Physical Examination Center, Qilu Hospital of Shandong University.

We divided our study into three phases (Figure S1). The first phase was designed to screen candidate reference genes in serum of 3 pooling samples (30 patients with colorectal adenocarcinoma, 25 patients with colorectal adenoma and 30 healthy controls, respectively) using Miseq sequencing. Only the miRNAs which showed no differential expression among the 3 pooling samples were selected as candidate reference genes. The second phase was designed to evaluate the reliability of candidate reference genes for normalization by qPCR assays in an independent cohort of 125 samples (including 45 colorectal adenocarcinoma patients, 40 colorectal adenoma patients and 40 healthy controls) and to select the most suitable reference genes. Then the third phase was designed to validate the selected reference genes in another independent cohort of 128 colorectal adenocarcinoma patients, 60 colorectal adenoma patients and 60 healthy controls.

All the diagnosis of patients with colorectal adenocarcinoma and adenoma were confirmed by histopathology or histobiopsy. Tumors were staged according to the tumor-node-metastasis (TNM) staging system of Union for International Cancer Control (UICC). Healthy controls were recruited from a large pool of individuals seeking a routine health checkup who showed no

Table 1. Clinicopathological data for patients and healthy controls.

Characteristics		Screening phase	Selection phase	Validation phase
Healthy control	number	30	40	60
Age	mean±SD	54.37±16.24	56.38±15.40	56.30±16.82
Sex	Male(%)	18(60.0)	23(57.5)	37(61.7)
	Female(%)	12(40.0)	17(42.5)	23(38.3)
Colorectal adenoma	number	25	40	60
Age	mean±SD	54.60±17.36	57.30±14.09	57.81±14.42
Sex	Male(%)	16(64.0)	26(65.0)	35(58.3)
	Female(%)	9(36.0)	14(35.0)	25(41.7)
Location of adenoma	Colon(%)	18(72.0)	28(70.0)	46(76.7)
	Rectum(%)	7(28.0)	12(30.0)	14(23.3)
Colorectal adenocarcinoma	number	30	45	128
Age	mean±SD	56.56±19.67	57.0±17.21	55.62±16.43
Sex	Male(%)	19(63.3)	25(55.6)	73(57.0)
	Female(%)	11(36.7)	20(44.4)	55(43.0)
Location of tumor	Colon(%)	22(73.3)	32(71.1)	89(69.5)
	Rectum(%)	8(26.7)	13(28.9)	39(30.5)
Tumor size	≤3cm(%)	4(13.3)	9(20.0)	22(17.2)
	>3cm(%)	26(86.7)	36(80.0)	106(82.8)
Differentiation	Well(%)	4(13.3)	7(15.6%)	19(14.8)
	Moderate(%)	19(63.3)	25(55.5)	81(63.3)
	Poor(%)	5(16.7)	10(22.2)	28(21.9)
	N/A(%)	2(6.7)	3(6.7)	
TNM Stage	StageI(%)	11(36.7)	16(35.5)	32(25.0)
	StageII(%)	8(26.6)	12(26.6)	32(25.0)
	StageIII(%)	5(16.7)	7(15.6)	32(25.0)
	StageIV(%)	4(13.3)	7(15.6)	32(25.0)
	N/A(%)	2(6.7)	3(6.7)	

doi: 10.1371/journal.pone.0083025.t001

evidence of disease and matched to the patients by age, sex (Table 1). All blood samples of patients were collected before any therapeutic procedures such as surgery, chemotherapy, and radiotherapy.

Serum preparation

5ml venous blood was collected from each subject. The whole blood was separated into serum and cellular fractions by centrifugation at 4000 rpm for 10min, followed by 12000 rpm high-speed centrifugation for 15 min to completely remove cell debris. The entire process was finished within 2h after blood extraction and the supernatant serum was stored at -80°C until analysis.

Miseq sequencing

For Miseq sequencing, equal volumes of serum from 30 patients with colorectal adenocarcinoma, 25 patients with colorectal adenoma and 30 healthy controls with similar age and sex distributions were pooled separately. Total RNA was

extracted and purified using the miRNeasy Mini Kit (Qiagen, Valencia, CA) according to the manufacturer's protocols. Briefly, a pair of adaptors was ligated to the 3' and 5' ends of miRNA, and miRNA molecules were amplified by qPCR to establish cDNA library. The quality of library was evaluated using KAPA qPCR kit by the two criteria: (a) the concentration of cDNA was more than 1nM, (b) showing no dimer contamination. Then the purified cDNA was used directly for sequencing analysis using Miseq sequencer (Illumina, San Diego, CA, USA) according to the manufacturer's instructions. After a bioinformatics analysis, the remaining miRNAs were searched against the miRBase 17.0 to identify known mature miRNAs.

Candidate reference genes

For Miseq sequencing, the final reads of each miRNA in each pooled sample were determined by normalization with the total reads of all called miRNAs in the sample. The miRNAs were selected as candidate reference genes that satisfied the following criteria: (a) having at least 50 copies in the three pooling samples, (b) showing no differential expression among three groups ($P > 0.05$).

qPCR analysis

qPCR was performed on ABI PRISM 7500 Sequence Detection System using the SYBR PrimeScript miRNA QPCR Kit (Takara Bio Inc). The reverse transcription (RT) reaction consisted of 10 μ l of 2 \times miRNA Reaction Buffer Mix, 2 μ l of miRNA Primescript RT Enzyme Mix, 2 μ l of 0.1% BSA and 3 μ l serum that mixed with 3 μ l serum buffer (2.5% Tween 20, 50 mmol/L Tris and 1mmol/L EDTA [14]). The reactions were incubated at 37°C for 60min, followed by 85°C for 5s and 4°C for 60min. The cDNA was centrifugated at 10000 rpm. for 10min. The PCR reaction consisted of 12.5 μ l of SYBR Premix Ex TaqII, 0.5 μ l of Dyell, 2 μ l of 5 μ M of forward primer, 1 μ l of 10 μ M of Uni-miR qPCR Primer, 7 μ l of ddH₂O and 2.0 μ l of template cDNA into a total volume of 25 μ l. The reactions were incubated at 95°C for 30s, followed by 45 cycles of 95 °C for 5s and 57°C for 34s. To verify that the used primer pair produced only a single product, a dissociation protocol was added after thermocycling, determining dissociation of the PCR products from 65°C to 95°C. All reactions were performed in triplicate and the Cq value was determined using the fixed threshold setting. We used three criteria to identify candidate reference genes from the qPCR data: (a) the miRNA must be expressed in all samples; (b) mean Cq of the miRNA must be below 35; (c) there is no evidence for differential expression among the three groups.

Data analysis

Statistical analysis was performed with SPSS 17.0 and Minitab 15 software and $P < 0.05$ was considered statistically difference. One-way ANOVA test was used to compare the expression of candidate reference genes among colorectal adenocarcinoma, adenoma and healthy controls. Average values of triplicate Cq values were converted to relative quantities for GeNorm and Normfinder software to select the most stable reference genes. To calculate the expression of

the target miR-92a-3p relative to suitable reference gene, $2^{-\Delta\Delta C_t}$ method was used. Mann-Whitney U test was used to compare the expression of miR-92a-3p among the three clinical groups.

Results

Selection of candidate reference genes by Miseq sequencing

Sequencing analysis was performed using Illumina's Miseq sequencer and the clean readouts were compared with the miRBase 17.0 database. Of the 740 serum miRNAs that were scanned by sequencing, 348, 303 and 283 miRNAs were detectable (more than 10 copies) in colorectal adenocarcinoma patients, colorectal adenoma patients and healthy controls, respectively. For Miseq sequencing, the final reads of each miRNA in each pooled sample were determined by normalization with total reads of all called miRNAs. There were 17, 32 and 25 miRNAs having more than 50 copies and showing no differential expression among pooling group of colorectal adenocarcinoma and colorectal adenoma, colorectal adenocarcinoma and healthy controls, colorectal adenoma patients and healthy controls, respectively ($P > 0.05$). According to the criteria described in "Material and Methods", a list of 13 miRNAs was considered as candidate reference genes (Table 2).

Confirmation of candidate reference genes by qPCR

The above-mentioned 13 candidate reference genes and U6 were included in further confirmation phase because U6 was frequently used as reference gene in previous studies. According the three criteria described in the method, five miRNAs (miR-151a-3p, miR-4446-3p, miR-221-3p, miR-93-5p and miR-3184-3p) were not detected in all samples that were excluded from further analysis. Using the Cq values, there was no evidence for differential expression of any of the nine candidate reference genes among three groups ($P > 0.05$) (Table 3 and Figure 1). The candidate reference genes displayed a wide expression range, with Cq values between 23.21 and 44.68 (Table 3). Expression of miR-197-3p and miR-26a-5p was relatively low, with median Cq more than 35 that were also excluded from further stability analysis.

Expression stability analysis of selected reference genes by GeNorm and NormFinder

Variable stability of selected reference genes (miR-103b, miR-484, miR-16-5p, miR-3615, miR-18a-3p, miR-191-5p and U6) was evaluated using two algorithms: geNorm and NormFinder. The ranking of reference genes by the two programs is summarized in Table 4. Lower stability values characterize greater stability. GeNorm generates a gene stability value (M) and ranks candidate reference genes according to average expression stability accompanied by stepwise exclusion of the least stable gene until the two most stable genes remained (Figure 2). It also generates a normalization factor (NF) which is used to determine the optimal number of reference genes for normalization. This factor is calculated using the variable V as the pairwise

Table 2. Details of candidate reference genes and target miRNA.

Name	Mature miRNA sequence	Copies			
		colorectal adenocarcinoma	colorectal adenoma	healthy control	
Candidate reference gene	miR-103b	UCAUAGCCCUGUACAAUGCUGCU	10191	9121	10070
	miR-484	UCAGGCUCAGUCCCUCCGGAU	3984	4222	3655
	miR-16-5p	UAGCAGCACGUAUUUUGGCG	3426	3228	3266
	miR-3615	UCUCUCGGCUCUCGCGGCUC	2099	1915	2077
	miR-18a-3p	ACUGCCUAAGUGCUCUUCUGG	1557	1504	1494
	miR-197-3p	UUCACCACUUCUCCACCAGC	808	814	794
	miR-191-5p	CAACGGAAUCCCAAAGCAGCUG	520	524	507
	miR-151a-3p	CUAGACUGAAGCUCUUGAGG	157	167	160
	miR-26a-5p	UUCAAGUAAUCCAGGAUAGGCU	142	133	139
	miR-4446-3p	CAGGGCUGGCAGUGACAUGGGU	113	109	104
	miR-221-3p	AGCUACAUUGUCUGUGGGUUUC	69	65	71
	miR-93-5p	CAAAGUCGUGUUCGUCAGGUAG	52	58	55
	miR-3184-3	AAAGUCUCGCUCUCGCCUCA	53	52	58
Target miRNA	miR-92a-3p	UAUUGCACUUGUCCGGCCUGU	53088	19481	9440

doi: 10.1371/journal.pone.0083025.t002

Table 3. Cq values of candidate reference genes.

Name	Cq Range	Cq Min	Cq Max	Mean Cq±SD			P value
				Colorectal adenocarcinoma	Colorectal adenoma	Healthy control	
miR-103b	10.05	29.34	39.39	33.10±2.78	34.04±3.83	34.29±2.72	0.306
miR-484	11.41	26.98	38.39	32.24±2.63	32.94±3.43	32.50±2.96	0.732
miR-16-5p	10.17	27.48	37.65	32.12±2.63	31.83±3.04	33.20±2.39	0.535
miR-3615	11.89	27.38	39.27	30.98±3.36	32.15±4.02	32.61±3.12	0.242
miR-18a-3p	10.54	29.01	39.55	32.33±3.20	33.82±2.89	34.18±2.15	0.5
miR-197-3p	16.36	27.64	44	37.17±3.89	35.02±3.45	38.25±2.52	0.135
miR-191-5p	3.49	28.37	31.86	30.34±1.02	30.40±5.78	30.59±0.68	0.793
miR-26a-5p	18.9	25.78	44.68	38.89±5.20	37.59±5.99	39.93±2.30	0.588
U6	6.15	23.21	29.36	25.55±1.84	26.34±1.72	27.01±1.40	0.199

doi: 10.1371/journal.pone.0083025.t003

variation (V_n/V_{n+1}) between two sequential NF_s (NF_n and NF_{n+1}). The number of reference genes is considered optimal when the V value achieves the lowest. GeNorm recommended the use of four (miR-191-5p, miR-16-5p, U6 and miR-18a-3p) of the seven most stable genes for optimal normalization. NormFinder and geNorm both identified miR-191-5p as the most stably expressed reference genes and selected miR-191-5p and U6 as the most stable pair of reference gene.

Validation of suitable reference genes in a cohort samples

In order to further validate the stability of miR-191-5p, U6 and combination of these two genes as reference genes, we applied an independent cohort study comprised of 128 colorectal adenocarcinoma patients, 60 colorectal adenoma patients and 60 healthy controls. The expression level of miR-191-5p, U6 and combination of miR-191-5p and U6 (mean) in three groups and four stages of colorectal adenocarcinoma group were shown in Figure 3 and Figure 4. Using the Cq values of each validate reference gene, there

was no evidence for differential expression among the three groups and four stages of colorectal adenocarcinoma.

Effect of suitable reference genes on relative quantity of serum miR-92a-3p

We evaluated miR-92a-3p expression using combination of miR-191-5p and U6 (mean) as reference gene in the serum of colorectal adenocarcinoma, colorectal adenoma and healthy controls. Our results indicated that serum miR-92a-3p was significantly higher in colorectal adenocarcinoma patients than colorectal adenoma patients and healthy controls ($P<0.001$). The difference in miR-92a-3p level remained significant between colorectal adenoma patients and healthy controls ($P<0.001$) (Figure 5). The expression of miR-92a-3p increased with the progress of TNM staging (Figure 6). There was significant difference of serum miR-92a-3p between Stage I and Stage III/IV, between Stage II and Stage III/IV, between Stage III and Stage IV. The difference of serum miR-92a-3p between Stage I and Stage II was not found ($P=0.66$).

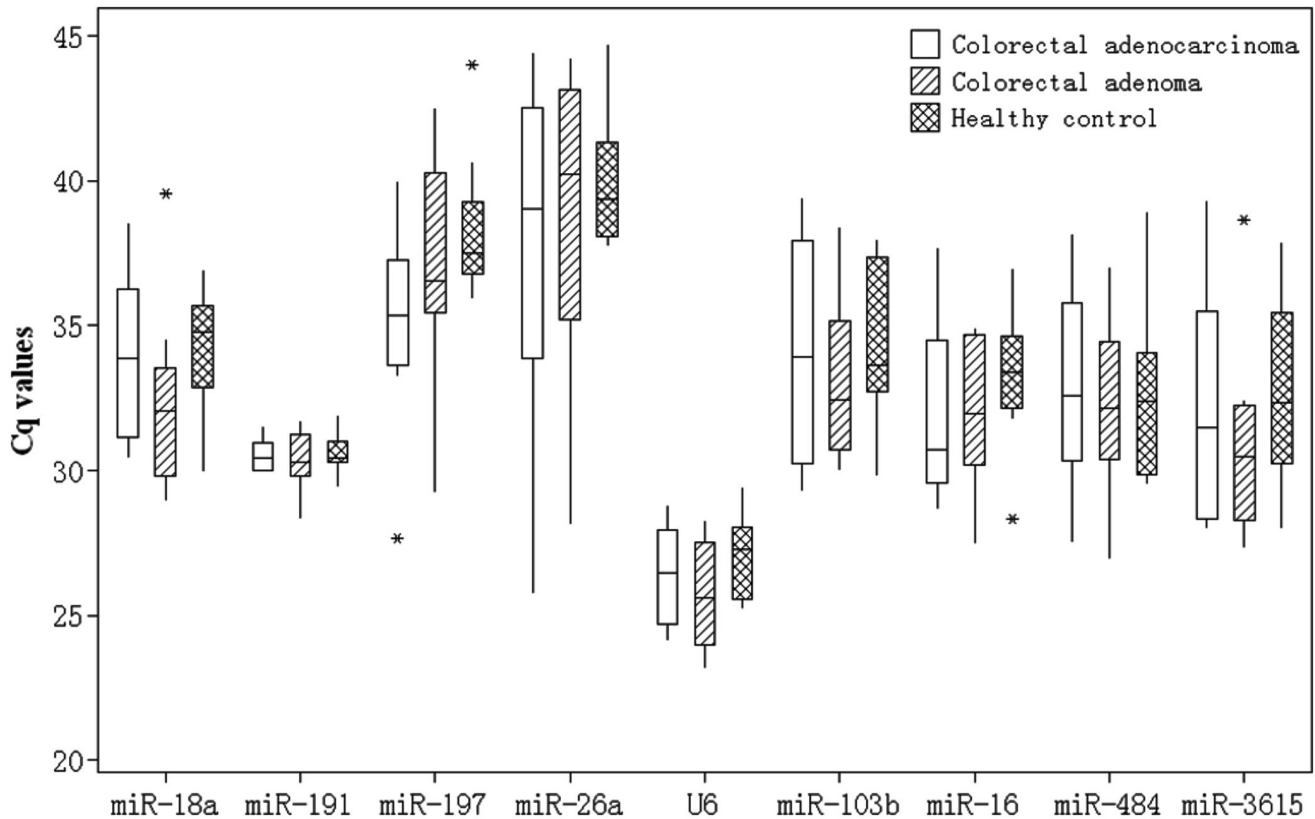


Figure 1. Cq values of candidate reference genes. Cq values of candidate reference genes in colorectal adenocarcinoma, colorectal adenoma and healthy control groups. No significant difference was found among three groups ($P>0.05$, ANOVA).

doi: 10.1371/journal.pone.0083025.g001

Discussion

Serum microRNAs as a class of novel promising non-invasive biomarkers for cancer diagnosis have been demonstrated by many studies [15-17]. They are stable, reproducible, easy to access and potentially cancer specific. General technologies for the analysis of miRNAs expression include chip-based microarrays, next generation sequencing and qPCR. High-throughput methods such as microarray and sequencing enable simultaneous detection of hundreds of miRNAs in a single sample. However, the meaningful interpretation of such large datasets has to be validated by method of qPCR [18,19]. The reliable results of qPCR demand appropriate normalization to minimize variation that can mask or exaggerate meaningful changes. The utility of reference genes as endogenous control is the most common method for normalizing qPCR data of miRNAs expression [20].

In a previous study, miRNA-484 and miRNA-191 were identified as the most stable pair of endogenous control for serum miRNAs detection in breast cancer [21]. Another report recommended the combination of miR-16 and miR-93 as suitable reference genes for serum miRNAs analysis in gastric cancer patients and healthy controls [22]. Kah [11] demonstrated in colorectal tissue that despite a relatively large sample size, when inappropriate reference genes were used

Table 4. Ranking and best combination of candidate reference genes based on expression stability calculated by NormFinder and geNorm.

Rank	NormFinder		geNorm	
	gene	stability	gene	Stability(M)
1	miR-191-5p	0.494	miR-191-5p	0.677
2	miR-18a-3p	0.515	U6	0.715
3	U6	0.528	miR-16-5p	0.742
4	miR-16-5p	0.592	miR-18a-3p	0.854
5	miR-103b	0.605	miR-103b	0.925
6	miR-484	0.706	miR-484	1.156
7	miR-3615	0.792	miR-3615	1.187
Best combination	miR-191-5p/U6	0.363	miR-191-5p/U6	0.277

doi: 10.1371/journal.pone.0083025.t004

for normalization, a true biological difference in expression between tumor and normal was not detected. These results highlight the importance of selecting appropriate reference genes in miRNAs expression study. To our knowledge, this is the first report detailing identification and validation of suitable reference genes for normalization of serum miRNAs qPCR data for colorectal adenocarcinoma.

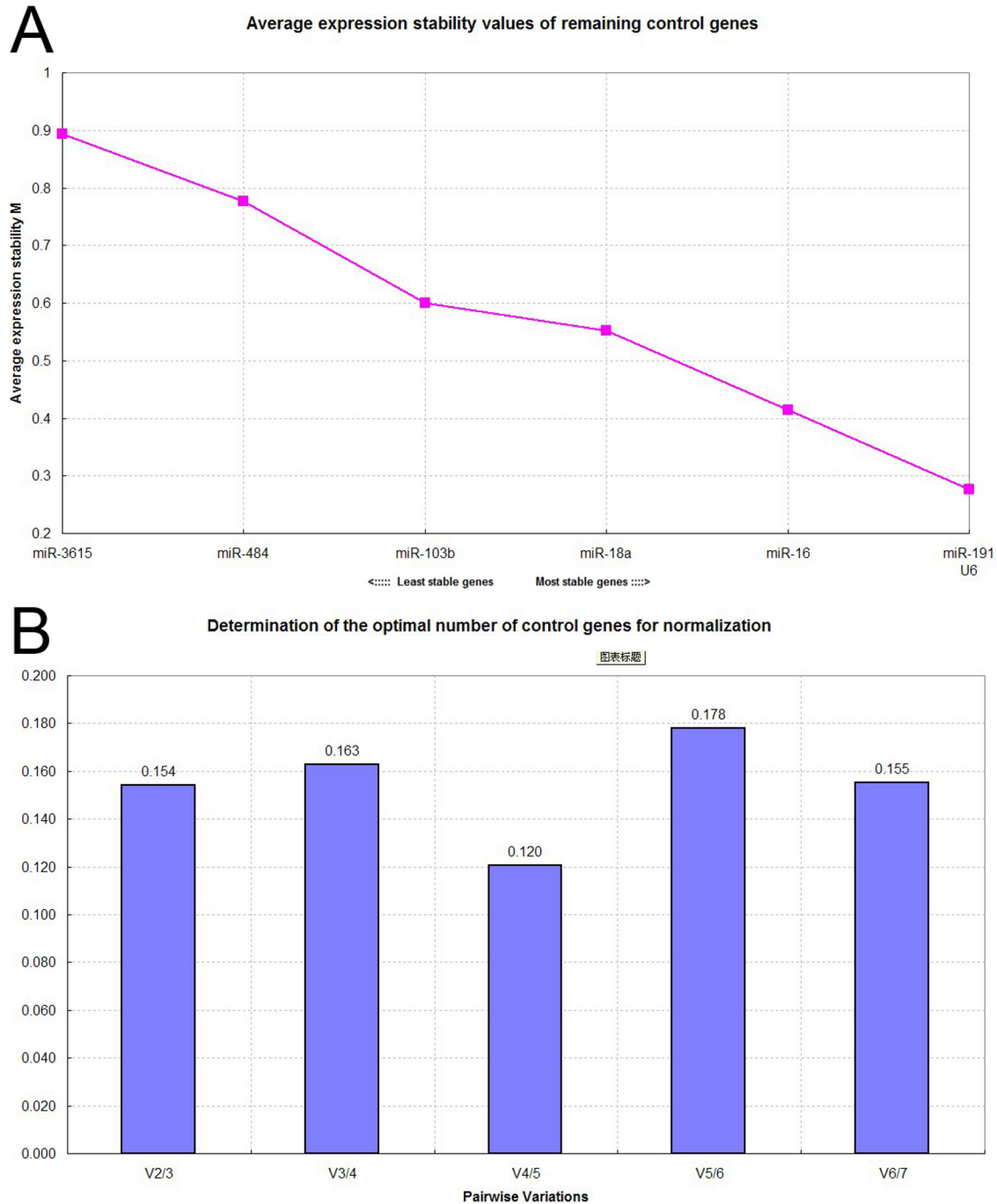


Figure 2. GeNorm analysis of candidate reference genes. (A) Ranking of candidate reference genes according to average expression stability. The least stable gene with the highest stability measure, M were excluded in a stepwise manner until two most stable genes remained: miR-191-5p and U6. (B) Determination of optimal number of reference genes for normalization. The geNorm output file indicated that optimal normalization of gene expression could be achieved using the top four most stable reference genes.

doi: 10.1371/journal.pone.0083025.g002

We firstly profiled pooled serum of colorectal adenocarcinoma, colorectal adenoma and healthy controls and selected a list of 13 miRNAs as candidate reference genes. U6

belongs to a class of small non-coding RNA and cannot be sequenced by Miseq technology in our designed program for miRNAs sequencing. Therefore, U6 didn't appear in the

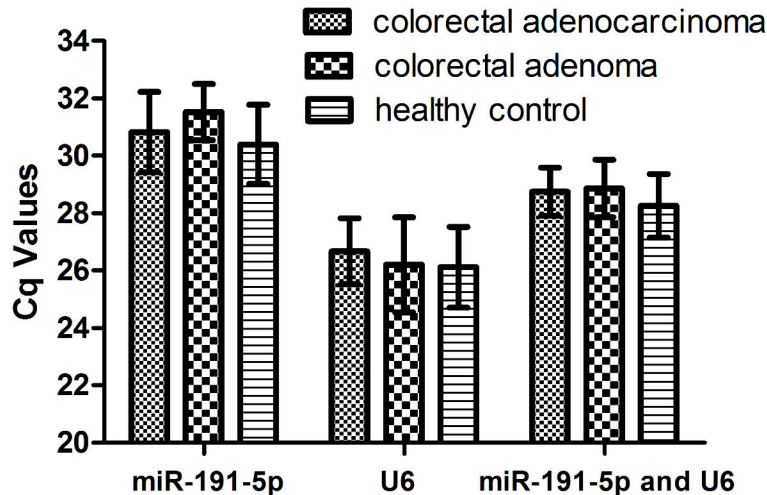


Figure 3. Cq values of miR-191-5p, U6, combination of miR-191-5p and U6 in three groups. No significant difference was found within 3 reference genes among the three groups ($P>0.05$, ANOVA).

doi: 10.1371/journal.pone.0083025.g003

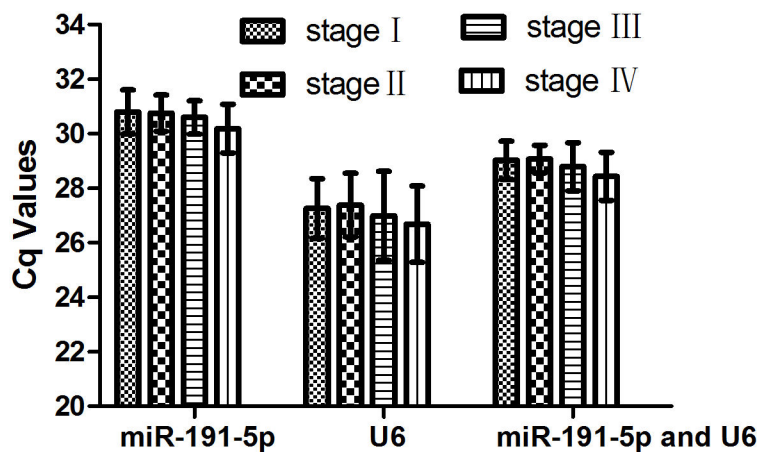


Figure 4. Cq values of miR-191-5p, U6, combination of miR-191-5p and U6 in different stages of colorectal adenocarcinoma group. No significant difference was found within 3 reference genes among the four stages ($P>0.05$, ANOVA).

doi: 10.1371/journal.pone.0083025.g004

screening phase. It was previously used as endogenous control for almost all the tissue miRNAs analysis. However, using U6 to normalize serum miRNAs levels is controversial because it was reported not suitable for serum miRNAs normalization in several studies while chosen as stable reference gene in other studies [23,24]. It was demonstrated that U6 was stable in CRC patients and healthy controls in our preliminary test. Thus, U6 and above-mentioned 13 miRNAs were included in further confirmation by qPCR. According to the three criteria that we have demonstrated in our results, 7 miRNAs (miR-151a-3p, miR-4446-3p, miR-221-3p, miR-93-5p and miR-3184-3p, miR-197-3p and miR-26a-5p) were excluded from further stability analysis. Then variable stability of the other 6 miRNAs (miR-103b, miR-484, miR-16-5p, miR-3615, miR-18a-3p and miR-191-5p) and U6 were evaluated using two

algorithms: geNorm and NormFinder. The geNorm is a pair wise comparison model while NormFinder is an ANOVA-based model. They are two distinct statistical models that have been widely used for selecting optimal reference genes in normalizing gene expression. One important difference between the geNorm and NormFinder model is that NormFinder allows designation of sample groups and determines both intragroup and intergroup variation [25]. GeNorm recommended the use of four most stable genes (miR-191-5p, miR-16-5p, U6 and miR-18a-3p) for optimal normalization. Both NormFinder and geNorm both identified miR-191-5p as the most stably expressed reference gene and selected miR-191-5p and U6 as the most stable pair of reference gene.

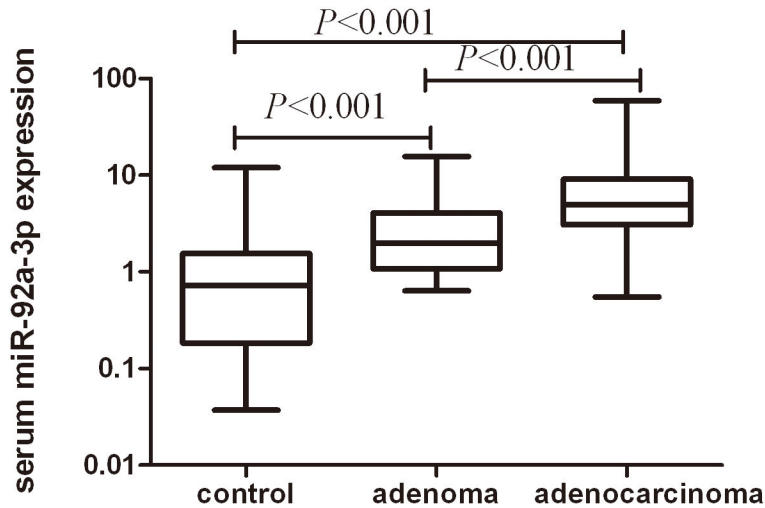


Figure 5. Serum miR-92a-3p expression in three groups. Our results indicated that serum miR-92a-3p was significantly higher in colorectal adenocarcinoma patients than colorectal adenoma patients and healthy controls ($P < 0.001$). The difference in miR-92a levels remained significant between colorectal adenoma patients and healthy controls ($P < 0.001$).

doi: 10.1371/journal.pone.0083025.g005

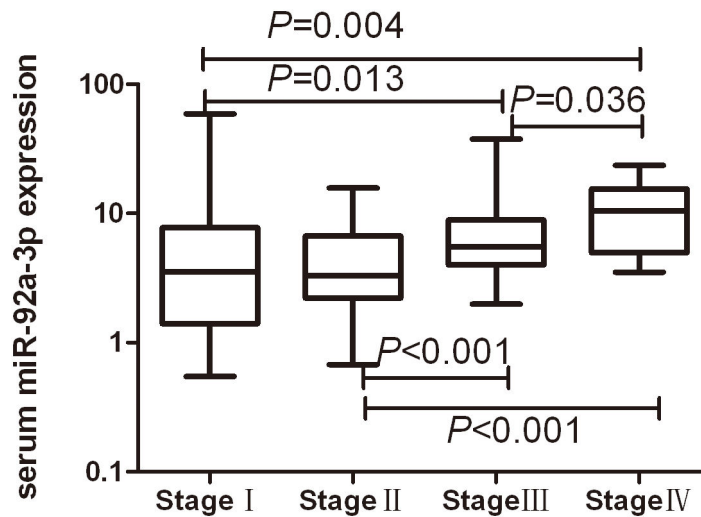


Figure 6. Serum miR-92a-3p expression in in different stages of colorectal adenocarcinoma group. The expression of miR-92a-3p increased with the progress of TNM staging. There was significant difference of serum miR-92a-3p between Stage I and Stage III/IV, between Stage II and Stage III/IV, between Stage III and Stage IV. The difference of serum miR-92a-3p between Stage I and Stage II was not found ($P = 0.66$).

doi: 10.1371/journal.pone.0083025.g006

In order to validate the selected reference genes, an independent large cohort was applied to evaluate the stability of reference genes. Our result demonstrated that there was no evidence for differential expression of miR-191-5p, U6 and combination of miR-191-5p and U6 (mean) among the three groups. Furthermore, serum miRNAs levels can vary highly depending on the tumor stages. Thus, we also evaluated the stability of reference gene in different stages of colorectal adenocarcinoma. The results indicated that there was no evidence for differential expression of miR-191-5p, U6 and

combination of miR-191-5p and U6 (mean) among the four stages.

Several studies evaluated the effect of selected reference genes as endogenous control by analyzing one or certain target miRNAs which has been shown to be upregulated or downregulated in previous reports [22,26]. It has been reported that miR-92a was up-regulated in plasma of patients with colorectal cancer and could be used as a marker for colorectal cancer screening in previous studies [2,27]. Furthermore, in our sequencing result, miR-92a-3p was higher in colorectal

adenocarcinoma group than in colorectal adenoma group ($P=2.964e-323$) and in healthy controls ($P=1.482e-323$). There was also significant difference of miR-92a-3p between colorectal adenoma group and healthy controls ($P=2.47e-323$). Therefore, we selected miR-92a as the target miRNA to test whether it has the similar discrimination effect among healthy, colorectal adenoma, and colorectal adenocarcinoma patients when using combination of miR-191-5p and U6 as reference genes. Our results were not completely consistent with the previous studies. Using combination of miR-191-5p and U6 as reference gene, miR-92a-3p could not only efficiently discriminate healthy controls, colorectal adenoma and colorectal adenocarcinoma patients but also find the minor difference between different stages of colorectal adenocarcinoma which were not found normalized to miR-16 and RNU6B in previous study [2,27].

It has been reported that using of more than one reference genes increases the accuracy of quantization comparing to a

single reference gene. Thus, we propose that combination of miR-191-5p and U6 could be used as reference genes for serum microRNAs qPCR study in patients with colorectal adenocarcinoma, adenoma and healthy controls.

Supporting Information

Figure S1. Workflow chart of the study design. (TIF)

Author Contributions

Conceived and designed the experiments: GXZ CXW. Performed the experiments: GXZ. Analyzed the data: GXZ. Contributed reagents/materials/analysis tools: HYW XZ YMY LLW. Wrote the manuscript: GXZ. Study supervision: LTD JL. Collection of samples: WL ALQ YML.

References

- Oberg AL, French AJ, Sarver AL, Subramanian S, Morlan BW et al. (2011) miRNA Expression in Colon Polyps Provides Evidence for a Multihit Model of Colon Cancer. *PLOS ONE* 6: e20465. doi:10.1371/journal.pone.0020465. PubMed: 21694772.
- Huang Z, Huang D, Ni S, Peng Z, Sheng W et al. (2010) Plasma microRNAs are promising novel biomarkers for early detection of colorectal cancer. *Int J Cancer* 127: 118–126. doi:10.1002/ijc.25007. PubMed: 19876917.
- Bartley AN, Yao H, Barkoh BA, Ivan C, Mishra BM et al. (2011) Complex Patterns of Altered MicroRNA Expression during the Adenoma-Adenocarcinoma Sequence for Microsatellite-Stable Colorectal Cancer. *Clin Cancer Res* 17: 7283-7293. doi: 10.1158/1078-0432.CCR-11-1452.
- Levin B, Lieberman DA, McFarland B, Smith RA, Brooks D, et al. (2008) Screening and surveillance for the early detection of colorectal cancer and adenomatous polyps, 2008: a joint guideline from the American Cancer Society, the US Multi-Society Task Force on Colorectal Cancer, and the American College of Radiology. *CA Cancer J Clin* 58: 130–160.
- Zhang C, Wang C, Chen X, Yang C, Li K et al. (2010) Expression profile of microRNAs in serum: a fingerprint for esophageal squamous cell carcinoma. *Clin Chem* 56: 1871–1879. doi:10.1373/clinchem.2010.147553. PubMed: 20943850.
- Liu R, Zhang C, Hu Z, Li G, Wang C et al. (2011) A five-microRNA signature identified from genome-wide serum microRNA expression profiling serves as a fingerprint for gastric cancer diagnosis. *Eur J Cancer* 47: 784–791. doi:10.1016/j.ejca.2010.10.025. PubMed: 21112772.
- Chen X, Hu Z, Wang W, Ba Y, Ma L et al. (2012) Identification of ten serum microRNAs from a genome-wide serum microRNA expression profile as novel noninvasive biomarkers for non small cell lung cancer diagnosis. *Int J Cancer* 130: 1620–1628. doi:10.1002/ijc.26177. PubMed: 21557218.
- Derveaux S, Vandesompele J, Hellemans J (2010) How to do successful gene expression analysis using real-time PCR. *Methods* 50: 227–230. doi:10.1016/j.ymeth.2009.11.001. PubMed: 19969088.
- Schmittgen TD, Livak KJ (2008) Analyzing real-time PCR data by the comparative C(T) method. *Nat Protoc* 3: 1101–1108. doi:10.1038/nprot.2008.73.
- Genovesi LA, Anderson D, Carter KW, Giles KM, Dallas PB (2012) Identification of suitable endogenous control genes for microRNA expression profiling of childhood medulloblastoma and human neural stem cells. *BMC Res Notes* 5: 507. doi:10.1186/1756-0500-5-507. PubMed: 22980291.
- Chang KH, Mestdagh P, Vandesompele J, Kerin MJ, Miller N (2010) MicroRNA expression profiling to identify and validate reference genes for relative quantification in colorectal cancer. *BMC Cancer* 10: 173. doi:10.1186/1471-2407-10-173. PubMed: 20429937.
- Kreth S, Heyn J, Grau S, Kretzschmar HA, Egensperger R et al. (2010) Identification of valid endogenous control genes for determining gene expression in human glioma. *Neuro Oncol* 12: 570–579. doi:10.1093/neuonc/nop072. PubMed: 20511187.
- Zhu HT, Dong QZ, Wang G, Zhou HJ, Ren N et al. (2012) Identification of Suitable Reference Genes for qPCR Analysis of Circulating microRNAs in Hepatitis B Virus-Infected Patients. *Mol Biotechnol* 50: 49-56. doi:10.1007/s12033-011-9414-6. PubMed: 21567136.
- Asaga S, Kuo C, Nguyen T, Terpenning M, Giuliano AE et al. (2011) Direct serum assay for microRNA-21 concentrations in early and advanced breast cancer. *Clin Chem* 57: 84-91. doi:10.1373/clinchem.2010.151845. PubMed: 21036945.
- Jones CI, Zabolotskaya MV, King AJ, Stewart HJ, Horne GA et al. (2012) Identification of circulating microRNAs as diagnostic biomarkers for use in multiple myeloma. *Br J Cancer* 107: 1987-1996. doi:10.1038/bjc.2012.525. PubMed: 23169280.
- Murakami Y, Toyoda H, Tanahashi T, Tanaka J, Kumada T et al. (2012) Comprehensive miRNA expression analysis in peripheral blood can diagnose liver disease. *PLOS ONE* 7: e48366. doi:10.1371/journal.pone.0048366. PubMed: 23152743.
- Mo MH, Chen L, Fu Y, Wang W, Fu SW (2012) Cell-free Circulating miRNA Biomarkers in Cancer. *J Cancer* 3: 432-448. doi:10.7150/jca.4919. PubMed: 23074383.
- Chen X, Ba Y, Ma L, Cai X, Yin Y et al. (2008) Characterization of microRNAs in serum: a novel class of biomarkers for diagnosis of cancer and other diseases. *Cell Res* 18: 997–1006. doi:10.1038/cr.2008.282. PubMed: 18766170.
- Bargaje R, Hariharan M, Scaria V, Pillai B (2010) Consensus miRNA expression profiles derived from interplatform normalization of microarray data. *RNA* 16: 16-25. doi:10.1261/rna.1688110. PubMed: 19948767.
- Lardizabal MN, Nocito AL, Daniele SM, Ornella LA, Palatnik JF et al. (2012) Reference Genes for Real-Time PCR Quantification of MicroRNAs and Messenger RNAs in Rat Models of Hepatotoxicity. *PLOS ONE* 7: e36323. doi:10.1371/journal.pone.0036323. PubMed: 22563491.
- Hu Z, Dong J, Wang LE, Ma H, Liu J et al. (2012) Serum microRNA profiling and breast cancer risk: the use of miR-484/191 as endogenous controls. *Carcinogenesis* 33: 828–834. doi:10.1093/carcin/bgs030. PubMed: 22298638.
- Song J, Bai Z, Han W, Zhang J, Meng H et al. (2012) Identification of Suitable Reference Genes for qPCR Analysis of Serum microRNA in Gastric Cancer Patients. *Dig Dis Sci* 57: 897–904. doi:10.1007/s10620-011-1981-7. PubMed: 22198701.
- Peltier HJ, Latham GJ (2008) Normalization of microRNA expression levels in quantitative qPCR assays: Identification of suitable reference RNA targets in normal and cancerous human solid tissues. *RNA* 14: 844-852. doi:10.1261/rna.939908. PubMed: 18375788.
- Wang Y, Tang N, Hui T, Wang S, Zeng X et al. (2013) Identification of endogenous reference genes for RT-qPCR analysis of plasma microRNAs levels in rats with acetaminophen-induced hepatotoxicity. *J Appl Toxicol*. [Epub ahead of print] Available online at: [10.1002/jat.2864](https://doi.org/10.1002/jat.2864).

25. Li YL, Ye F, Hu Y, Lu WG, Xie X (2009) Identification of suitable reference genes for gene expression studies of human serous ovarian cancer by real-time polymerase chain reaction. *Anal Biochem* 394: 110-116. doi:10.1016/j.ab.2009.07.022. PubMed: 19622337.
26. Zhu HT, Dong QZ, Wang G, Zhou HJ, Ren N et al. (2012) Identification of Suitable Reference Genes for qPCR Analysis of Circulating microRNAs in Hepatitis B Virus-Infected Patients. *Mol Biotechnol* 50: 49–56. doi:10.1007/s12033-011-9414-6. PubMed: 21567136.
27. Ng EK, Chong WW, Jin H, Lam EK, Shin VY et al. (2009) Differential expression of microRNAs in plasma of patients with colorectal cancer: a potential marker for colorectal cancer screening. *Gut* 58: 1375-1381. doi:10.1136/gut.2008.167817. PubMed: 19201770.