Identification and validation of microRNAs as endogenous controls for quantitative polymerase chain reaction in plasma for stable coronary artery disease

Yuejuan Zhang¹*, Wenxian Tang²*, Ling Peng³, Jinqiang Tang¹, Zhaokai Yuan³

¹Department of Biochemistry and Molecular Biology, College of Medicine, Hunan University of Chinese Medicine, Changsha, Hunan, China
²College of Life Science, Hunan Normal University, Changsha, Hunan, China
³Research Institute of TCM Diagnosis, College of Traditional Chinese Medicine, Hunan University of Chinese Medicine, Changsha, Hunan, China

Abstract

Background: Circulating microRNAs (miRNAs) have been proved to serve as biomarkers for diagnosis and assessment of prognosis of coronary artery disease (CAD). Reverse transcription quantitative polymerase chain reaction (RT-qPCR) is a widely-used technique to estimate expression levels of circulating miRNAs. Selection of optimal endogenous control (EC) remains critical to obtain reliable qPCR data of miRNAs expression. However, reference controls for normalization of circulating miRNA in CAD are still lacking. The purpose of this study was to identify stably expressed miRNAs to normalize RT-qPCR data derived from plasma in stable CAD.

Methods: We identified 10 stably expressed candidate ECs by combining miRNA microarray screening and literature screening. These 10 candidate ECs were estimated by RT-qPCR and the data were analyzed by NormFinder and BestKeeper algorithm.

Results: Two most stable ECs were identified as EC candidates and they were subsequently validated in another larger cohort. The 2 candidates were also validated by normalizing the expression levels of miR-21. In general, they were superior to the commonly used reference gene RNU6 in quantification cycle (Cq) value, stability value and normalization effect.

Conclusions: Our results demonstrated that miR-6090 and miR-4516 can be used as reference genes for plasma miRNA analysis in stable CAD. (Cardiol J 2016; 23, 6: 694–703)

Key words: stable coronary artery disease, circulating microRNA, reverse transcription quantitative polymerase chain reaction (RT-qPCR), reference genes, normalization

Introduction

Coronary artery disease (CAD) remains a major public health issue. The growing social and economic burdens have stimulated the interest of searching novel biomarkers for disease diagnostics, prognostication and monitoring. MicroRNAs (miRNAs) are small non-coding RNA molecules which have key post-transcriptional regulatory roles in many diseases [1]. Dysregulation of miRNAs has

Address for correspondence: Prof. Zhaokai Yuan, Research Institute of TCM Diagnosis, College of Traditional Chinese Medicine, Hunan University of Chinese Medicine, Changsha, Hunan, China, e-mail: yzk20151@163.com

*Contributed equally

Received: 02.03.2016 Accepted: 25.08.2016
been detected in a wide range of diseases, such as cardiovascular disease [2], cancer [3], and tissue injury [4]. In the past few years, scientists have found that miRNA can be rapidly released from tissues into the circulation with the development of pathology, and exist in a highly stable, cell-free form in body fluids such as plasma [5], serum [6], and urine [7]. These findings have inspired a great interest in using extracellular circulating miRNAs as non-invasive biomarkers for molecular diagnostics and prognostics [8].

Reverse transcription quantitative polymerase chain reaction (RT-qPCR) is a widely-used approach to estimate expression levels of circulating miRNAs [9]. Accurate quantitation of miRNA levels is necessary for their use as biomarkers and in determining their functions, which can be affected by sample collection, sample storage, RNA isolation, RT and PCR efficiency. A suitable normalization can minimize measurement errors and technical variability in experiments. Therefore, the choice of normalization method plays an instrumental role and needs to be strongly considered. However, to date, there is no consensus on normalization of circulating miRNA in qPCR amplification.

One of the main challenges in quantifying circulating miRNA expression is that the concentration of total RNA extracted from plasma is often close to, or below, the detection limit of NanoDrop spectrophotometer (< 50 ng), making it difficult to measure and equalize miRNA input levels before analysis by RT-qPCR. Thus, standard methods for measurement of the RNA yield and quality are inappropriate for biofluid samples. Endogenous reference miRNA [10–12], exogenous spike-in artificial synthetic oligonucleotide [13], and the mean global expression [14] have been used to be described in normalization of circulating miRNAs. The selection of suitable endogenous controls (ECs), with which to normalize RT-qPCR data, is currently the most commonly used method for the determination of miRNA expression levels. Circulating miRNAs expression is greatly influenced by the different pathology, so there is no reference miRNA which can be suitable for all types of samples. Thus, researchers have to choose appropriate reference genes according to different study populations. To the best of our knowledge, currently, there is no consensus on reference miRNAs for qPCR analysis of plasma miRNAs in CAD. Herein, we identified and characterized 2 stable reference genes for normalizing plasma miRNAs in stable CAD patients and healthy controls.

Methods

Study cohort and sample collection

Stable CAD patients (50–75 years old) were recruited at the First Hospital of Hunan University of Chinese Medicine, Changsha, China. Blood from healthy participants (50–75 years old) was also collected during physical examinations performed at the First Hospital of Hunan University of Chinese Medicine, Changsha, China. Stable CAD was defined according to 2012 ACCF/AHA practice guideline for stable ischemic heart disease. Patients presented typical angina, history of CAD and/or positive electrocardiography (ECG) stress test, and were confirmed by coronary angiography to have a > 70% coronary artery stenosis. The serum levels of creatine kinase of the subjects were normal. The control subjects had a normal ECG and no history of arterial hypertension, angina, CAD or peripheral artery disease. General exclusion criteria were autoimmune or malignant diseases, acute infections and severe hepatic or renal diseases. Written informed consents were obtained from all participants for the use of their venous blood samples in this study. Ethical approval was granted by the ethics committee of the Hunan University of Chinese Medicine. Relevant clinical characteristics and sample sets were shown in Table 1.

Venous blood samples (5 mL per subject) were collected from each donor into EDTA-containing tubes and processed for plasma isolation within 3 h post-collection. To obtain plasma, samples were centrifuged at 3000 rpm for 10 min at room temperature, followed by 8000 rpm for 10 min at 4°C to completely remove cell debris. The supernatants were transferred into RNase-free tubes, and the plasma samples were stored at −80°C prior to RNA extraction. Plasmas with pink/red discoloration were considered to be hemolyzed and were excluded.

Isolation of total RNA

Total RNA was extracted and purified using miRNeasy® Mini kit (Qiagen, German), following the manufacturer’s instructions. RNA concentrations were determined with a NanoDrop 1000 (Thermo Scientific, Wilmington, USA). In general, RNA concentrations 8–15 ng/μL, and 100 ng of total RNA per sample were used in the subsequent miRNA microarray.

Plasma sample of 300 μL was mixed with TRIzol®LS reagent (Invitrogen, USA) with a ratio of 1:3 in 1.5 mL microcentrifuge tube, and incu-
miRNA microarray hybridization

To assess the expression level of plasma miRNA, microarrays from Agilent Technologies were used according to manufacturer’s protocols.

In brief, 100 ng of total RNA was dephosphorylated with calf intestinal alkaline phosphatase (CIP) at 37°C for 30 min, followed by denaturing with dimethyl sulfoxide (DMSO). Then the 3’ end of dephosphorylated single-stranded RNA (including miRNA) was ligated with cyanine 3-cytidine bis-phosphate (pCp) by T4 RNA ligase. Micro Bio-Spin 6 columns were used to purify the labeled RNA through removing DMSO and free cyanine3-pCp. Human miRNA slides were hybridized with the purified Cy3-labeled miRNA probes in hybridization oven at 55°C, 10 rpm for 20 h. After the hybridization, slides were washed in Agilent Gene Expression wash buffer with Triton X-102. Slides were scanned by Agilent Microarray Scanner and Feature Extraction software10.7 (Agilent Technologies, US) with default settings. Raw data were normalized by Quantile algorithm, Gene Spring Software 12.6 (Agilent Technologies, US).

RT-qPCR analysis

Reverse transcription was carried out using an all-in-one™ miRNA first-strand cDNA synthesis kit (Genecopoeia, USA) according to manufacturer’s protocols. A universal Oligo-dT adaptor primer was used. In brief, the reaction mixture consisted of 5 μL 5 × reaction buffers, 1 μL RTase Mix, 1 μL 2.5 U/μL Poly(A) polymerase, and 18 μL total RNA solution, in a total volume of 25 μL. Mixture was incubated at 37°C for 60 min and 85°C for 5 min.

Quantitative polymerase chain reaction was immediately performed by using an all-in-one™ miRNA qPCR Kit (GeneCopoeia, USA) according to the manufacturer’s instructions. The expression levels of miRNA were confirmed with a SYBR-based qPCR using individual miRNA-
specific primers (GeneCopoeia, USA). The RT-qPCR was performed in triplicate for each miRNA, and a non-template control was included on the same plate. The qPCR reaction mixture consisted of 10 μL 2 x all-in-one qPCR Mix, 0.5 μL all-in-one miRNA qPCR primer, 0.5 μL universal adaptor primer, 0.4 μL 50 × ROX reference dye, 7.6 μL RNase-free Water and 1 μL cDNA. Amplification was performed on the StepOnePlus™ Real-Time PCR System (Applied Biosystems, USA).

**Data analysis**

Statistical analyses were performed with SPSS 17.0 statistical software. The unpaired student’s t-test (for parametric data) or Mann-Whitney U-test (for non-parametric data) was used for two-group comparisons. All tests were performed two-sided and a significance level of p < 0.05 was considered statistically significant. Selection of optimal reference gene was conducted using two commonly used software programs: NormFinder [15] and BestKeeper [16].

**Results**

Detection of the linear relationship between Cq values and plasma volumes in qPCR amplification

To evaluate the range of detectable quantification cycle (Cq) values, we first extracted total RNA from different volumes of plasma ranging from 50 μL to 400 μL (n = 5, volume equal to 50, 100, 200, 300 and 400 μL, respectively). Then we examined miR-6090 in those samples and plotted the regression curve. The RT-qPCR assay showed a good linear relationship (R^2 = 0.9694) when the Cq values were between 25 and 45 (Fig. 1). This result proves the method of measurement was appropriate for this study. We chose 300 μL plasma as the proper volume for subsequent RT-qPCR experiments according to the experiment result.

Literature screening of additional candidate reference genes

We conducted a PubMed search of the published literature to identify traditionally utilized reference genes in miRNA research. We narrowed the search strategy by limiting our consideration of articles by publication date (from 2008 to 2015) and language (English). Keywords included: “microRNA/miRNA” and “reference gene/housekeeping gene/internal control/normalize/normalization”. Finally, one small nuclear RNA (RNU6 [17–19]) and five housekeeping miRNAs (hsa-miR-16-5p [20], hsa-let-7d-5p [10], hsa-miR-484 [21], hsa-miR-191-5p [22], and hsa-miR-423 [23]) were selected as candidate reference genes based on the recording frequency during literature screening. They were all commonly used reference genes in analyzing tissue/cell miRNA. Some reports also indicate that RNU6, miR-484, let-7d, miR-16, and miR-191 can be utilized in some serum samples.

Selection of the most stable reference genes by Agilent miRNA arrays

We screened stable plasma miRNAs by Agilent miRNA arrays. Sixteen plasma samples were analyzed, including 8 healthy and 8 CAD patients. Candidate miRNAs were chosen if they fulfilled the following criteria: (1) express in all samples, share similar properties in patients and controls (p > 0.05); (2) highly expressed, have comparable abundance to other miRNAs (the average hybridization intensity > 9). According to these criteria, 4 miRNAs (hsa-miR-6090, hsa-miR-4516, hsa-miR-6089, hsa-miR-3960) were selected as candidate reference genes (Fig. 2).
Ten candidate reference genes (hsa-miR-6090, hsa-miR-4516, hsa-miR-6089, hsa-miR-3960, hsa-miR-484, hsa-miR-16-5p, hsa-let-7d-5p, hsa-miR-191-5p, hsa-miR-423, RNU6) were further evaluated by RT-qPCR assay in 21 stable CAD patients and 21 healthy controls (Fig. 3). Expression of hsa-let-7d-5p, hsa-miR-191-5p hsa-miR-423 and RNU6 were relatively low, with median Cq from 32 to 39, and even were not detected in some plasma samples. Whereas hsa-miR-484, hsa-miR-3960, hsa-miR-4516 and hsa-miR-6089 were moderately abundant with median Cq of 31.77, 33.23, 31.50 and 29.98, respectively. hsa-miR-6090 and hsa-miR-16-5p were most abundant, with a median Cq of 28.71 and 28.45, respectively. Thus, hsa-let-7d-5p, hsa-miR-191-5p and hsa-miR-423 were excluded from further analysis because of their low expression. Considering RNU6 was widely used in tissue/cell miRNA qPCR experiment, we

**Figure 2.** The average expression values (hybridization intensities ± standard deviations) of the selected miR-NAs were plotted. Significance was calculated by t-test or Mann-Whitney U-test. No differences were found between miRNAs of the patients and healthy controls (p > 0.05).

**Figure 3.** Primary quantification cycle (Cq) values of candidate reference miRNAs in all samples. No significant differences were found between miRNAs of the patients and healthy controls (p > 0.05).
reserved it in the last analysis. The NormFinder and BestKeeper algorithms were performed to rank the rest of candidate reference genes according to their expression stabilities. There was no evidence for differential expression of candidate reference miRNAs between stable CAD patients and healthy controls (p > 0.05).

hsa-miR-6090 and hsa-miR-4516 had the lowest stability values which mean the most stable expression levels and thus were selected as the best combination of reference genes (Table 2). In contrast, hsa-miR-484, hsa-miR-16-5p and RNU6, the most commonly used reference genes for miRNA RT-qPCR normalization, ranked behind hsa-miR-6090 and hsa-miR-4516 (Table 2). The above results indicate hsa-miR-6090 was the most stable reference gene. Moreover, RNU6 was the least stable reference gene.

The identified optimal reference gene, hsa-miR-6090 and hsa-miR-4516, were subsequently validated in a large sample set containing 90 healthy controls and 90 stable CAD patients. As shown in Figure 4, the expression levels of hsa-miR-6090 and hsa-miR-4516 remained constant across CAD patients and controls. No significant differences were detected.

**Table 2.** Evaluation of expression stability of candidate reference miRNAs in stable coronary artery disease patients and healthy controls.

<table>
<thead>
<tr>
<th>Rank</th>
<th>miRNA</th>
<th>NormFinder Stability</th>
<th>BestKeeper Stability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Patients</td>
<td>Controls</td>
</tr>
<tr>
<td>1</td>
<td>miR-6090/4516</td>
<td>0.003</td>
<td>0.005</td>
</tr>
<tr>
<td>2</td>
<td>hsa-miR-6090</td>
<td>0.008</td>
<td>0.004</td>
</tr>
<tr>
<td>3</td>
<td>hsa-miR-4516</td>
<td>0.009</td>
<td>0.014</td>
</tr>
<tr>
<td>4</td>
<td>hsa-miR-6089</td>
<td>0.027</td>
<td>0.015</td>
</tr>
<tr>
<td>5</td>
<td>hsa-miR-3960</td>
<td>0.030</td>
<td>0.019</td>
</tr>
<tr>
<td>6</td>
<td>hsa-miR-484</td>
<td>0.031</td>
<td>0.039</td>
</tr>
<tr>
<td>7</td>
<td>hsa-miR-16-5p</td>
<td>0.040</td>
<td>0.065</td>
</tr>
<tr>
<td>8</td>
<td>RNU6</td>
<td>0.067</td>
<td>0.049</td>
</tr>
</tbody>
</table>

Quantification of plasma miRNAs is significantly influenced by different normalization approaches

We have tested the accuracy of hsa-miR-6090 and hsa-miR-4516 as reference genes by using exogenous miRNA (spiked-in cel-miR-39) as a target. cel-miR-39 was spiked into two groups of plasma samples (2 μL per sample) after the addition of denaturing solution. Then, total RNA was isolated and the relative expression levels of cel-miR-39 were calculated by normalization with hsa-miR-6090, hsa-miR-16-5p, hsa-miR-4516, hsa-miR-484 and RNU6, individually. Normalization with hsa-miR-6090 produced the most consistent results compared with the initial input (Fig. 5). The results indicate that normalization to hsa-miR-6090 is a superior normalization method as it corrects technical variations better than existing methods and achieves accurate identification where others do not.

Furthermore, endogenous miR-21 was selected as targets because they have been shown to be elevated in aortic stenosis patients [24]. We assessed the relative expression levels of miR-21 in the plasma of patients and controls by hsa-miR-6090, hsa-miR-4516, and RNU6, respectively. The miRNA was considered to be significantly up regulated when a mean fold change > 2. Normalization with hsa-miR-6090 or hsa-miR-4516 revealed significant up regulated of miR-21 in plasma from patients compared with controls (Fig. 6). Normalization with RNU6 revealed no significant differences in miR-21 levels in plasma from patients versus controls (Fig. 6).

The results further indicate that different normalization strategies greatly influence the consequence: the use of stable genes for normalization will improve sensitivity and reproducibility, whereas the choice of unstable reference genes can lead to inaccurate results.

**Discussion**

More and more evidence indicate circulating miRNAs play important roles in pathological
Figure 4. Expression levels of miR-6090/miR-4516 in large numbers of plasma samples; A. The expression levels of miR-6090/miR-4516 were measured in the plasma from stable coronary artery disease patients (n = 90) and healthy controls (n = 90). Each point represents the mean of duplicate measurements; PmiR — sample from a patient; CmiR — sample from control subject; CV — coefficient of variation; B. The expression values were showed by mean quantification cycle (Cq) values ± standard deviations; significances were calculated by t-test or Mann-Whitney U-test; no significant differences were found between miRNAs from the patients and healthy controls (p > 0.05).

Figure 5. Effects of different normalization methods on the expression levels of plasma miRNAs. Plasma samples from stable coronary artery disease patients and healthy controls were divided into two groups (n = 5 in each group). 2 μL artificial miRNA cel-39 was spiked into 300 μL plasma of each group. The levels of artificial miRNA were assessed by reverse transcription quantitative polymerase chain reaction and were normalized to miR-16, miR-6090, miR-4516, miR-484, and RNU6, respectively. Significance was calculated by t-test or U test.
Expression levels of miR-21 were measured in plasma from stable coronary artery disease patients (n = 30) and healthy controls (n = 30) by quantitative polymerase chain reaction and were normalized to miR-6090, miR-4516 and RNU6, respectively. Relative levels were calculated using the $2^{-\Delta\Delta Cq}$ method and were presented as mean fold changes ± standard errors. Significance was calculated by t-test and Mann-Whitney U test.

Figure 6. Expression levels of miR-21 were measured in plasma from stable coronary artery disease patients (n = 30) and healthy controls (n = 30) by quantitative polymerase chain reaction and were normalized to miR-6090, miR-4516 and RNU6, respectively. Relative levels were calculated using the $2^{-\Delta\Delta Cq}$ method and were presented as mean fold changes ± standard errors. Significance was calculated by t-test and Mann-Whitney U test.

processes and can serve as candidate biomarkers in multiple conditions. qPCR is an important technology to studying circulating miRNAs. The high sensitivity of qPCR demands a suitable reference gene to correct the non-biological variation.

An ideal endogenous control gene should be expressed at a constant level across all samples, exhibit relatively stable expression levels between samples and groups, and have no known association with the disease under investigation. Traditionally used endogenous reference genes for tissue/cell miRNAs were not suitable to normalize circulating miRNA levels because they are not miRNAs. The efficiency of their extraction, reverse transcription and PCR amplification may differ from that of circulating miRNAs. Some research even indicated that these genes varied significantly in serum and were unstable even after short-term storage [10]. RNU6 is a most commonly used reference gene in tissue/cell samples. But in our study, its expression level in plasma was very low (Fig. 3), and could not be detected in some plasma samples. Its expression level was least stable comparing with other candidate controls (Table 2). These results confirmed recent findings that RNU6 was not stably expressed in plasma/serum samples and is not a suitable control [6, 25]. miR-16 is also a universal reference gene because of its high expression level. But some recent studies reported up-regulation of miR-16 in plasma in some diseases, such as autoimmune thyroid disease [26], experimental sepsis [27] and gastric cancer [28]. Other research demonstrated that miR-16 is highly expressed in red blood cells so that any hemolysis of blood samples strongly affects their level in plasma [29]. In our research, we also found variable expression of miR-16 between individuals, and miR-16 showed the lowest stability in the candidate miRNAs we detected (Table 2). Synthetic non-human (e.g., C. elegans) miRNAs is also usually used as spike-in controls to normalize circulating miRNA. However, spike-in controls can only correct the technical variability in RNA extraction and PCR efficiency, but cannot correct for variability in sample collection and storage [10]. Indeed, the spiked exogenous C. elegans miRNAs, including cel-miR-39, cel-miR-238 and cel-miR-54, did not significantly improve assay precision in some reports [30]. Therefore, spike-in controls are also not an ideal reference gene.

Circulating miRNAs themselves share similar properties, such as stability and purification procedure. So it may be the most suitable method for normalzation. In our study, 10 candidate reference miRNAs were selected based on literature screenimg and miRNA arrays. RT-qPCR results indicated that let-7d, miR-191, miR-423 were not good reference genes because of their low abundances (Fig. 3). The stabilities of the rest miRNAs were further investigated using the NormFinder and BestKeeper methods. Whereas there were slightly different in the result of NormFinder and BestKeeper, miR-6090 was ranked as the most stable reference gene by both NormFinder and BestKeeper (Table 2). The effect of different normalization strategies was compared. When the data were normalized to miR-6090, miR-4516, significant differences were detected between patients and controls. However, when the data were normalized to RNU6, there was no significant difference (Fig. 6).

Our study shows that miR-6090 and miR-4516 could serve as reference miRNAs for qPCR analysis of plasma miRNA in stable CAD patients and healthy controls. Additionally, it is an essential component in accurately analyzed qPCR data.

We noticed that sample set 4 was not very large. While the assessment of a larger number of subjects would clearly be beneficial, the current sample sizes may nevertheless be adequate to address the conclusion that miR-6090 and miR-4516 are better controls than RNU6. A good reference gene should meet three main criteria: 1) be de-
tectable in all samples, 2) exhibit stable levels, and 3) have similar performance in controls and patients. Our results show that miR-6090/4515 meet all these criteria and perform better than RNU6 not only in the experiment from sample set 4, but also in experiments from sample set 2 and 3. Another possible limitation of this study is that we only detected the expression of reference gene in stable CAD. Since other types of CAD, such as acute coronary syndrome, may significantly alter miRNAs release into plasma, they need to be discussed separately. The possibility of miR-6090/4516 use in other types of CAD deserves further exploration in the future. The third limitation is we isolated RNAs by using TRIzol LS reagent. Furthermore, we found miR-6090 performed better (with lower Cq values) by using miRNeasy Mini kit. So, if miR-6090 can be used as reference gene in the future, we recommend extracting RNA by miRNeasy Mini kit.

Conclusions

In conclusion, we found that the traditionally-used EC, RNU6, are not suitable for circulating miRNA normalization because of their low expression (mean Cq value = 34.65) and low stability (0.067 or 1.92 in patients, 0.049 or 2.08 in controls) calculated by NormFinder and BestKeeper algorithm. Finally, our results suggest two miRNAs, miR-6090 and miR-4516, can be used as reference genes for stable CAD plasma miRNA analysis.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Nos. 81302883) and the National Science Foundation of Hunan Province (Nos.14JJ4056). The microarray experiments were performed by Bo Hao Bio-tech, Shanghai, China.

Conflict of interest: None declared

References


