Identification of miR-23a as a novel microRNA normalizer for relative quantification in human uterine cervical tissues

Yuanming Shen[1], Yang Li[1], Feng Ye[1], Fenfen Wang[1], Xiaoyun Wan[2], Weiguo Lu[2] and Xing Xie[2,3]

[1]Women's Reproductive Health Laboratory of Zhejiang Province,
 Women's Hospital
[2]Women's Reproductive Health Laboratory of Zhejiang Province,
Department of Gynecologic Oncology
Women's Hospital
School of Medicine
Zhejiang University
Hangzhou, 310006, China
[3]Corresponding author: Tel, 86-571-87061501-1001;
Fax, 86-571-87036290; E-mail, xiex@zju.edu.cn
DOI 10.3858/emm.2011.43.6.039

Abstract

Quantitative real-time RT-PCR (RT-qPCR) is being widely used in microRNA expression research. However, few reports detailed a robust identification and validation strategy for suitable reference genes for normalisation in microRNA RT-qPCR studies. The aim of this study was to identify the most stable reference gene(s) for quantification of microRNA expression analysis in uterine cervical tissues. A microarray was performed on 6 pairs of uterine cervical tissues to identify the candidate reference genes. The stability of candidate reference genes was assessed by RT-qPCR in 23 pairs of uterine cervical tissues. The identified most stable reference genes were further validated in other cohort of 108 clinical uterine cervical samples: (HR-HPV- normal, n = 21; HR-HPV+ normal, n = 19; cervical intraepithelial neoplasia [CIN], n = 47; cancer, n = 21), and the effects of normalizers on the relative quantity of target miR-424 were assessed. In the array experiment, miR-26a, miR-23a, miR-200c, let-7a, and miR-1979 were identified as candidate reference genes for subsequent validation. MiR-23a was identified as the most reliable reference gene followed by miR-191. The use of miR-23a and miR-191 to normalize expression data enabled detection of a significant deregulation of miR-424 between normal, CIN and cancer tissue. Our results suggested that miR-23a and miR-191 are the optimal reference microRNAs that can be used for normalization in profiling studies of cervical tissues; miR-23a is a novel microRNA normalizer.

Introduction

Uterine cervical cancer is one of the most common cancers in women worldwide, with an estimated global incidence of 470000 new cases and approximately 233000 deaths per year (Munoz *et al.*, 2003; Matsukura and Sugase, 2008). The causal relationship between high-risk human papillomavirus (HR-HPV) infection and cervical cancer has been well documented. However, HPV infection alone is insufficient to induce the malignant transformation, and other unidentified host genetic variations are also involved (Thomison *et al.*, 2008). The identification of such genetic alterations is of considerable importance because altered genes would be candidate targets for screening, diagnosis and treatment of cervical cancer. The microRNA represents an important new class of regulatory biomolecules that play fundamental global roles in human biology, including development, differentiation, apoptosis, metabolism, viral infection and tumor genesis (Wu *et al.*, 2007; Thomison *et al.*, 2008), suggesting that it very likely to play a key role in the occurrence and development of the HPV-induced cervical cancer. In recent years, differences in the miRNome (defined as the full complement of microRNAs in a genome) have been detected in normal versus pathologic cervical tissues or in the same tissues at different stages of differentiation (Martin *et al.*, 2006; Lui *et al.*, 2007; Thomison *et al.*, 2008; Wang *et al.*, 2009; Hu *et al.*, 2010; Pereira *et al.*, 2010). These differences frequently occur in tumor-specific microRNA signatures, which are very helpful to diagnose the origin of the neoplasia and, sometimes, also specific tumor subtypes. Those specific microRNAs are frequently located in cancer-related genomic regions,

which include fragile sites at or near HPV integration sites as well as common breakpoint regions in or near oncogenes or tumour suppressor genes (Bartel, 2004; Esquela-Kerscher and Slack, 2006; Thomison *et al.*, 2008). All of the findings indicate that microRNAs play an important role in cervical carcinogenesis, ranging from HPV infection to cancer. A better understanding of the various tumor-specific microRNA in cervical cancer may not only reveal novel pathogenic mechanism of HPV but also identity potential targets both for cervical cancer screening, diagnosis, and therapy.

To use microRNAs as a biomarker in clinic, it is critical to standardize microRNA testing and make it reproducible in any diagnostic laboratory (Bartel, 2004; Negrini *et al.*, 2007). Over the past several years, a number of high-throughput approaches had been used to quantifying microRNAs in clinical samples, such as Northern blotting, microarrays, a modified invader assay, a bead-based flow cytometric assay, and quantitative real-time PCR (RT-qPCR) (Beezhold *et al.*, 2010). In terms of accuracy and specificity, RT-qPCR has become an increasingly popular tool in microRNA expression detection because it can detect low copy number precursor and mature microRNA with high sensitivity and specificity (Daud and Scott, 2008; Beezhold *et al.*, 2010). However, the accuracy of RT-qPCR gene expression analysis depends largely on a proper normalization (Tricarico *et al.*, 2002; Vandesompele *et al.*, 2002; Andersen *et al.*, 2004; Li *et al.*, 2009). The purpose of normalization is to reduce the technical variation within a dataset, enabling a better appreciation of the biological variation and improving the repeatability among the various laboratories. Selection of optimal normalizers for proper normalization is critical in microRNA RT-qPCR data analysis (Peltier and Latham, 2008). The microRNAs represent perhaps only 0.01% of the mass amount of total RNA in a sample, and this fraction can vary significantly across different samples (Liang *et al.*, 2007). Further, the extraction efficiency of microRNAs from samples is quite different from longer RNAs. In addition, microRNAs expressions have a greater tissue-specific (Bargaje *et al.*, 2010). Thus, microRNAs pose a significant challenge for normalization; the normalizer(s) should not only mirror any wholesale changes in the global microRNAs population but also have similar purification properties as microRNAs.

Unfortunately, up to date, only few candidate reference microRNAs (let-7a, miR-103, miR-16, miR-191, miR-26a, U6 snRNA, and 5S rRNA) have been reported, the selection of reference genes to normalize microRNA levels is still rather empirical and no suitable reference genes have been identified in human cervical tissues to date. This study aimed to select optimal references for normalizing RT-qPCR assays of microRNA expression in cervical tissues and use it to identify the effective microRNAs as diagnostic and prognostic biomarkers of cervical cancer. Firstly, to identify the candidate reference genes, an 875 microRNA expression profiling was done in six pairs of cervical tumor and normal tissues. Following microarray data analysis, the stability of those candidate reference genes was further assessed by RT-qPCR in 46 cervical tissues with an expanded panel of candidate reference microRNAs (miR-103 and miR-191) and two small nucleolar RNAs (U6 snRNA and 5S rRNA). Finally the further validation of the most stable reference gene(s) was subsequently carried out in other larger cohort of 108 clinic cervical tissues, and relative quantities of target miR-424 using four different normalisers (miR-23a, miR-191, u6, and combination of miR-23a and miR-191) were calculated to evaluate the effect of reference genes on relative expression of a target microRNA .

Results

Expression levels of candidate microRNA normalizers

A cohort of 23 cervical normal (HR-HPV negative) and 23 cancer (HR-HPV positive) tissues was randomly selected to validate the chip's data and evaluate the expression patterns of the 9 candidate reference genes. Distribution of data was determined using the Anderson-Darling normality test and parametric tests, all candidate microRNA normalizers showed a normal distribution pattern. Anderson-Darling normality test is a variation of the Kolmogorov-Smirnov (K-S) test that uses a sample's "*P*-value" to measure whether it is "normal." If the P-value is less than the generally-accepted standard of 0.05, the null hypothesis is likely to be false and differences between the samples are likely to exist. The threshold cycle (Cq) is the amplification cycle number at which the fluorescence generated within a reaction rises above a defined threshold fluorescence. Mean Cq values per molecules and the range of Cq values for each molecule were shown in Table 1. Among these genes, U6 snRNA was the most abundant transcript with mean ($\pm$ SD) Cq values of 8.60 $\pm$ 2.38, and let-7a was the lowest expressed gene with mean ($\pm$ SD) Cq values of 23.67 $\pm$ 0.92. The maximum and minimum expression range was 5.71 cycles for miR-200c and 2.04 for miR-23a, respectively. Five to ninth Candidate microRNA normalizers [U6 (*P* = 0.39), 5s (*P* = 0.9), miR-23a (*P* = 0.5), miR-103 (*P*

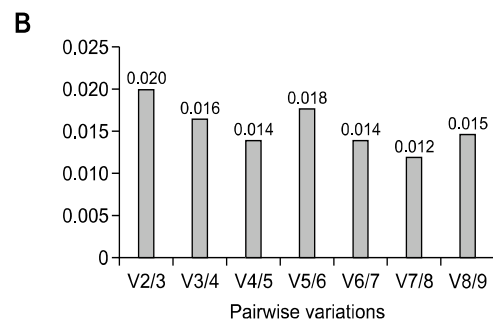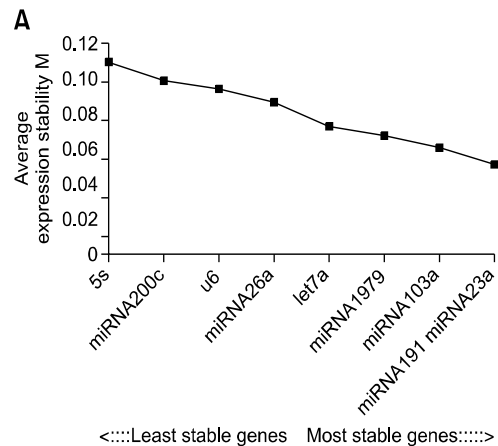**Table 1.** Mean Cq values per molecules and the range of Cq values for each molecules

Gene symbol	Cq Range	Cq Min	Cq Max	Mean Cq $\pm$ sd
miR-200c	5.17	16.07	21.23	18.61 $\pm$ 1.31
miR-26a	4.82	15.98	20.80	17.86 $\pm$ 1.10
miR-23a	2.04	15.96	18.00	16.82 $\pm$ 0.51
miR-1979	3.88	16.55	20.43	18.37 $\pm$ 0.91
miR-191	2.37	18.52	20.90	19.58 $\pm$ 0.53
miR-103	4.05	21.82	25.87	22.98 $\pm$ 0.79
let-7a	3.78	21.88	25.66	23.67 $\pm$ 0.92
U6 snRNA	2.38	7.43	9.81	8.60 $\pm$ 2.38
5S rRNA	4.86	10.20	15.06	12.53 $\pm$ 1.12
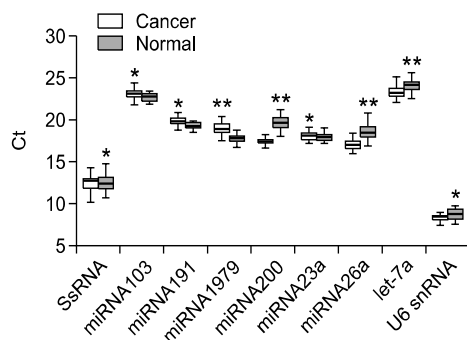
Figure 1. Expression levels of candidate reference genes in normal and cervical cancer samples. Values were given as real-time PCR cycle threshold numbers (Cq values). Boxes (black star, normal; striated, cancer) represented the lower and upper quartiles with medians. No significant difference ($P > 0.05$, t-test) was found within 5 reference genes between tumor and normal tissues (*), while a significant difference ($P > 0.05$, bonferroni correction t-test) was found within 4 reference genes between tumor and normal tissues (**).

Figure 2. GeNorm analysis of the candidate reference molecules. Results were presented according to the output file of the geNorm program. (A) Stepwise exclusion of the least stable molecules by calculating the average expression stability measure M. The value of M was calculated for each molecule, and the least stable molecule with the highest M value was automatically excluded for the next calculation round. The x-axis from left to right indicated the ranking of the reference molecules according to expression stability and the y-axis indicates the stability measure, M. (B) Determination of the optimal number of reference molecules for normalization.

= 0.12), and miR-191 ($P = 0.1$)] were equivalently expressed between the normal and tumor groups, while the expressions of other four Candidate microRNA normalizers (miR-1979, let7a, miR-200C, and miR-26a) were significantly increased in malignant samples (all $P < 0.001$), as shown in Figure 1.

Analysis of candidate microRNA normalizer stability

The stability of candidate microRNA and small RNA reference genes were further analyzed using two statistical analysis tools, geNorm and Norm-Finder. The analysis by geNorm showed that all studied genes reached a high expression stability with low M values, below the default limit of M = 1.5 (Figure 2A). The optimal reference genes for normalizing gene expression data between normal and cancer tissues were miR-23a and miR-191 followed by miR-103. After excluding four candi-

date microRNA normalizers that differed in expression between normal and cancer tissues, the geNorm analysis was repeated. The most stable genes were still miR-23a, miR-191, and miR-103 which were consistent with the previous findings. To determine the suitable number of normalizers required for RT-qPCR date normalization, geNorm calculates the pair wise variation (Vn/Vn + 1) between sequential normalization factors (NF) (NFn and NFn + 1). The combination of miR-23a, miR-191, miR-103, and let-7a was suitable for normalizing gene expression data yielding a V4/5 value of 0.14, which is lower than the cutoff value 0.15 (Figure 2B). Combined miR-23a, miR-191, and miR-103 yield a variation value of 0.16. When samples were independently considered using the NormFinder software, the result appeared to be similar to the one determined by geNorm. Thus, we found that miR-23a achieved the best stability values and the

Table 2. Candidate reference molecules for normalization listed according to their expression stability calculated by the geNorm and NormFinder software.

| Rank	NormFinder		geNorm	
	Gene name	Stability value	Gene name	Stability value
1	miR-23a	0.176	microRNA23a	0.09
2	5S rRNA	0.228	microRNA191	0.092
3	miR-191	0.233	microRNA103	0.093
4	U6 snRNA	0.306	let7a	0.098
5	miR-26a	0.311	microRNA1979	0.112
6	miR-103	0.374	microRNA26a	0.114
7	let7a	0.425	U6 snRNA	0.118
8	miR-200C	0.581	microRNA200C	0.119
9	miR-1979	1.280	5S rRNA	0.147
Best combination	miR-23a and miR-191		miR-23a and miR-191	

best combination of two genes was that of miR-191 and miR-23a (Table 2).

Validation the suitable reference gene(s) in clinical cervical samples

In order to further validate the stability of miR-23a, miR-191, and combination of these two genes in clinical cervical specimens as reference gene, we applied a model comprised of four groups representing the full spectrum of clinic cervical carcinogenesis samples: HR-HPV negative normal tissues (n = 21), HR-HPV positive normal tissues (n = 19), cervical intraepithelial neoplasia tissues (CIN1, n = 11; CIN2-3, n = 36), and invasive cancer tissues (n = 21). There was no overlap between these 108 samples used for the validation experiment and the 60 samples used in the initial experiment. The expression levels of U6, miR-23a, miR-191, and combination of miR-23a and miR-191 in these four groups of clinic cervical samples are shown in Figure 3. Using the Cq values of each validated reference gene, there was no evidence for differential expression between HR-HPV negative normal, HR-HPV positive normal, CIN and cancer tissues ($P < 0.05$; ANOVA). Additionally, the expression of these candidate genes did not depend on age (correlation coefficient = 0.121 to 0.245; $P = 0.102$ to 1.000), histological type (correlation coefficient = 0.103 to 0.312; $P = 0.365$ to 1.000), FIGO stage (correlation coefficient = 0.136 to 0.213; $P = 0.147$ to 0.932), and tumor degree (correlation coefficient = 0.093 to 0.232; $P = 0.254$ to 0.957).
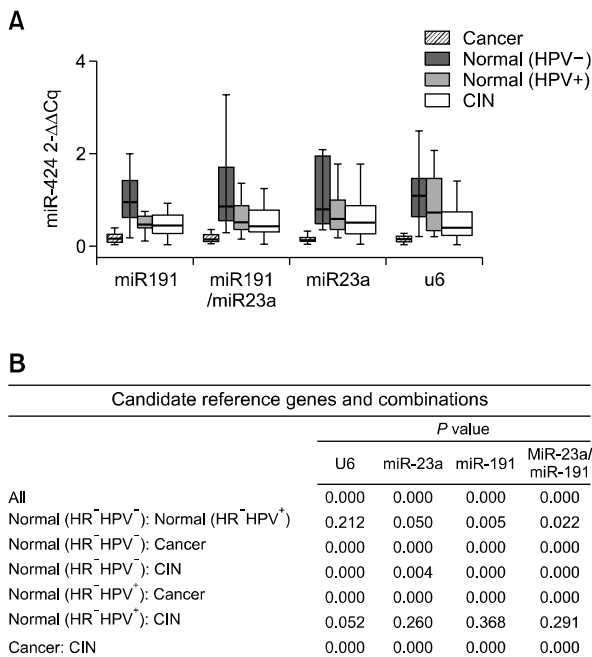
Effect of reference genes on relative expression of target microRNAs

In the microRNA microarray data, the expression of miR-424 is significant difference between HR-
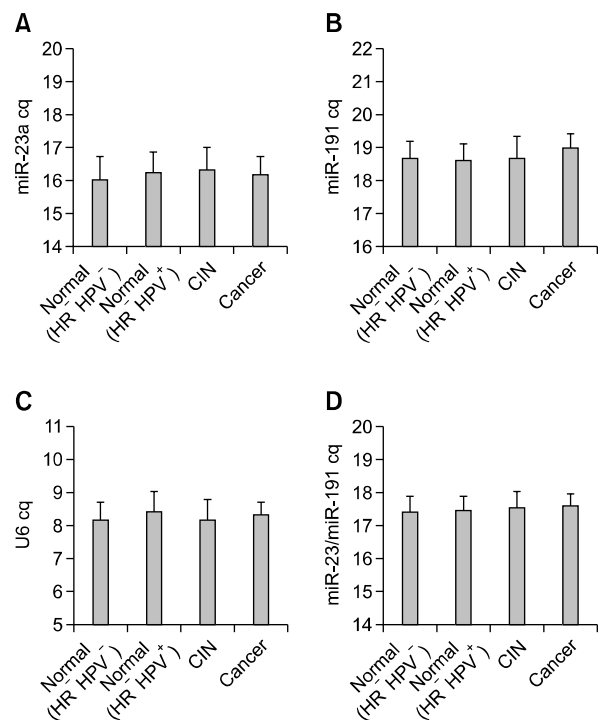
Figure 3. Expression levels of U6, miR-23a, miR-191, and combination of miR-23a and miR-191 in four groups of clinical cervical samples. Values are given as real-time PCR cycle threshold numbers (Cq values). No significant difference ($P > 0.05$, t-test) was found within 4 reference genes (A: miR-23a, B: miR-191, C: U6, D: miR-23a/miR-191) between four groups ($P > 0.05$, ANOVA)

HPV negative cervical normal tissues and HR-HPV positive cancer tissues. This difference expression of miR-424 in different types of cervical tissues suggests that it can be exploited as novel biomarkers for cervical cancer diagnostics and therapy. So, relative quantities of target miR-424 in those full spectrum of clinic cervical carcinogenesis samples (HR-HPV negative normal, HR-HPV positive normal, CIN, and invasive cancer) using

A

B

	P value			
Candidate reference genes and combinations	U6	miR-23a	miR-191	MiR-23a/ miR-191
All	0.000	0.000	0.000	0.000
Normal (HR⁻HPV⁻): Normal (HR⁻HPV⁺)	0.212	0.050	0.005	0.022
Normal (HR⁻HPV⁻): Cancer	0.000	0.000	0.000	0.000
Normal (HR⁻HPV⁻): CIN	0.000	0.004	0.000	0.000
Normal (HR⁻HPV⁺): Cancer	0.000	0.000	0.000	0.000
Normal (HR⁻HPV⁺): CIN	0.052	0.260	0.368	0.291
Cancer: CIN	0.000	0.000	0.000	0.000

Figure 4. Effect of reference gene choice on relative expression of target miR-424. Relative expression of miR-424 in a full spectrum of clinic cervical samples (cancer, n = 21; normal with HR⁻HPV⁻, n = 21; normal with HR⁻HPV⁺, n = 19 and CIN, n = 47) using 4 different normalisers: (A) Values were given as real-time PCR cycle threshold numbers (Cq values). Boxplots depicted median lines, inter-quartile range boxes. Error bar represented range of values. (B) Relative expression of target miR-424 between cervical cancer, normal (HR⁻HPV⁻), normal (HR⁻HPV⁺), and CIN tissues normalised to different reference genes with *P* values indicated.

different normalisers were further studied, the result are shown in Figure 4A with *P* values indicated in Figure 4B. After normalization to four different references (U6, miR-23a, miR-191, and the combination of miR-23a and miR-191), the choice of reference gene(s) did not influence the relative quantity of miR-424 between normal, CIN and cancer tissues, but did affect the relative quantity of miR-424 between HR-HPV negative normal and HR-HPV positive normal cervical tissues. When miR-23a, miR-191 or the combination of miR-23a and miR-191 were used as reference gene(s), a significant difference was detected between HR-HPV negative and HR-HPV positive normal tissues in the expression of miR-424. While when U6 was used as a reference gene, no significant difference was detected between HR-HPV negative and HR-HPV positive normal tissues in the expression of miR-424.

Discussion

Recently accumulated evidences supports a role

for microRNAs as diagnostics and prognostic biomarkers of human cancers (Negrini *et al.*, 2007; Szafranska *et al.*, 2007). When open the way to translating microRNA biology to eventual clinical application, three main aspect should be considered: 1) identified the specific microRNAs as a biomarker; 2) be detected stability in tissue or blood; 3) an efficient and reliable RT-qPCR assays to quantify the expression of microRNAs. The advent of high-throughput techniques made it easy to select the tumor-specific microRNAs as biomark. Additionally, microRNAs are more stable than mRNAs and, therefore, they can be more accurately detected in tissues and in plasma/serum (Negrini *et al.*, 2007). However, microRNA expression has higher tissue specificity, the capability of microRNAs to regulate multiple targets within the same pathway can amplify their biological effects, and even relatively small changes in microRNA expression may be biologically and clinically significant (Calin and Croce, 2006; Szafranska *et al.*, 2007; Bargaje *et al.*, 2010; Beezhold *et al.*, 2010).Thus, veracious detection of microRNA expression is very crucial and microRNA normalizers are the gold standard to normalize the microRNA fraction. RT-qPCR is more quantitative and sensitive than other high-throughput assays which has become an increasingly popular tool in microRNA expression detection. Meanwhile, it is also an optimal tool to validate microarray expression profiles and quantify microRNA of interest identified from those analyses. However, the selection of reference RNA targets to normalize microRNA levels is still rather empirical in microRNAs RT-qPCR assays, and several different kinds of microRNA or small RNA, such as 5S, U6, 18S, miR-16, and let-7a, have been described in microRNA RT-qPCR studies for normalization (Peltier and Latham, 2008; Wang *et al.*, 2009; Bargaje *et al.*, 2010; Zhou *et al.*, 2010). There are few reports on rigorous experimental judgment for those selected targets up to date, to our knowledge. Peltier & Latham et al evaluated systematically normalization targets in microRNA RT-qPCR assays in normal and cancerous human solid tissues and revealed that miR-191 and miR-103 were highly consistent in expression across 13 normal tissues and 5 pairs of distinct tumor/ normal adjacent tissues, and these microRNAs were statistically superior to the most commonly used reference RNAs used in microRNA RT-qPCR experiments, such as 5S rRNA, U6 snRNA, or total RNA (Peltier and Latham, 2008). It was also the first study that described systematic assessment of candidate microRNA normalizers for the normalization of microRNA RT-qPCR data, but it did not include human cervical tissues.

This is the first study detailing identification and validation of suitable reference genes for normalisation of microRNA RT-qPCR in human cervical tissues. Firstly we profiled the expression of 875 microRNAs on 12 cervical tissues (6 HR-HPV negative normal, 6 HR-HPV positive invasive cancer) by a microarray (LC science) and selected the candidate reference genes. The large and constantly growing amount of array-based expression profiling data represents an excellent source for identification of novel candidate normalizers (Bargaje *et al.*, 2010). Using a microRNA microarray scan in the six pairs of cervical tissue, we screened 3 novel candidate microRNA normalizers (miR-23a, miR-200c, and miR-1979) with enhanced abundance and stability among cervical cancer and normal cervical samples, which were not reported before to our knowledge. Secondly, in order to exclude between-run variations, a cohort of 23 cervical normal (HR-HPV negative) and 23 cancer tissues (HR-HPV positive) was randomly selected to validate the chip's data and evaluate the expression patterns of 9 candidate reference genes (5 were derived from the microRNA microarray data and 4 were based on their frequent use as normalizers for microRNA RT-qPCR studies). Both NormFinder and geNorm were used to identify the most stable normalisers. 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 gene in normalizing gene expression. One important difference between the geNorm and NormFinder model is that NormFinder allows designation of sample groups and determines both intra-group and inter-group (i.e., tumor versus normal) variation (Andersen *et al.*, 2004; Li *et al.*, 2009). In the current study, we identified 9 candidate normalizers and found all of them to reach a high expression stability with low M values, below the default limit of M = 1.5. Both geNorm and NormFinder identified miR-23a and miR-191 as the most stable pair of normalisers in cervical tissues. Additionally, a normalization strategy applying multiple reference genes is also suggested by some other authors. In the present study, we found that the combination of miR-23a, miR-191, miR-103, and let-7a was suitable for normalizing gene expression data yielding a V4/5 value of 0.14 and the combination of miR-23a, miR-191 and miR-103 yield a variation value of 0.16 (Figure 2B). Though GeNorm proposes a pair wise variation of 0.15 as the cutoff under which the inclusion of an additional reference gene is unnecessary, the cutoff of 0.15 should not be considered in a strict sense, but rather as guidance to determine the optimal number of reference genes. Sometimes the observed trend can be equally informative, and using the two-three best reference genes is, in most cases, a valid normalization strategy. Thus this study suggested that the combination of miR-191 and miR-23a was best reference genes for microRNA detection in cervical cancer samples. Although all 3 candidate microRNA normalizers reached a high expression stability (M < 1.5), the expressions of miR-200c and miR-1979 were significantly increased in malignant samples with a very low *P* value (*P* < 0.001). Thus, miR-200c and miR-1979 should be excluded and miR-23a was the most stable microRNA for normalizing gene expression data between normal and cancerous cervical tissues.

In the further validation step, the applicability of the two recommended reference genes miR-191, miR-23a, and the combination of them in extended clinic cervical samples were further studied. To elucidate influence of HPV infection and disease type, we applied a larger cohort including four groups (HR-HPV negative normal, HR-HPV positive normal, CIN, and invasive cancer) representing the full spectrum of cervical carcinogenesis. True reference genes are non-functional in the disease process, and should remain stable expression throughout all stages of the disease. Our results showed that miR-191 and miR-23a remained stable expression in the larger cohort of variable type of cervical diseases, suggesting that miR-23a and miR-191 are the suitable reference genes in all clinic cervical lesion samples.

Previously, Kah *et al.* (Chang *et al.*, 2010) demonstrated in colorectal tissue that despite a relatively large sample size, when inappropriate reference genes were used for normalization, a true biological difference in expression between tumor and normal was not detected. In our study, miR-424 are differentially expressed in HR-HPV negative normal versus HR-HPV positive cervical cancer tissues in our microarray profile, providing the role of miR-424 as a diagnostic and prognostic biomarker. So, relative quantities of target miR-424 using different normalisers were further studied to illustrate that inappropriate use of reference genes can significantly influence the results of microRNAs biomarker quantitation. With the use of the miR-191, miR-23a, or the combination of two reference genes, significant difference of target miR-424 was detected between HR-HPV negative and positive normal tissues, while with the use of U6 no significant difference of target miR-424 was detected. These results highlight the importance of selecting appropriate and validated reference genes clinically.

In conclusion, our results in the study indicate that miR-23a, miR-191, and the combination of two genes are the optimal reference microRNAs that can be used as microRNA normalizers for microRNA RT-qPCR detect in clinic cervical tissues, of them; miR-23a is a novel microRNA normalizer. The clinical and pathologically diverse essence of the tissues used in this study should be of interest to a broad spectrum of the cervical cancer research community. A further study is needed to confirm their clinic application in other cancer.

Methods

Patients and samples

This study was approved by the Ethical Committee for Clinical Research of Women's Hospital, School of Medicine, Zhejiang University China. We collected a total of 168 tissue specimens, including 69 normal cervical tissues, 47 cervical intraepithelial neoplasia (CIN) tissues, 50 cervical cancer tissues between September, 2008 and September, 2009 in Women's Hospital, School of Medicine, Zhejiang University. Normal cervical tissues were derived from the patients who underwent hysterectomy due to nonmalignant disease. Cervical cancer tissues were obtained from primary untreated patients under colposcopy. Tumor stage and grade were determined according to the International Federation of Gynecology and Obstetrics standards (FIGO). Clinicopathological data for these patients was showed in Supplemental Data Table 1.

Informed consents were obtained by all patients. All the samples were immediately snap frozen and stored in liquid nitrogen (-180°C). To obtain homogeneous and histological well-characterized samples for RNA analyses, the nature of the tissue and its specified composition were determined by an experienced pathologist. HPV detection was performed using a newly commercially available HPV Geno-Array test kit (Hybribio Limited, Hong Kong). All cancer samples were high-risk human papillomavirus (HR-HPV) positive and contained at least 80% tumor tissues without necrosis. Normal cervical samples including 50 HR-HPV negative and 19 HR-HPV positive were verified to be free of any cervical lesions. CIN samples were verified to contain at least 80% CIN tissues, among them 42 were HR-HPV positive (9 cases of CIN 1 and 33 of CIN 2-3) and 5 were HR-HPV negative (3 of CIN 1 and 2 of CIN 2-3).

RNA extraction and quality

Preserved tissue samples (50-100 mg) were ground into a fine powder while still nitrogen-frozen, and total RNA, containing microRNA was extracted from homogenized tissues using 1 ml TRIZOL$^{\circledR}$ reagent (Invitrogen, Carlsbad, CA) following the manufacturer's instructions. The quantity and purity of total RNA extracted were estimated monitoring both the absorbance at 260 nm and ratios 260/280 nm by Nanodrop (ND-1000UV-Vis Spectrophotometer; NanoDrop Technologies). The quality of RNA was evaluated by 1% agarose gel electrophoresis. Intact rRNA sub-

units (28S and 18S) were observed on the gel indicating minimal degradation of the RNA. The criterion to include RNA samples was 260/280 $\approx$ 2 (1.8 to 2.1) and 28S/18S ratio $\geq$ 1.7.

microRNA microarray

From the liquid nitrogen frozen tissues, we randomly selected cervical normal (HR-HPV negative) and cancer tissues each 6 cases, sent to the United States LC science companies for microRNA isolation, quality control, chip hybridization (μParaflo$^{\circledR}$ Microfluidic Biochip Technology) and microarray data analysis after succeeding in extracting the total RNA .

cDNA synthesis and real-time RT-PCR

The microRNA RT assays used 500 pg total RNA samples, mixed with 2.0 μl 10 $\times$ RT Buffer (TaKaRa Biotechnology), 1.0 μl dNTPs 2.5mM each (TaKaRa), 1.0 μl 5 $\times$ RT Primer (invitrogen), 0.2 μl RNase Inhibitor Protein 40 U/ml (TaKaRa), 0.5 μl wt-MMLV-RT 100 U/μl (Ambion), in a final volume of 10 μl, and incubated on a 48-well GeneAmp PCR System 9700 at 16°C for 30 min, 42°C for 30 min, then 85°C for 5 min. Similarly, U6 snRNA and 5S rRNA small nuclear RNA were quantified using its PCR reverse primer (invitrogen) for RT reaction respectively.

Following the RT step, SYBR green real-time PCR was performed with an Applied Biosystems 7900 HT Fast Real-time PCR system using the SYBR Premix Ex Taq (perfect real time) (TaKaRa Biotechnology). The PCR volume was 20 μl, containing 2 μl cDNA, 10μl SYBR green I, 0.4 μl 50 $\times$ ROX Standard, 1.0 μl 10 $\times$ PCR forward and reverse primer mix (invitrogen). The following cycling began with template denaturizing and hot start Taq activation at 95°C for 10 S, then 40 cycles of 95°C for 10 sec, and 60°C for 30 sec. A melting curve was constructed for each primer pair to confirm product specificity. In chip validation step, all 46 samples were measured in the same run in order to avoid inter-run variations. In the further validation step, inter-run calibrators (IRC) were used to remove the run-to-run difference by normalizing the cq value.

The PCR amplification efficiencies and correlation coefficients (R^2) of each primer pair (Supplemental Data Table 2) were calculated according to standard methods curves using the equation $E = (10-1/slope-1) \times 100$ by serial dilutions. Six serial dilutions (from 1 dilution to 100,000 dilutions) were set up from our sample groups to determine Cq values and reaction efficiencies for all primer pairs. Standard curves were generated for each oligonucleotide pairs using the Cq value versus log input of cDNA.

Putative reference RNA targets and primer design

Nine putative reference RNA targets were selected for investigation to identify the most suitable normalizer in microRNA RT-qPCR studies in cervical cancer (Supplemental Data Table 2). These include 5 microRNAs (miR-23a, miR-200c, miR-26a, miR-1979, and let7a) which were

chosen based on a cervical microRNA microarray expression data, two microRNAs (miR-191 and miR-103) which had be reported to be the most stable reference RNA targets in microRNA RT-qPCR studies (Peltier and Latham, 2008), 5S rRNA (121 nt) and U6 snRNA owing to their purported expression stability and use in several published microRNA RT-qPCR studies (Vandesompele *et al.*, 2002; Peltier and Latham, 2008; Kheirelseid *et al.*, 2010). We filtered the microRNA microarray expression data (form 6 HR-HPV negative cervical normal tissues and 6 HR-HPV positive cancer tissues) according to three criteria: (1) microRNA must be highly expressed in most of the samples; (2) microRNA must be consistently expressed (measured by the modified z-score), with a small coefficient of variation (CV); (3) only one representative microRNA was considered from a given family. Modified z-score = χ-μ / σ (χ refers to the average signal value of microRNAs on the chip in all samples, μ is the average of χ, σ is a standard error of a certain microRNA expression levels in all samples). The CV is a statistic for comparing the degree of variation between microRNA, even if the mean expressions are drastically different from each other. The CV equals the standard deviation divided by the mean (expressed as a percentage). The calculated Modified z-scores and CVs for all microRNA were ranked (Supplemental Data Table 3).

The microRNA target specific stem-loop RT primers and specific PCR forward primers were designed with laser-gene 6.0 software according to the sequences obtained from miRBase database (http://microrna.sanger.ac.uk). The microRNA reverse PCR primers and small RNA (5S rRNA and U6 snRNA) RT and PCR primers were obtained from literatures. Details of primers for each of the 9 evaluated reference RNA targets were shown in Supplemental Data Table 2. The specificity of the primer sequences was confirmed by BLAST searches. The amplification efficiencies of RT-qPCR (E) of 9 primer pairs ranged from 87.8% to 94.9%, and correlation coefficients (R^2) ranged from 0.994 to 0.998.

Amplicon cloning and sequencing

The specificity of each microRNA SYB green assay for the corresponding target was assessed by amplicon cloning followed by DNA sequencing. The PCR amplicons were run on 2.0 ethidium-bromide agarose gel followed by UV-visualization and gel-purification by CSsuper Tissue/Cell Genomic DNA Isolation Kit (Chaoshi Biotechnologies co, shanghai, China). Each PCR amplicon was ligated TA-cloning vectors following manufacturer's instructions. The ligation products were then used to transform DH5alpha competent cell. After transformation, two positive clones for each microRNAs were sequenced using T7 primes and Applied Biosystems technology on ABI Prism 310 DNA sequencer (Applied Biosystems). The obtained sequences were confirmed by BLAST alignment (http:// blast.ncbi.nlm.nih.gov /bl2seq/wblast2.cgi).

Data analysis

Two different gene normalization algorithms were utilized in this work, geNorm and NormFinder. GeNorm (vs. 3.5

available on the internet http://medgen.ugent.be/genorm/) is a Visual Basic Application for Excel based on a pair-wise comparison model. It calculates a gene-stability measure (M) for all candidate reference molecules in a given set of samples and determines the most reliable pair of reference molecules, showing greatest stability of expression ratio across samples. It also calculates a normalization factor (NF) required to determine the optimal number of reference molecules for accurate normalizations. NormFinder, an excel-add-in, based on an ANOVA mathematical model to estimate intra- and inter-group variation and provides a ranking of the tested molecules based on a direct measure of both overall expression variation and the variation between sample subgroups of candidate reference molecules (Li *et al.*, 2009). NormFinder provides the single most stable reference molecule and reference molecule pair where the stability of the latter is greater than that of the single reference molecule. Prior to geNorm and NormFinder analysis, Cq values were converted into Q.rel values (E = $-\Delta$Cq), as detailed above.

Statistical analyses were performed with SPSS 15.0 statistical software, *P* values < 0.05 were considered statistically significant. One-way ANOVA were used to compare calibrator-scaled reference molecules quantities between normal, CIN and malignant tissue groups. The Anderson-Darling normality test and parametric tests were used to determined distribution of data. Difference in variance between genes was assessed using Bartlett's test.

Supplemental data

Supplemental Data include three tables and can be found with this article online at http://e-emm.or.kr/article/article_files/SP-43-6-05.pdf.

References

Andersen CL, Jensen JL, Orntoft 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;64:5245-50

Bargaje R, Hariharan M, Scaria V, Pillai B. Consensus miRNA expression profiles derived from interplatform normalization of microarray data. RNA 2010;16:16-25

Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 2004;116:281-97

Beezhold KJ, Castranova V, Chen F. Microprocessor of microRNAs: regulation and potential for therapeutic intervention. Mol Cancer 2010;9:134

Calin GA, Croce CM. MicroRNA signatures in human cancers. Nat Rev Cancer 2006;6:857-66

Chang KH, Mestdagh P, Vandesompele J, Kerin MJ, Miller N. MicroRNA expression profiling to identify and validate reference genes for relative quantification in colorectal cancer. BMC Cancer 2010;10:173

Daud II, Scott ME. Validation of reference genes in cervical cell samples from human papillomavirus-infected and -uninfected women for quantitative reverse trans-cription-PCR assays. Clin Vaccine Immunol 2008;15:1369-73

Esquela-Kerscher A, Slack FJ. Oncomirs - microRNAs with a role in cancer. Nat Rev Cancer 2006;6:259-69

Hu X, Schwarz JK, Lewis JJ, Huettner PC, Rader JS, Deasy JO, Grigsby PW, Wang X. A microRNA expression signature for cervical cancer prognosis. Cancer Res 2010;70:1441-8

Kheirelseid EA, Chang KH, Newell J, Kerin MJ, Miller N. Identification of endogenous control genes for normalisation of real-time quantitative PCR data in colorectal cancer. BMC Mol Biol 2010;11:12

Li YL, Ye F, Hu Y, Lu WG, Xie X. Identification of suitable reference genes for gene expression studies of human serous ovarian cancer by real-time polymerase chain reaction. Anal Biochem 2009;394:110-6

Liang Y, Ridzon D, Wong L, Chen C. Characterization of microRNA expression profiles in normal human tissues. BMC Genomics 2007;8:166

Lui WO, Pourmand N, Patterson BK, Fire A. Patterns of known and novel small RNAs in human cervical cancer. Cancer Res 2007;67:6031-43

Martin CM, Astbury K, O'Leary JJ. Molecular profiling of cervical neoplasia. Expert Rev Mol Diagn 2006;6:217-29

Matsukura T, Sugase M. Pitfalls in the epidemiologic classification of human papillomavirus types associated with cervical cancer using polymerase chain reaction: driver and passenger. Int J Gynecol Cancer 2008;18:1042-50

Munoz N, Bosch FX, de Sanjose S, Herrero R, Castellsague X, Shah KV, Snijders PJ, Meijer CJ. Epidemiologic classification of human papillomavirus types associated with cervical cancer. N Engl J Med 2003;348:518-27

Negrini M, Ferracin M, Sabbioni S, Croce CM. MicroRNAs in human cancer: from research to therapy. J Cell Sci 2007;120:1833-40

Peltier HJ, Latham GJ. Normalization of microRNA expression levels in quantitative RT-PCR assays: identifi-cation of suitable reference RNA targets in normal and cancerous human solid tissues. RNA 2008;14:844-52

Pereira PM, Marques JP, Soares AR, Carreto L, Santos MA. MicroRNA expression variability in human cervical tissues. PLoS One 2010;5:e11780

Szafranska AE, Davison TS, John J, Cannon T, Sipos B, Maghnouj A, Labourier E, Hahn SA. MicroRNA expression alterations are linked to tumorigenesis and non- neoplastic processes in pancreatic ductal adenocarcinoma. Oncogene 2007;26:4442-52

Thomison JR, Thomas LK, Shroyer KR. Human papillo-mavirus: molecular and cytologic/histologic aspects related to cervical intraepithelial neoplasia and carcinoma. Hum Pathol 2008;39:154-66

Tricarico C, Pinzani P, Bianchi S,Paglierani M, Distante V, Pazzagli M, Bustin SA, Orlando C. Quantitative real-time reverse transcription polymerase chain reaction: normali-zation to rRNA or single housekeeping genes is inappro-priate for human tissue biopsies. Anal Biochem 2002; 309:293-300

Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol 2002;3: RESEARCH0034

Wang X, Wang HK, McCoy JP, Banerjee NS, Rader JS, Broker TR, Meyers C, Chow LT, Zheng ZM. Oncogenic HPV infection interrupts the expression of tumor- suppressive miR-34a through viral oncoprotein E6. RNA 2009;15:637-47

Wu W, Sun M, Zou GM, Chen J. MicroRNA and cancer: Current status and prospective. Int J Cancer 2007;120: 953-60

Zhou X, Chen X, Hu L, Han S, Qiang F, Wu Y, Pan L, Shen H, Li Y, Hu Z. Polymorphisms involved in the miR-218-LAMB3 pathway and susceptibility of cervical cancer, a case-control study in Chinese women. Gynecol Oncol 2010;117:287-90