Genomics 92 (2008) 122–128

Contents lists available at ScienceDirect

Genomics

journal homepage: www.elsevier.com/locate/ygeno

Comparison of normalization methods with microRNA microarray

You-Jia Hua [a,b,c,1], Kang Tu [a,c,1], Zhong-Yi Tang [a,c,1], Yi-Xue Li [a,*], Hua-Sheng Xiao [a,b,*]

[a] Bioinformatics Center, The Center of Functional Genomics, Key Lab of System Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, People's Republic of China
[b] National Engineering Center for Biochip at Shanghai, Shanghai 201203, People's Republic of China
[c] Graduate School of the Chinese Academy of Sciences, Shanghai 200031, People's Republic of China

ARTICLE INFO

Article history:
Received 6 December 2007
Accepted 1 April 2008
Available online 2 June 2008

Keywords:
microRNA microarray
Normalization
Print-tip loess

ABSTRACT

MicroRNAs (miRNAs) are a group of RNAs that play important roles in regulating gene expression and protein translation. In a previous study, we established an oligonucleotide microarray platform to detect miRNA expression. Because it contained only hundreds of probes, data normalization was difficult. In this study, the microarray data for eight miRNAs extracted from inflamed rat dorsal root ganglion (DRG) tissue were normalized using 15 methods and compared with the results of real-time polymerase chain reaction. It was found that the miRNA microarray data normalized by the print-tip loess method were the most consistent with results from real-time polymerase chain reaction. Moreover, the same pattern was also observed in 14 different types of rat tissue. This study compares a variety of normalization methods and will be helpful in the preprocessing of miRNA microarray data.

Introduction

MicroRNAs (miRNAs), a large family of small, ~22-nt, noncoding RNAs, have been identified by cloning or prediction in genomes of dozens of species. Relevant information has been published in a database [1]. MiRNAs regulate a large number of genes in animals and plants. In vertebrates, miRNAs mostly repress the translation of target genes by binding to 3′ untranslated regions, and sometimes cleave the mRNAs of those genes [2,3]. However, in plants, almost all of the miRNAs cleave their target mRNAs, while a few repress transcription [4,5]. MiRNAs are very important regulators of such biological processes as development [6,7], cellular differentiation [8,9], and tumor generation [10,11]. Many techniques have been used to study miRNA expression, such as microarray, RT-PCR [12], Northern blotting [13], and in situ hybridization. MiRNA microarray has been found to be a global analysis tool for detecting miRNA expression. There have been many microarray experiments on the relationship between miRNAs and metabolism, cancer, development, cell fate acquisition, and tissue differentiation; however, in most of these studies, analysis was accompanied by little or no normalization. For example, Liu and Calin et al. [14–16] used the per-chip 50th percentile method to normalize each of their miRNA microarrays on its median; Baskerville and Bartel [17], Liang et al. [18], and Thomson et al. [12] simply performed background signal subtraction on their miRNA microarray data. For the study described here, an established, robust, microarray-based technique [13] was used to measure the expression of 172 miRNAs in DRG after CFA-induced inflammation and 14 rat normal tissues over the time course of DRG inflammation. We chose a number of miRNAs and compared their microarray expression, as normalized using 15 methods, with the real-time PCR data. The results indicate that miRNA microarray data normalized with the print-tip loess method are highly consistent with real-time PCR results.

Results

Rat miRNA microarray development and the data on rat DRG from CFA-induced inflammation model and different normal rat tissues

A rat miRNA microarray was developed that contained 172 rat miRNA precursor sequences and 14 control miRNAs. All probes were 40 nt long, and located close to the 3′ end of each miRNA precursor. Most of the probes contained mature miRNA sequences. For all microarray slides, RNA samples were labeled with Cy5; Cy3-tagged spike-in oligonucleotides were used for internal normalization. The rat miRNA microarray was used to study miRNA expression of rat DRG from complete Freund's adjuvant (CFA)-induced inflammation model animals and normal rat tissues. Two sets of miRNA microarray data were obtained. One comprised 14 rat tissues, and the other included the time course of CFA-induced rat DRG inflammation. Experiments were repeated two and four times, respectively. Real-time PCR was used to validate the miRNA microarray data. A total of eight miRNAs (rno-mir-103-2, rno-mir-128b, rno-mir-135b, rno-mir-140, rno-mir-

Abbreviations: miRNA, microRNA; RT-PCR, reverse transcription polymerase chain reaction; DRG, dorsal root ganglion; CFA, complete Freund's adjuvant.

* Corresponding authors. H.-S. Xiao is to be contacted at National Engineering Center for Biochip at Shanghai, Shanghai 201203, People's Republic of China.
E-mail addresses: yxli@sibs.ac.cn (Y.-X. Li), huasheng_xiao@shbiochip.com (H.-S. Xiao).

[1] Y.-J.H., K.T., and Z.-Y.T. contributed equally to this work.

0888-7543/$ – see front matter.
doi:10.1016/j.ygeno.2008.04.002

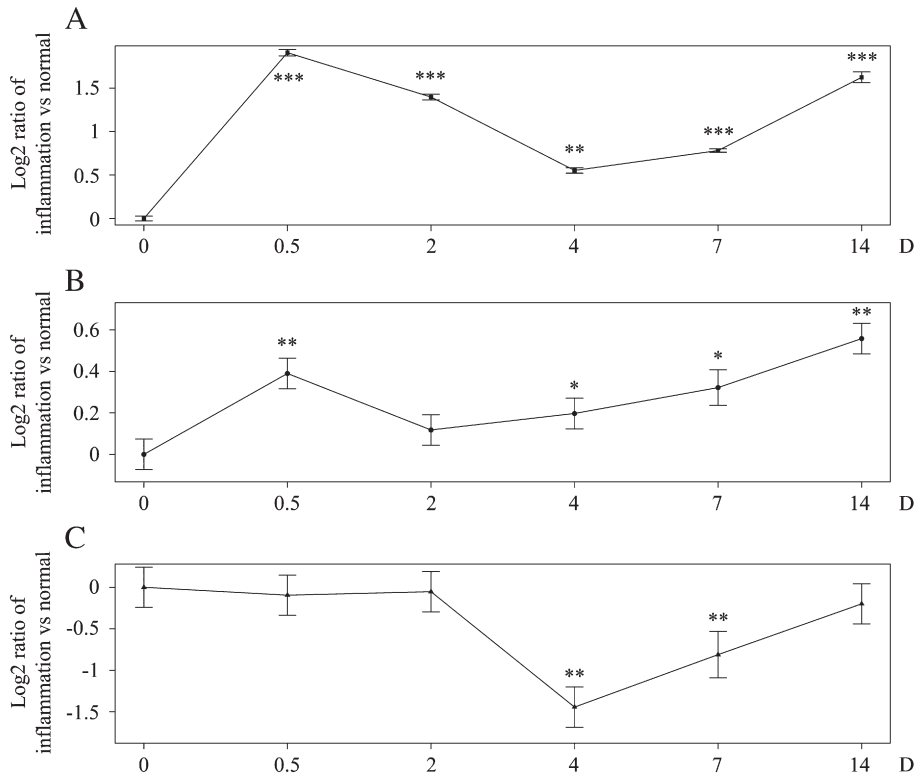

Fig. 1. In the CFA-induced inflammation model, the log 2 ratio of the relative expression level of rno-mir-128b in (A) real-time PCR data, (B) print-tip loess-normalized microarray data, and (C) non-normalized microarray data. *$P<0.05$; **$P<0.01$; ***$P<0.001$.

143, rno-mir-148b, rno-mir-200b, and rno-mir-203) were selected to test the accuracy of microarrays.

After background subtraction, the signal of each miRNA was averaged. Coefficients of correlation between microarray replicates were greater than 0.9. The average signal ranged from 1016 to 2945, and average background ranged from 205 to 308. A probe set with a signal-to-background ratio greater than 3 was considered "present." The present call rate among all the microarrays ranged from 36 to 74%.

Comparison of results obtained using 15 methods for normalization of miRNA microarray data with real-time PCR data

We compared the raw microarray data for the CFA model with real-time PCR data. The results revealed that the correlation between the non-normalized microarray data and the real-time PCR data was quite low (Fig. 1), ranging from −0.66 to 0.54 (Table 1). The raw intensities of the positive and negative controls could not be separated completely by hierarchical clustering (Figs. 2A and C). As shown in Figs. 2B and D, after normalization, positive and negative controls were almost completely separated from each other. This result indicates the importance of appropriate normalization for miRNA microarrays.

Next, we compared the performance of 15 normalization methods, using the real-time PCR data as the "gold standard." Both Pearson and Spearman coefficients of correlation between the normalized microarray data and the real-time PCR results were calculated for each normalization method (Fig. 3). Fig. 3A illustrates that for miRNA-203, Pearson's coefficient of correlation between real-time PCR and microarray data normalized by print-tip loess was the highest. This result was confirmed by the results for all the other miRNAs tested, for which the average correlation coefficient was 0.4 (Fig. 3B). Table 1 lists all Pearson's correlation coefficients. Among the 15 normalization

Table 1
Pearson's correlation coefficients between real-time PCR data and data obtained with 15 normalization methods for eight miRNAs

Method	mir-140	mir-128b	mir-103-2	mir-135b	mir-148b	mir-143	mir-200b	mir-203
Print-tip loess	0.14	0.77	−0.26	0.89	0.14	0.49	0.43	0.66
None	−0.66	−0.03	0.09	0.54	−0.03	−0.26	0.26	0.49
Median	0.09	0.89	0.09	0.54	−0.31	0.26	0.54	0.49
Loess	0.14	0.66	0.60	0.20	0.09	0.37	0.31	0.49
TwoD	0.20	0.71	−0.03	0.83	−0.77	0.37	0.49	0.49
ScalePrintTipMAD	0.03	0.77	−0.26	0.89	0.09	0.49	0.43	0.66
vsn	0.07	0.68	−0.2	0.66	0.09	0.49	0.43	0.66
cy5.none	−0.54	−0.49	−0.31	0.43	−0.37	−0.31	0.31	0.20
cy5.quantiles	−0.20	−0.66	−0.60	−0.09	−0.09	−0.31	0.31	0.43
cy5.qua−ntiles.robust	−0.20	−0.66	−0.60	−0.09	−0.09	−0.31	0.31	0.43
cy5.qspline	−0.37	−0.60	−0.54	0.09	0.14	−0.43	0.43	0.66
cy5.loess	−0.20	−0.66	−0.60	−0.09	−0.09	−0.31	0.31	0.43
cy5.vsn	−0.2	−0.66	−0.54	0.09	−0.09	−0.31	0.31	0.45
cy5. housekeeping	−0.09	−0.43	−0.37	0.14	0.14	−0.20	0.20	0.54
Logratio.housekeeping	0.31	0.43	0.31	−0.43	0.14	0.26	−0.26	−0.37

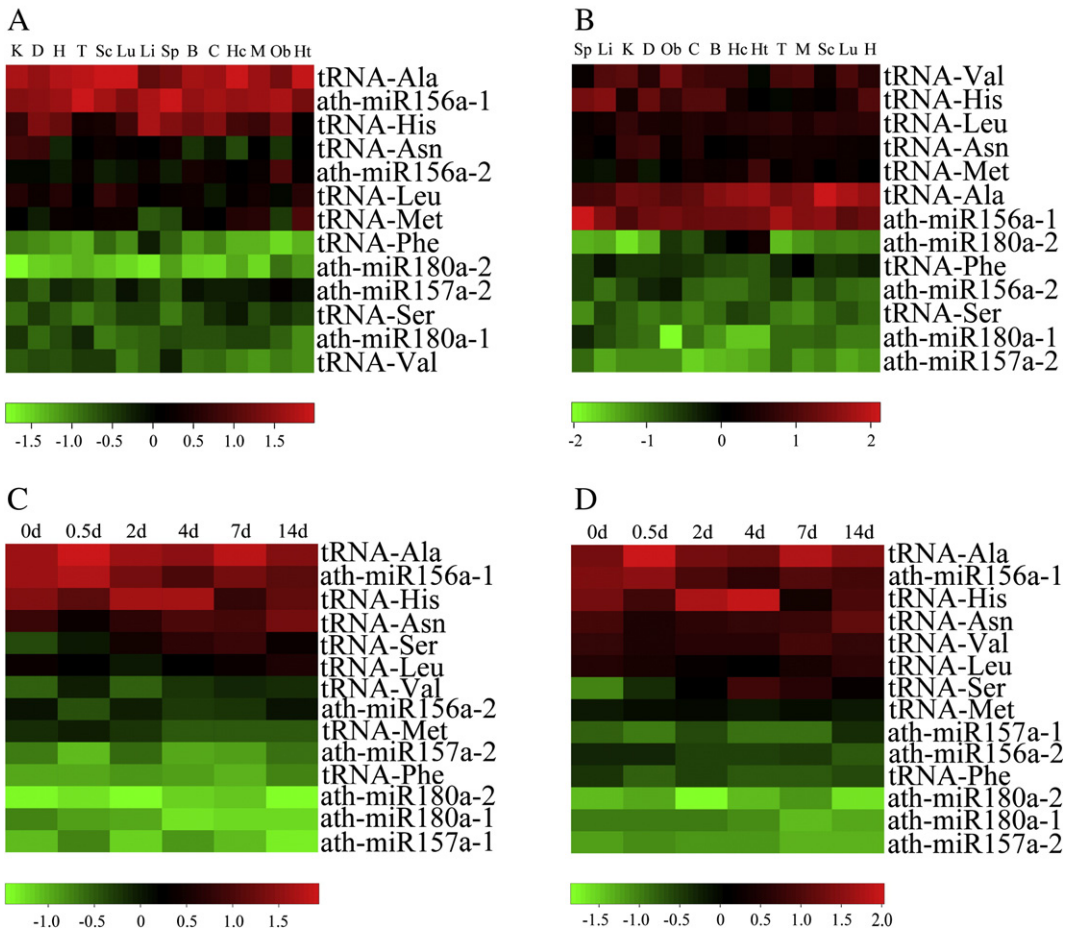

Fig. 2. Clustering of microarray control signals from: (A) raw data in miRNA tissue expression profiles; (B) print-tip loess-normalized data in miRNA tissue expression profiles; (C) raw data for time course of CFA-induced inflammation of DRG; and (D) print-tip loess-normalized data for time course of CFA-induced inflammation of DRG. Red color denotes high expression, and green color denotes low expression. Probes beginning with "tRNA" are positive controls, and probes beginning with "ath" are negative controls. B. brain stem; C, cortex; D, DRG; H, heart; Hc, hippocampus; Ht, hypothalamus; K, kidney; Li, liver; Lu, lung; M, muscle; Ob, olfactory bulb; Sc, spinal cord; Sp, spleen; T, testicle.

methods, 8 were designed for two-channel microarrays and 7 for one-channel microarrays. Fig. 4 illustrates that, on the whole, the two-channel normalization methods were clearly better than the one-channel methods. This means that that Cy3 channel, which consists of spike-in heterogeneous oligonucleotides, is very important for system correlation, and should be used in normalization procedures. As a positive correlation between the Cy3 and Cy5 signals on each spot is generally expected, it may be necessary to use the Cy5/Cy3 ratio instead of raw intensities (Fig. 3). Among the eight two-channel normalization methods, print-tip loess had the highest correlation (Fig. 3 and Table 1). For example, in the CFA model, rno-miR-128b was markedly upregulated, especially on Days 0.5 and 14 after CFA injection, as shown in the print-tip loess-normalized microarray data, as well as in the real-time PCR data (Fig. 1). However, in the non-normalized microarray data, rno-miR-128b appeared to be slightly downregulated, especially on Day 4 (Fig. 1). Details of the technique of print-tip loess normalization are given in Fig. 5. There were a total of six subarrays or blocks (2 rows×3 columns) in each microarray. The three columns were technical triplicates. Each *M* value is normalized by subtracting the corresponding value on the tip-group loess curve from the raw data. The normalized values are the log ratios after subtraction of the residuals of the print-tip loess regression [10], suggesting there was an *M* value excursion with respect to the *A* value for most spots in each microarray before normalization (Fig. 4A), and there was also a two-channel signal system error on each spot with respect to its corresponding block (Fig. 4A). This system error for each block was well eliminated from raw data by print-tip loess (Fig. 4B), and the hypothesis of loess normalization was valid for each print-tip block.

To validate the effect of the print-tip loess normalization method, we analyzed the expression of one miRNA (rno-mir-203), which was measured in 14 rat normal tissues using both microarray and real-time PCR (Fig. 5). Apparently, print-tip loess normalization increased data comparability between the two platforms, as can be seen in Fig. 5. Expression of rno-miR-203 was low in olfactory bulb and heart, among 14 tissues, as indicated by both the print-tip loess-normalized microarray data and the real-time PCR data. However, in the non-normalized microarray data, the miRNA appeared to be highly expressed in these two tissues. This shows that print-tip loess normalization can efficiently correct systemic bias in miRNA microarrays.

Discussion

Microarray is a powerful tool for high-throughput detection of gene and miRNA expression. However, miRNA microarray has some unique characteristics such as much fewer spots, so the normalization methods commonly used for other types of microarrays (e.g., whole-genome gene expression microarray) may not be appropriate. Several articles discussing this problem have been published. The aim of this study was to evaluate a variety of available normalization methods and choose the one that performs best on miRNA microarray.

In the study described in this article, we designed the miRNA microarray probes and labeling method according to Liu [14]. The probes of the miRNA microarray were based on the sequences of

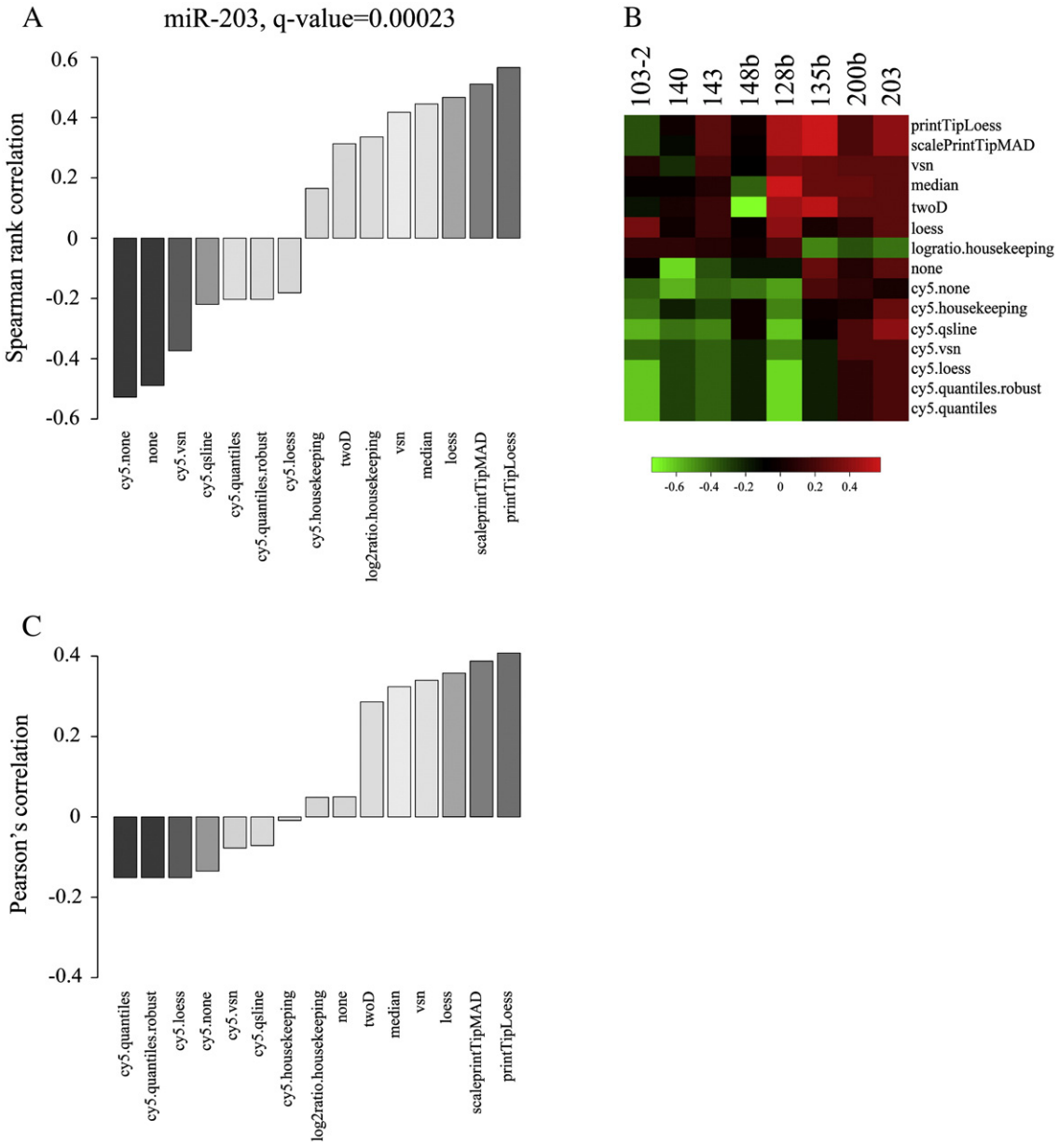

Fig. 3. Spearman's rank correlation coefficients and Pearson's correlation coefficients, which were calculated for the 15 normalization methods (including no normalization) and real-time PCR. (A) Spearman's rank correlation coefficients of rno-mir-203 expression level were sorted by their values. The *x* axis denotes the type of method, and the *y* axis shows the value of each Spearman's rank relative coefficient. (B) Clustering of the Pearson's correlation coefficients of expression level to eight miRNAs in the microarray. (C) Results of sorting the average relative coefficients of all the miRNAs in (B) by their expression level, reflecting the average coincidence between microarray data after normalization and real-time PCR data for eight miRNAs. The *x* axis denotes the normalization method, and the *y* axis shows the average value of the Pearson's correlation coefficients for eight miRNAs.

miRNA precursors, which included mature sites. This means that the microarray could detect precursor and mature miRNAs. Our probes had undergone BLAST alignment to the rat Refseq database, avoiding or reducing nonspecific hybridization to other RNA molecules. Our previous study indicated that mRNA has little cross-hybridization effect on the miRNA microarray [13].

We observed low consistency between non-normalized microarray data and real-time PCR data in this study, suggesting that direct use of microarray data without normalization is unreliable.

We compared 15 normalization methods using microarray data and real-time PCR data. The results for both data sets showed that two-channel data normalization is better than one-channel or no normalization, and also demonstrated that Cy3 channel (signals of spike-in oligonucleotides for internal control) is very important for normalization. This is because unwanted spot effects, such as probe concentration, shape, and size, can be eliminated by using the two-channel intensities together.

There are many normalization methods for two-channel microarray data, such as loess, median, and positive control. Positive control normalization uses the signals of positive controls (also called "housekeeping genes") as a standard for normalization. It is based on the hypothesis that the expression level of each housekeeping gene should be invariable in different tissues or under different environmental conditions. But this hypothesis is not always valid, because the expression of some housekeeping genes may vary in different tissues. The median method adjusts the median value of the Cy5/Cy3 log 2 ratio of all the microarrays to 0. It can eliminate systematic bias in signals between microarrays, but cannot eliminate the bias on each microarray [20]. However, the loess method, which is a nonparametric regression method, can efficiently eliminate the systematic bias in

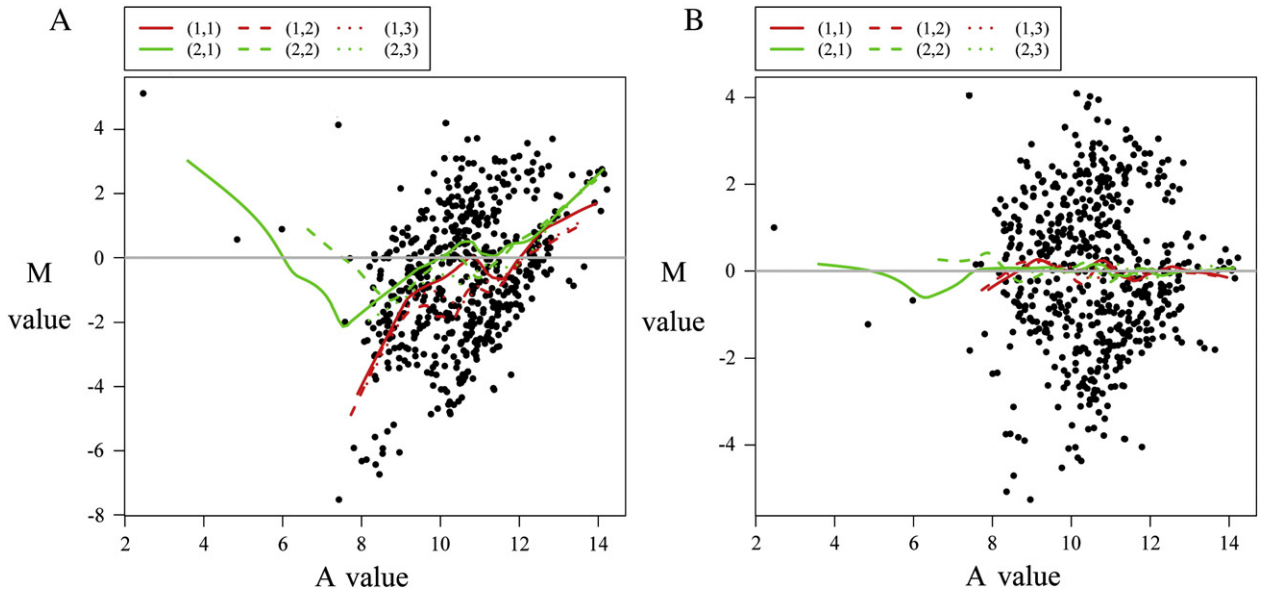

Fig. 4. (A) Before normalization and (B) after print-tip loess normalization. Each spot denotes the *M* value (A) and *A* value (B) of each signal, and each curve denotes the loess regression curve of each block (or subarray) in the array. Six blocks (2×3) were marked as their row number followed by their column number. Then the *M* value of each spot was checked against the regression curve.

signals on each microarray, but is not fit for between-array normalization [20]. Print-tip loess is a well-tested, general-purpose normalization method that has provided good results on a wide range of microarrays [25]. Another improved method, scalePrintTipMAD, theoretically based on scale normalization, has a high requirement for "scale consistency." Despite the characteristics (such as much fewer spots), miRNA microarray is processed in the same way as other oligonucleotide microarrays: fabrication, reverse transcription of

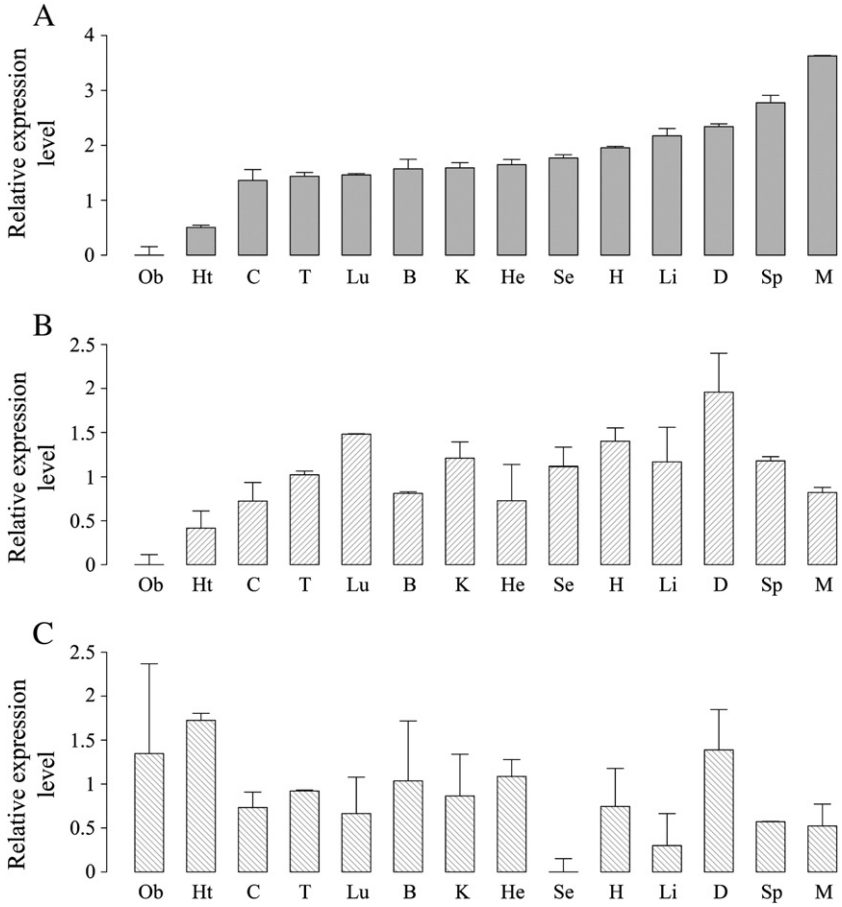

Fig. 5. Relative expression level of rno-mir-203 in rat tissue expression profiles, in (A) real-time PCR data, (B) print-tip loess-normalized microarray data, and (C) non-normalized microarray data. B, brain stem; C, cortex; D, DRG; H, heart; Hc, hippocampus; Ht, hypothalamus; K, kidney; Li, liver; Lu, lung; M, muscle; Ob, olfactory bulb; Sc, spinal cord; Sp, spleen; T, testicle.

samples, and hybridization. Because of its universality, print-tip loess may perform better in miRNA microarray than other methods.

Print-tip loess performed better than all the other normalization methods on our data sets. The fact that print-tip loess is better than the median and loess methods (Fig. 3C) illustrates that miRNA microarray has two characteristics: (1) there is a system excursion of log ratio relative to the *A* value; (2) there is a system excursion with respect to each block. The method of scalePrintTipMAD, which additionally requires "scale consistency" in different print-tip groups, does not have as good an effect as print-tip loess. In general, fewer spots may lead to lower consistency. So this method is not fit for miRNA microarray because of the limited number of probes.

Materals and methods

Tissue preparation and total RNA isolation

A total of 70 adult male Sprague–Dawley rats (body weight, 200–250 g) were used to prepare the DRG tissues from the CFA-induced inflammation model animals. The subcutaneous injection of 200 μL of CFA was made with a sterile tuberculin syringe into the palmar surface of the terminal phalanx of the third digit of the left hindpaw of Sprague–Dawley rats. The rats were allowed to survive 0.5, 2, 4, 7, and 14 days (10 rats per group). Subcutaneous injections and postinjection animal care were carried out in accordance with the policy of the Society for Neuroscience (USA) on the use of animals in neuroscience research and the guidelines of the Committee for Research and Ethic Issues of the International Association for the Study of Pain. The experiments were approved by the Committee of Use of Laboratory Animals and Common Facility, Institute of Neuroscience, Chinese Academy of Sciences. We kept the animals under deep anesthesia for ~1 h after the CFA injection to minimize pain. All animals were kept in a standard environment with close monitoring and postinjection care. Animals with inflammation and 10 normal rats were anesthetized with sodium pentobarbital (60 mg/kg), and the tissues were dissected.

A total of 10 Sprague–Dawley male rats (body weight, 200–250 g) were used to prepare 14 types of normal tissues. Seven neural tissues (olfactory bulb, cortex, hippocampus, brain stem, hypothalamus, spinal cord, and DRG) and seven nonneural tissues (heart, lung, muscle, spleen, testicle, kidney, and liver) were collected from each rat.

Total RNAs of all the samples were extracted with Trizol (Invitrogen, Grand Island, NY, USA) according to the manufacturer's protocol with the following modifications: threefold ethanol was add to the supernatant for precipitation; and after RNA isolation, the washing step with ethanol was not performed.

MiRNA microarray

A rat miRNA microarray was used to profile miRNA expression in DRG and other tissues. A total of 172 rat miRNA precursor sequences with annotated active sites were selected for oligonucleotide design. These sequences corresponded to rat miRNAs published in the miRNA Registry (http://www.sanger.ac.uk/Software/Rfam/mirna; v7.0, accessed July 2005). These miRNA microarrays contain gene-specific oligonucleotide probes generated from 172 rat miRNAs and 14 control miRNAs (8 rat tRNAs for positive control and 6 *Arabidopsis thaliana* miRNAs for negative control). BLAST alignment was performed for all of the sequences with the corresponding genome at http://www.ncbi.nlm.nih.gov, and the hairpin structures were analyzed at http://www.bioinfo.rpi.edu/applications/mfold/old/rna. All probes were 40 nt long, and were dissolved in 150 mM phosphate acid buffer (pH 7.5–8.0). The final concentration of the probes was 25 pmol/μL. Thereafter, a certain concentration of spike-in heterogeneous oligonucleotide sequence was interfused in all solutions, including both probes and controls. Fullmoon Biosystem oligonucleotide slides (Fullmoon Biosystem, Sunnyvale, CA, USA) were used, and the miRNA microarray was fabricated with a GeneMachine OmniGrid 100 Microarrayer (Gene Machine, Rochester, MN, USA) in 1×2-pin and 12×8-spot configurations of each subarray in triplicate. For each microarray, there were six subarrays arranged in two rows and three columns (in triplicate for each probe). The humidity was 75%, and the temperature was 20 °C. After printing, slides were hydrated over night in saturated salt solution, and then UV crosslinked at 600 mJ/cm^2 (CL1000, UVP LLC, Upland, CA, USA).

Ten micrograms of total RNA was added to the reverse transcript reaction mix in a final volume of 11.5 μL, containing 1 μg of [3′-(N)8-(A)3-Cy5-5′] oligonucleotide primer. The mixture was incubated for 10 min at 70 °C and chilled on ice. With the mixture on ice, 2 μL of 10× first-strand buffer, 1 μL of 5 mM unlabeled dNTP mix, 1.5 μL of 1 mM Cy5-dCTP, 1 μL of RNase inhibitor, and 3 μL of SuperScript II RNaseH$^-$ reverse transcriptase (200 units/μL, Invitrogen) were mixed; the final volume was 20 μL. The mixture was incubated for 2 h at 42 °C and then for 10 min at 70 °C. After incubation for first-strand cDNA synthesis, 2 μL of 2.5 N NaOH was added to the first-strand reaction mix and the reaction was incubated at 37 °C for 15 min to denature the RNA/DNA hybrids and degrade RNA templates. Then, 10 μL of 2 N Hepes was added to neutralize the reaction mix. The cDNA targets were purified with the QIAquick Nucleotide Removal Kit (Qiagene, Valencia, CA, USA). The slides were hybridized in 6× SSPE/5× Denhardt with 5 μg Cy3-tagged complementary sequence of spike in heterogeneous oligonucleotide, which would be used as the standard for data normalization at 42 °C for 16 h, and then washed in Lotion I (2× SSC/0.5% SDS) at 42 °C for 15 min, Lotion II (1×SSC/0.1% SDS) at 42 °C for 10 min, Lotion III (0.1× SSC) at room temperature for 5 min, and deionized distilled water at room temperature for 1–2 min. Processed slides were scanned with an Agilent Scanner (Santa Clara, CA, USA) with the laser set to 633 and 545 nm, at power 80 and PMT 100 settings, and a scan resolution of 10 μm.

Real-time quantitative PCR

Real-time quantitative PCR was performed according to standard protocols on an Applied Biosystem 7000 Sequence Detection System (Applied Biosystems, Foster City, CA USA). Five micrograms of total RNA from each sample was reverse transcribed to cDNA. Three microliters of a 1/20 dilution of cDNA in water was added to 12.5 μL of the 2× SYBR green PCR master mix (Applied Biosystems), 0.5 μL of Rox (Applied Biosystems), 5 pmol of each primer, and water to bring the final volume to 25 μL. The reactions were amplified for 15 s at 95 °C and 1 min at 60 °C for 45 cycles. The thermal denaturation protocol was run at the end of the PCR to determine the number of products present in the reaction. U6 snRNA (U6) was used as an internal control. All reactions were run in triplicate and included no template and no reverse transcription as negative controls for each gene. The cycle number at which the reaction crossed an arbitrarily placed threshold (C_T) was determined for each gene, and the relative amount of each miRNA to U6 RNA was described using $2^{-}\Delta^{C}{}_{T}$, where $\Delta C_T = (C_{T\ \mathrm{miRNA}} - C_{T\ \mathrm{U6RNA}})$.

5′ and 3′ primers

rno-mir-203
Forward: 5′-CTGGTCCAGTGGTTCTTAACAGT-3′
Reverse: 5′-GGTCTAGTGGTCCTAAACATTTCA-3′
rno-mir-140
Forward: 5′-TGTCTCTCTCTGTGTCCTGCCA-3′
Reverse: 5′-TATCCTGTCCGTGGTTCTACCCTG-3′
rno-mir-135b,
Forward: 5′-CTGCTGTGGCCTATGGCTTTT-3′
Reverse: 5′-TAGCCCATGGCTTTTAGCCCT-3′;
rno-mir-143
Forward: 5′-CAGTGCTGCATCTCTGGTCAGT-3′
Reverse: 5′-TCCCTTCCTGAGCTACAGTGCT-3′
rno-mir-128b
Forward: 5′-GGCCGATGCACTGTAAGAGAGT-3′
Reverse: 5′-AGACCGGTTCACTGTGAGACCT-3′
rno-mir-103-2
Forward: 5′-CAGCTTCTTTACAGTGCTGCC-3′
Reverse: 5′-GGTTCTTTCATAGCCCTGTACAAT-3′
rno-mir-148b
Forward: 5′-CAGGCACTCTTAGCATTTGAGG-3′
Reverse: 5′-CGAGACAAAGTTCTGTGATGCA-3′
rno-mir-200b
Forward: 5′-CAGCCGTGGCCATCTTACT-3′
Reverse: 5′- CTCCGCCGTCATCATTACC-3′
U6
Forward: 5′-CTCGCTTCGGCAGCACA-3′
Reverse: 5′-AAGGCAGCAGGTCGTATAGT-3′

Data analysis

Our microarrays were hybridized with Cy5-labeled RNA samples and Cy3-tagged spike in oligonucleotide sequence as internal controls, simultaneously. After microarray scanning (Agilent scanner) and image reading (ImaGene), background was subtracted from signal for each spot. As only Cy5 channel signal was related to the experimental aim, both the two-channel normalization methods (using both Cy3 and Cy5) and one-channel methods (using Cy5 only) were tested. Each normalization method was performed by calling corresponding functions in R Bioconductor [19,23]. Two-channel data normalization methods included: global median centering (median) [20], global intensity-dependent location normalization (loess) [20], two-dimensional spatial location normalization (twoD) [20], within-print-tip-group intensity-dependent location normalization (print-tip loess) [20], within-print-tip-group intensity-dependent location normalization followed by within-print-tip-group scale normalization using the median absolute deviation (scalePrintTipMAD) [20], positive control normalization (log ratio.housekeeping), global transformation using variance stabilizing normalization (vsn), and no normalization (none). One-channel data normalization methods included: quantile normalization (cy5.quantiles) [21], cubic splines normalization (cy5.qspine) [22], local polynomial regression fitting normalization (cy5.loess) [23], robust quantile normalization (cy5.quantiles.robust) [23], positive control normalization (cy5.housekeeping), global transformation using variance stabilizing normalization (cy5.vsn), and no normalization (cy5.none). All these methods were evaluated by calculating Pearson and Spearman [24] coefficients of correlation between the normalized microarray data and the real-time PCR data, respectively.

Acknowledgments

We thank Xu Zhang's lab (Laboratory of Sensory System, Institute of Neuroscience, Shanghai Institute Biological Science, Chinese Academy of

Sciences) for providing all of the RNA samples. This work was supported by the "863" program (2006AA020704) and by the "National Basic Research Program of China" (2006CB910700, 2004CB720103, 2004CB518606, 2003CB715901).

References

[1] S. Griffiths-Jones, The microRNA Registry, Nucleic Acids Res. 32 (2004) D109–D111.
[2] P.H. Olsen, V. Ambros, The lin-4 regulatory RNA controls developmental timing in *Caenorhabditis elegans* by blocking LIN-14 protein synthesis after the initiation of translation, Dev. Biol. 216 (1999) 671–680.
[3] K. Seggerson, L. Tang, E.G. Moss, Two genetic circuits repress the *Caenorhabditis elegans* heterochronic gene lin-28 after translation initiation, Dev. Biol. 243 (2002) 215–225.
[4] M.W. Rhoades, et al., Prediction of plant microRNA targets, Cell 110 (2002) 513–520.
[5] G. Tang, B.J. Reinhart, D.P. Bartel, P.D. Zamore, A biochemical framework for RNA silencing in plants, Genes Dev. 17 (2003) 49–63.
[6] B.J. Reinhart, et al., The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*, Nature 403 (2000) 901–906.
[7] A.L. Abbott, et al., The let-7 MicroRNA family members mir-48, mir-84, and mir-241 function together to regulate developmental timing in *Caenorhabditis elegans*, Dev. Cell Biol. 9 (2005) 403–414.
[8] C.Z. Chen, L. Li, H.F. Lodish, D.P. Bartel, MicroRNAs modulate hematopoietic lineage differentiation, Science 303 (2004) 83–86.
[9] M. Lagos-Quintana, R. Rauhut, W. Lendeckel, T. Tuschl, Identification of novel genes coding for small expressed RNAs, Science 294 (2001) 853–858.
[10] C.Z. Chen, MicroRNAs as oncogenes and tumor suppressors, N. Engl. J. Med. 353 (2005) 1768–1771.
[11] L. He, et al., A microRNA polycistron as a potential human oncogene, Nature 435 (2005) 828–833.
[12] T.D. Schmittgen, J. Jiang, Q. Liu, L. Yang, A high-throughput method to monitor the expression of microRNA precursors, Nucleic Acids Res. 32 (2004) e43.
[13] J.J. Zhao, et al., Genome-wide microRNA profiling in human fetal nervous tissues by oligonucleotide microarray, Childs Nerv. Syst. 22 (2006) 1419–1425.
[14] C.G. Liu, et al., An oligonucleotide microchip for genome-wide microRNA profiling in human and mouse tissues, Proc. Natl. Acad. Sci. U. S. A. 101 (2004) 9740–9744.
[15] G.A. Calin, et al., MicroRNA profiling reveals distinct signatures in B cell chronic lymphocytic leukemias, Proc. Natl. Acad. Sci. U. S. A. 101 (2004) 11755–11760.
[16] G.A. Calin, et al., Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers, Proc. Natl. Acad. Sci. U. S. A. 101 (2004) 2999–3004.
[17] S. Baskerville, D.P. Bartel, Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes, RNA 11 (2005) 241–247.
[18] R.Q. Liang, et al., An oligonucleotide microarray for microRNA expression analysis based on labeling RNA with quantum dot and nanogold probe, Nucleic Acids Res. 33 (2005) e17.
[19] R Development Core Team, R: A language and environment for statistical computing, R Foundation for Statistical Computing, 2006.
[20] Y.H. Yang, S. Dudoit, P. Luu, T.P. Speed, Normalization for cDNA microarray data.
[21] B.M. Bolstad, R.A. Irizarry, M. Astrand, T.P. Speed, A comparison of normalization methods for high density oligonucleotide array data based on variance and bias, Bioinformatics 19 (2003) 185–193.
[22] C. Workman, et al., A new non-linear normalization method for reducing variability in DNA microarray experiments, Genome Biol. 3 (2002) (research0048).
[23] R.C. Gentleman, et al., Bioconductor: open software development for computational biology and bioinformatics, Genome Biol. 5 (2004) R80.
[24] M. Hollander, D.A. Wolfe, Nonparametric statistical inference, 1973.
[25] G.K. Smyth, T. Speed, Normalization of cDNA microarray data, Methods 31 (2003) 265–273.