

Global analysis of microRNA target gene expression reveals that miRNA targets are lower expressed in mature mouse and *Drosophila* tissues than in the embryos

Zhenbao Yu*, Zhaofeng Jian¹, Shi-Hsiang Shen, Enrico Purisima¹ and Edwin Wang^{1,*}

Mammalian Cell Genetics Group and ¹Computational Chemistry & Biology Group, Biotechnology Research Institute, National Research Council Canada, Montreal, Quebec, H4P 2R2, Canada

Received June 29, 2006; Revised September 27, 2006; Accepted October 19, 2006

ABSTRACT

MicroRNAs (miRNAs) are non-coding small RNAs of ~22 nt that regulate the gene expression by base pairing with target mRNAs, leading to mRNA cleavage or translational repression. It is currently estimated that miRNAs account for ~1% of predicted genes in higher eukaryotic genomes and that up to 30% of genes might be regulated by miRNAs. However, only very few miRNAs have been functionally characterized and the general functions of miRNAs are not globally studied. In this study, we systematically analyzed the expression patterns of miRNA targets using several public microarray profiles. We found that the expression levels of miRNA targets are lower in all mouse and *Drosophila* tissues than in the embryos. We also found miRNAs more preferentially target ubiquitously expressed genes than tissue-specifically expressed genes. These results support the current suggestion that miRNAs are likely to be largely involved in embryo development and maintaining of tissue identity.

INTRODUCTION

MicroRNAs (miRNAs), encoded in the chromosomal DNA and transcribed as longer stem-loop precursors, termed pri-miRNAs, are non-coding small (21–23 nt) RNAs that regulate the expression of target mRNAs [reviewed in (1–4)]. Upon transcription, pri-miRNA is converted to mature miRNA duplex through sequential processing by RNase III family of endonucleases Droscha and Dicer (3,4). One strand of the processed duplex is incorporated into a silencing complex and guided to target sequences by base pairing [reviewed in (5,6)]. This results in the cleavage of target mRNAs or repression of their productive translation (5,6).

In the past few years, several hundred miRNAs were identified in animals and plants. It is currently estimated that miRNAs account for ~1% of predicted genes in higher eukaryotic genomes (7).

Despite the large number of identified miRNAs, only a handful of them have been functionally characterized. For example, lin-4 and let-7 regulate the timing of larval development in *Caenorhabditis elegans* (8,9). Lsy-6 and miR-273 act sequentially to control the left/right asymmetric gene expression in *C.elegans* chemosensory neurons (10). Bantam promotes cell proliferation and inhibits apoptosis in *Drosophila* (11). MiR-14 suppresses cell death and regulates fat metabolism (12). MiR-181 potentiates B-cell differentiation (13). These findings, together with the complicated expression patterns and large number of predicted targets, imply that miRNAs may regulate a broad range of physiological and developmental processes.

Identification of the targets of each miRNA is crucial for understanding the biological function of miRNAs. Accumulating empirical evidence has revealed the importance of the 5-terminal segment of miRNAs with 6–8 nt in length, called ‘seed’ region, for miRNA function (14–17). For example, systematical single nucleotide mutation studies demonstrated that base pairing of miRNAs to their targets with 7 nt at the 5-terminus of miRNAs from position 2 to position 8 is essential and sometimes sufficient for miRNAs to knockdown their target expression (14). Based on these discoveries, several computational methods have been developed to search for miRNA targets (18–27). Most of these methods have been biologically validated and proved to be very efficient and accurate. The accuracy of these methods has also been proved by large scale gene expression profile studies (28,29). In one study, Lim *et al.* (28) reported that transfections of miR-1 and miR-124 into HeLa cells, respectively caused down-regulation of large numbers of target mRNAs and majority (76 and 88%, respectively) of down-regulated mRNAs showed a segment with 6 nt complementary to the 5'-terminus of the transfected miRNAs

*To whom correspondence should be addressed. Tel: +514 496 6377; Fax: +514 496 6319; Email: zhenbao.yu@nrc.ca

*Correspondence may also be addressed to Edwin Wang. Tel: +514 496 0914; Fax: +514 496 5143; Email: edwin.wang@cnrc-nrc.gc.ca

(the 'seed' sequence). In another study, Krutzfeldt *et al.* (29) demonstrated that knockdown of miRNA-122 by intravenous administration of miRNA 'antagomirs' led to up-regulation and down-regulation of a large number of genes in liver. They found that the 3'-untranslated regions (3'-UTRs) of up-regulated genes are strongly enriched in miRNA-122 'seed'-match motifs, whereas down-regulated genes are depleted in these motifs (29).

These methods have yielded a large number of candidate targets in both plants and animals. The estimated human miRNA targets can account for up to one-third of human genes (23). The diversity and abundance of miRNA targets reflect that miRNAs and their targets appear to form a complex regulatory network. For example, a single miRNA can regulate hundreds of mRNAs and a single mRNA can be targeted by several different miRNAs.

Systematical analysis of gene expression profiles has been proved to be valuable for studies on diverse biological processes (30–33). Based on its biochemical function, the biological functions of a miRNA should depend on the combination of its action to each of all its targets for their expression. Accordingly, Farh *et al.* (34), Stark *et al.* (35) and Sood *et al.* (36) have recently found a clear correlation of the expression of miRNAs and their targets through analysis of mRNA target expression profiles and *in situ* hybridization. In this study, we undertook a global analysis of the expression of mRNA targets in human, mouse and *Drosophila* using several public gene expression datasets (37–39). To our surprise, we found that the average expression levels of the total targets of all miRNAs are significantly different in distinct tissues compared to the expression of the total genes. For example, we found that the expression levels of miRNA targets are significantly lower in all mouse mature tissues and *Drosophila* later life stages than in the embryos. We also found that miRNA targets are more ubiquitously expressed.

MATERIALS AND METHODS

Stand-alone Java programs or Perl scripts were used where necessary to facilitate the following analyses.

Datasets used in this study

The datasets used in this study include two complete lists of human miRNA targets published by Lewis *et al.* (23) and Krek *et al.* (21), one complete list of human and mouse miRNA targets published by John *et al.* (19), two complete lists of *Drosophila* miRNA targets published by Enright *et al.* (18) and Stark *et al.* (26), a microarray expression dataset for more than 10 000 human genes across 52 tissues and cell lines (38), a microarray expression dataset for nearly 40 000 known and predicted mRNAs in 55 mouse tissues (39) and a microarray expression data for nearly one-third of all *Drosophila* genes during the whole life-cycle (37).

All miRNA target datasets were downloaded from the most recently updated websites. The datasets published by Lewis *et al.* (23) and Krek *et al.* (21) contain human miRNA targets but not mouse miRNA target. To get mouse miRNA targets, we mapped the human miRNA targets to the orthologous genes in mouse genome. To do so, we first converted the

human gene IDs from the datasets published by Lewis *et al.* (23) and Krek *et al.* (21) to NCBI GeneIDs using ID Converter (<http://idconverter.bioinfo.cipf.es/>). To obtain the corresponding mouse orthologous gene pairs, we queried the HomoloGene data file downloaded from HomoloGene (<ftp://ftp.ncbi.nih.gov/pub/HomoloGene/>) using the human genes.

For the human gene microarray dataset (38), we removed some samples which derived from the same or similar tissues in order to avoid over-representation of these samples for the analysis. For example, several brain samples were present in the dataset, we removed most of them. Totally 41 human tissues were analyzed.

As the mouse microarray gene chips were designed to detect 40 000 known and predicted mouse mRNAs (39), it is expected that many predicted mRNAs should not be expressed. Totally 21 622 genes could be detected in at least one tissue in the microarray gene chips. The rest (almost half) of them were not detected in any tissues. Therefore, only 21 622 genes were analyzed in this study.

Ranking of miRNA target genes

To study the correlation of the expression of miRNAs and their targets, we analyzed the microarray expression data containing ~10 000 genes published by Johnson *et al.* (38) across 41 human tissues by ranking each gene over all tissues according to its expression level in the respective tissue as described previously (28). A lower rank number represents less expression level. For example, if a gene is expressed less in a defined tissue than in any other tissues, the rank number of this gene in this tissue is 1. Similarly, rank number 41 means that the expression level of a gene in the tissue is higher than that in any other tissues. We then mapped miRNA targets to the genes in the microarray dataset. The miRNA targets we used in this analysis are from the datasets published by Lewis *et al.* (22) and John *et al.* (19), respectively. We only chose the miRNAs, which have more than 50 targets found in the microarray dataset for further analysis. On average, 180 targets per miRNA could be found in the microarray dataset. For each miRNA, in each tissue, we obtained a rank number of each of its targets.

To facilitate a more global view, we grouped the genes into two sets, one with lower half of rank numbers and the other with higher half. When there is an odd number of a rank, the middlemost rank is excluded. In our specific example with 41 tissue samples, one set consists of ranks 1–20 and the other from 22–41, excluding rank 21. For any miRNA in any tissue, we counted the total number of its targets within lower-rank set and divided it by the total number of its targets within higher-rank set to yield the calculated rank ratio, RR. For example, for a miRNA in the 41-tissue set, $RR = N_{\text{Rank } 1-20} / N_{\text{Rank } 22-41}$. The RR value is an indicator of the preferential tissue expression of a given miRNA's target genes. An RR value greater than one means that the majority of expressed targets of a miRNA in this tissue have a lower expression level than the median level of expression of the miRNA's targets across all the tissues. If the RR value of a miRNA is greater in a particular tissue than that in any others, the expression level of the targets of this miRNA in this tissue is very likely to be the lowest among the 41 human tissues. We also did the same analysis for total genes presented in

the microarray dataset. The RR value of target genes for each miRNA in a tissue was normalized by the RR value of total genes in the same tissue and then plotted as a function of tissues and miRNAs, respectively. The RR value provides a global descriptor of the tissue distribution of a miRNA's target genes rather than the expression levels of individual genes. It does not provide gene-specific information but allows the extraction of global trends of a group of genes (miRNA target genes) amid the noisy data for individual genes. This method was also used to analyze a microarray dataset containing 55 mouse samples. In this case, $RR = N_{\text{Rank } 1-27} / N_{\text{Rank } 29-55}$.

Comparison of the expression level of miRNA targets in the embryos of mouse and *Drosophila* with that in their mature tissues

To explore the difference of the expression level of the total targets of all available miRNAs in mouse mature tissues from that in 12.5-day mouse embryo, we compared the absolute expression value of each gene between tissues and the embryo to determine if a gene has a higher or lower expression level in a tissue than in embryo. We then calculated the ratio of lower-expressed targets to the higher-expressed targets in each tissue (termed R_{miRNA}).

To determine the statistical significance of the observation, we performed resampling statistical tests. A more detailed explanation of randomization tests was described previously (40). In each test, we randomly picked up the same number of genes as the number of miRNA targets from total genes. We calculated the ratio of lower-expressed genes to higher-expressed genes in this sub-pool and defined it as R_{random} . We performed the randomly picking test for 5000 tests and counted the number of times (N) when $R_{\text{random}} > R_{\text{miRNA}}$. We defined P -value as $N/5000$. If $P < 0.05$, we thought that the number of lower-expressed miRNA targets is significantly more than that of higher-expressed miRNA targets. We rejected the hypothesis if the $P < 0.05$.

We used three sets of miRNA targets, respectively, for this analysis. They are the mouse miRNA dataset published by John *et al.* (19) and the human miRNA datasets published by Lewis *et al.* (23) and Krek *et al.* (21), respectively. Human miRNA targets were converted to mouse orthologous for the analysis. We performed correlation analysis of the results obtained from each of the three datasets and calculated the correlation coefficient using R (<http://www.r-project.org/>).

A similar method was used to compare the expression of miRNA targets in the different periods of *Drosophila* life cycle with that in the 23–24 h embryo using a microarray dataset published by Arbeitman *et al.* (37), except that the ratio of number of miRNA targets whose expression level in a defined period is 2-fold lower than that in 23–24 h embryo to that of miRNA targets whose expression level is two times higher than that in 23–24 h embryo was represented. The statistical analysis for this set was conducted as described above.

Analysis of the tissue-specificity of miRNA target expression

Using the microarray database published by Zhang *et al.* (39), in which 21 622 mouse genes including 2276 miRNA targets

published by John *et al.* (19) were expressed in at least one tissue, we analyzed the tissue-specificity of miRNA target expression. The mouse genes are classified into 55 groups according to the number of tissues (1–55) in which a gene was expressed. The total numbers of miRNA targets and total genes in each group were counted, respectively. The percentage of miRNA targets to the total genes in each group was calculated and compared to the percentage (10.52%) of the total miRNA targets (2276) to the total genes (21 622) for determining if the miRNA targets are enriched in or excluded from the respective group.

RESULTS

Expression level of miRNA targets is tissue-dependent

Since the function of a miRNA depends on the combination of its actions to all of its targets for their expression, to understand the global role of these numerous miRNAs, we undertook a global analysis of the expression of mRNA targets in human, mouse and *Drosophila* using several public datasets. We first analyzed the microarray expression data containing ~10 000 genes over 41 human tissues published by Johnson *et al.* (38). We compared the relative expression level of the total targets of individual miRNAs across the 41 human tissues. Since each miRNA has many targets and the absolute expression levels of these targets are very different, the contribution of each target to the result should be different if we simply use the average expression level of these targets for comparison. To make each target equally contribute to the comparison, we ranked each gene over 41 human tissues according its expression levels in the respective tissues (see Materials and Methods). A lower rank number means a lower expression level. For each miRNA, in each tissue, we counted the total number of its targets (23) at each rank position (Supplementary Table S1). By comparing the distribution of the rank number of the targets between different tissues, we could find the relative expression levels of the total targets of a miRNA in each tissue compared to other tissues. This method could avoid the effect of the bias of the absolute expression levels of the miRNA targets on the analysis. Figure 1A shows a typical result for the distribution of the rank number of miR-128a targets (23) in liver and brain. In liver, the number of miR-128a targets with a lower rank number is obviously more than that of those with a higher rank number. In contrast, in brain, the result is reversed. This suggests that the overall expression level of the targets of miR-128a in liver is lower than that in brain. To obtain a quick overview, we grouped the targets into two sets, one with rank numbers from 1–20 and the other with rank numbers from 22–41 (see inset in Figure 1A). We then calculated the RR value (see Materials and Methods), $N_{\text{Rank } 1-20} / N_{\text{Rank } 22-41}$. A higher RR value of a miRNA in a defined tissue means that more targets of the miRNA have a lower expression level (lower ranking number) in the tissue and less targets of the miRNA have a higher expression level (higher ranking number) than that in most tissues (more than half of the 41 tissues). A RR value more than one means that the total number of the targets with a rank number lower than 21 is more than the total number of the targets with a rank number higher than 21, and suggests

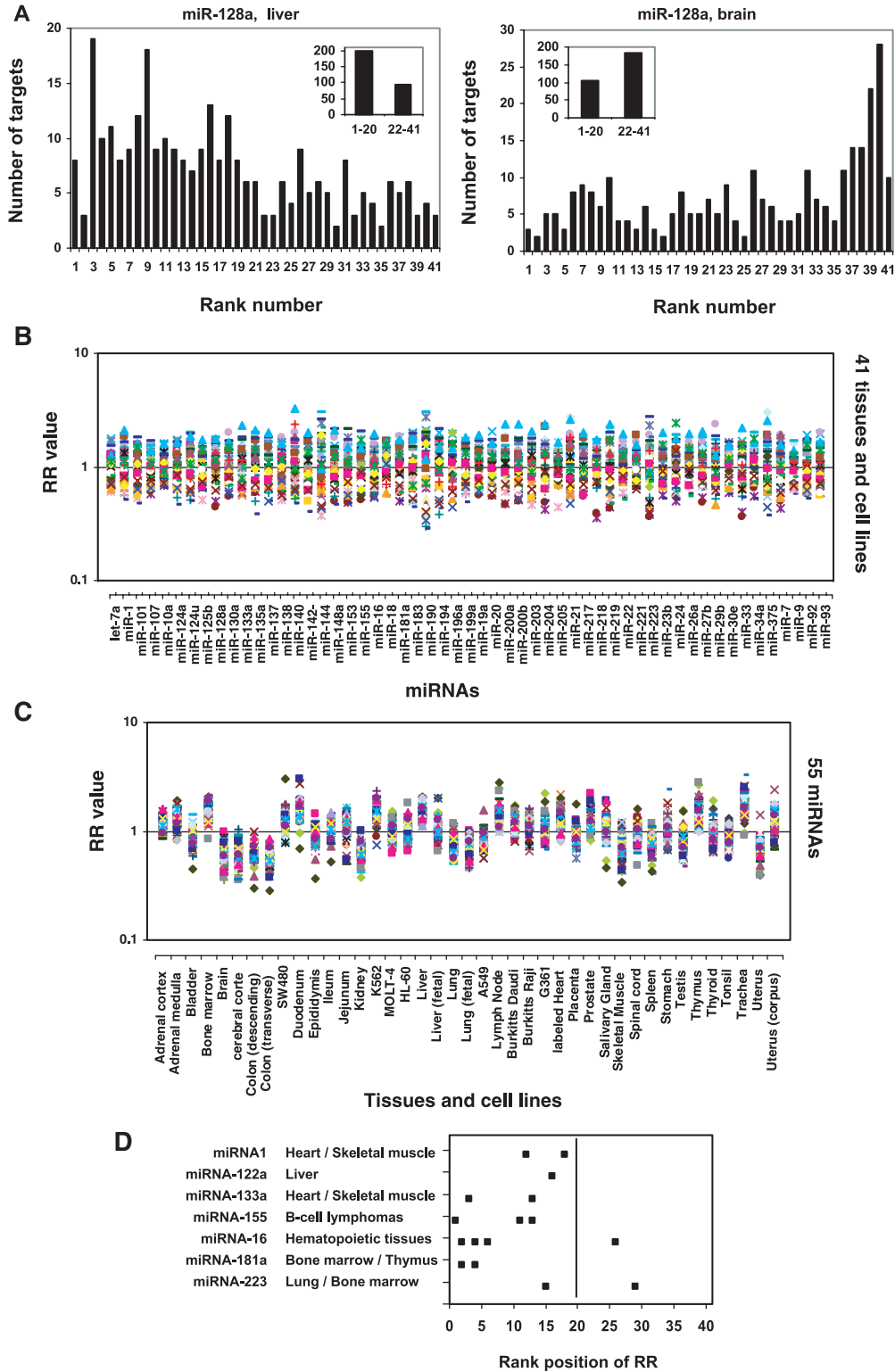


Figure 1. Ranking of the expression levels of miRNA targets in human tissues. Each miRNA target was ranked over 41 human tissue samples as described in the ‘Materials and Methods’ according to its expression level. The RR means the ratio of the total number ($N_{\text{Rank } 1-20}$) of the miRNA targets with a lower rank number (between 1 and 20) to the total number ($N_{\text{Rank } 22-41}$) of the miRNA targets with a higher rank number (between 22 and 41). The RR value of the target genes of each miRNA in a tissue is normalized by the RR value of the total genes in the same tissue. (A) Ranking results of the 295 targets of miR-128a in liver and brain. (B) The RR values of the 55 miRNAs are plotted as a function of human tissues and cell lines. (C) The RR values over the 41 tissues and cell lines are plotted as a function of miRNAs. (D) For each miRNA, the RR values of its targets were ranked across the 41 samples from higher value to lower value. The miRNAs shown are highly specifically expressed in the indicated tissues as demonstrated previously for miRNA-1 and miRNA-133 in references (51,54); miRNA-122 in reference (60); miRNA-16, miRNA-142, miRNA-181 and miRNA-223 in reference (61); miRNA-155 in references (52,53). The rank positions in the indicated tissues were obtained and shown in the figure.

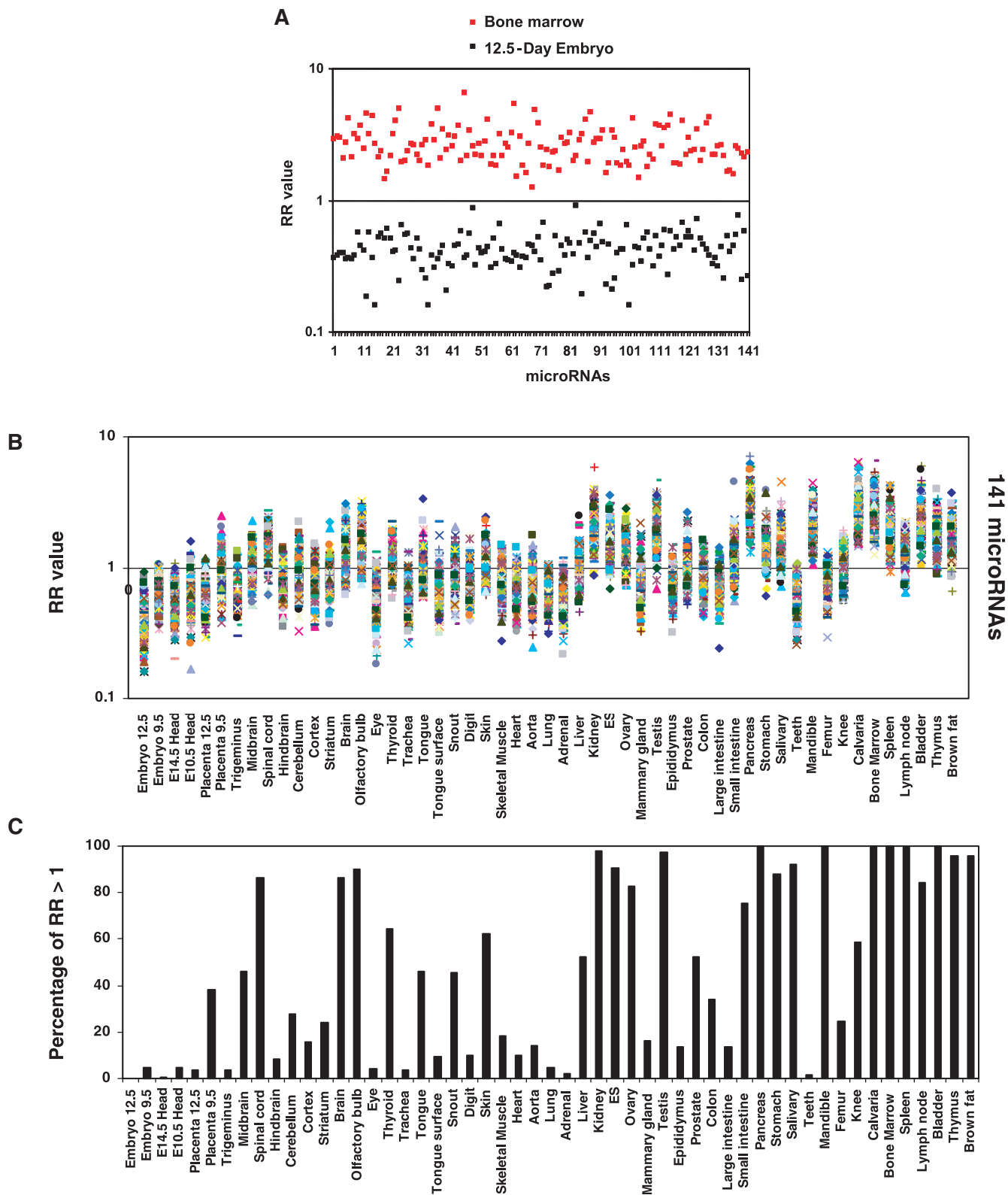


Figure 2. Ranking of the expression levels of miRNA targets in mouse tissues. Each miRNA target gene was ranked over 55 mouse samples according to its relative expression level in the respective sample as described in the ‘Materials and Methods’. The RR means the ratio of the total number ($N_{\text{Rank } 1-27}$) of the miRNA targets with a lower rank number (between 1 and 27) to the total number ($N_{\text{Rank } 29-55}$) of the miRNA targets with a higher rank number (between 29 and 55). The RR value of the target genes of each miRNA in a tissue is normalized by the RR value of the total genes in the same tissue. (A) Comparison of the target gene expressions (RR values) of the 141 miRNAs in bone marrow and 12.5-day embryo. (B) RR values over 141 miRNAs are plotted as a function of mouse tissues. (C) The percentage of miRNAs, of which the total number ($N_{\text{Rank } 1-27}$) of the targets with lower rank number (between 1 and 27) is more than the number ($N_{\text{Rank } 29-55}$) of the targets with higher rank number (between 29 and 55) (RR value >1).

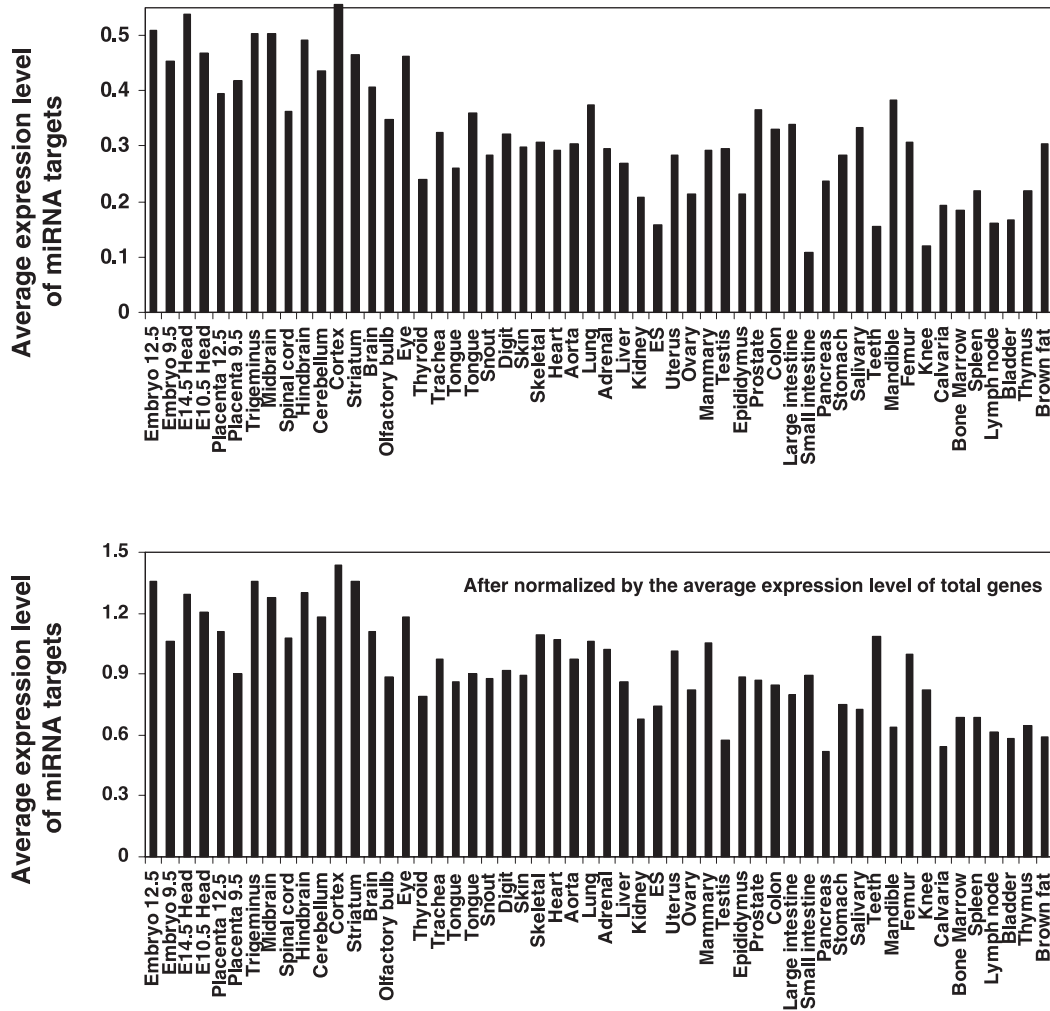
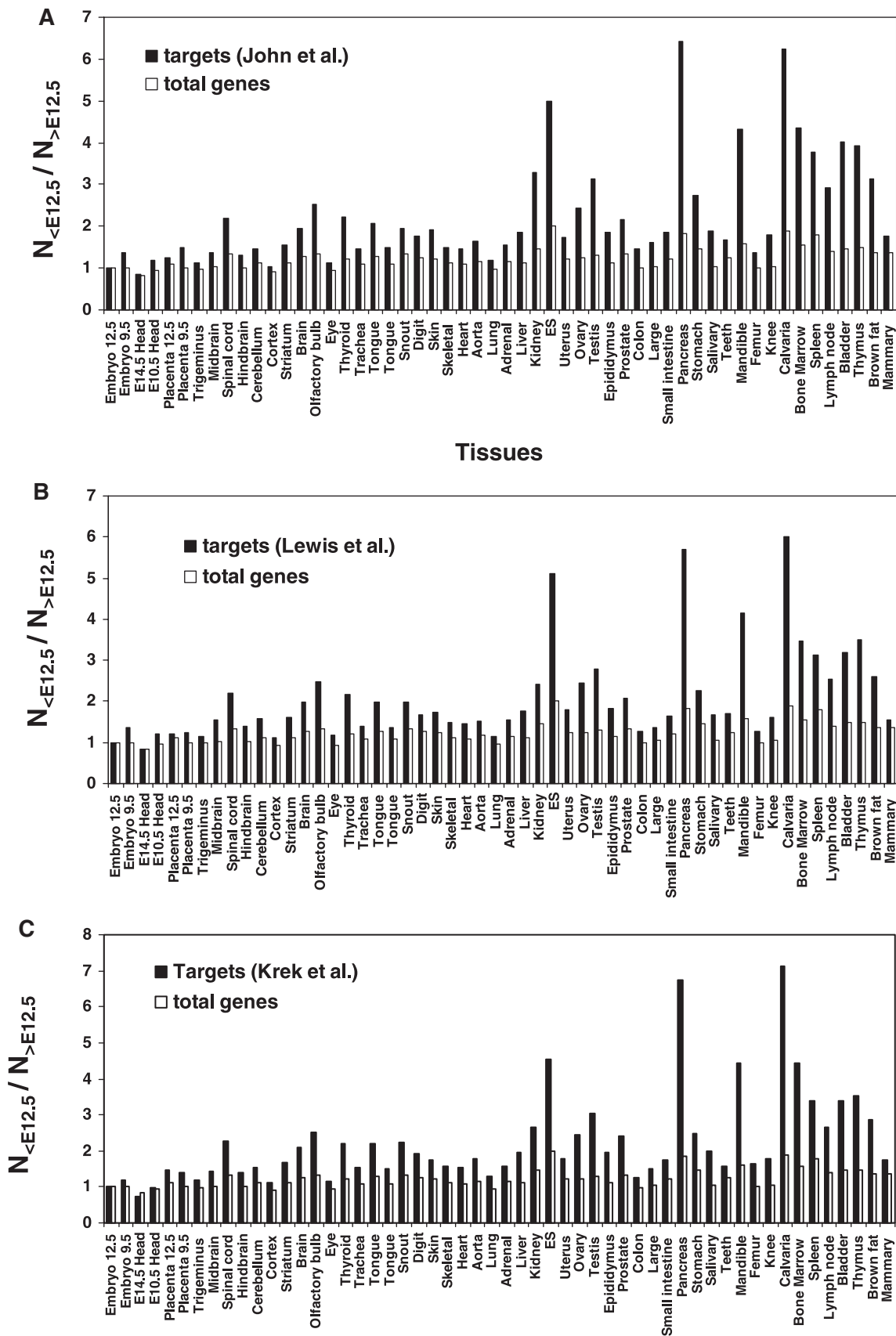


Figure 3. Overall expression levels of miRNA targets in mouse tissues. The average expression levels of the miRNA targets and the total genes in each mouse tissue were calculated using a microarray dataset (39), which contains 21 622 mouse genes including 2276 miRNA targets published by John *et al.* (19).

that the general expression level of the targets of a miRNA in the tissue is most likely to be lower than the median expression level of the targets in all tissues. For example, the RR value for miR-128a is 2.1 (197 targets/92 targets) in liver and 0.57 (104 targets/184 targets) in brain, suggesting a lower expression level of the miR-128a targets in liver than that in brain. Totally, we analyzed 55 miRNAs, each of which have at least 50 targets presented in the microarray dataset (average 180 targets/miRNA), across 41 tissues. We also did the same analysis for total genes present in the microarray dataset. The RR values are shown in Supplementary Table S1. The RR value of the target genes for each miRNA in a tissue was normalized by the RR value of the total genes in the same tissue and then plotted as a function of miRNAs and tissues, respectively (Figure 1B and C). As expected, for each individual miRNA, the RR values in different tissues are equally distributed around one (the number of the tissues with a RR value more than one is similar to the number of the tissues with a RR value less than one) (Figure 1B).

We next studied the correlation of the expression of miRNAs and their targets using several public microRNA

microarray expression datasets. Because of the low accuracy of the miRNA microarray expression profiling data, we did not find obvious correlation of the expression of miRNAs and their target expression (data not shown). For example, the miRNAs microarray expression profiles published by different researchers are not consistent (41–50). Therefore, we next focused on the miRNAs, which were highly specifically expressed in one or a few tissues and whose expression was highly confirmed by different researchers through northern blot analysis and/or cloning method (51–54). To do so, for each miRNA, we further ranked the RR values of their targets across the 41 tissues (from high value to low value) and found the rank position for the tissues in which the miRNA is highly specifically expressed. As shown in Figure 1D, most miRNAs have a low rank number of their target RR values in the tissues where they are specifically expressed. This suggests that the expression of miRNAs has a negative correlation with their target expression. However, we noted that two miRNAs (miRNA-124a and miRNA-125b) specifically expressed in brain do not show negative correlation with their targets. For example, the RR values of miRNA-124a and miRNA-125b in brain ranked at 38 and



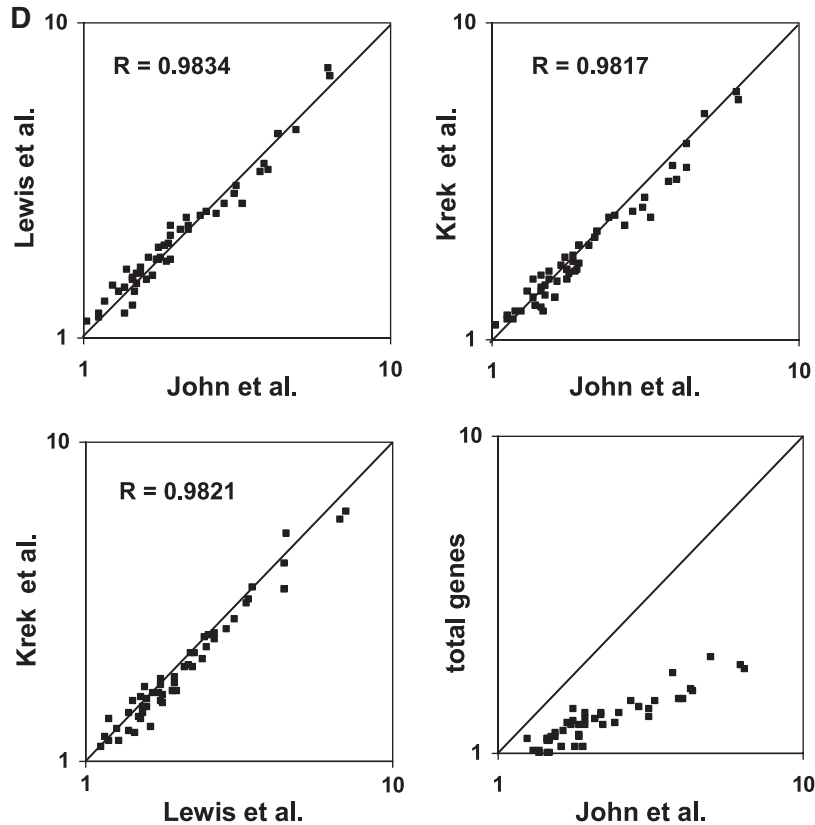


Figure 4. Comparison of the expression levels of miRNA target genes in mouse tissues with that in 12.5-day embryo. (A–C) A microarray dataset containing 21 622 genes over 55 mouse samples published by Zhang *et al.* (39) was used in this analysis. In each individual tissue, the expression level of each gene was compared with that in 12.5-day embryo. The number of the miRNA targets whose expression level in a given tissue is lower than that in 12.5-day embryo ($N_{<E12.5}$) is divided by the number of miRNA targets with an expression level in the same tissue higher than that in 12.5-day embryo ($N_{>E12.5}$). Three sets of miRNA targets published by John *et al.* (19), Lewis *et al.* (23) and Krek *et al.* (21), respectively were analyzed. Each of them contains more than two thousand miRNA targets found in the microarray dataset. As control, the same calculation was made for total genes. (D) Correlation between the data ($N_{<E12.5}/N_{>E12.5}$) obtained using three different miRNA target datasets.

35 of the 41 tissues, respectively. Interestingly, Sood *et al.* found that for many miRNAs, their predicted targets are up-regulated in neuronal tissues compared with most other tissue types, without considering the expression patterns of the miRNAs (36). We obtained a similar result (see below). However, it is not clear if the unexpected higher expression level of the targets of miRNA-124a and miRNA-125b in brain, where the miRNAs are highly expressed, is caused by the general high expression level of miRNA targets in this tissue.

When we looked at the distribution of the RR values in each tissue (Figure 1C), to our surprise, we found a dramatic difference between different tissues. In some tissues, the preponderance of miRNAs have a $RR > 1$. Conversely, in some tissues, $RR < 1$ for an overwhelming fraction of the miRNAs. This suggests that the overall expression level of the total targets of all miRNAs is quite lower in some tissues and higher in others. For example, in bone marrow, 54 of the 55 miRNAs have a RR value more than one, whereas in brain, none of them has a RR value more than one, indicating that the overall expression level of miRNA targets is lower in bone marrow but higher in brain. Similar results were obtained when using the miRNA targets published by John *et al.* (19) (Supplementary Figure S1).

Expression levels of miRNA targets are lower in mature tissues than in embryos in both mammalian and fly

We next analyzed the expression of miRNA targets in 55 mouse samples using the gene expression profile data published by Zhang *et al.* (39) and the dataset of microRNA targets published by John *et al.* (19). Similar to what we found in human tissues, the RR value for all of the 141 miRNAs in mouse bone marrow is greater than one (Figure 2A), suggesting that the expression level of miRNA targets in this tissue is obviously lower than the median level across the other tissues. A similar result was observed in other hematopoietic cells-rich or lymphocytes-rich tissues, such as thymus, spleen and lymph node (Figure 2B and C). Most importantly, the expression levels of miRNA targets in embryo, embryo head and placenta are significantly higher than that in most mature tissues (Figure 2B and C). For example, in 12.5-day embryo, all of the 141 miRNAs have a RR value below one.

To further confirm the observation, we directly compared the average expression value of the total 2276 miRNA targets in the mouse tissues. Consistent with the result above derived from ranking analysis, the average expression level of miRNA target genes is higher in 12.5-day embryo than that in any mature tissues except for cerebral cortex (Figure 3). For example, the levels in bone marrow, spleen, thymus

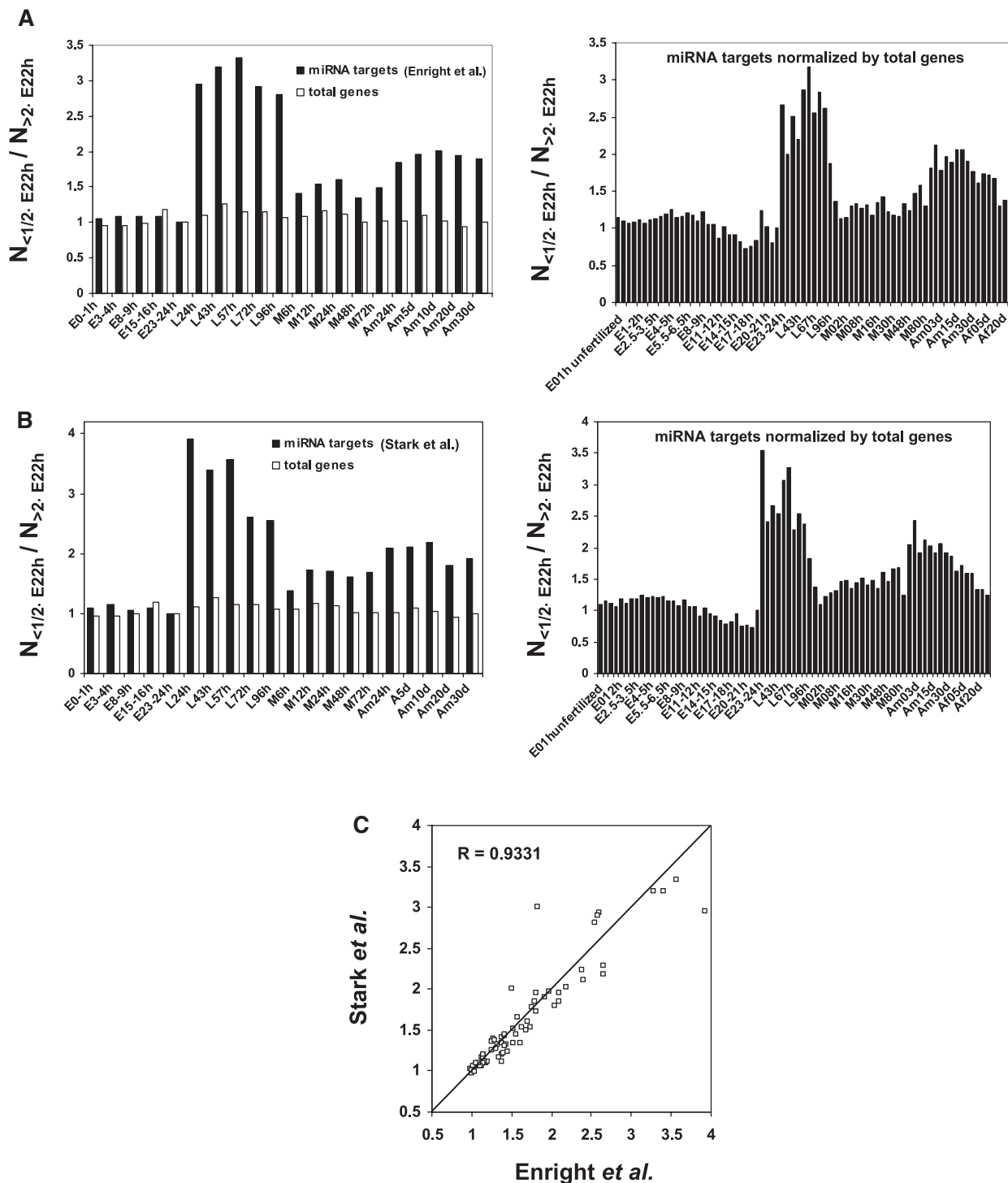


Figure 5. Analysis of the expression of miRNA targets in *Drosophila*. (A and B) A published microarray dataset (37) was analyzed for the comparison of the expression levels of the miRNA targets during different periods of *Drosophila* life cycle. The total number of miRNA targets whose expression level is two times lower than that in 23–24 h embryo ($N_{<1/2 \times E23-24 h}$) was divided by the total number of those with expression level two times higher than that in 23–24 h embryo ($N_{>2 \times E23-24 h}$). Two sets of *Drosophila* miRNA targets published by Enright *et al.* (18) and Stark *et al.* (18,18,26), respectively were analyzed. As control, the expression levels of the total genes were also analyzed by the same way. E, embryo; L, larva; M, pupae; Am, adult male and Af, adult female. (C) Correlation between the data ($N_{<1/2 \times E23-24 h} / N_{>2 \times E23-24 h}$) obtained using two different miRNA target datasets.

and lymph node are significantly (more than two times) lower than in 12.5-day embryo, whereas, the levels in neuronal tissues, such as whole brain, middle brain, hindbrain, cerebellum and trigeminus are slightly lower than in

12.5-day embryo or even higher (cerebral cortex) than in 12.5-day embryo. This consists with Sood *et al.*'s observation (36) and our observation in human brain. This analysis further confirmed that, miRNA target expression levels are

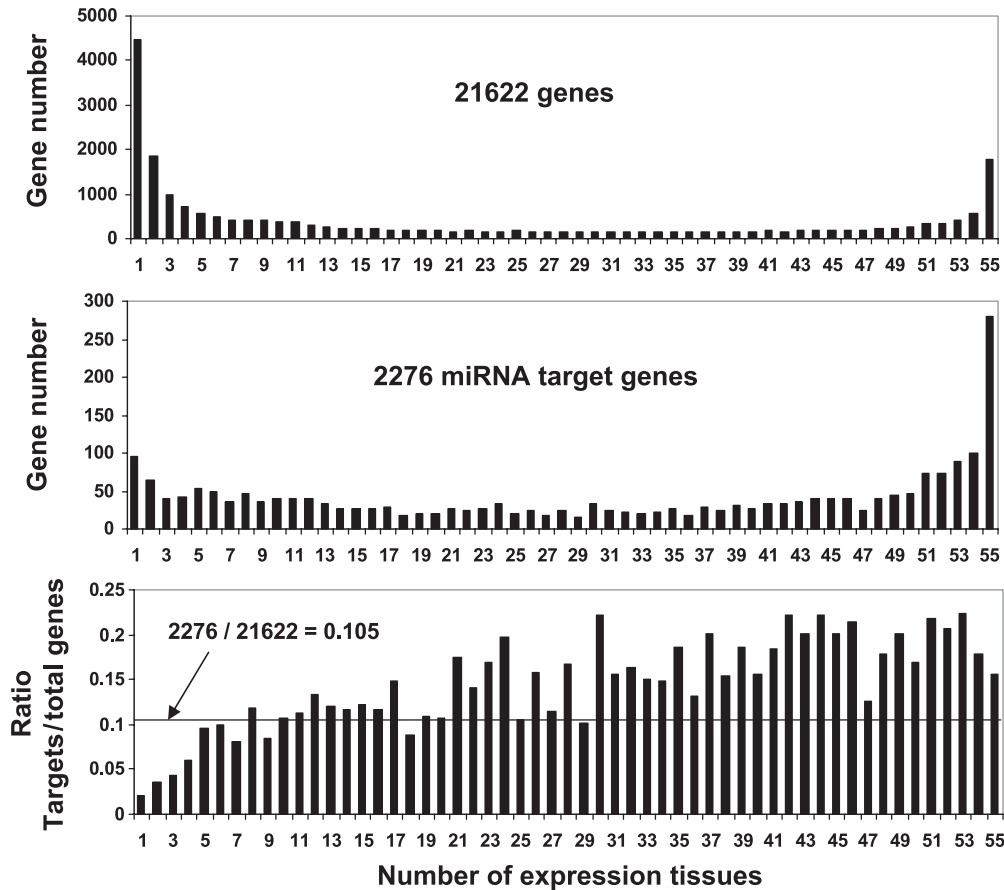


Figure 6. MiRNAs more frequently target ubiquitously expressed genes than tissue-restrictively expressed genes. A microarray dataset published by Zhang *et al.* (39), which contains 21 622 genes including 2276 miRNA targets (19), was used in this analysis as described in the ‘Materials and Methods’. The ratio of the miRNA target genes to the total genes expressed in an indicated number of tissues was calculated.

significantly lower in most mature mouse tissues, except for the neuronal tissues, than in embryos.

We then focused on the comparison of mouse mature tissues to embryo. To do so, we counted the total number of miRNA targets whose expression level is lower in a given tissue than that in 12.5-day embryo ($N_{<E12.5}$) and divided it by the total number of miRNA targets whose expression level is higher than that in 12.5-day embryo ($N_{>E12.5}$). As shown in Figure 4A, in all mature tissues, the lower expressed target number is more than higher expressed target number ($N_{<E12.5}/N_{>E12.5} > 1$). As a control, we carried out the same calculations for all genes. We see that, for all tissues, the $N_{<E12.5}/N_{>E12.5}$ value of total genes is lower than that of miRNA targets (Figure 4A). Resampling statistical tests (see ‘Materials and Methods’ for details) demonstrated that the difference is significant ($P < 0.0002$ for almost all of the tissues, Supplementary Table S2). To further confirm the observation, we performed the same analysis with the total miRNA targets published by Lewis *et al.* (23) and Krek *et al.* (21), respectively. We found the similar patterns (Figure 4B and C). Figure 4D shows that the data obtained using the miRNA targets published by the three groups respectively, are highly correlated. This supports both the quality of the original data and our analysis method.

To determine if the observed expression pattern is conserved in other species, we analyzed the published gene

expression profile over 75 stages of the whole life cycle of *Drosophila* (37). As shown in Figure 5A and B, compared to 23–24 h embryo, while the ratio of lower expressed miRNA targets to higher expressed miRNA targets remains the same during embryo period, it dramatically increases starting from larval periods and lasting to adulthood ($P < 0.0002$ for all larva and male adult, see Supplementary Table S3 for more details). There is a clear correspondence ($R^2 = 0.93$) between the data calculated using the target sets published by Enright *et al.* (18) and Stark *et al.* (26), respectively (Figure 5C). This data supports the important role of miRNAs for determining the timing of tissue differentiation during larva period of *Drosophila* development and maintaining the tissue identity during the adulthood.

It should be noted that the human microarray dataset (38) we used in this study does not contain human embryo and a large scale gene expression profile containing human embryo is not available. Consequently, we could not perform the comparison of the expression levels of human miRNA targets between human mature tissues and human embryos.

MiRNAs more frequently target ubiquitously expressed genes than tissue-specific genes

We analyzed the tissue-specificity of miRNA target expression using the microarray data representing 21 622 mouse

genes (39) including 2276 predicted miRNA targets (19). Each of these genes was expressed in at least one of the 55 mouse tissues (39). Both the 21 622 genes and the 2276 miRNA targets were classified into 55 groups according to the number of tissues (between 1 and 55) in which the gene was expressed. We counted the numbers of miRNA targets and total genes in each group, respectively. To determine if the miRNA targets are enriched in or excluded from some groups, we compared the percentage of miRNA targets to the total genes in each group with the percentage (10.5%) of the total miRNA targets (2276) to the total genes (21 622). As shown in Figure 6, among the genes that are expressed ubiquitously, the targets of miRNAs are over-represented. For example, in genes found in groups 45–55 (i.e. genes found in almost all the tissues), ~20% are miRNA targets, around twice the fraction of miRNA targets in the whole list of genes (10.5%). In contrast, among genes that are expressed in a small number of tissues, miRNA targets are under-represented. For example, among the genes that are specifically expressed in only 1–4 tissues, the fraction of miRNA targets present is about half or less than that in the general gene population.

DISCUSSION

In this study, we found that the expression level of miRNA targets in most mature tissues is significantly lower than that in embryos in both mammalian and fly and that miRNAs more frequently target ubiquitously expressed genes than tissue-specific genes. These findings support the recent observations that some miRNAs play a most important role in driving tissue terminal differentiation and maintaining tissue identity. Previous studies suggest that 10 to 30% of human genes are potential miRNA targets (19,22). However, analysis of the specific gene ontology (GO) molecular function classification among the predicted targets could not reveal any specific biological functions of animal miRNAs since the animal miRNA targets populated many GO functional categories (19,22). Only ~13% of mammalian miRNA targets predicted by Lewis *et al.* were involved in development according to the GO biological process categories (22). Failing to predict the functions of miRNA targets through GO analysis may be simply caused by the evolving stage of the classification of GO function categories. Alternatively, on principle, the functions of miRNAs could not be predicted by the GO function categories of their targets because the expression and therefore the functions of their targets are proposed to be turned down but not induced by miRNA expression and GO analysis can only tell the function of a group of genes in a GO function category when they are expressed or up-regulated but not that when they are down-regulated. Our studies demonstrate that statistical analysis of the expression of miRNA targets may reflect the global functions of miRNAs. However, the biological significance of the observation is not clear right now. One proposal is that the statistically lower expression level of miRNA targets in matured tissues than in embryo may reflect the important role of miRNAs for embryo development and for maintaining of the identity of matured tissues.

Up to now, the molecular mechanism determining the lower expression level of miRNA targets in mature tissues than that

in embryos remains to be elucidated. One potential reason is that the miRNA expression level is lower in embryos. This is true for zebrafish. Recently, Wienholds *et al.* have reported that most zebrafish miRNAs were not detected during early development (55). However, the results regarding to the global expression patterns of miRNAs during mouse or *Drosophila* embryo development are currently not available although the differential expression of miRNAs in different tissues were clearly demonstrated. Another possibility is that the activity of miRNA machinery is lower in embryos than in other tissues. For example, Yang *et al.* (56) reported that dicer, an important protein for miRNA biogenesis and miRNA function, starts expression in 7-day old mouse embryos and remains stable through 17-day embryos. However, maternal Dicer protein is probably present from fertilization (57–59). In addition, a possible delay from miRNA processing to target gene inactivation might also contribute to the high level expression of miRNA targets in embryos.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

Funding to pay the Open Access publication charges for this article was provided by National Research Council Canada.

Conflict of interest statement. None declared.

REFERENCES

- Ambros, V. (2001) microRNAs: tiny regulators with great potential. *Cell*, **107**, 823–826.
- Bartel, D.P. (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, **116**, 281–297.
- Cullen, B.R. (2004) Transcription and processing of human microRNA precursors. *Mol. Cell*, **16**, 861–865.
- Kim, V.N. (2005) MicroRNA biogenesis: coordinated cropping and dicing. *Nat. Rev. Mol. Cell. Biol.*, **6**, 376–385.
- Carmell, M.A. and Hannon, G.J. (2004) RNase III enzymes and the initiation of gene silencing. *Nat. Struct. Mol. Biol.*, **11**, 214–218.
- Meister, G. and Tuschl, T. (2004) Mechanisms of gene silencing by double-stranded RNA. *Nature*, **431**, 343–349.
- Griffiths-Jones, S. (2004) The microRNA registry. *Nucleic Acids Res.*, **32**, D109–D111.
- Lee, R.C., Feinbaum, R.L. and Ambros, V. (1993) The *C.elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell*, **75**, 843–854.
- Reinhart, B.J., Slack, F.J., Basson, M., Pasquinelli, A.E., Bettinger, J.C., Rougvie, A.E., Horvitz, H.R. and Ruvkun, G. (2000) The 21-nucleotide *let-7* RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature*, **403**, 901–906.
- Johnston, R.J. and Hobert, O. (2003) A microRNA controlling left/right neuronal asymmetry in *Caenorhabditis elegans*. *Nature*, **426**, 845–849.
- Brennecke, J., Hipfner, D.R., Stark, A., Russell, R.B. and Cohen, S.M. (2003) bantam encodes a developmentally regulated microRNA that controls cell proliferation and regulates the proapoptotic gene *hid* in *Drosophila*. *Cell*, **113**, 25–36.
- Xu, P., Vernooy, S.Y., Guo, M. and Hay, B.A. (2003) The *Drosophila* microRNA *Mir-14* suppresses cell death and is required for normal fat metabolism. *Curr. Biol.*, **13**, 790–795.
- Chen, C.Z., Li, L., Lodish, H.F. and Bartel, D.P. (2004) MicroRNAs modulate hematopoietic lineage differentiation. *Science*, **303**, 83–86.
- Brennecke, J., Stark, A., Russell, R.B. and Cohen, S.M. (2005) Principles of microRNA-target recognition. *PLoS Biol.*, **3**, e85.

15. Doench, J.G. and Sharp, P.A. (2004) Specificity of microRNA target selection in translational repression. *Genes Dev.*, **18**, 504–511.
16. Kiriakidou, M., Nelson, P.T., Kouranov, A., Fitziev, P., Bouyioukos, C., Mourelatos, Z. and Hatzigeorgiou, A. (2004) A combined computational-experimental approach predicts human microRNA targets. *Genes Dev.*, **18**, 1165–1178.
17. Kloosterman, W.P., Wienholds, E., Ketting, R.F. and Plasterk, R.H. (2004) Substrate requirements for let-7 function in the developing zebrafish embryo. *Nucleic Acids Res.*, **32**, 6284–6291.
18. Enright, A.J., John, B., Gaul, U., Tuschl, T., Sander, C. and Marks, D.S. (2003) MicroRNA targets in *Drosophila*. *Genome Biol.*, **5**, R1.
19. John, B., Enright, A.J., Aravin, A., Tuschl, T., Sander, C. and Marks, D.S. (2004) Human MicroRNA targets. *PLoS Biol.*, **2**, e363.
20. Kiriakidou, M., Nelson, P.T., Kouranov, A., Fitziev, P., Bouyioukos, C., Mourelatos, Z. and Hatzigeorgiou, A. (2004) A combined computational-experimental approach predicts human microRNA targets. *Genes Dev.*, **18**, 1165–1178.
21. Krek, A., Grun, D., Poy, M.N., Wolf, R., Rosenberg, L., Epstein, E.J., MacMenamin, P., da Piedade, I., Gunsalus, K.C., Stoffel, M. et al. (2005) Combinatorial microRNA target predictions. *Nat. Genet.*, **37**, 495–500.
22. Lewis, B.P., Shih, I.H., Jones-Rhoades, M.W., Bartel, D.P. and Burge, C.B. (2003) Prediction of mammalian microRNA targets. *Cell*, **115**, 787–798.
23. Lewis, B.P., Burge, C.B. and Bartel, D.P. (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell*, **120**, 15–20.
24. Rhoades, M.W., Reinhart, B.J., Lim, L.P., Burge, C.B., Bartel, B. and Bartel, D.P. (2002) Prediction of plant microRNA targets. *Cell*, **110**, 513–520.
25. Robins, H., Li, Y. and Padgett, R.W. (2005) Incorporating structure to predict microRNA targets. *Proc. Natl Acad. Sci. USA*, **102**, 4006–4009.
26. Stark, A., Brennecke, J., Russell, R.B. and Cohen, S.M. (2003) Identification of *Drosophila* MicroRNA targets. *PLoS Biol.*, **1**, E60.
27. Wang, X.J., Reyes, J.L., Chua, N.H. and Gaasterland, T. (2004) Prediction and identification of *Arabidopsis thaliana* microRNAs and their mRNA targets. *Genome Biol.*, **5**, R65.
28. Lim, L.P., Lau, N.C., Garrett-Engele, P., Grimson, A., Schelter, J.M., Castle, J., Bartel, D.P., Linsley, P.S. and Johnson, J.M. (2005) Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature*, **433**, 769–773.
29. Krutzfeldt, J., Rajewsky, N., Braich, R., Rajeev, K.G., Tuschl, T., Manoharan, M. and Stoffel, M. (2005) Silencing of microRNAs *in vivo* with 'antagomirs'. *Nature*, **438**, 685–689.
30. Inaoka, H., Fukuoka, Y. and Kohane, I.S. (2006) Lower expression of genes near microRNA in *C. elegans* germline. *BMC Bioinformatics.*, **7**, 112.
31. Laule, O., Hirsch-Hoffmann, M., Hruz, T., Gruissem, W. and Zimmermann, P. (2006) Web-based analysis of the mouse transcriptome using Genevestigator. *BMC Bioinformatics.*, **7**, 311.
32. Oishi, K., Amagai, N., Shirai, H., Kadota, K., Ohkura, N. and Ishida, N. (2005) Genome-wide expression analysis reveals 100 adrenal gland-dependent circadian genes in the mouse liver. *DNA Res.*, **12**, 191–202.
33. Wang, H., Hubbell, E., Hu, J.S., Mei, G., Cline, M., Lu, G., Clark, T., Siani-Rose, M.A., Ares, M., Kulp, D.C. et al. (2003) Gene structure-based splice variant deconvolution using a microarray platform. *Bioinformatics.*, **19**, i315–i322.
34. Farh, K.K., Grimson, A., Jan, C., Lewis, B.P., Johnston, W.K., Lim, L.P., Burge, C.B. and Bartel, D.P. (2005) The widespread impact of mammalian MicroRNAs on mRNA repression and evolution. *Science*, **310**, 1817–1821.
35. Stark, A., Brennecke, J., Bushati, N., Russell, R.B. and Cohen, S.M. (2005) Animal MicroRNAs confer robustness to gene expression and have a significant impact on 3'-UTR evolution. *Cell*, **123**, 1133–1146.
36. Sood, P., Krek, A., Zavolan, M., Macino, G. and Rajewsky, N. (2006) Cell-type-specific signatures of microRNAs on target mRNA expression. *Proc. Natl Acad. Sci. USA*, **103**, 2746–2751.
37. Arbeitman, M.N., Furlong, E.E., Imam, F., Johnson, E., Null, B.H., Baker, B.S., Krasnow, M.A., Scott, M.P., Davis, R.W. and White, K.P. (2002) Gene expression during the life cycle of *Drosophila melanogaster*. *Science*, **297**, 2270–2275.
38. Johnson, J.M., Castle, J., Garrett-Engele, P., Kan, Z., Loerch, P.M., Armour, C.D., Santos, R., Schadt, E.E., Stoughton, R. and Shoemaker, D.D. (2003) Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays. *Science*, **302**, 2141–2144.
39. Zhang, W., Morris, Q.D., Chang, R., Shai, O., Bakowski, M.A., Mitsakakis, N., Mohammad, N., Robinson, M.D., Zirngibl, R., Somogyi, E. et al. (2004) The functional landscape of mouse gene expression. *J. Biol.*, **3**, 21.
40. Wang, E. and Purisima, E. (2005) Network motifs are enriched with transcription factors whose transcripts have short half-lives. *Trends Genet.*, **21**, 492–495.
41. Babak, T., Zhang, W., Morris, Q., Blencowe, B.J. and Hughes, T.R. (2004) Probing microRNAs with microarrays: tissue specificity and functional inference. *RNA*, **10**, 1813–1819.
42. Barad, O., Meiri, E., Avniel, A., Aharonov, R., Barzilai, A., Bentwich, I., Einav, U., Gilad, S., Hurban, P., Karov, Y. et al. (2004) MicroRNA expression detected by oligonucleotide microarrays: system establishment and expression profiling in human tissues. *Genome Res.*, **14**, 2486–2494.
43. Baskerville, S. and Bartel, D.P. (2005) Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes. *RNA*, **11**, 241–247.
44. Houbaviy, H.B., Murray, M.F. and Sharp, P.A. (2003) Embryonic stem cell-specific microRNAs. *Dev. Cell*, **5**, 351–358.
45. Krichevsky, A.M., King, K.S., Donahue, C.P., Khrapko, K. and Kosik, K.S. (2003) A microRNA array reveals extensive regulation of microRNAs during brain development. *RNA*, **9**, 1274–1281.
46. Liu, C.G., Calin, G.A., Meloon, B., Gamlie, N., Sevignani, C., Ferracin, M., Dumitru, C.D., Shimizu, M., Zupo, S., Dono, M. et al. (2004) An oligonucleotide microchip for genome-wide microRNA profiling in human and mouse tissues. *Proc. Natl Acad. Sci. USA*, **101**, 9740–9744.
47. Miska, E.A., Alvarez-Saavedra, E., Townsend, M., Yoshii, A., Sestan, N., Rakic, P., Constantine-Paton, M. and Horvitz, H.R. (2004) Microarray analysis of microRNA expression in the developing mammalian brain. *Genome Biol.*, **5**, R68.
48. Nelson, P.T., Baldwin, D.A., Scarce, L.M., Oberholtzer, J.C., Tobias, J.W. and Mourelatos, Z. (2004) Microarray-based, high-throughput gene expression profiling of microRNAs. *Nat. Meth.*, **1**, 155–161.
49. Sempere, L.F., Freemantle, S., Pitha-Rowe, I., Moss, E., Dmitrovsky, E. and Ambros, V. (2004) Expression profiling of mammalian microRNAs uncovers a subset of brain-expressed microRNAs with possible roles in murine and human neuronal differentiation. *Genome Biol.*, **5**, R13.
50. Thomson, J.M., Parker, J., Perou, C.M. and Hammond, S.M. (2004) A custom microarray platform for analysis of microRNA gene expression. *Nature Meth.*, **1**, 47–53.
51. Chen, J.F., Mandel, E.M., Thomson, J.M., Wu, Q., Callis, T.E., Hammond, S.M., Conlon, F.L. and Wang, D.Z. (2006) The role of microRNA-1 and microRNA-133 in skeletal muscle proliferation and differentiation. *Nature Genet.*, **38**, 228–233.
52. Eis, P.S., Tam, W., Sun, L., Chadburn, A., Li, Z., Gomez, M.F., Lund, E. and Dahlberg, J.E. (2005) Accumulation of miR-155 and BIC RNA in human B cell lymphomas. *Proc. Natl Acad. Sci. USA*, **102**, 3627–3632.
53. Kluiver, J., Poppema, S., de Jong, D., Blokzijl, T., Harms, G., Jacobs, S., Kroesen, B.J. and van den, B.A. (2005) BIC and miR-155 are highly expressed in Hodgkin, primary mediastinal and diffuse large B cell lymphomas. *J. Pathol.*, **207**, 243–249.
54. Zhao, Y., Samal, E. and Srivastava, D. (2005) Serum response factor regulates a muscle-specific microRNA that targets Hand2 during cardiogenesis. *Nature*, **436**, 214–220.
55. Wienholds, E., Kloosterman, W.P., Miska, E., Alvarez-Saavedra, E., Berezikov, E., de Bruijn, E., Horvitz, H.R., Kauppinen, S. and Plasterk, R.H. (2005) MicroRNA expression in zebrafish embryonic development. *Science*, **309**, 310–311.
56. Yang, W.J., Yang, D.D., Na, S., Sandusky, G.E., Zhang, Q. and Zhao, G. (2005) Dicer is required for embryonic angiogenesis during mouse development. *J. Biol. Chem.*, **280**, 9330–9335.
57. Bernstein, E., Kim, S.Y., Carmell, M.A., Murchison, E.P., Alcorn, H., Li, M.Z., Mills, A.A., Elledge, S.J., Anderson, K.V. and Hannon, G.J. (2003) Dicer is essential for mouse development. *Nature Genet.*, **35**, 215–217.

58. Giraldez,A.J., Cinalli,R.M., Glasner,M.E., Enright,A.J., Thomson,J.M., Baskerville,S., Hammond,S.M., Bartel,D.P. and Schier,A.F. (2005) MicroRNAs regulate brain morphogenesis in zebrafish. *Science*, **308**, 833–838.
59. Wienholds,E., Koudijs,M.J., van Eeden,F.J., Cuppen,E. and Plasterk,R.H. (2003) The microRNA-producing enzyme Dicer1 is essential for zebrafish development. *Nature Genet.*, **35**, 217–218.
60. Lagos-Quintana,M., Rauhut,R., Yalcin,A., Meyer,J., Lendeckel,W. and Tuschl,T. (2002) Identification of tissue-specific microRNAs from mouse. *Curr. Biol.*, **12**, 735–739.
61. Chen,C.Z., Li,L., Lodish,H.F. and Bartel,D.P. (2004) MicroRNAs modulate hematopoietic lineage differentiation. *Science*, **303**, 83–86.