Deposited research article

Global analysis of microRNA target gene expression reveals the potential roles of microRNAs in maintaining tissue identity

Zhenbao Yu, Zhaofeng Jian, Shi-Hsiang Shen, Enrico Purisima, and Edwin Wang

Addresses: Biotechnology Research Institute, National Research Council of Canada, Montréal, Québec, H4P 2R2, Canada.

Correspondence: Edwin Wang. E-mail: edwin.wang@cnrc-nrc.gc.ca. Zhenbao Yu. E-mail: zhenbao.yu@nrc.ca

Posted: 19 December 2005

Genome ***Biology*** 2005, **6:**P14

The electronic version of this article is the complete one and can be found online at http://genomebiology.com/2005/6/13/P14

Received: 13 December 2005

This is the first version of this article to be made available publicly. This article was submitted to *Genome* ***Biology*** for peer review.

Global analysis of microRNA target gene expression reveals the potential roles of microRNAs in maintaining tissue identity

Zhenbao Yu[*] , Zhaofeng Jian, Shi-Hsiang Shen, Enrico Purisima, and Edwin Wang[*]

Biotechnology Research Institute, National Research Council of Canada, Montréal, Québec, H4P 2R2, Canada

[*]Correspondence should be addressed to E.W. (edwin.wang@cnrc-nrc.gc.ca) and Z.Y. (zhenbao.yu@nrc.ca)

Abstract

Background: MicroRNAs are non-coding small RNAs of ~22 nucleotides that regulate the gene expression by base-paring with target mRNAs, leading to mRNA cleavage or translational repression. It is currently estimated that microRNAs account for ~ 1% of predicted genes in higher eukaryotic genomes and that up to 30% of genes might be regulated by microRNAs. However, only very few microRNAs have been functionally characterized and the general functions of microRNAs are not globally studied.

Results: We systematically analyzed the expression patterns of microRNA targets using several public microarray profiles and found that the expression levels of microRNA targets are significantly lower in all mouse and *Drosophila* tissues than in the embryos and that microRNA targets are dramatically excluded from the tissue-specifically expressed gene groups.

Conclusion: These results strongly suggest that the global functions of microRNAs are largely involved in driving tissue differentiation and maintaining tissue identity rather than in tissue-specific physiological functions. In addition, these findings imply that disruption of microRNA functions might cause delineation of differentiated cells, a crucial step towards carcinogenesis.

Background

MicroRNAs (miRNAs), encoded in the chromosomal DNA and transcribed as longer stem-loop precursors, termed pri-miRNAs, are non-coding small (21-23 nucleotide) 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 RNaseIII family of endonucleases Drosha 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 [7-18]. It is currently estimated that miRNAs account for ~ 1% of predicted genes in higher eukaryotic genomes [19].

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 *C. elegans* [20,21]. Lsy-6 and miR-273 act sequentially to control the left/right asymmetric gene expression in *C. elegans* chemosensory neurons [22]. Bantam promotes cell proliferation and inhibits apoptosis in *Drosophila* [23]. MiR-14 suppresses cell death and regulates fat metabolism [24]. MiR-181 potentiates B-cell differentiation [25]. 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 nucleotides in length, called “seed” region, for miRNA function [26-29]. For example, systematical single nucleotide mutation studies demonstrated that base-pairing of miRNAs to their targets with 7 nucleotides at the 5-terminus of miRNAs from position 2 to position 8 is essential and sometimes sufficient for miRNAs to knockdown their target expression [26]. Based on these discoveries, several computational methods have been developed to search for miRNA targets [30-39]. 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 [40,41]. In one study, Lim et al. [40] 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 nucleotides complementary to the 5’-terminus of the transfected miRNAs (the “seed” sequence). In another study, Krutzfeldt et al. [41] demonstrated that knockdown of miRNA-122 by intravenous administration of miRNA “antagomirs” led to upregulation and downregulation of a large number of genes in liver. They found that the 3’-untranslated regions of upregulated genes are strongly enriched in miRNA-122 “seed”-match motifs, whereas downregulated genes are depleted in these motifs [41].

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 [35]. 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.

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. Theoretically, the tissues with low level of the expression of the targets of a miRNA are probably the tissues in which the miRNA is functionally involved. Systematical analysis of gene expression profiles has been proved to be valuable for studies on diverse biological processes [42-48]. 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 [49-51]. We found that the expression levels of miRNA targets are significantly lower in all mouse and *Drosophila* tissues than in the embryos. We also found that the percentage of the number of tissue-specifically expressed miRNA targets is significantly lower than that of ubiquitously expressed miRNA targets. These findings strongly suggest that miRNAs play a most important role in driving tissue terminal differentiation and particularly in maintaining tissue identity rather than in determining or regulating tissue-specific physiological functions.

Results

Expression level of miRNA targets is tissue-dependent

Since miRNA function depends on the combination of its actions to each of all 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. [50]. We compared the relative expression level of the total targets of individual miRNAs across the 41 human tissues. For each miRNA, we could find the tissues in which its functions may be involved by searching for the tissues which have lower expression level of its total targets. Since each miRNA has many targets and the absolute expression levels of these targets are very different, to make each target equally contribute to the comparison, we first ranked each gene over 41 human tissues according its expression levels in the respective tissues (see methods). A lower rank number means a lower expression level. For each miRNA, in each tissue, we counted the number of its targets [35] at each rank position (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 [35] 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 miR-

128a targets 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 Methods), $N_{Rank1-20}/N_{Rank\ 22-41}$. A higher RR value means lower expression level of the miRNA targets. A RR value more than one suggests that the expression level of the targets of a miRNA in a 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 55 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 Table S1. 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 miRNAs and tissues respectively (Figure 1b and 1c). 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). For each miRNA, the tissues with highest RR values could be found from this figure and Table S1, and they are most likely to be the tissues in which this mRNA is functionally involved.

However, 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 miRNA targets is quite depleted in some tissues and enhanced 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 in bone marrow is lower than that in brain no matter which miRNA it is. Similar results were obtained when using the miRNA targets published by John et al. [31] (Figure S1).

Expression levels of miRNA targets are lower in differentiated 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. [51] and the dataset of microRNA targets published by John et al. [31]. 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 hematopoitic cells-rich or lymphocytes-rich tissues, such as thymus, spleen and lymph node (Figure 2b and 2c). Interestingly, we found an obvious correlation of the distribution pattern of RR values and the cluster of tissue property (Figure 2b and 2c). Most importantly, the expression levels of miRNA targets in embryo, embryo head and placenta are significantly higher than that in other tissues (Figure 2b and 2c). 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. Consisting 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 tissues except for the 14.5-day embryo head and cortex (Figure 3). For example, the levels in bone marrow, spleen, thymus and lymph node are respectively more than two times lower than that in 12.5-day embryo.

We then focused on the comparison of mouse 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 tissues except for 14.5-day embryo head, 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). Resembling statistical tests (see "Methods" for details) demonstrated that the difference is significant ($P<0.0002$ for almost all of the tissues, Table S2). To further confirm the observation, we performed the same analysis with the total miRNA targets published by Lewis et al. [35] and Krek et al. [33] respectively. We found the similar patterns (Figure 4b and 4c). Figure 4D shows that the data obtained using each set of the miRNA targets published by each of the three groups are highly correlated. This supports both the quality of the original data and our analysis method.

Taken together, the overall expression level of miRNA targets in every differentiated mouse tissues is lower than that in mouse embryo, suggesting that miRNAs may play an important role for determining the fate of tissue differentiation during embryo development, and maintaining the tissue identity in the later stage.

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* [49]. As shown in Figure 5a and 5b, 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 Table S3 for more details). There is a clear correspondence ($R = 0.93$) between the data calculated using the target sets published by Enright et al. [30] and Stark et al. [30,30,38] respectively (Figure 5c). This data strongly suggest that miRNAs play important roles 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 [50] 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 tissues and human embryos.

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

To determine if miRNAs are involved in tissue-specific physiological functions, we analyzed the tissue-specificity of miRNA target expression using the microarray data representing 21622 mouse genes [51] including 2276 predicted miRNA targets [31]. Each of these genes was expressed in at least one of the 55 mouse tissues [51]. Both the 21622 genes and the 2276 miRNA targets were classified into 55 groups according to the number of tissues (between 1 to 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 (21622). 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), approximately 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. Since tissue-specifically expressed genes are mostly involved in tissue-specific physiological functions, this observation suggests that miRNAs are preferably involved in determining and maintaining tissue identity rather than playing a tissue-specific physiological role. By

that we mean the targets of miRNAs are common to most tissues. Among the different tissues, varying combinations of these common genes are suppressed by the miRNA resulting in the specific tissue type. As an analogy, the common genes are like a starting block of marble with the miRNAs being the sculptor that chisels away everything that is not needed for the final figure.

Discussion

In this study, we developed a method for the analysis of the global expression patterns of miRNA targets. In this method, we considered the bias of the absolute expression levels of each miRNA targets and limited its effect on the calculation by ranking each gene over all samples according its expression levels in the respective tissues and counting the total number of targets at each ranking position. By comparing the RR values defined in this study between each samples, we are able to find the tissues in which miRNA targets are less expressed. Based on the biochemical function of miRNAs, the tissues with lowest expression levels of its targets are likely to be the tissues in which a miRNA is functionally involved. Therefore, in this study, we first provide a tool for the prediction of miRNA functions through analysis of their target expression.

More importantly, we found that the expression level of miRNA targets in differentiated 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 strongly suggest that miRNAs play a most important role in driving tissue terminal differentiation and maintaining tissue identity rather than in determining or regulating tissue-specific physiological functions. Previous studies suggest that 10% to 30% of human genes are potential miRNA targets [31,34]. 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 [31,34]. Only ~13% of mammalian miRNA targets predicted by Lewis et al. were involved in development

according to the GO biological process categories [34]. 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 can reflect the global functions of miRNAs. The statistically lower expression level of miRNA targets in matured tissues than in embryo demonstrates that miRNAs play an important role not only for determining tissue fate during embryo development but also for maintaining identity and preventing dedifferentiation of matured tissues.

Our studies also demonstrate that the overall expression levels of miRNA targets could be clustered in large part according to their anatomic locations or physiological functions. For example, the average expression levels of miRNA targets in lymphoid and myeloid tissues (lymph node, thymus, spleen and bone marrow) are much lower than that in most of other tissues. This result indicates that miRNAs might also play a very important role for the differentiation of hematopoitietic lineage from myeloid and lymphoid progenitors in addition to that for the differentiation of tissues from embryo cells.

Up to now, the molecular mechanism determining the lower expression level of miRNA targets in differentiated 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 [52]. 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 [53-62]. Another possibility is that the activity of miRNA machinery is lower in embryos than in other tissues. For example, Yang et al. [63] reported that dicer, an important protein for both miRNA biogenesis and miRNA function, starts expression in 7-day old mouse embryos and remains stable through 17-day embryos .

In this study, we also found that miRNA targets are significantly enriched in ubiquitously expressed genes and largely excluded from tissue-restrictively expressed genes. This result implies that miRNAs play less important role for tissue-specific physiological functions. Instead, the major biological function of miRNAs is to determine the fate of tissue differentiation and maintain the tissue identity.

Our study also indicates that reduction of miRNA expression might cause delineation of differentiated cells, a crucial step towards carcinogenesis. This is consistent with the recent discoveries [64,65]. For example, more than 50% of human miRNAs are located in chromosome regions involved in human cancers [64]. Most recently, miRNA expression

profiles has revealed that most of miRNAs had lower expression levels in tumors compared with normal tissues [65]. Hence, understanding the global role of miRNAs in maintaining lineage of differentiated tissues and cells has great impact on the studies of miRNAs in cancer genetics.

Methods

Datasets used in this study

The datasets used in this study include two complete lists of human miRNA targets published by Lewis et al. [35] and John et al. [31], three complete lists of mouse miRNA targets published by John et al. [31] , Lewis et al. [35] and Krek et al. [33], two complete lists of *Drosophila* miRNA targets published by Enright et al.[30] and Stark et al. [30,30,38], a microarray expression dataset for more than 10,000 human genes in 41 tissues and cell lines [50], a microarray expression dataset for nearly 40,000 known and predicted mRNAs in 55 mouse tissues [51] and a microarray expression data for nearly one-third of all *Drosophila* genes during the whole life cycle [49,51].

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 over 41 human tissues published by Johnson et al. [50] by ranking each gene over all tissues according to its expression level in the respective tissue as described previously [40]. 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 a tissue is higher than that in any other tissues. For each miRNA, we collected all of its predicted targets presented in the microarray dataset and obtained the rank number of each of its targets in any given tissues. The miRNA targets we used in this analysis are

from the datasets published by Lewis et al. [34] and John et al. [31] respectively. On average, 180 targets per miRNA could be found in the microarray dataset.

To facilitate a more global view, we also 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 ranks, 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, $RR = N_{Rank\ 1\text{-}20}/N_{Rank\ 22\text{-}41}$. 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 an 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 = $N_{Rank\ 1\text{-}27}/N_{Rank\ 29\text{-}55}$.

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

To compare the expression level of miRNA targets in mouse tissues with that in mouse embryo, we counted the total numbers of miRNA targets whose expression level is lower and higher respectively in a defined tissue than that in 12.5-day embryo, and calculated the ratio of lower-expressed targets to the higher-expressed targets. A ratio more than one means that the number of lower-expressed targets is more than higher-expressed targets. To obtain the statistical significance of this ratio, we performed resampling statistical tests. In each resampling, we randomly sub-pooled the same number of genes as the number of miRNA targets from the pool of total genes that are lower- and higher-expressed genes in a defined tissue to the embryo. We calculated the ratio of lower-expressed genes to the higher-expressed genes in this sub-pool and defined it as R_{random}. We tested the null hypothesis $R_{random} \geq R_{mirna}$ by performing 5,000 times of resampling tests. We rejected the hypothesis if $p < 0.05$.

A similar method was used to compare the expression of miRNA targets in the different periods of *Drosophila* life cycle with 23-24 h embryo using a microarray dataset published by Arbeitman et al. [49,51], except that the ratio of number of miRNA targets whose expression level in a defined period is two-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. [51], which contains 21622 mouse genes including 2276 predicted miRNA targets, we analyzed the tissue-specificity of miRNA target expression. The mouse genes are classified into 55 groups according to the number of tissues (1 to 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 (21622) for determining if the miRNA targets are enriched in or excluded from the respective group.

Reference List

1. Ambros V: **microRNAs: tiny regulators with great potential.** *Cell* 2001, **107:**823-826.

2. Bartel DP: **MicroRNAs: genomics, biogenesis, mechanism, and function.** *Cell* 2004, **116:**281-297.

3. Cullen BR: **Transcription and processing of human microRNA precursors.** *Mol Cell* 2004, **16:**861-865.

4. Kim VN: **MicroRNA biogenesis: coordinated cropping and dicing.** *Nat Rev Mol Cell Biol* 2005, **6:**376-385.

5. Carmell MA, Hannon GJ: **RNase III enzymes and the initiation of gene silencing.** *Nat Struct Mol Biol* 2004, **11:**214-218.

6. Meister G, Tuschl T: **Mechanisms of gene silencing by double-stranded RNA.** *Nature* 2004, **431:**343-349.

7. Lagos-Quintana M, Rauhut R, Lendeckel W, Tuschl T: **Identification of novel genes coding for small expressed RNAs.** *Science* 2001, **294:**853-858.

8. Lau NC, Lim LP, Weinstein EG, Bartel DP: **An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans.** *Science* 2001, **294:**858-862.

9. Lee RC, Ambros V: **An extensive class of small RNAs in Caenorhabditis elegans.** *Science* 2001, **294:**862-864.

10. Lagos-Quintana M, Rauhut R, Yalcin A, Meyer J, Lendeckel W, Tuschl T: **Identification of tissue-specific microRNAs from mouse.** *Curr Biol* 2002, **12:**735-739.

11. Bentwich I, Avniel A, Karov Y, Aharonov R, Gilad S, Barad O, Barzilai A, Einat P, Einav U, Meiri E et al.: **Identification of hundreds of conserved and nonconserved human microRNAs.** *Nat Genet* 2005, **37:**766-770.

12. Berezikov E, Guryev V, van de BJ, Wienholds E, Plasterk RH, Cuppen E: **Phylogenetic shadowing and computational identification of human microRNA genes.** *Cell* 2005, **120:**21-24.

13. Grad Y, Aach J, Hayes GD, Reinhart BJ, Church GM, Ruvkun G, Kim J: **Computational and experimental identification of C. elegans microRNAs.** *Mol Cell* 2003, **11:**1253-1263.

14. Jones-Rhoades MW, Bartel DP: **Computational identification of plant microRNAs and their targets, including a stress-induced miRNA.** *Mol Cell* 2004, **14:**787-799.

15. Kim J, Krichevsky A, Grad Y, Hayes GD, Kosik KS, Church GM, Ruvkun G: **Identification of many microRNAs that copurify with polyribosomes in mammalian neurons.** *Proc Natl Acad Sci U S A* 2004, **101:**360-365.

16. Lagos-Quintana M, Rauhut R, Meyer J, Borkhardt A, Tuschl T: **New microRNAs from mouse and human.** *RNA* 2003, **9:**175-179.

17. Lim LP, Lau NC, Weinstein EG, Abdelhakim A, Yekta S, Rhoades MW, Burge CB, Bartel DP: **The microRNAs of Caenorhabditis elegans.** *Genes Dev* 2003, **17:**991-1008.

18. Sunkar R, Girke T, Jain PK, Zhu JK: **Cloning and characterization of microRNAs from rice.** *Plant Cell* 2005, **17:**1397-1411.

19. Griffiths-Jones S: **The microRNA Registry.** *Nucleic Acids Res* 2004, **32:**D109-D111.

20. Lee RC, Feinbaum RL, Ambros V: **The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14.** *Cell* 1993, **75:**843-854.

21. Reinhart BJ, Slack FJ, Basson M, Pasquinelli AE, Bettinger JC, Rougvie AE, Horvitz HR, Ruvkun G: **The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans.** *Nature* 2000, **403:**901-906.

22. Johnston RJ, Hobert O: **A microRNA controlling left/right neuronal asymmetry in Caenorhabditis elegans.** *Nature* 2003, **426:**845-849.

23. Brennecke J, Hipfner DR, Stark A, Russell RB, Cohen SM: **bantam encodes a developmentally regulated microRNA that controls cell proliferation and regulates the proapoptotic gene hid in Drosophila.** *Cell* 2003, **113:**25-36.

24. Xu P, Vernooy SY, Guo M, Hay BA: **The Drosophila microRNA Mir-14 suppresses cell death and is required for normal fat metabolism.** *Curr Biol* 2003, **13:**790-795.

25. Chen CZ, Li L, Lodish HF, Bartel DP: **MicroRNAs modulate hematopoietic lineage differentiation.** *Science* 2004, **303:**83-86.

26. Brennecke J, Stark A, Russell RB, Cohen SM: **Principles of microRNA-target recognition.** *PLoS Biol* 2005, **3:**e85.

27. Doench JG, Sharp PA: **Specificity of microRNA target selection in translational repression.** *Genes Dev* 2004, **18:**504-511.

28. Kiriakidou M, Nelson PT, Kouranov A, Fitziev P, Bouyioukos C, Mourelatos Z, Hatzigeorgiou A: **A combined computational-experimental approach predicts human microRNA targets.** *Genes Dev* 2004, **18:**1165-1178.

29. Kloosterman WP, Wienholds E, Ketting RF, Plasterk RH: **Substrate requirements for let-7 function in the developing zebrafish embryo.** *Nucleic Acids Res* 2004, **32:**6284-6291.

30. Enright AJ, John B, Gaul U, Tuschl T, Sander C, Marks DS: **MicroRNA targets in Drosophila.** *Genome Biol* 2003, **5:**R1.

31. John B, Enright AJ, Aravin A, Tuschl T, Sander C, Marks DS: **Human MicroRNA targets.** *PLoS Biol* 2004, **2:**e363.

32. Kiriakidou M, Nelson PT, Kouranov A, Fitziev P, Bouyioukos C, Mourelatos Z, Hatzigeorgiou A: **A combined computational-experimental approach predicts human microRNA targets.** *Genes Dev* 2004, **18:**1165-1178.

33. Krek A, Grun D, Poy MN, Wolf R, Rosenberg L, Epstein EJ, MacMenamin P, da P, I, Gunsalus KC, Stoffel M et al.: **Combinatorial microRNA target predictions.** *Nat Genet* 2005, **37:**495-500.

34. Lewis BP, Shih IH, Jones-Rhoades MW, Bartel DP, Burge CB: **Prediction of mammalian microRNA targets.** *Cell* 2003, **115:**787-798.

35. Lewis BP, Burge CB, Bartel DP: **Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets.** *Cell* 2005, **120:**15-20.

36. Rhoades MW, Reinhart BJ, Lim LP, Burge CB, Bartel B, Bartel DP: **Prediction of plant microRNA targets.** *Cell* 2002, **110:**513-520.

37. Robins H, Li Y, Padgett RW: **Incorporating structure to predict microRNA targets.** *Proc Natl Acad Sci U S A* 2005, **102:**4006-4009.

38. Stark A, Brennecke J, Russell RB, Cohen SM: **Identification of Drosophila MicroRNA targets.** *PLoS Biol* 2003, **1:**E60.

39. Wang XJ, Reyes JL, Chua NH, Gaasterland T: **Prediction and identification of Arabidopsis thaliana microRNAs and their mRNA targets.** *Genome Biol* 2004, **5:**R65.

40. Lim LP, Lau NC, Garrett-Engele P, Grimson A, Schelter JM, Castle J, Bartel DP, Linsley PS, Johnson JM: **Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs.** *Nature* 2005, **433:**769-773.

41. Krutzfeldt J, Rajewsky N, Braich R, Rajeev KG, Tuschl T, Manoharan M, Stoffel M: **Silencing of microRNAs in vivo with 'antagomirs'.** *Nature* 2005.

42. Cheung VG, Spielman RS, Ewens KG, Weber TM, Morley M, Burdick JT: **Mapping determinants of human gene expression by regional and genome-wide association.** *Nature* 2005, **437:**1365-1369.

43. Huminiecki L, Wolfe KH: **Divergence of spatial gene expression profiles following species-specific gene duplications in human and mouse.** *Genome Res* 2004, **14:**1870-1879.

44. Morley M, Molony CM, Weber TM, Devlin JL, Ewens KG, Spielman RS, Cheung VG: **Genetic analysis of genome-wide variation in human gene expression.** *Nature* 2004, **430:**743-747.

45. Rifkin SA, Houle D, Kim J, White KP: **A mutation accumulation assay reveals a broad capacity for rapid evolution of gene expression.** *Nature* 2005, **438:**220-223.

46. Rodwell GE, Sonu R, Zahn JM, Lund J, Wilhelmy J, Wang L, Xiao W, Mindrinos M, Crane E, Segal E et al.: **A transcriptional profile of aging in the human kidney.** *PLoS Biol* 2004, **2:**e427.

47. Stolc V, Gauhar Z, Mason C, Halasz G, van Batenburg MF, Rifkin SA, Hua S, Herreman T, Tongprasit W, Barbano PE et al.: **A gene expression map for the euchromatic genome of Drosophila melanogaster.** *Science* 2004, **306:**655-660.

48. Stuart JM, Segal E, Koller D, Kim SK: **A gene-coexpression network for global discovery of conserved genetic modules.** *Science* 2003, **302:**249-255.

49. Arbeitman MN, Furlong EE, Imam F, Johnson E, Null BH, Baker BS, Krasnow MA, Scott MP, Davis RW, White KP: **Gene expression during the life cycle of Drosophila melanogaster.** *Science* 2002, **297:**2270-2275.

50. Johnson JM, Castle J, Garrett-Engele P, Kan Z, Loerch PM, Armour CD, Santos R, Schadt EE, Stoughton R, Shoemaker DD: **Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays.** *Science* 2003, **302:**2141-2144.

51. Zhang W, Morris QD, Chang R, Shai O, Bakowski MA, Mitsakakis N, Mohammad N, Robinson MD, Zirngibl R, Somogyi E et al.: **The functional landscape of mouse gene expression.** *J Biol* 2004, **3:**21.

52. Wienholds E, Kloosterman WP, Miska E, Alvarez-Saavedra E, Berezikov E, de Bruijn E, Horvitz HR, Kauppinen S, Plasterk RH: **MicroRNA expression in zebrafish embryonic development.** *Science* 2005, **309:**310-311.

53. Babak T, Zhang W, Morris Q, Blencowe BJ, Hughes TR: **Probing microRNAs with microarrays: tissue specificity and functional inference.** *RNA* 2004, **10:**1813-1819.

54. Barad O, Meiri E, Avniel A, Aharonov R, Barzilai A, Bentwich I, Einav U, Gilad S, Hurban P, Karov Y et al.: **MicroRNA expression detected by oligonucleotide microarrays: system establishment and expression profiling in human tissues.** *Genome Res* 2004, **14:**2486-2494.

55. Baskerville S, Bartel DP: **Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes.** *RNA* 2005, **11:**241-247.

56. Houbaviy HB, Murray MF, Sharp PA: **Embryonic stem cell-specific MicroRNAs.** *Dev Cell* 2003, **5:**351-358.

57. Krichevsky AM, King KS, Donahue CP, Khrapko K, Kosik KS: **A microRNA array reveals extensive regulation of microRNAs during brain development.** *RNA* 2003, **9:**1274-1281.

58. Liu CG, Calin GA, Meloon B, Gamliel N, Sevignani C, Ferracin M, Dumitru CD, Shimizu M, Zupo S, Dono M et al.: **An oligonucleotide microchip for genome-wide microRNA profiling in human and mouse tissues.** *Proc Natl Acad Sci U S A* 2004, **101:**9740-9744.

59. Miska EA, Alvarez-Saavedra E, Townsend M, Yoshii A, Sestan N, Rakic P, Constantine-Paton M, Horvitz HR: **Microarray analysis of microRNA expression in the developing mammalian brain.** *Genome Biol* 2004, **5:**R68.

60. Nelson PT, Baldwin DA, Scearce LM, Oberholtzer JC, Tobias JW, Mourelatos Z: **Microarray-based, high-throughput gene expression profiling of microRNAs.** *Nat Methods* 2004, **1:**155-161.

61. Sempere LF, Freemantle S, Pitha-Rowe I, Moss E, Dmitrovsky E, Ambros V: **Expression profiling of mammalian microRNAs uncovers a subset of brain-expressed microRNAs with possible roles in murine and human neuronal differentiation.** *Genome Biol* 2004, **5:**R13.

62. Thomson JM, Parker J, Perou CM, Hammond SM: **A custom microarray platform for analysis of microRNA gene expression.** *Nat Methods* 2004, **1:**47-53.

63. Yang WJ, Yang DD, Na S, Sandusky GE, Zhang Q, Zhao G: **Dicer is required for embryonic angiogenesis during mouse development.** *J Biol Chem* 2005, **280:**9330-9335.

64. Calin GA, Sevignani C, Dumitru CD, Hyslop T, Noch E, Yendamuri S, Shimizu M, Rattan S, Bullrich F, Negrini M et al.: **Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers.** *Proc Natl Acad Sci U S A* 2004, **101:**2999-3004.

65. Lu J, Getz G, Miska EA, Alvarez-Saavedra E, Lamb J, Peck D, Sweet-Cordero A, Ebert BL, Mak RH, Ferrando AA et al.: **MicroRNA expression profiles classify human cancers.** *Nature* 2005, **435:**834-838.

Figure legends:

Figure 1

Ranking of the expression levels of miRNA targets in human tissues. Each miRNA target was ranked over 41 human tissues as described in the "Methods" according to its expression level. The RR means the ratio of the total number ($N_{Rank1-20}$) of the miRNA targets with a lower rank number (between 1 and 20) to the total number ($N_{Rank22-41}$) 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.

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 "Methods". The RR means the ratio of the total number ($N_{Rank1-27}$) of the miRNA targets with a lower rank number (between 1 and 27) to the total number ($N_{Rank29-55}$) 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 ($N_{Rank1-27}$) of the targets with lower rank number (between 1 and 27) is more than the number ($N_{Rank29-55}$) 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 [51] which contains 21622 mouse genes including 2276 miRNA targets published by John et al. [31].

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 21622 genes over 55 mouse samples published by Zhang et al. [51] 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. [31] , Lewis et al. [35] and Krek et al. [33] 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.

Figure 5

Analysis of the expression of miRNA targets in *Drosophila*. **(a, b)** A published microarray dataset [49,51] 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 ($N_{<1/2 \times E23\text{-}24h}$) was divided by the total number of those with expression level two times higher than that in 23-24 h embryo ($N_{>2 \times E23\text{-}24h}$). Two sets of *Drosophila* miRNA targets published by Enright et al.[30] and Stark et al. [30,30,38] 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 ($N_{<1/2 \times E23\text{-}24h}$/ $N_{>2 \times E23\text{-}24h}$) obtained using two different miRNA target datasets.

Figure 6

MiRNAs more frequently target ubiquitously expressed genes than tissue-restrictively expressed genes. A microarray dataset published by Zhang et al [51], which contains 21622 genes including 2276 miRNA targets [31], was used in this analysis as described in the "Methods". The ratio of the miRNA target genes to the total genes expressed in an indicated number of tissues was calculated.

Supporting Information

Figure S1

Ranking of the expression levels of miRNA targets published by John et al. [31] in human tissues. Each miRNA target was ranked over 41 human tissues as described in the "Methods" according to its expression level. The RR means the ratio of the total number ($N_{Rank1-20}$) of the miRNA targets with a lower rank number (between 1 and 20) to the total number ($N_{Rank22-41}$) 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)** The RR values of the 82 miRNAs are plotted as a function of human tissues and cell lines. **(b)** The RR values over the 41 tissues and cell lines are plotted as a function of miRNAs (1, let-7a; 2, let-7b; 3, let-7c; 4, let-7d; 5, let-7e; 6, let-7f; 7, let-7g; 8, let-7i; 9, mir-103; 10, mir-106b; 11, mir-107; 12, mir-122a; 13, mir-124a; 14, mir-125a; 15, mir-125b; 16, mir-128a; 17, mir-128b; 18, mir-129; 19, mir-130a; 20, mir-130b; 21, mir-132; 22, mir-133a; 23, mir-134; 24, mir-136; 25, mir-141; 26, mir-143; 27, mir-144; 28, mir-145; 29, mir-146; 30, mir-149; 31, mir-150; 32, mir-152; 33, mir-15a; 34, mir-15b; 35, mir-17_5p; 36, mir-18; 37, mir-181a; 38, mir-181b; 39, mir-183; 40, mir-184; 41, mir-185; 42, mir-187; 43, mir-188; 44, mir-189; 45, mir-194; 46, mir-195; 47, mir-196; 48, mir-197; 49, mir-198; 50, mir-199a; 51, mir-199a*; 52, mir-20; 53, mir-200c; 54, mir-204; 55, mir-205; 56, mir-206; 57, mir-210; 58, mir-211; 59, mir-212; 60, mir-214; 61, mir-217; 62, mir-22; 63, mir-222; 64, mir-24; 65, mir-25; 66, mir-27b; 67, mir-28; 68, mir-296; 69, mir-29a; 70, mir-29c; 71, mir-301;

72, mir-302; 73, mir-30a; 74, mir-31; 75, mir-320; 76, mir-321; 77, mir-34a; 78, mir-34c; 79, mir-361; 80, mir-9; 81, mir-93; 82, mir-98).

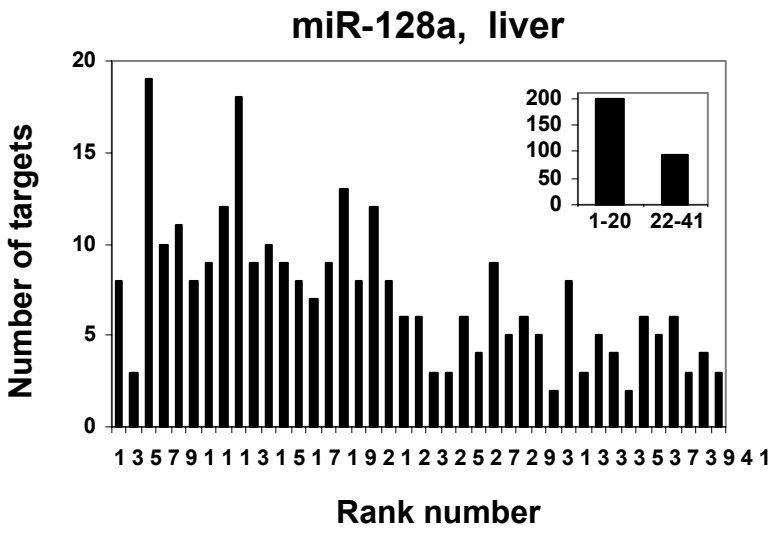
Table S1. Ranking of the expression of human miRNA targets and total genes

Table S2. The numbers of miRNA targets and total genes whose expression level is lower than that in mouse 12.5-day embryo

Table S3. The numbers of miRNA targets and total genes whose expression level is two fold lower than that in 23-24h embryo

Figure 1

a



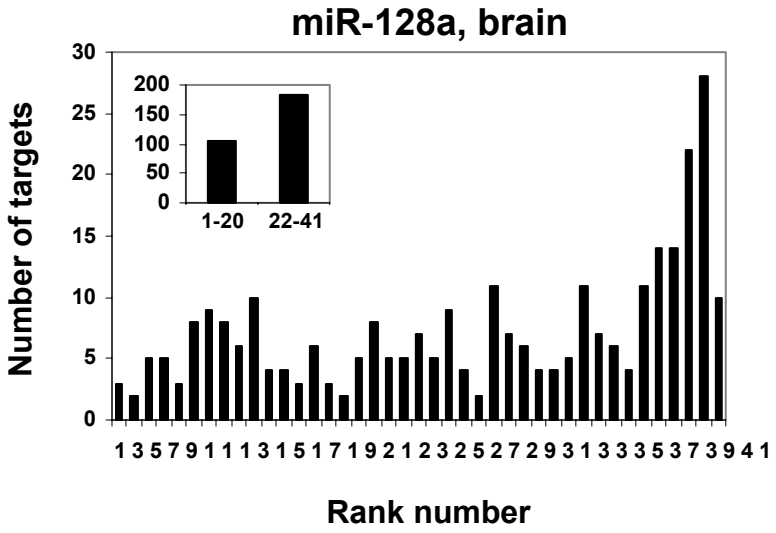


Figure 1a

Figure 2

Figure 1b and 1c

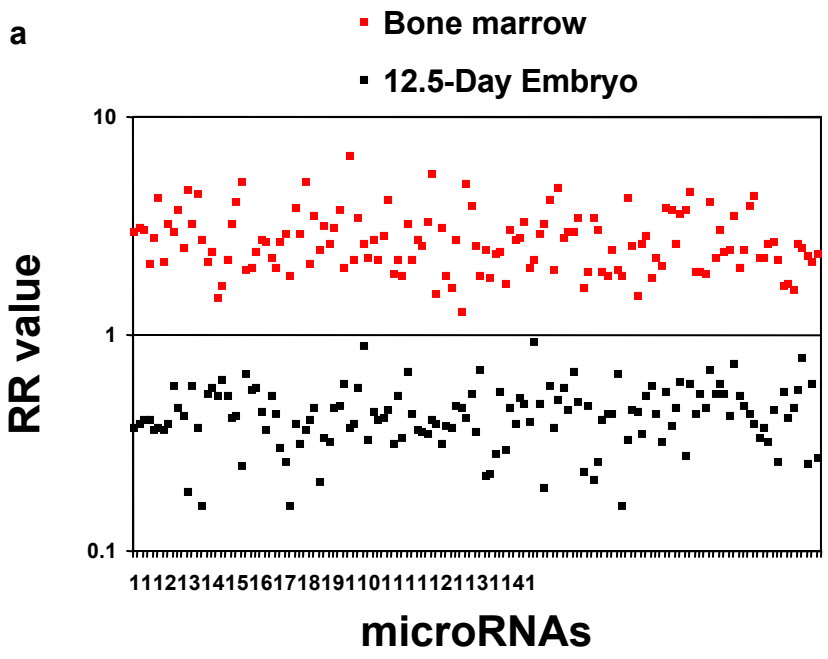

a
Bone marrow
12.5-Day Embryo
10
1
0.1
RR value
111213141516171819110111112113114 1
microRNAs

Figure 2a

Figure 4

Figure 2b and 2c

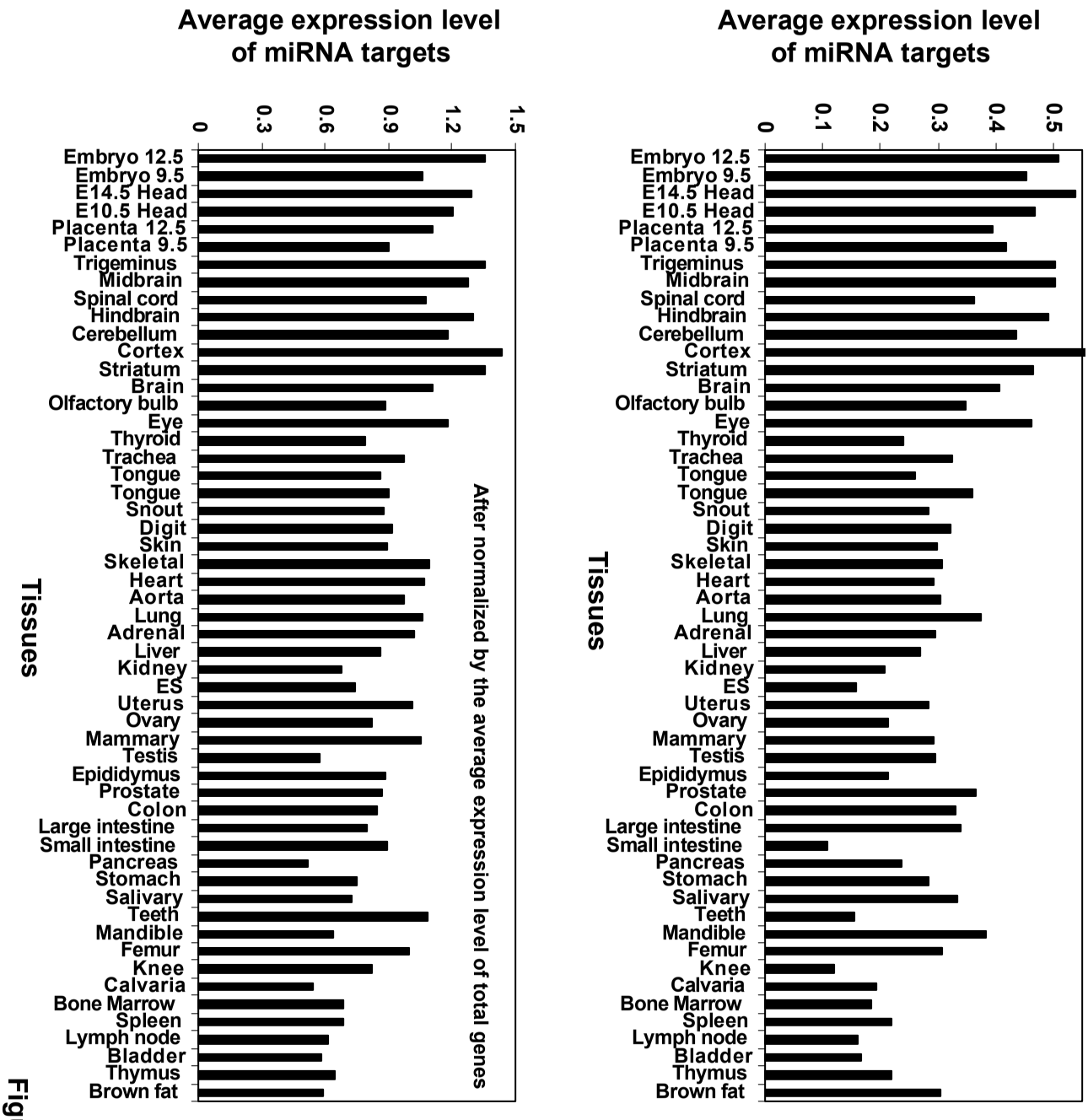

Average expression level of miRNA targets
0
0.1
0.2
0.3
0.4
0.5
Embryo 12.5
Embryo 9.5
E14.5 Head
E10.5 Head
Placenta 12.5
Placenta 9.5
Trigeminus
Midbrain
Spinal cord
Hindbrain
Cerebellum
Cortex
Striatum
Brain
Olfactory bulb
Eye
Thyroid
Trachea
Tongue
Tongue
Snout
Digit
Skin
Skeletal
Heart
Aorta
Lung
Adrenal
Liver
Kidney
ES
Uterus
Ovary
Mammary
Testis
Epididymus
Prostate
Colon
Large intestine
Small intestine
Pancreas
Stomach
Salivary
Teeth
Mandible
Femur
Knee
Calvaria
Bone Marrow
Spleen
Lymph node
Bladder
Thymus
Brown fat
Tissues
Average expression level of miRNA targets
0
0.3
0.6
0.9
1.2
1.5
After normalized by the average expression level of total genes
Tissues

Figure 3

Figure 6

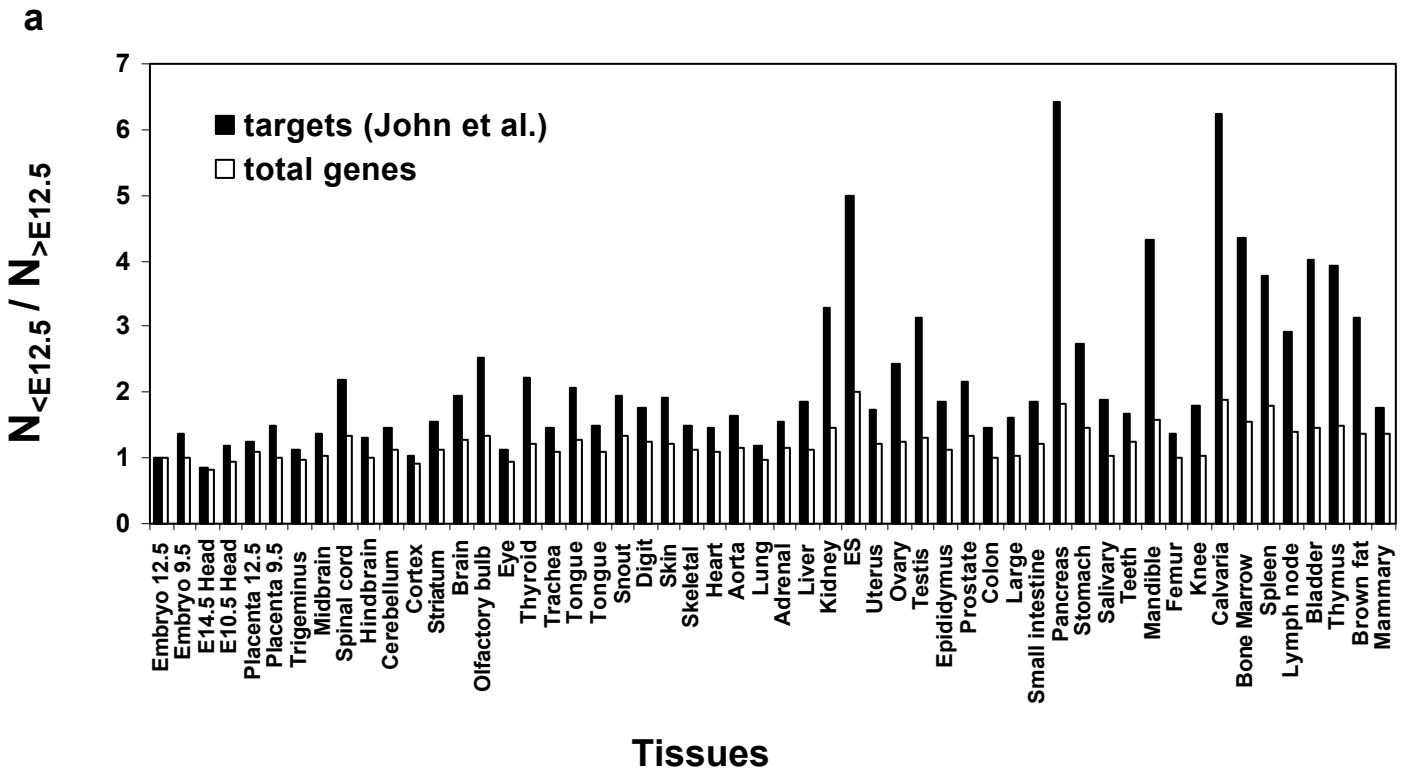


Figure 4a

Figure 7

b

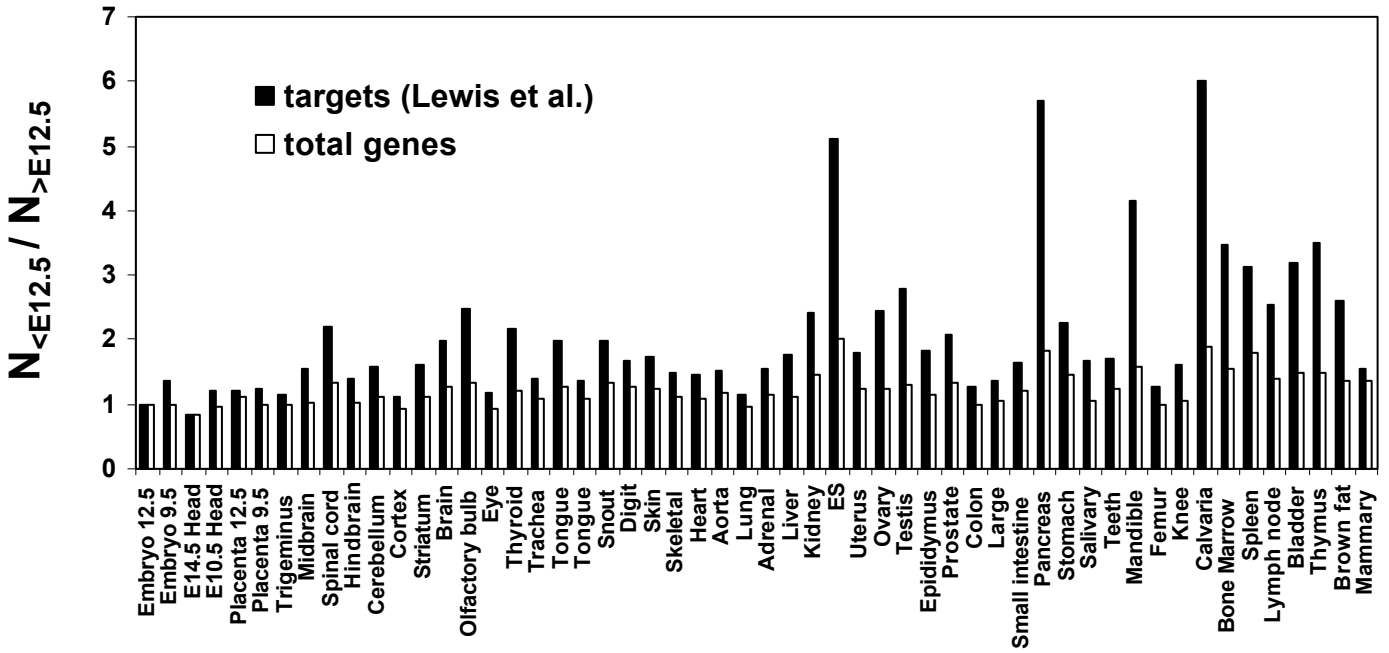


Figure 4b

Figure 8

c

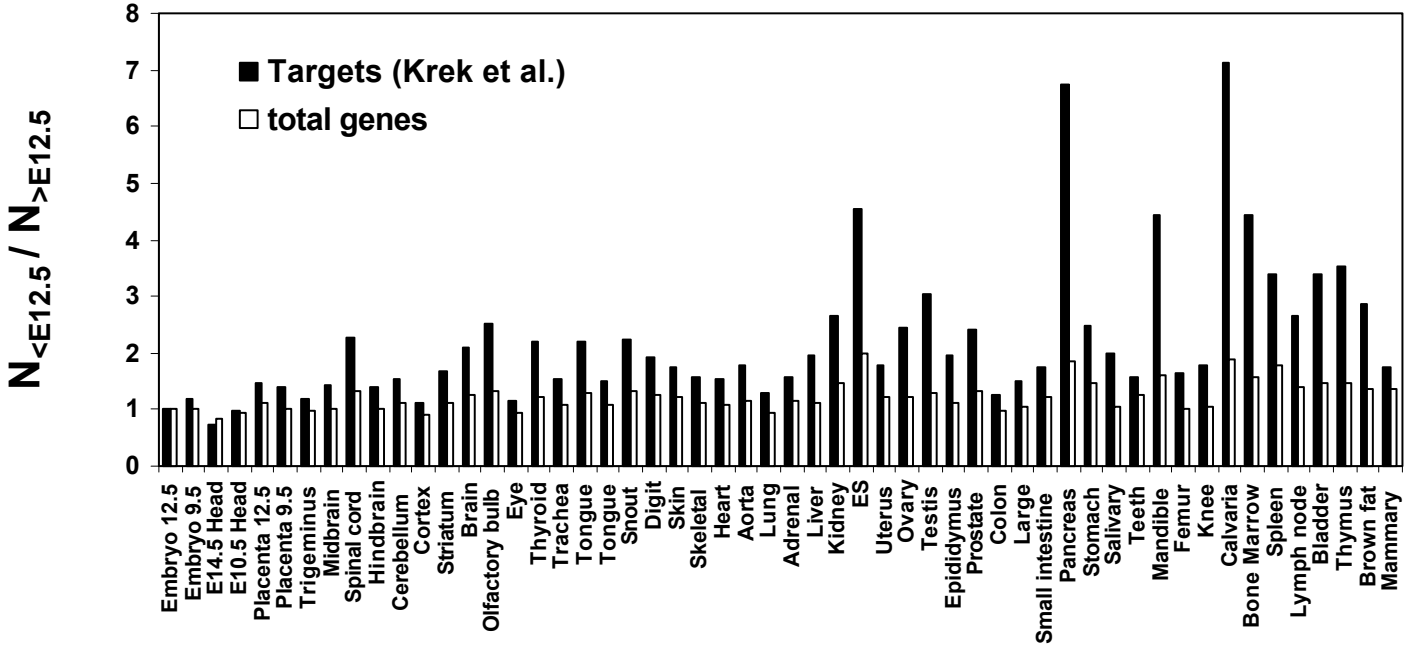


Figure 4c

D

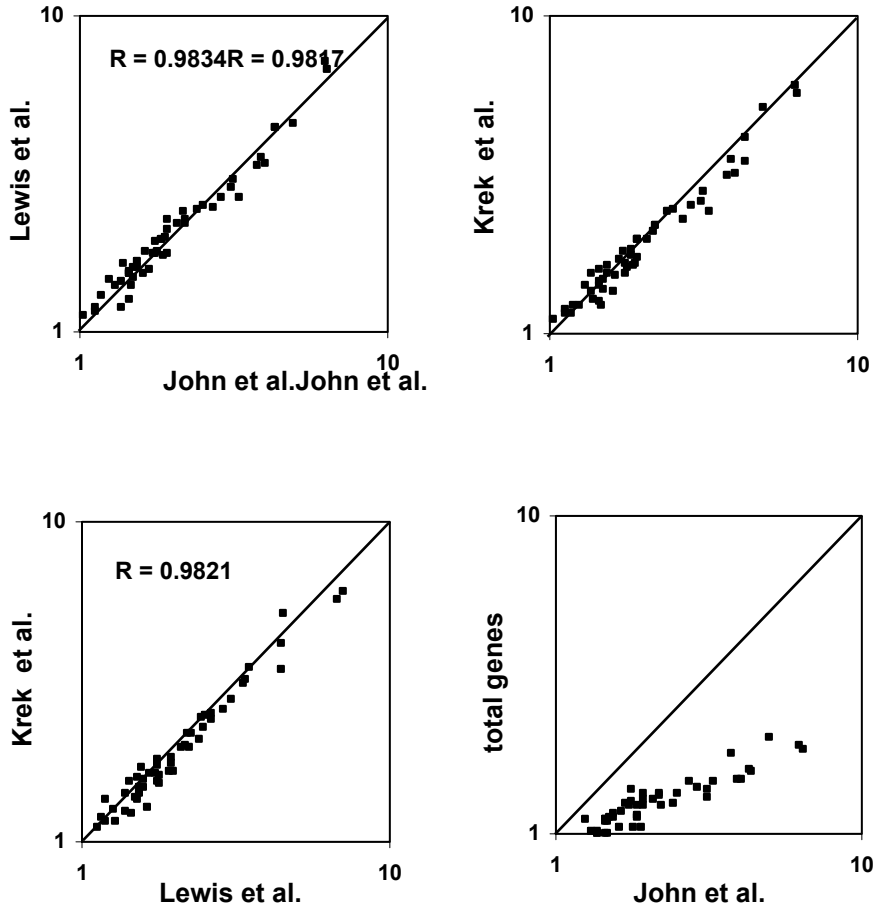
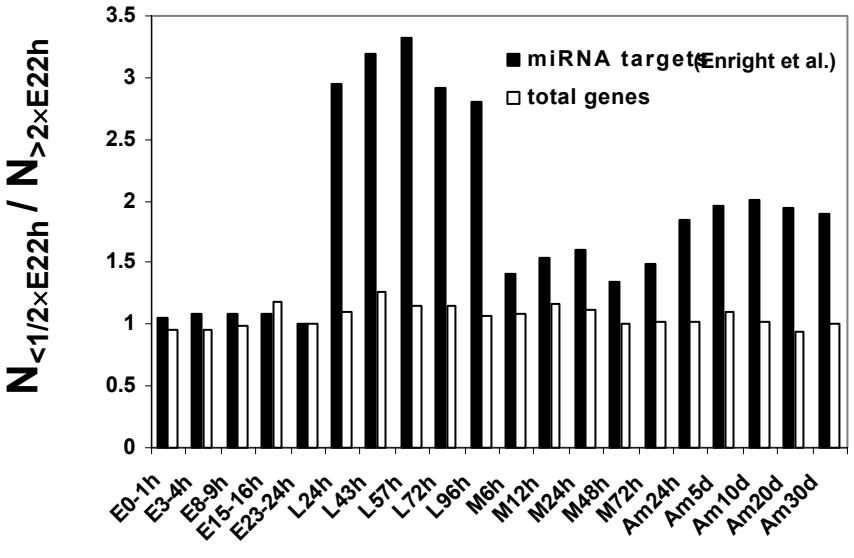

10
R = 0.9834R = 0.9817
Lewis et al.
1
1
10
John et al.John et al.
10
Krek et al.
1
1
10
10
R = 0.9821
Krek et al.
1
1
10
Lewis et al.
10
total genes
1
1
10
John et al.

Figure 4D

Figure 10

a



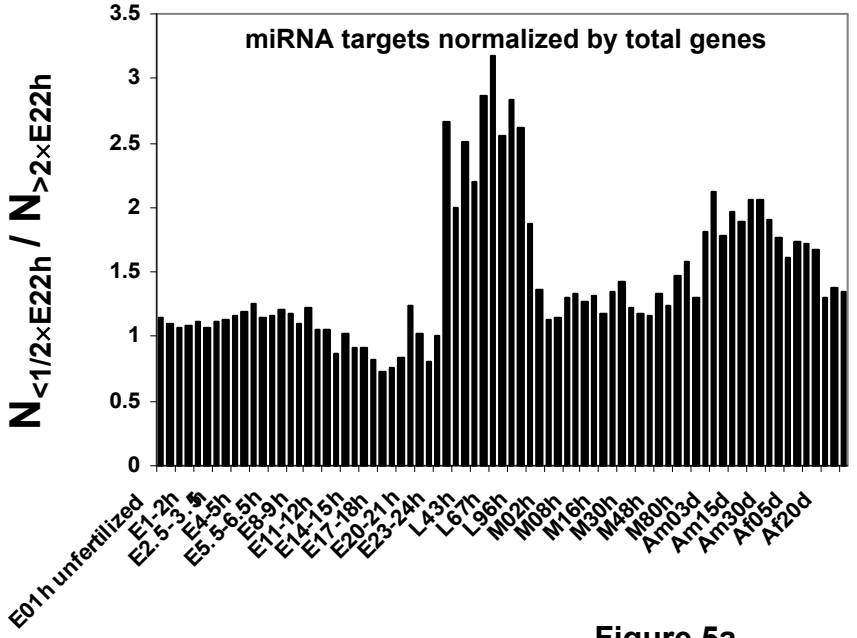
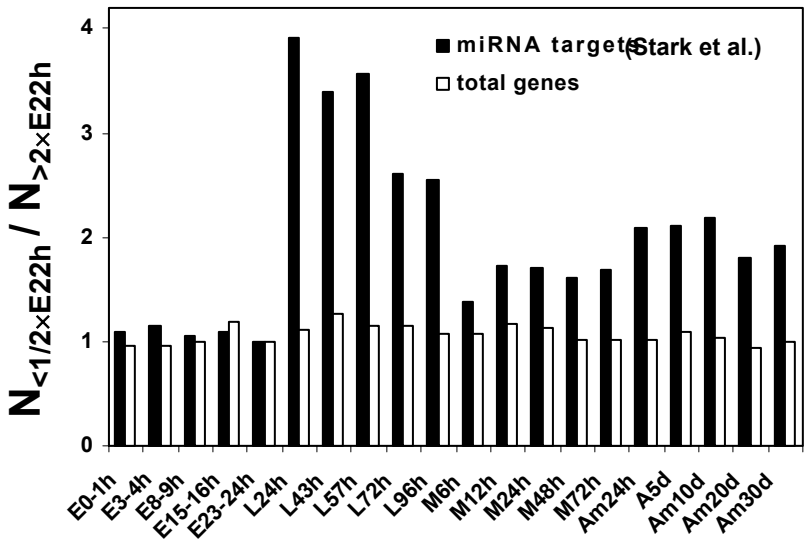


Figure 5a

b


miRNA targets (Stark et al.)
total genes
N<1/2×E22h / N>2×E22h
0
1
2
3
4
E0-1h
E3-4h
E8-9h
E15-16h
E23-24h
L24h
L43h
L57h
L72h
L96h
M6h
M12h
M24h
M48h
M72h
Am24h
A5d
Am10d
Am20d
Am30d

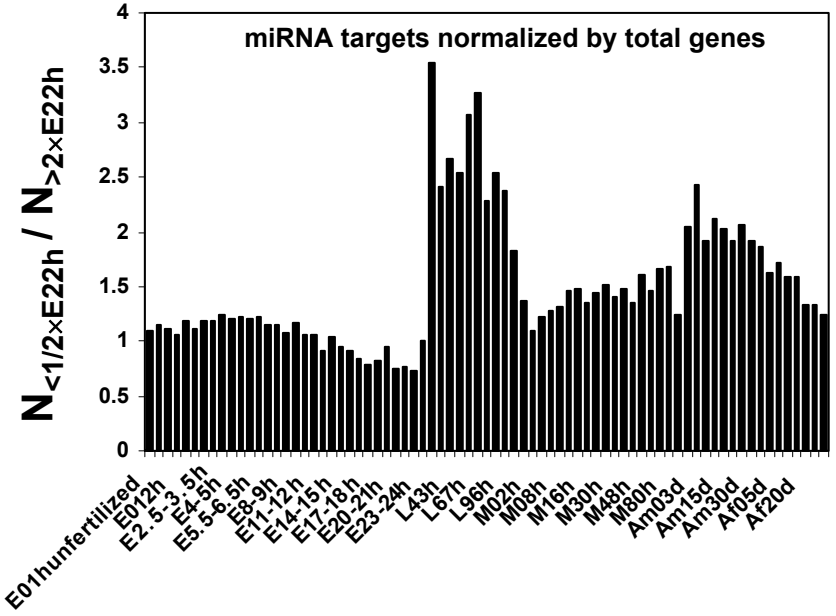

miRNA targets normalized by total genes
N<1/2×E22h / N>2×E22h
0
0.5
1
1.5
2
2.5
3
3.5
4
E01hunfertilized
E012h
E2.5-3.5h
E4-5h
E5.5-6.5h
E8-9h
E11-12h
E14-15h
E17-18h
E20-21h
E23-24h
L43h
L67h
L96h
M02h
M08h
M16h
M30h
M48h
M80h
Am03d
Am15d
Am30d
Af05d
Af20d

Figure 5b

c

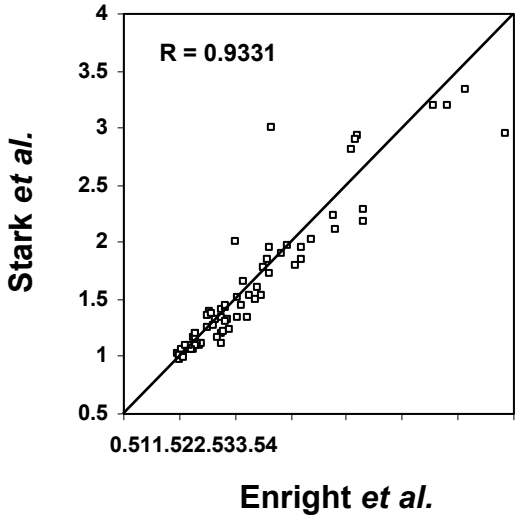

R = 0.9331
Stark *et al.*
Enright *et al.*
4
3.5
3
2.5
2
1.5
1
0.5
0.511.522.533.54

Figure 5c

Figure 13

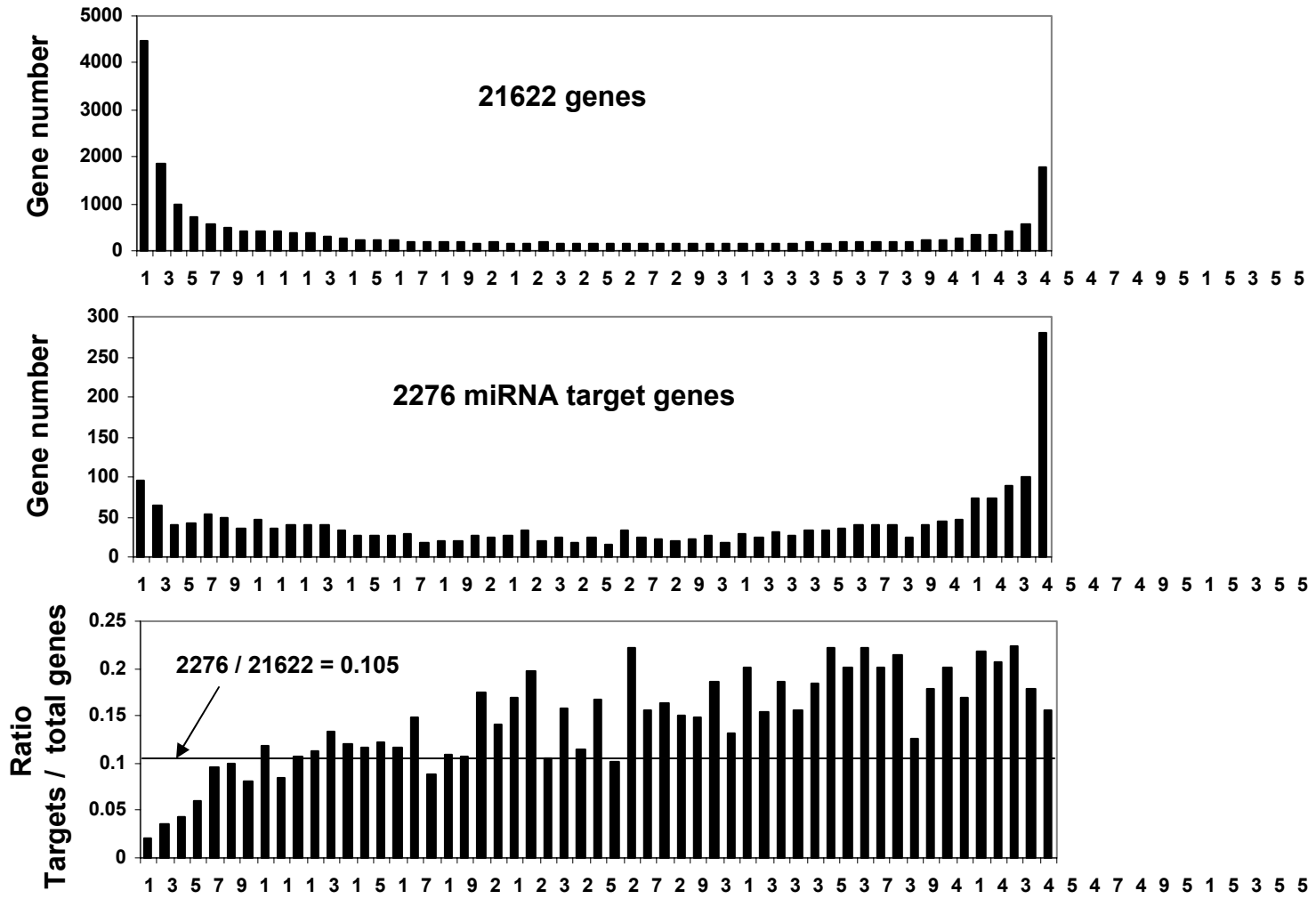
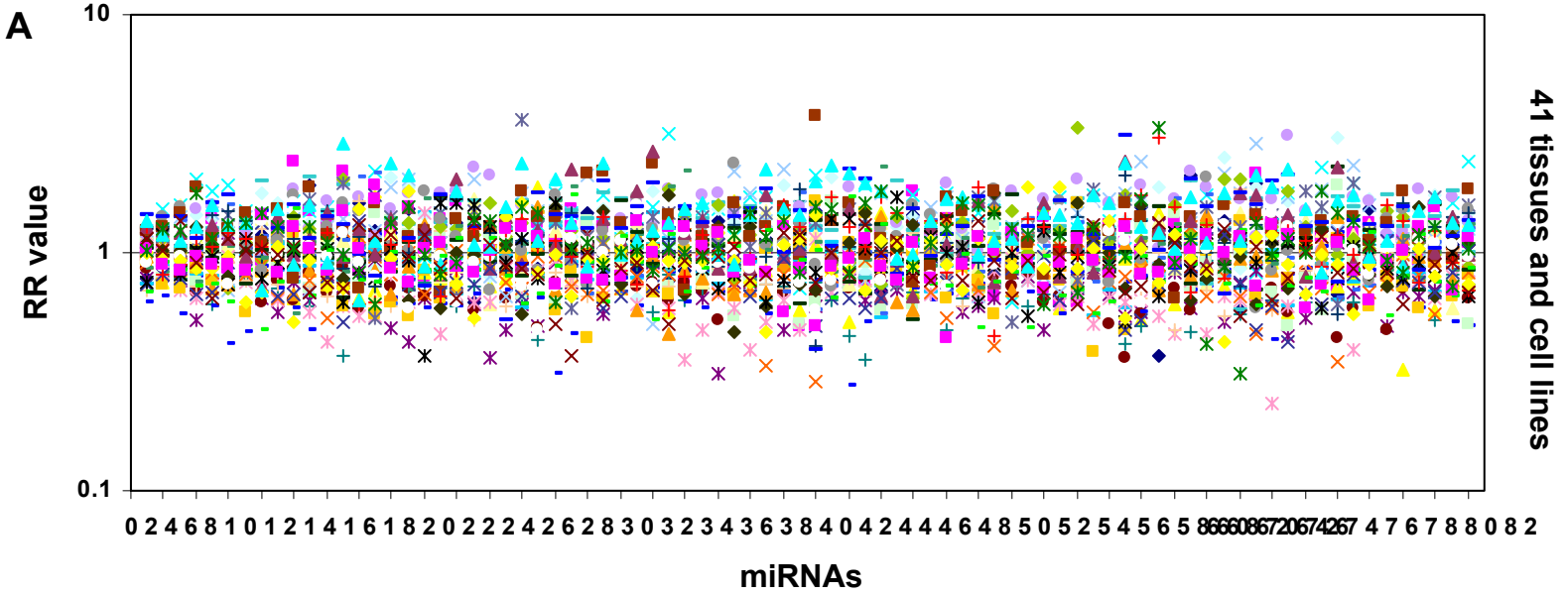


Figure 6

Figure 14



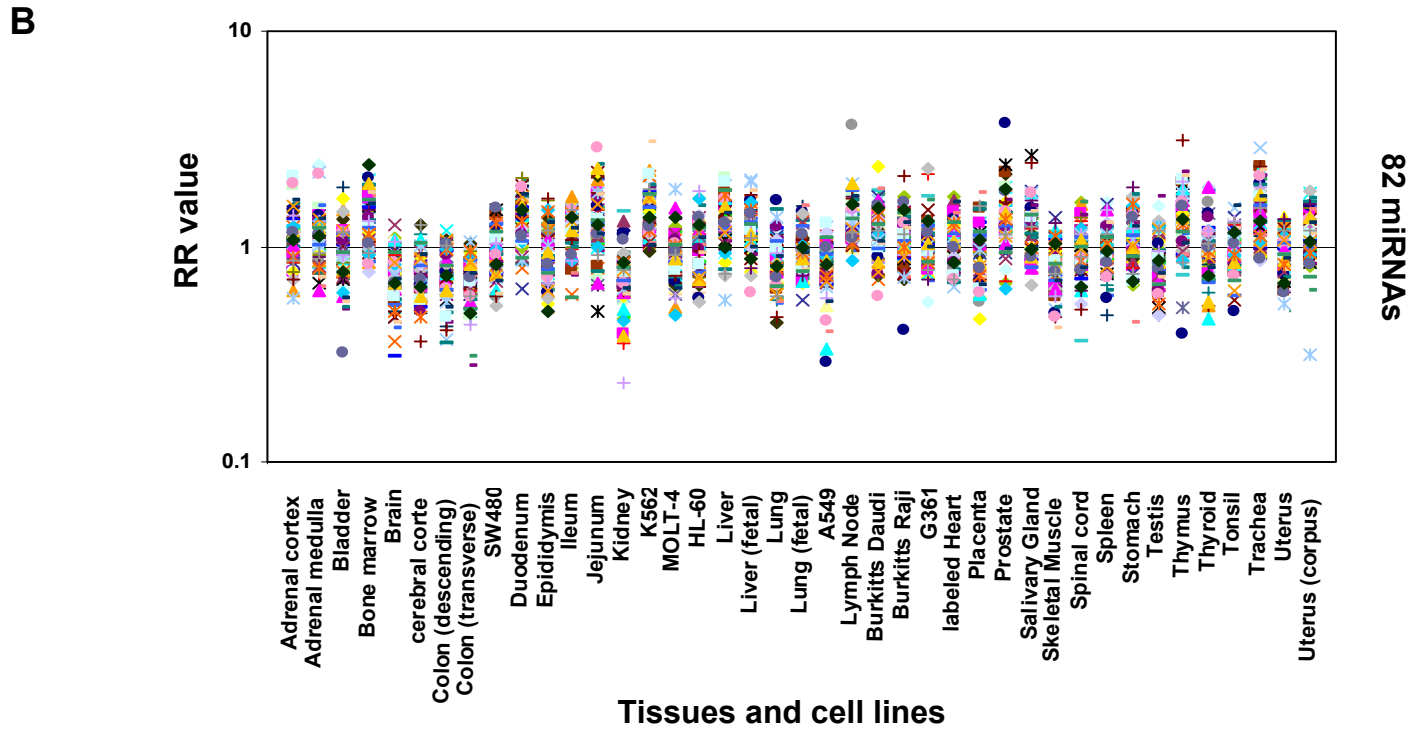


Figure S1

Additional files provided with this submission:

Additional file 3 : supporting Table 3.xls : 27Kb
http://genomebiology.com/imedia/2117590123887542/sup3.XLS
Additional file 2 : supporting Table 2.xls : 32Kb
http://genomebiology.com/imedia/9015394488875428/sup2.XLS
Additional file 1 : supporting Table 1.xls : 1460Kb
http://genomebiology.com/imedia/2071719576887542/sup1.XLS