Primary transcripts and expressions of mammal intergenic microRNAs detected by mapping ESTs to their flanking sequences

Jin Gu,* Tao He,* Yunfei Pei, Fei Li, Xiaowo Wang, Jing Zhang, Xuegong Zhang, Yanda Li

Bioinformatics Division, Tsinghua National Laboratory for Information Science and Technolgy, and Department of Automation, Tsinghua University, Beijing, 100084, China

Abstract

MicroRNAs (miRNAs) are a class of approximately 22-nt small RNAs that regulate posttranscriptional gene expression. Thousands of expressed sequence tags (ESTs) have been identified by using upstream 2500-nt and downstream 4000-nt flanking sequences to BLAST in the dbEST database. The cotranscription of the miRNAs and their flanking sequences covered by the matched ESTs is verified by RT-PCR. It directly reveals that a large portion of mammalian intergenic miRNAs are first transcribed as long primary transcripts (pri-miRNAs). Also, the transcripts' ranges of tens of pri-miRNAs are predicted by the EST-extension method. We then extracted the tissue-specific expression information from the annotations of the matched ESTs and established the expression profile of the studied miRNAs for tens of tissues. This provided a new way to establish the expression profiles of miRNAs. Results show that the human brain, lung, liver, and eye and the mouse brain, eye, and mammary gland are tissues in which enriched numbers of miRNAs are expressed.

Introduction

MiRNAs have attracted much attention in recent years because of their important roles in posttranscriptional regulation by binding to the complementary sequences in 3′ untranslated regions (UTRs) of the target genes (Ambros 2004; Bartel 2004). Hundreds of miRNAs have been identified by computational and experimental methods. To understand the transcription and expression of miRNAs,

---

*This authors contributed equally to this work.

Correspondence to: Yanda Li; E-mail: daulyd@tsinghua.edu.cn

many studies have been conducted. Rodriguez et al. (2004) suggest that the intragenic miRNAs are transcribed in parallel with their host genes. Lee et al. (2004) verified that several intergenic miRNAs are first transcribed as long primary transcripts (pri-miRNAs) by *pol*II. Resembling the mRNA structure, the pri-miRNAs contain 5′ caps and 3′-poly(A) tails (Cai et al. 2004; Lee et al. 2004). These long pri-miRNAs are then processed into pre-miRNAs and miRNAs. The question remains whether it is a universal mechanism of the intergenic miRNAs to be first transcribed as the long pri-miRNAs.

Single-pass sequencing of cDNA clones is an important method for gene transcription and expression analysis (Adams et al. 1991). These sequences, generally known as expressed sequenced tags (ESTs), have been produced for several mammals, especially human, mouse, and rat. Considering that over 10 million ESTs have rather universal coverage of the three mammals' transcriptomes, mapping ESTs to the flanking sequences of the intergenic miRNAs will provide much information about their transcription and expression. Smalheiser (2003) tried to study the transcription and function of miRNAs with ESTs. He used the approximately 100-nt precursor sequences of miRNAs (pre-miRNAs) to search for matched ESTs; less than half of the examined miRNAs (41 of 90) were matched in the database. Much longer flanking sequences are used in our study. About 76% of the investigated miRNAs (379 of 501) can find matched ESTs in their upstream 2500-nt and downstream 4000-nt flanking sequences. RT-PCR experiments were conducted to verify the cotranscription of the miRNAs and their flanking sequences covered by the matched ESTs. Although previous studies have proven the existence of long primary transcripts of several miRNAs, we provide the first direct evidence that a large portion of mammalian intergenic miRNAs are first tran-

scribed as the long pri-miRNAs. Also, using the EST-extension method, the primary transcripts are predicted for tens of miRNAs.

Studying the temporal and spatial expression of miRNAs is crucial for understanding their functions. For example, the *C. elegans* let-7 is specifically expressed in larval stage 3, stage 4, and adult, controlling the development from the larva to adult (Reinhart et al. 2000); the mouse miR-181 is preferentially expressed in the B-lymphoid cells of bone marrow, and it modulates the hematopoietic lineage differentiation (Chen et al. 2004). Many experimental methods, including Northern blot (Sempere et al. 2004; Valoczi et al. 2004), cloning (Lagos-Quintana et al. 2002, 2003), and microarray (Babak et al. 2004; Barad et al. 2004; Baskerville and Bartel 2005; Calin et al. 2004; Krichevsky et al. 2003; Liang et al. 2005; Liu et al. 2004; Miska et al. 2004; Sioud et al. 2004; Sun et al. 2004), have been devoted to studying the expression of miRNAs. These experimental methods are still expensive and time-consuming, so it will be beneficial if computational methods can help provide large-scale expression profiles of miRNAs based on existing data. Mapping of ESTs to the flanking sequences provides a new way to study the expression of the intergenic miRNAs. Because there are instances when miRNAs are initially transcribed as long primary transcripts, we can establish the expression profile by examining the tissue-specific and developmental-stage–specific information of the ESTs that are matched to the flanking sequences of the intergenic miRNAs (the flanking-EST method). The clustered miRNAs are shown to be coexpressed (Baskerville and Bartel 2004), providing further support for the coexpression of the miRNAs and their nearby ESTs. In this article we apply the flanking-EST method to all the reported human, mouse, and rat intergenic miRNAs and obtain their expression profiles.

Materials and methods

Data. The human, mouse, and rat miRNA sequences were downloaded from the miRNA Registry database (release 8.0) (ftp://ftp.sanger.ac.uk/pub/databases/Rfam/miRNA/) (Griffiths-Jones 2004). The human (build 35, May 2004 assembly), mouse (build 35, August 2005 assembly), and rat (RGSC v3.4, November 2004 assembly) genome sequences and annotations were downloaded from the NCBI website (ftp://ftp.ncbi.nih.gov/genomes/). All 836 listed mammalian miRNAs were mapped to the genomes using the MEGABLAST program of the BLAST package (Zhang et al. 2000). Then 501

intergenic miRNAs were identified by comparing the locations of the pre-miRNAs with the annotated regions in the three genomes. The upstream (–) 10,000-nt and downstream (+) 10,000-nt flanking sequences of these miRNAs were extracted from the genome sequences (the regions that overlapped with any annotated transcripts NM_XXXXX or XM_XXXXX were truncated). Mammalian ESTs were downloaded from the dbEST database (http://www.ncbi.nlm.nih.gov/dbEST/index.html).

Mapping of ESTs to the mammalian intergenic MiRNAs. To study the transcription of the mammalian intergenic miRNAs, the −0,000-nt to +10,000-nt flanking sequence of the miRNAs was then used to BLAST in the dbEST (according to organism) with the command "megablast -e 1e-100 -F "m L" -D 3" (MEGABLAST, Zhang et al. 2000). The stand-alone BLAST packages (v2.2.10) were obtained from the NCBI website (ftp://ftp.ncbi.nlm.nih.gov/blast/executables/). The ESTs with an aligned length less than 90% of their original length were filtered out. Then the remaining ESTs were used to BLAST against the genome sequences; the ESTs with multiple matched loci in the genome (a matched locus was defined as a locus with the aligned length larger than 90% of the EST length) were filtered out. After finishing above filtering, it was found that most of the matched ESTs were located between the −2500-nt and the +4000-nt regions and few ESTs could be matched outside this region (Fig. 1). Therefore, the the ESTs matched to the −2500-nt to +4000-nt flanking regions were chosen for the following analysis.

The accession numbers, library name, organs, tissues, and developmental stages were extracted from dbEST entries to establish the expression profile. A miRNA was regarded as expressed in a specific tissue if any of the matched ESTs of that miRNA were annotated as derived from that specific tissue.

To predict the pri-miRNAs by the EST-extension method, the EST anchors were first searched for all intergenic miRNAs. The EST anchors were defined as follows (Fig. 2): (1) the ESTs from the same clone whose genome locations cover the pre-miRNAs (Fig. 2A) and (2) the ESTs whose genome locations cover the pre-miRNAs (Fig. 2B). The pri-miRNAs were predicted by extending these anchors with other matched ESTs from the same library.

The miRNA locations, miRNA and their matched ESTs, and predicted pri-miRNAs are given in Supplementary Tables (http://bioinfo.au.tsinghua.edu.cn/member/~gujin/mirna_est/).
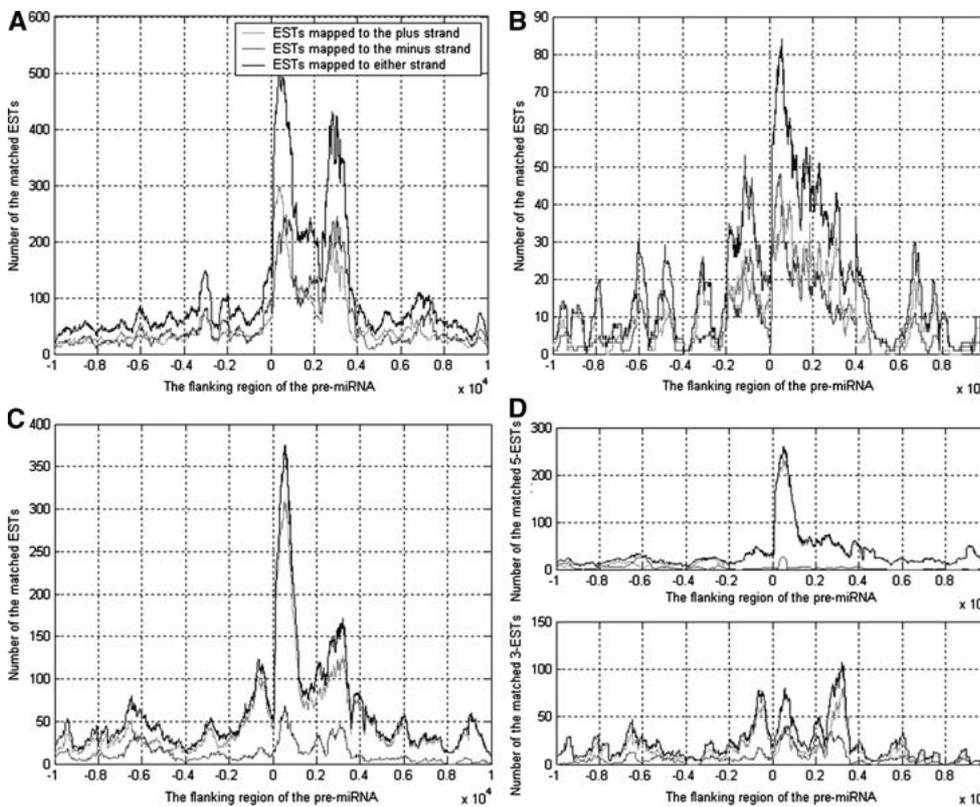
Fig. 1. The number of matched ESTs along the −10,000-nt to + 10,000-nt flanking region of pre-miRNAs. **A**) Human, (**B**) rat, (**C, D**) mouse.

RT-PCR analysis for MiRNA transcripts. RT-PCR experiments were conducted to verify that the matched ESTs in the flanking regions were from the same transcripts with the miRNA. Seven of the mouse miRNAs that were predicted to be expressed in brain, spleen, and kidney were used for the experiments. Sense (forward) and antisense (reverse) primers were designed at the 5′ and 3′ ends determined by the matched ESTs in the −2500-nt to +4000-nt region. Total RNA was isolated from the brain and kidney of the newborn BALB/c mice and from the brain and spleen of adult BALB/c mice with the RNeasy Mini Kit (Qiagen, Valencia, CA), respectively, following the manufacturer's protocol. First-strand cDNA was synthesized from the total RNA with SuperScript™ II RNase H- Reverse Transcriptase (Invitrogen, Carls-

bad, CA) with oligo(dT) as the primer. PCR was performed in accordance with standard procedures with 1 μM of each primer and 2.5 U LA Taq DNA polymerase (Takara, Otsu, Shiga, Japan) to amplify the miRNA transcripts. Control experiments were conducted by no reverse transcription reaction (−RT control) to verify that the amplified products were from the RNA and not from the genomic DNA contamination. No PCR product was observed in the absence of reverse transcriptase. The products were resolved by electrophoresis on 1% w/v agarose gel in TAE buffer (40 mmol/L Tris-acetate, 2 mmol/L Na2EDTA, 2H$_2$O) and stained by GoldView™ (SBS Genetech Beijing, China). Finally, DNA bands were quantified using a UVP-GDS-8000 system UV transilluminator and a LabWorks program (UVP, Upland, CA). The primers used to detect the long pri-miRNA transcripts were as follows: (1) mmu-mir-29b-2-29c, "sense: CTTACTTGGGTTCAGTTGT T," "antisense: ATC TGTCATACTTAGAGGAG C," range: −841 nt to +1246 nt; (2) mmu-mir-331, "sense: TACAGTCTTT GACCCAGTG," "antisense: GCAGATTATTCAG CCCTA," range: −1793 nt to + 2089 nt; (3) mmu-let-7d, "sense: GAAGGGTA CTGCTGGACT," "antisense: CAATGTTCCCA CTTACGA," range: −1438 nt to + 190 nt; (4) mmu-let-7c-2-7b, "sense: CCTCAA GAAGCCACAACA," "antisense: ATTTATACCC AGGTCCCAC," range:
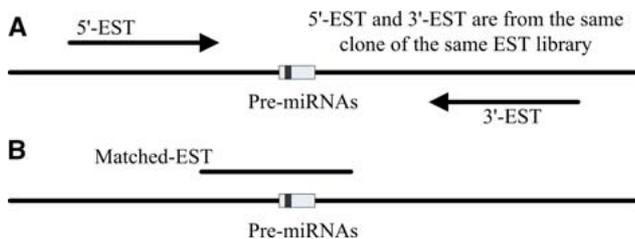
Fig. 2. The schematic for the two types of EST anchors. (**A**) The ESTs from the same clone cover the pre-miRNAs. (**B**) The ESTs cover the pre-miRNAs.

Table 1. Summary of the matched ESTs of the intergenic miRNAs

Organism	Number of intergenic miRNAs	Matched with flanking sequences		Matched with only precursor sequences	
		MiRNAs [a]	ESTs [b]	MiRNAs [c]	ESTs [d]
Human	204	176	3018	53	125
Mouse	162	125	1621	21	48
Rat	135	78	517	7	19
Total	**501**	**379**	**5156**	**81**	**192**

[a]The number of miRNAs whose flanking sequences can be matched by at least one EST sequence.
[b]The number of matched ESTs with the −2500-nt to + 4000-nt flanking sequences.
[c]The number of miRNAs whose precursor sequences can be matched by at least one EST sequence.
[d]The number of matched ESTs with the approximately 100-nt pre-miRNAs.

−758 nt to + 182 nt; (5) mmu-mir-143-145, "sense: CTTGAAATGGGTGGGTCT ATC," "antisense: TGAGTGGGTCAGCAACA GC," range: −1627 nt to +1685 nt; (6) mmu-mir-124-2, "sense: CGCAATA CAAAGAAGGAG," antisense: ACCAGGGTTCA GAAGACA," range: −626 nt to +4133 nt (the end was larger than 4000 nt because there were a few ESTs matched around 4000 nt; the range was extended to the end of these matched ESTs); (7) mmu-mir-181b-1, "sense: ACGCCGTTTCTCGG TTTA," "antisense: TGGGTTCCAGGCAAGTCC," range: +142 nt to +2985 nt.

Results

We focused on the intergenic miRNAs in this study because of the possible different transcription mechanisms between the intragenic and intergenic miRNAs. We first searched for matched ESTs in the −10,000-nt to +10,000-nt flanking sequences of the 501 intergenic pre-miRNAs of the three mammals (human, mouse, and rat). It was observed that significantly more ESTs could be matched with the flanking sequences by using only pre-miRNA sequences, and most of the matched ESTs were located within the region between the −2500-nt to +4000-nt flanking sequences of the miRNAs. Matched 5′-ESTs were enriched within the 0 to +1000-nt flanking region and most of them were matched to the same strand as the pre-miRNAs, which indicated that these ESTs were likely derived from the pri-miRNAs. However, the matched 3′-ESTs were distributed more broadly (Fig. 1). Accordingly, we used the −2500-nt to +4000-nt flanking sequences of the miRNAs to study their transcription with ESTs. Among the 501 studied mammalian intergenic miRNAs, 379 hit a total of 5156 ESTs by their flanking sequences compared with the fact that only 81 miRNA precursor sequences could find hits (with a total of 192 matched ESTs). The matched ESTs with flanking sequences and with only pre-miRNAs are summarized in Table 1. Among the 379 miRNAs

matched by ESTs with their flanking sequences, 197 have 1−5 matched ESTs per miRNA, 114 have 6−20 matched ESTs per miRNA, and 67 have more than 20 matched ESTs per miRNA.

Observing that many ESTs frequently can be matched to the flanking sequences of the intergenic miRNAs allows the possibility that it is a common mechanism of the intergenic miRNAs to be transcribed as the long (>1 kb) pri-miRNAs. RT-PCR experiments were conducted to verify the cotranscription of the intergenic miRNAs and their flank-
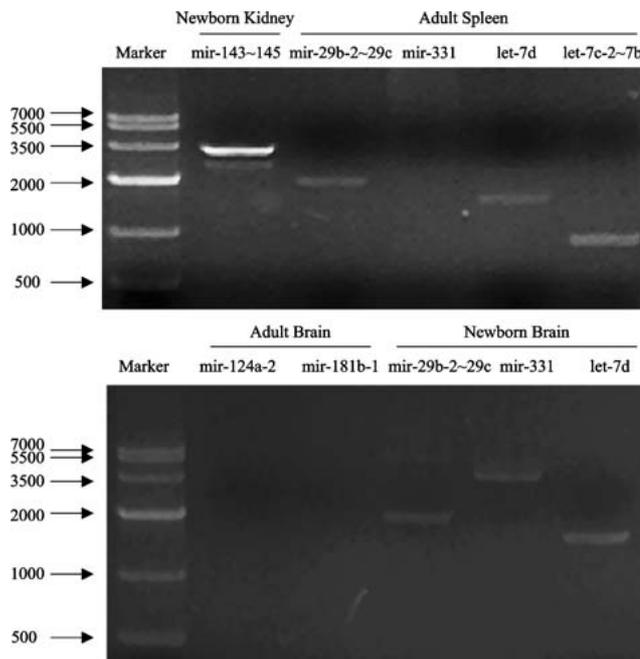
Fig. 3. RT-PCR results for verifying that those ESTs mapped near the miRNAs are from the same pri-miRNA transcription units. Three mouse miRNAs predicted by the matched ESTs to be expressed in newborn brain, one expressed in newborn kidney, two expressed in adult brain, and four expressed in adult spleen are selected for RT-PCR. In natal brain: mir-29b-29c, mir-331, let-7d; in natal kidney: mir-143-145; in adult brain: mir-124a-2, mir-181b-1; in adult spleen: mir-29b-29c, mir-331, let-7d, let-7c-2-7b.

**Table 2.** **The predicted mouse pri-miRNAs by the EST-extension method**

miRNA	Library name	Start	End	Type [a]
mmu-let-7b	Soares_mammary_gland_NbMMG	−800	206	Clone
	Soares mouse 3NbMS	−540	206	Clone
	NIH_BMAP_MHI2_S1	−312	206	Clone
	pBluescript Lion	−197	206	Single 3′
mmu-let-7c-2	Soares_mammary_gland_NbMMG	−84	922	Clone
mmu-let-7d	Barstead mouse myotubes MPLRB5	−512	589	Clone
	RIKEN full-length enriched, 12-day embryo male wolffian	−1003	569	Clone
	NCI_CGAP_PlTr1	−47	592	Single 3′
mmu-let-7f-1	NIH_MGC_166	−341	642	Single 5′
mmu-let-7i	RIKEN full-length enriched, 0-day neonate kidney	−106	646	Clone
	RIKEN full-length enriched mouse cDNA library, Rathke's pouches 14.5-day embryo	−66	326	Single 5′
mmu-mir-29b-2	Soares mouse 3NbMS	−133	602	Anti-sense
	NIH_MGC_166	−405	332	Single
mmu-mir-29c	Soares mouse 3NbMS	−640	95	Anti-sense
	RIKEN full-length enriched, visual cortex	−433	1279	Clone
mmu-mir-130a	RIKEN full-length enriched, 15-day embryo male testis	−466	3042	Clone
mmu-mir-138-2	NIH_BMAP_M_S1	−271	597	Clone
mmu-mir-142	Soares_mammary_gland_NMLMG	−195	405	Clone
	Soares_thymus_2NbMT	−214	215	All 5′
	Hematopoietic stem cells	−200	404	Single
mmu-mir-144	Soares mouse lymph node NbMLN	−340	132	Single 5′
mmu-mir-145	NCI_CGAP_Lu29	−410	939	All 5′
mmu-mir-154	NIA Mouse E10.5 whole embryo cDNA library (Long)	−342	71	Single 5′
mmu-mir-181b-1	RIKEN full-length enriched, 0-day neonate eyeball	−107	3215	Clone
	NIH_MGC_94	−88	724	Single 5′
	NIA Mouse E13.5 whole-embryo cDNA library (Long)	−83	432	Single 5′
mmu-mir-191	RIKEN full-length enriched, 0-day neonate head	−785	851	Clone
mmu-mir-196b	RIKEN full-length enriched, 2-day pregnant adult female	−366	1887	Clone
mmu-mir-223	RIKEN full-length enriched, adult male bone	−2405	820	Clone
	RIKEN full-length enriched, B6-derived CD11 + ve dendritic	−502	824	All 3′
	NIH_MGC_190	−522	86	Single 5′
mmu-mir-299	Kaestner ngn3 wt	178	−1038	Anti-sense
mmu-mir-331	Soares mouse embryo NbME13.5 14.5	−1161	952	Clone
	RIKEN full-length enriched mouse cDNA library, C57BL/6J	−1070	952	Clone
mmu-mir-351	NIH_MGC_169	−236	236	Single 5′
mmu-mir-363	Soares mouse lymph node NbMLN	−60	444	All 5′
mmu-mir-377	RIKEN full-length enriched, adult male pituitary gland	−866	992	Clone
	NCI_CGAP_Mam4	−325	388	All 5′
	Soares NMBP1	−39	604	Single 5′
mmu-mir-425	RIKEN full-length enriched, 0-day neonate head	−1243	393	Clone

[a]The explanations for different Type: Clone: The transcript is extended from both 5′- and 3′-ESTs of the same clone. Single 5′: The transcript is predicted from a single 5′-EST. Single 3′: The transcript is predicted from a single 3′-EST. All 5′: The transcript is extended from multiple 5′-ESTs. All 3′: The transcript is extended from multiple 3′-ESTs. Anti-sense: The predicted transcript is on the minus strand of the miRNA flanking sequence.

ing sequences covered by the matched ESTs. Three mouse miRNAs predicted by the matched ESTs to be expressed in newborn brain, one expressed in newborn kidney, two expressed in adult brain, and four expressed in adult spleen (including the three miRNAs predicted to be expressed in natal brain) were used for the experiments. Five of these seven EST-predicted pri-miRNA transcripts (7 of 10 experiments) were successfully identified by RT-PCR (Fig. 3). That suggests that a large portion of the mammalian intergenic miRNAs were first transcribed as the long pri-miRNAs. The three negative results may have been caused by the different devel-

opmental stages of the adult mice used for producing the matched ESTs and for the RT-PCR experiments or the low-expressed level of the transcripts.

To predict more precisely the length of the pri-miRNAs, the EST anchors were searched for each miRNA locus and extended in each EST library. Seventy pri-miRNAs were successfully predicted. Thirty-three of 70 predicted transcripts were extended from both the 5′- and 3′-ESTs of the same clone and 11 of them were derived from full-length cDNA libraries. The lengths of predicted transcripts ranged from several hundred to several thousand nucleotides (Table 2).
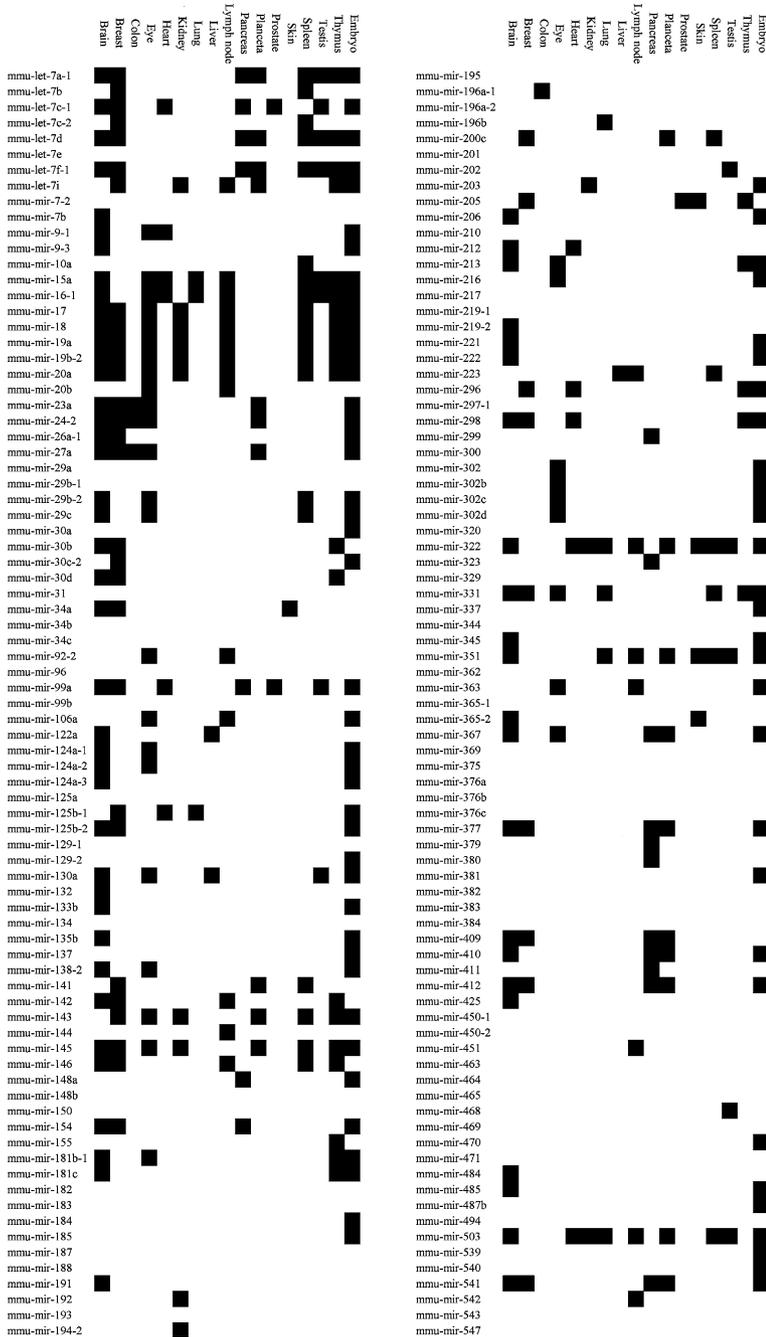
Fig. 4. Expression profiles of the mammalian intergenic miRNAs by the flanking-EST method. The black block indicates that the matched ESTs of the specific miRNA are annotated as expressed in the specific tissue.

The RT-PCR experiments showed that the tissue-specific and developmental-stage–specific information of the matched ESTs was useful in studying the expression of the corresponding miR-NA. An expression map of the mouse intergenic miRNAs was created by examining the tissue-specific information of the matched ESTs (Fig. 4). A miRNA was regarded as being expressed in a specific tissue if any of the matched ESTs of that miRNA were annotated as deriving from that specific tissue. However, Fig. 4 is meant for easy visual display be-cause in the same tissue the developmental stage of the matched ESTs may be different. The detailed miRNAs vs. ESTs list in the Supplementary Tables provides much more information.

With the flanking matched ESTs as evidence of the expression of the pri-miRNAs, the expression profiles of the miRNAs were studied for tens of tissues. Results show that the human brain, lung, liver, and eye and the mouse brain, eye, and breast are tissues in which enriched numbers of miRNAs are expressed. Because the number of ESTs in different
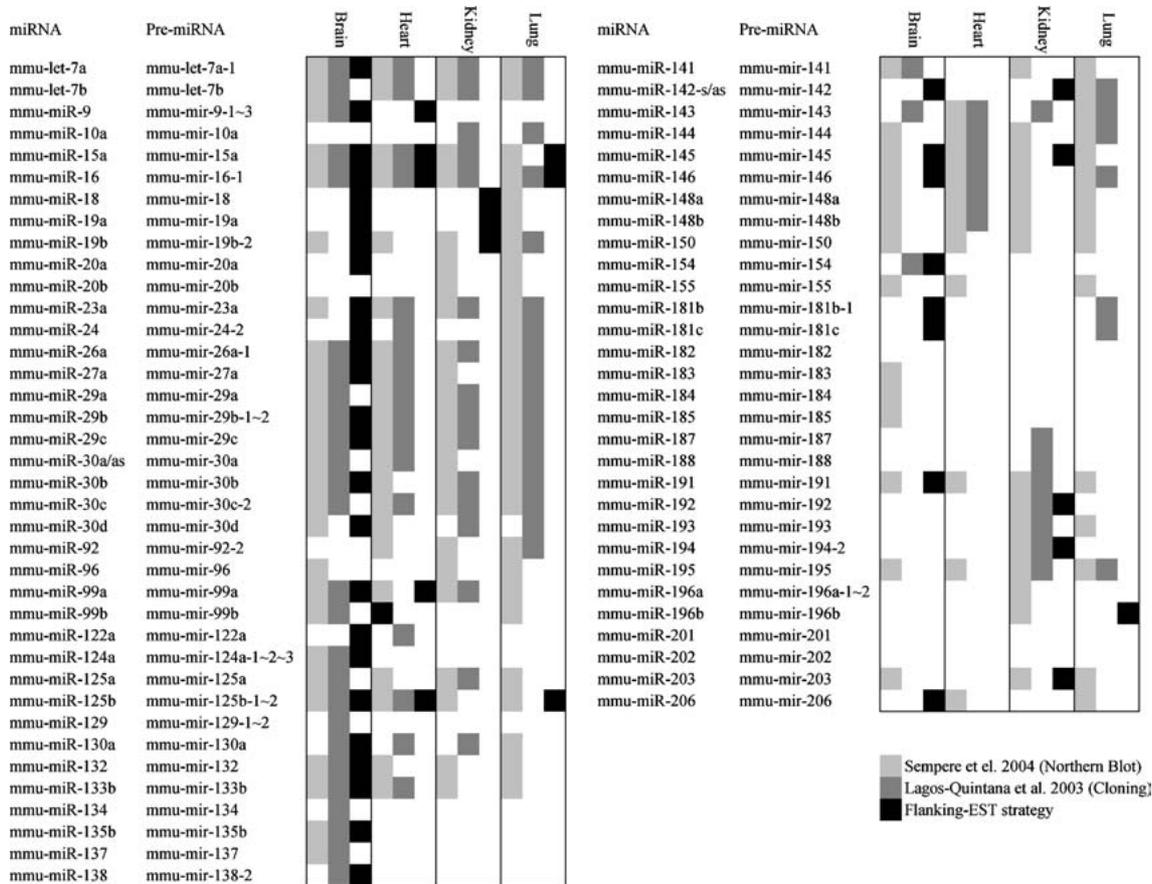
Fig. 5. The comparison of the expression profiles obtained by different methods. The expression profiles in the four mouse tissues, brain, heart, kidney, and lung, are compared. The existence of the miRNAs in the tissues identified by different methods are shown in different colors: light gray for Northern blot (Sempere et al. 2004), dark gray for cloning (Lagos-Quintana et al. 2003), and black for the flanking-EST method.

tissues can be very different in the database, it is necessary to check whether the tissue distribution of the miRNAs in the profiles is correlated with the abundance of ESTs in the database. Comparing the tissue distributions of the miRNAs with that of the ESTs in 15 human tissues, it was found that the miRNA abundance estimated by the flanking-EST strategy did not correlate with the abundance of ESTs. According to the expression profile based on the flanking-EST evidence, human brain, lung, eye, and liver are the more miRNA-diverse tissues. It has already been reported that the brain is as a miRNA-abundant tissue (Sempere et al. 2004). Besides brain, we found several other tissues that are also miRNA-rich, although their total numbers of ESTs in the database are not as large as that in brain.

Lagos-uintana et al. (2003) detected 131 miRNAs in 16 different mouse tissues by cloning. Sempere et al. (2004) characterized 119 mammalian miRNAs by Northern blot. These results were compared with the profiles from the flanking-EST method of 68 mouse intergenic miRNAs (including 75 unique precursors)

in brain, heart, kidney, and lung. These tissues and the miRNAs are examined in both their work and ours (Fig. 5). The comparison indicates that the profiles in the brain tissue agree well, but larger variations exist in the other three tissues. The overall false-positive result was estimated as 30% (15 of 50) if the results by Northern blot are regarded as the standard. The EST-based method can be a complement to the microarray, Northern blot, and cloning methods.

Discussion

In this article we found that a large portion of the mammalian intergenic miRNAs are first transcribed as long pri-miRNAs. The existence of the pri-miR-NAs has been successfully predicted by mapping the ESTs to the long flanking sequences. Although the number of the matched ESTs is much smaller than in the mRNA case, tens of pri-miRNAs are predicted by the EST-extension method. It is noted that 24% of the intergenic miRNAs still cannot find any matched EST in their flanking sequences. This can be

the result of either the low abundance of miRNAs or some different mechanism of the miRNA transcription. Only 81 pre-miRNAs are covered by ESTs, but many more ESTs are matched by the flanking sequences. Two explanations can be given for this phenomenon: (1) the quick processing of pre-miRNAs from primary transcripts. Most inserted cDNAs cover only the downstream of the pre-miRNAs (this also explains the peak right after the loci of the pre-miRNAs for the matched 5′-ESTs); (2) the ESTs are derived from the 5′ and 3′ ends of the inserted clones, while many pre-miRNAs are located in the middle of the long primary transcripts.

Using the flanking-EST method, the expression profiles are quickly and simply established for hundreds of miRNAs in different tissues. Although the current EST data for many tissues are not sufficient to complete the expression analysis in all tissues, the obtained profiles contain intact expression data for tissues that people are more concerned with such as brain, eye, and breast. The profile indicates the presence of the primary transcripts of the miRNAs in that tissue. They should be correlated with the expression of the mature miRNAs in general but apparent differences can exist as a result of regulated processing of the miRNAs and different stabilities of the miRNAs and their primary transcripts. The differences between the pri-miRNA profiles and the pre/mature-miRNA profiles suggest potential processing and regulatory steps.

Acknowledgments

References

1. Adams MD, Kelley JM, Gocayne DJ, Dubnick M, Polymeropoulos MH, et al. (1991) Complementary DNA sequencing: expressed sequence tags and human genome project. Science 252, 1651–1656
2. Ambros V (2004) The functions of animal microRNAs. Nature 431, 350–355
3. Babak T, Zhang W, Morris Q, Blencowe BJ, Hughes TR (2004) Probing microRNAs with microarrays: tissue specificity and functional inference. RNA 10, 1813–1819
4. Barad O, Meiri E, Avniel A, Aharonov R, Barzilai A, et al. (2004) MicroRNA expression detected by oligonucleotide microarrays: system establishment and expression profiling in human tissues. Genome Res 14, 2486–2494
5. Bartel DP (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 116, 281–297
6. Baskerville S, Bartel DP (2005) Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes. RNA 11, 241–247
7. Cai X, Hagedorn CH, Cullen BR (2004) Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. RNA 10, 1957–1966
8. Calin GA, Liu CG, Sevignani C, Ferracin M, Felli N, et al. (2004) MicroRNA profiling reveals distinct signatures in B cell chronic lymphocytic leukemias. Proc Natl Acad Sci USA 101, 11755–11760
9. Chen CZ, Li L, Lodish HF, Bartel DP (2004) MicroRNAs modulate hematopoietic lineage differentiation. Science 303, 83–86
10. Griffiths-Jones S (2004) The microRNA Registry. Nucleic Acids Res 32, D109–111
11. Krichevsky AM, King KS, Donahue CP, Khrapko K, Kosik KS (2003) A microRNA array reveals extensive regulation of microRNAs during brain development. RNA 9, 1274–1281
12. Lagos-Quintana M, Rauhut R, Yalcin A, Meyer J, Lendeckel W, et al. (2002) Identification of tissue-specific microRNAs from mouse. Curr Biol 12, 735–739
13. Lagos-Quintana M, Rauhut R, Meyer J, Borkhardt A, Tuschl T (2003) New microRNAs from mouse and human. RNA 9, 175–179
14. Lee Y, Kim M, Han J, Yeom KH, Lee S, et al. (2004) MicroRNA genes are transcribed by RNA polymerase II. EMBO J 23, 4051–4060
15. Liang RQ, Li W, Li Y, Tan CY, Li JX, et al. (2005) An oligonucleotide microarray for microRNA expression analysis based on labeling RNA with quantum dot and nanogold probe. Nucleic Acids Res 33, e17
16. Liu CG, Calin GA, Meloon B, Gamliel N, Sevignani C, et al. (2004) An oligonucleotide microchip for genome-wide microRNA profiling in human and mouse tissues. Proc Natl Acad Sci USA 101, 9740–9744
17. Miska EA, Alvarez-Saavedra E, Townsend M, Yoshii A, Sestan N, et al. (2004) Microarray analysis of microRNA expression in the developing mammalian brain. Genome Biol 5, R68
18. Reinhart BJ, Slack FJ, Basson M, Pasquinelli AE, Bettinger JC, et al. (2000) The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. Nature 403, 901–906
19. Rodriguez A, Griffiths-Jones S, Ashurst JL, Bradley A (2004) Identification of mammalian microRNA host genes and transcription units. Genome Res 14, 1902–1910
20. Sempere LF, Freemantle S, Pitha-Rowe I, Moss E, Dmitrovsky E, et al. (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

21. Sioud M, Rosok O (2004) Profiling microRNA expression using sensitive cDNA probes and filter arrays. Biotechniques 37, 574–576578–580

22. Smalheiser NR (2003) EST analyses predict the existence of a population of chimeric microRNA precursor-mRNA transcripts expressed in normal human and mouse tissues. Genome Biol 4, 403

23. Sun Y, Koo S, White N, Peralta E, Esau C, et al. (2004) Development of a micro-array to detect human and mouse microRNAs and characterization of expression in human organs. Nucleic Acids Res 32, e188

24. Valoczi A, Hornyik C, Varga N, Burgyan J, Kauppinen S, et al. (2004) Sensitive and specific detection of microRNAs by northern blot analysis using LNA-modified oligonucleotide probes. Nucleic Acids Res 32, e175

25. Zhang Z, Schwartz S, Wagner L, Miller W (2000) A greedy algorithm for aligning DNA sequences. J Comput Biol 7, 203–214