

Biogenesis of intronic miRNAs located in clusters by independent transcription and alternative splicing

PRADEEP RAMALINGAM,¹ JAYANTH KUMAR PALANICHAMY,^{1,2} ANAND SINGH,¹ PRERNA DAS,¹ MOHITA BHAGAT,¹ MUZAFFER AHMAD KASSAB,¹ SUBRATA SINHA,^{1,3} and PARTHAPRASAD CHATTOPADHYAY^{1,4}

¹Department of Biochemistry, All India Institute of Medical Sciences, New Delhi 110029, India

ABSTRACT

miRNAs are generally classified as “intergenic” or “intronic” based upon their genomic location. Intergenic miRNAs are known to be transcribed as independent transcription units, while intronic miRNAs are believed to be processed from the introns of their hosting transcription units and hence share common regulatory mechanisms and expression patterns with its host gene. Recent reports in the literature suggest that some intronic miRNAs, which do not show concordance in expression with their respective host genes, might be transcribed and regulated as independent transcription units. However, there is no direct evidence for the existence of independently transcribed intronic miRNA in humans to date. We have characterized the full-length primary transcripts (pri-miRNAs) of three human intronic miRNAs—miR 106b, miR 93, and miR 24-1—by RNA ligase-mediated RACE and show that human intronic miRNA can indeed be transcribed as independent transcription units. Also, clustered miRNAs are generally believed to arise from a common primary transcript and are expected to have similar expression profiles. However, we have identified several novel alternatively spliced transcripts by RT-PCR, each of which harbors a single pre-miRNA from a cluster of closely located intronic miRNAs. We show that these transcripts represent unique pri-miRNAs for each of these clustered miRNAs. We also report the identification of conserved splice acceptor signals which are responsible for maturation of these novel splice variants. Our results suggest that alternative splicing might play a role in uncoupling the expression of clustered miRNAs from each other, which otherwise are generally believed to be co-transcribed and co-expressed.

Keywords: miRNA biogenesis; intronic miRNA; alternative splicing; clustered miRNA; miRs 106b–93–25; miRs 23b–27b–24-1

INTRODUCTION

MicroRNAs (miRNAs) constitute a family of short (~21 nt) noncoding RNAs which fine-tune gene expression in a multitude of physiological and pathological processes through post-transcriptional and transcriptional gene-silencing mechanisms (Bartel 2004; Kim et al. 2008). Based on their genomic location, miRNAs are generally classified as “intergenic” or “intronic.” Intergenic miRNAs are situated in previously unannotated regions of the human genome and are transcribed from their own unique promoters as long primary transcripts (pri-miRNAs) by RNA POL2 (Lee et al. 2004). For intronic miRNAs, the hnRNA of their respective host genes are considered to be their pri-miRNAs and the subse-

quent steps of biogenesis are the same as for intergenic miRNAs (Lee et al. 2002; Kim and Kim 2007). However, the key difference in the biogenesis of intergenic and intronic miRNAs lies in their mechanisms of transcriptional regulation; while intergenic miRNAs have their own promoters, intronic miRNAs are believed to share promoters with their host genes.

Based upon the canonical pathway of intronic miRNA biogenesis, expression of intronic miRNAs is expected to be in concordance with the expression of their respective host genes. Indeed, initial studies which compared the expression of intronic miRNA–host gene mRNA pairs suggested that this hypothesis holds true for several intronic miRNAs (Rodriguez 2004; Baskerville and Bartel 2005). However, there are several recent reports in which discordance in expression of some intronic miRNAs and their respective

²Present address: Department of Pathology & Laboratory Medicine, David Geffen School of Medicine at UCLA, Los Angeles, CA 90095-1732, USA

³Present address: National Brain Research Center, Manesar, Gurgaon, Haryana 122050, India

⁴Corresponding author

E-mail parthoaiims@hotmail.com

Article published online ahead of print. Article and publication date are at <http://www.rnajournal.org/cgi/doi/10.1261/rna.041814.113>.

© 2013 Ramalingam et al. This article is distributed exclusively by the RNA Society for the first 12 months after the full-issue publication date (see <http://rnajournal.cshlp.org/site/misc/terms.xhtml>). After 12 months, it is available under a Creative Commons License (Attribution-NonCommercial 3.0 Unported), as described at <http://creativecommons.org/licenses/by-nc/3.0/>.

host genes has been observed (Wang et al. 2009a; Radfar et al. 2011). It is likely that these intronic miRNAs might be transcribed from unique transcription units having their own promoters independent of their host genes. Here in our study, we sought to ascertain the possibility of independent transcription of intronic miRNAs. The majority of human miRNAs, both intronic and intergenic, are situated in clusters in the genome. miRNAs located in clusters are generally believed to share a common primary transcript and hence expected to be co-regulated and co-expressed (Altuvia et al. 2005; Kim et al. 2009). Hence, if the expression of the individual members of an intronic miRNA cluster is not concordant with each other, it would again provide indirect evidence for the possibility that such miRNAs can be transcribed and regulated as independent transcription units. We first sought to establish whether such discordance in expression exists by studying certain intronic miRNA–host mRNA pairs by performing qPCR. For this purpose, we chose the condition of hypoxia to study the expression of certain reported hypoxia-regulated miRNAs (HRMs) in relation to the expression of their host genes.

Here, we show that there is discordance in expression of clustered intronic miRNAs miRs 106b–93–25 and miRs 23b–27b–24-1 with respect to each other and with respect to the expression of their host genes, *Mini-Chromosome Maintenance Complex subunit 7 (MCM7)* and *Aminopeptidase-O (AP-O)* respectively, under hypoxic conditions, which suggests the possibility of independent transcription of these intronic miRNAs. We have identified the full-length pri-miRNAs corresponding to three human intronic miRNAs, miR 106b, miR 93, and miR 24-1, by RNA ligase-mediated RACE (RLM-RACE), which confirms independent transcription of these miRNAs within their host genes. We have also identified novel splice variants encoding intronic miRNAs 25 and 93 in which alternative splicing events within the miRNA-harboring intron of their host gene, *MCM7*, result in the production of unique pri-miRNAs for each of these clustered miRNAs. Our results suggest that both independent transcription and alternative splicing might be responsible for uncoupling the expression of clustered miRNAs from each other as well as with their host genes. Based on these results, we propose an alternative model of intronic miRNA biogenesis.

RESULTS

Expression of intronic miRNAs with respect to their host genes

In order to study the concordance in expression of clustered intronic miRNAs and their host genes, we chose the condition of hypoxia to compare the expression of certain reported HRMs in relation to their host genes' expression. Several studies have reported the existence of HRMs, miRNAs that undergo up- or down-regulation under conditions of decreased

oxygen tension (Kulshreshtha et al. 2008). Kulshreshtha et al. (2006) identified 27 HRMs which are up-regulated under hypoxic conditions in MCF7 cell line by microarray analysis. Three of these HRMs, namely miRs 93, 23b, and 24-1, are intronic miRNAs situated in two different genomic miRNA clusters.

miR 93 is a member of the miRNA cluster miR 106b–93–25 located in the thirteenth intron of the gene *MCM7*. miR 93 was reported to be up-regulated after hypoxia induction in MCF7 cells (Kulshreshtha et al. 2006). However, recent reports by Pires et al. (2010) and Hubbi et al. (2011) indicate that *MCM7*, the host gene of miR 93, is down-regulated under hypoxia in several cell lines. In order to investigate this apparent dichotomy, we selected intronic miRNAs 106b–93–25 and their host gene *MCM7* to study their expression by qPCR under hypoxic conditions in MCF7 cells. Similarly, miR 23b and miR 24-1 were reported to be up-regulated under hypoxia (Kulshreshtha et al. 2006). miR 23b and miR 24-1 are members of the intronic miRNA cluster miR 23b–27b–24-1 located in the fourteenth intron of its host gene *AP-O*. However, miR 27b was not reported to be up-regulated under hypoxia unlike miRs 23b and 24-1, which suggests discordance in expression within this miRNA cluster (Kulshreshtha et al. 2006).

As the earlier studies employed microarrays to quantify miRNA expression under hypoxia, we studied the expression of the six selected intronic miRNAs (miRs 106b–93–25 and miRs 23b–27b–24-1) under hypoxic conditions in MCF7 cell line by Real-Time PCR in order to confirm their hypoxia responsiveness. Hypoxia induction was confirmed by studying the expression of established *HIF-1α* responsive genes like *Carbonic Anhydrase-9 (CA9)*, *Vascular Endothelial Growth Factor (VEGF)*, and *Phosphoglycerate Kinase-1 (PGK-1)*, which are known to be up-regulated at the transcriptional level under hypoxia (Supplemental Fig. 1; De Schutter et al. 2005). In the miR 106b–93–25 cluster, miR 106b and miR 25 were observed to be up-regulated, while miR 93 did not show any significant change in expression under hypoxia (Fig. 1A). Their host gene *MCM7* was observed to be down-regulated under hypoxia (Fig. 1A). In the miR 23b–27b–24-1 cluster, all three miRNAs were observed to be up-regulated under hypoxia albeit to different extents (Fig. 1B). However, their host gene *AP-O* did not show significant change in expression under hypoxia (Fig. 1B). Though the change in expression of these intronic miRNAs under hypoxia was not as dramatic as observed with hypoxia responsive genes like *CA9*, the asynchronous expression among the cluster members under hypoxia and discordance in expression with respect to their host genes suggested the possibility that miRNAs belonging to these two clusters might be transcribed from independent transcription units.

Aside from their expression under hypoxia, several lines of evidence in literature also suggest the possibility of independent transcription for these intronic miRNAs. Sikand et al. (2009) have shown that miRNAs belonging to miR 106b–

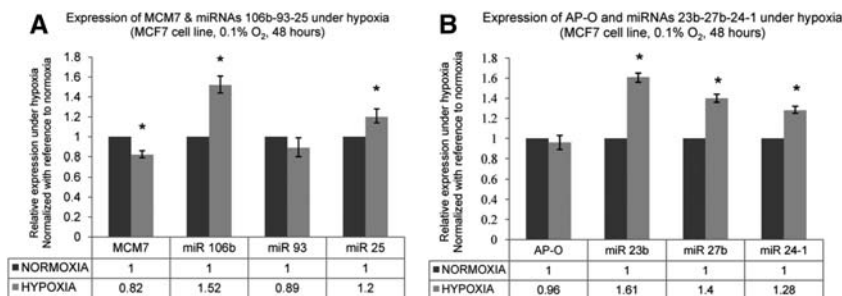


FIGURE 1. Expression of intronic miRNAs and their host genes under hypoxia. (A) Expression of miRNAs 106b, 93, and 25 and the host gene *MCM7* under hypoxia. MCF7 cells were maintained under normoxia (20% O₂) and hypoxia (0.1% O₂) for 48 h. The values represent relative expression under hypoxia as compared with normoxia. *MCM7* expression normalized with reference to the expression of housekeeping genes *18s rRNA* and *PPIA* (Materials and Methods). miRNA expression normalized with reference to expression of *U6 snRNA*. Error bars represent standard error (SEM) of the fold changes. (*) $P < 0.01$. (B) Expression of miRNAs 23b, 27b, and 24-1 and the host gene *AP-O* under hypoxia. The values represent expression under hypoxia as compared with normoxia. (*) $P < 0.01$.

93–25 cluster had a poor correlation with the expression of their host gene *MCM7* in PC-3 cells. Similarly, according to the study by Wang et al. (2009a), the Pearson's correlation coefficients for the co-expression of miRs 23b and 24-1 and their host gene *AP-O* are only 0.24 and 0.09 respectively, suggesting that these miRNA might be independently transcribed. Sempere et al. (2004) have shown that expression of miRNAs in the cluster of 23b–27b–24 was loosely correlated with each other. While miRs 23b and 27b were observed to be abundantly expressed in human brain, heart, and skeletal muscle, the expression of miR 24-1 was not detected in any of these tissues, suggesting discordance in expression among the cluster members. In order to ascertain the possibility of independent transcription, we proceeded to perform 5' and 3' RLM-RACE for these intronic miRNAs.

Identification and confirmation of pri-miRNAs of intronic miRNAs by Drosha knockdown and RLM-RACE

The identification of the full-length sequences of pri-miRNA is technically challenging because of the low abundance of these transcripts under physiological conditions (Lee et al. 2002, 2004; Chien et al. 2011). Consequently, the pri-miRNA sequences of only ~11 human miRNAs (intergenic) have been successfully identified to date out of the ~1500 human miRNAs (Lee et al. 2004; Chien et al. 2011). This is because Drosha cleaves the pre-miRNA soon after its transcription, resulting in extremely short half-lives of pri-miRNAs. One suggested way to circumvent this problem is to increase the steady state levels of pri-miRNAs by prior Drosha knockdown (Lee et al. 2004; Chien et al. 2011). We utilized this strategy before performing 5' and 3' RLM-RACE in MCF7 cells. Drosha knockdown was confirmed by performing qPCR of Drosha mRNA expression after transfection of siRNA targeting Drosha mRNA in MCF7 and HeLa cells (Supplemental Fig. 2A). In order to confirm whether this

decrease in Drosha mRNA levels translated into a decrease in Drosha activity at a functional level, we performed RT-PCR to study the expression of pri-miRNAs, the natural substrates for Drosha.

Generally, PCR quantitation of pri-miRNA is performed according to the method proposed by Schmittgen et al. (2004) in which the forward primer is designed to bind ~100 bp upstream of the precursor miRNA sequence (Supplemental Fig. 2B). However, the 5' end might not extend up to 100 bp for all pri-miRNAs (e.g., the experimentally identified TSS [transcription start site] of pri-miR 23a is located just 15 bp upstream of its pre-miRNA) (Lee et al. 2004). Also, the upstream region corresponding to the primer binding site might be spliced out in

certain pri-miRNAs. In order to overcome these issues, we propose a novel strategy to quantify pri-miRNAs by RT-PCR (Supplemental Fig. 2C). We have designed the primers in such a way that both the forward and reverse primers bind within the stem-loop sequence of the precursor miRNA sequence. Though this primer pair can amplify both the pre-miRNA as well as the pri-miRNA, the amplification of pre-miRNA (which lack poly(A) tails) is prevented by using oligo-dT primed reverse transcription. By employing this strategy, we studied the expression of pri-miRNAs of four arbitrarily selected miRNAs, namely 23b, 24-1, 330, and 26a1, by RT-PCR after Drosha knockdown in MCF7 cells. We observed an accumulation of all the four pri-miRNAs after Drosha knockdown (Supplemental Fig. 2D). This provided confirmation of Drosha knockdown at a functional level.

In order to perform RACE, we designed forward and reverse primers targeting the pre-miRNA sequences of each of the six selected intronic miRNA (Supplemental Fig. 2E, F). To verify the binding specificity and efficiency of these primers, we used another set of primers designed in the upstream exons of the host gene flanking the miRNA-harboring intron. The sequences for these primers (for *MCM7* gene) were obtained from Kim and Kim (2007). For these primers, we have retained the same labeling terminology used by Kim and Kim (2007) in order to maintain coherence and ease of comparison. We successfully identified the full-length pri-miRNA sequences of three intronic miRNAs, namely miRs 106b, 93, and 24-1 (Fig. 2A,B). miRs 106b and 93 share a common primary transcript, referred henceforth as PRI 106b/93, whose TSS lies 79 bp upstream of miR 106b (Fig. 2A; Supplemental Fig. 2G). The 3' end of this pri-miRNA extends 336 bp downstream from miR 106b. The full-length sequence of PRI 106b/93 is 437 nt in length and is hosted entirely within intron 13 of *MCM7*. PRI 106b/93 is capped at its 5' end (as it was identified using RLM-RACE, which detects only capped transcripts) and has a poly(A) tail at its 3'

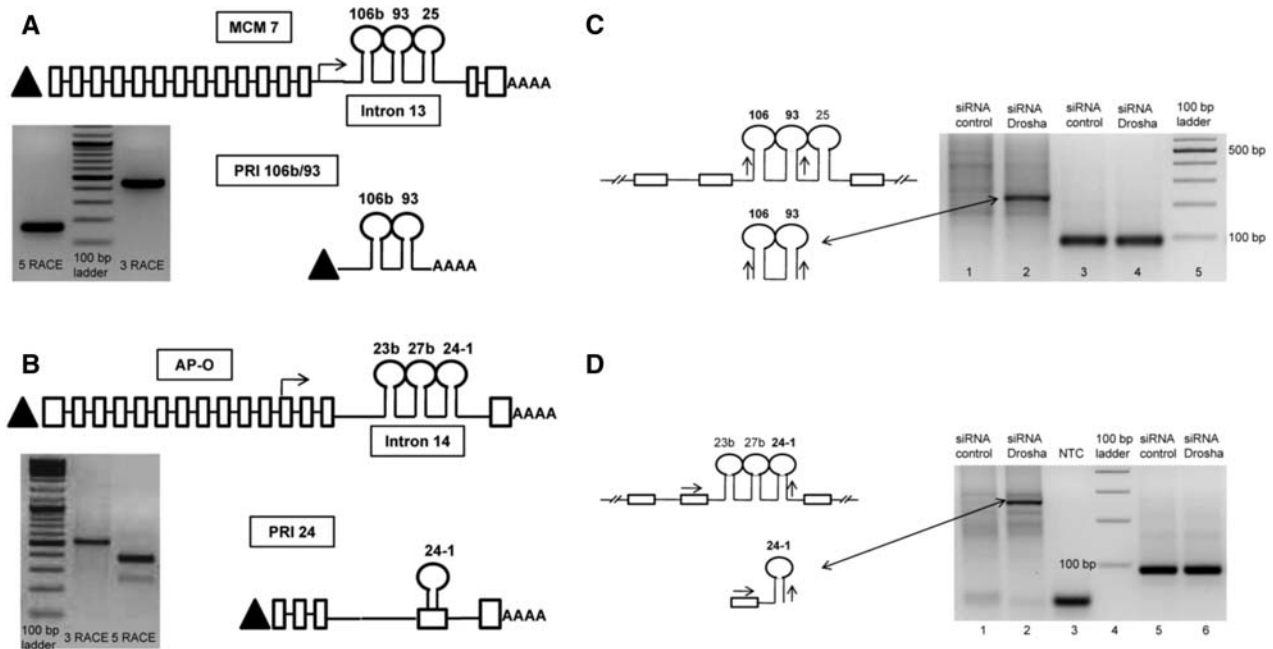


FIGURE 2. Identification of pri-miRNAs encoding intronic miRNAs. (A) Schematic representation of PRI 106b/93 in relation to the host gene *MCM7*. The gel image shows the 5' RACE and 3' RACE products of miRNA 106b which were sequenced (see Supplemental Fig. 2E,F for primer RACE primer design). The upper panel represents the exon–intron structure of the host gene *MCM7*. Boxes represent exons, while the interconnecting lines represent introns. Triangle (▲) represents 5' 7-methyl guanosine cap, while “AAAA” represents the 3' poly(A) tail. The stem–loop structures represent the pre-miRNAs. The arrow represents the TSS of the pri-miRNA with respect to the host gene mRNA. The lower panel represents the gene structure of PRI 106b/93. The sequence of PRI 106b/93 is shown in Supplemental Figure 2G. (B) Schematic representation of PRI 24 in relation to the host gene *AP-O*. The gel image shows the RACE products of miRNA 24-1. The upper panel represents the exon–intron structure of the host gene *AP-O*. The lower panel represents the gene structure of pri-miRNA of miR 24-1. The precursor sequence of miR 24-1, which is part of intron 14 of the host gene *AP-O*, becomes part of exon 4 of PRI 24. The sequence of PRI 24 is shown in Supplemental Figure 2H. (C) Confirmation of the novel transcript PRI 106b/93 as pri-miRNA. (Lanes 1,2) RT-PCRs performed with 3-106b OUT FWD as forward primer and 5-93 OUT REV as reverse primer (schematic of primer design shown on left upper panel). Lane 2 in the figure shows accumulation of PRI 106b/93 transcript upon Drosha knockdown in MCF7 cells (schematic on left lower panel). (Lanes 3,4) Expression of housekeeping gene *PPIA* used as loading control. (D) Confirmation of the novel transcript PRI 24 as pri-miRNA. (Lanes 1,2) RT-PCRs performed with *AP-O* EXON FWD and 5-24 OUT REV primer (schematic on left upper panel). (Lane 3) No template control (labeled as NTC) for this PCR. The figure shows accumulation of the PRI 24 transcript upon Drosha knockdown in MCF7 cells (schematic on left lower panel). (Lanes 5,6) Expression of housekeeping gene *PPIA* used as loading control.

end, confirming that it is transcribed by RNA POL2. The TSS of the host gene *MCM7* lies 7867 bp upstream of miR 106b. Hence, PRI 106b/93 serves as an example of independently transcribed pri-miRNA of intronic miRNAs 106b and 93. Interestingly, both miRs 106b and 93 share identical seed sequences and are reported to have common target genes like p21 (Kim et al. 2009).

The pri-miRNA corresponding to miR 24-1, referred to as PRI 24, is 890 nt in length and is composed of five exons and four introns (Fig. 2B; Supplemental Fig. 2H). The TSS of PRI 24 resides ~25 kb upstream of genomic locus of miR 24-1, whereas the TSS of host gene *AP-O* resides ~356 kb upstream of miR 24-1. Hence, the protein-coding *AP-O* mRNA and the miRNA-encoding PRI 24 arise from two different transcription units independent of each other. On comparing the exon composition of *AP-O* mRNA with PRI 24 pri-miRNA, we observe that four out of the five exons of PRI 24 (i.e., exons 1, 2, 3, and 5) are similar to exons 12–15 of *AP-O* (Fig. 2B). However, exon 4 of PRI 24 is unique and is derived from alternative splicing within intron 14 of *AP-O* in such a manner

that the entire precursor sequence of miR 24-1 is included in its sequence. It is interesting to note that the precursor sequences of miRNAs 23b and 27b (which lie 716 bp and 479 bp upstream of pre-miR 24-1) are excluded from the sequence of PRI 24 as a result of this splicing event. Hence, PRI 24 serves as an example of a pri-miRNA produced by independent transcription within its host gene and alternative splicing within the miRNA-harboring intron. PRI 24 was also observed to be expressed in HeLa cells, suggesting that this pri-miRNA does not represent a tissue-specific splicing event (Supplemental Fig. 2I). Importantly, the exon–intron structure of PRI 24 suggests that 5' ends of pri-miRNAs identified by primer extension techniques might result in ambiguous TSS predictions as such techniques assume the lack of introns between the primer binding site and the transcription termination site.

In order to ascertain whether any splice variant of *AP-O* harbored the precursor sequence of miR 24, we performed RT-PCR with a forward primer designed to bind within exon 9 of *AP-O* and reverse primer designed to bind within

the precursor sequence of miR 24-1 (Supplemental Fig. 2J). Exon 9 is a conserved exon present in all the protein-coding splice variant mRNAs of *AP-O*. However, we observed no amplicons in the RT-PCR performed with these primers. This confirms PRI 24 as a unique pri-miRNA for miR 24.

In order to confirm whether a novel transcript could behave as a pri-miRNA, it must satisfy the following criteria: (a) The transcript should harbor the entire precursor sequence (pre-miRNA) of the corresponding mature miRNA; (b) it should preferably not have protein-coding potential (though this criteria is not absolutely mandatory because of increasing evidence of bifunctional RNAs) (Ulveling et al. 2011); (c) Drosha knockdown should result in the accumulation of these transcripts, as pri-miRNAs are natural substrates for Drosha cleavage (Gregory et al. 2004). The full-length sequences of PRI 106b/93 and PRI 24 did not have a continuous ORF in any of their three reading frames as predicted by the Transeq algorithm (www.ebi.ac.uk/Tools/st/emboss_transeq/) (Supplemental Fig. 2G,H). This precludes protein-coding potential and suggests that these transcripts are noncoding RNAs. Both PRI 106b/93 and PRI 24 were observed to be accumulated upon Drosha knockdown, suggesting that these transcripts can be considered as genuine pri-miRNAs of miRs 93, 106b, and 24-1 (Fig. 2C,D). These pri-miRNAs represent the first instance of human intronic miRNAs being transcribed from unique transcription units independent of their host genes.

Identification of novel splice variants encoding intronic miRNAs

Since the pri-miRNAs corresponding to intronic miRNAs 25, 23b, and 27b could not be characterized by RACE, we proceeded to study, by RT-PCR, whether their pri-miRNAs were also involved in alternative splicing events akin to PRI 24. For instance, in the case of miR 25, the forward primer was designed to bind in the upstream exon (i.e., exon 13 of *MCM7*) flanking the miRNA-harboring intron (labeled as F1 in Supplemental Fig. 2E) while the reverse primer would bind within the precursor sequence of the miRNA of interest (labeled as R1 and R2 in Supplemental Fig. 2E). If alternative splicing occurred, one would expect to see at least two amplicons in the RT-PCR: a larger amplicon corresponding to the host gene *hnRNA* and a smaller amplicon corresponding to the alternatively spliced pri-miRNA. With this strategy, we checked for the presence of splice variants corresponding to the selected intronic miRNAs. The larger amplicons of expected sizes were observed when PCR was performed either with MCF7 genomic DNA (Supplemental Fig. 3A) or oligo-dT primed MCF7 cDNA (Fig. 3A). However, with cDNA as the template, we also observed novel amplicons of smaller size (~100 bp) both in the case of miR 93 and miR 25 (Fig. 3A). Sequencing confirmed that these were indeed specific amplicons and represented novel alternatively spliced transcript variants (Supplemental Fig. 3B,C) and are

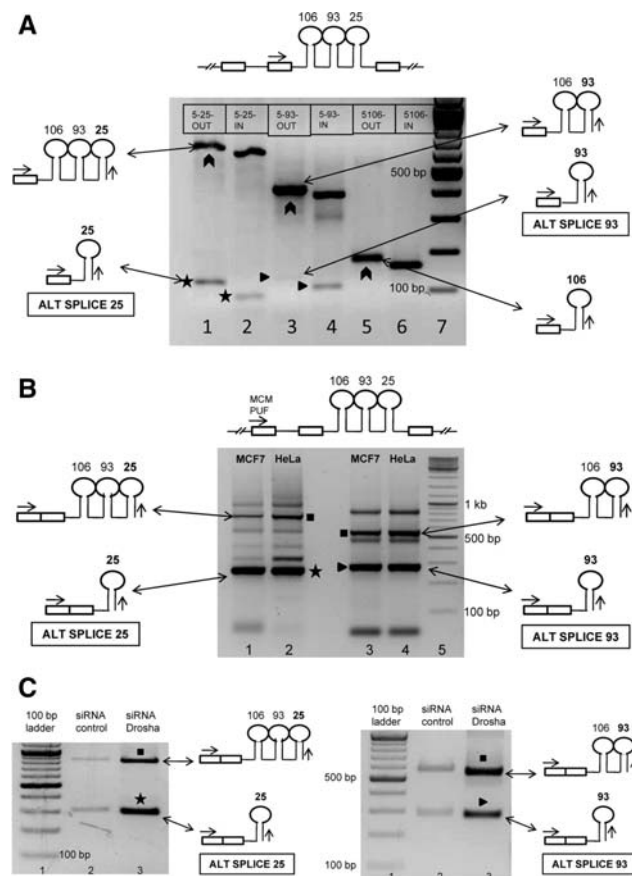


FIGURE 3. Identification of novel alternatively spliced pri-miRNAs encoding intronic miRNAs. (A) Identification of novel alternatively spliced transcripts corresponding to intronic miRNAs 106b–93–25. RT-PCR performed with oligo dT primed MCF7 cDNA as template. The common forward primer (MCM EXON FWD) is represented by the arrow in the schematic shown above the gel image. Lane 1 represents the PCR performed with 5-25-OUT as reverse primer, while lane 2 represents the PCR performed with 5-25-IN as reverse primer (refer to Supplemental Fig. 2E for primer design). Lanes 3 and 4 correspond to PCRs performed with 5' RACE OUT and IN primers of miR 93, while lanes 5 and 6 correspond to miR 106b 5' RACE OUT and IN primers. Boxes within the gel image represent the reverse primers used in the PCR. The chevrons represent the amplicons corresponding to the host gene *hnRNA*. The stars (*) represent the alternatively spliced transcript ALT SPLICE 25, while triangles (▲) represent ALT SPLICE 93. (B) Confirmation of inclusion of upstream *MCM7* exon in the novel alternatively spliced transcripts. Forward primer (MCM PUF) is common to all the PCRs. (Lanes 1,2) PCRs performed with 5-25 OUT as reverse primer using oligo dT primed cDNA of MCF7 and HeLa. (Lanes 3,4) PCRs performed with 5-93 OUT as reverse primer. Star (*) represents ALT SPLICE 25. The schematic representation shows inclusion of the upstream exon of its host gene. The triangle (▲) represents ALT SPLICE 93. The squares (■) represent amplicons corresponding to the partially spliced host gene mRNA. The other bands could not be sequenced and probably represent nonspecific amplicons and primer dimers. The sequences of the ALT SPLICE 25 and ALT SPLICE 93 are shown in Supplemental Figure 3B,C. (C) Confirmation of novel transcripts ALT SPLICE 25 and ALT SPLICE 93 as pri-miRNAs. RT-PCR was performed as shown in B. The agarose gel electrophoresis images show accumulation of the novel transcripts ALT SPLICE 25 (marked by a star, *) and ALT SPLICE 93 (marked by a triangle, ▲) upon Drosha knockdown in MCF7 cells. The figure also shows accumulation of partially spliced host gene mRNA (marked by squares, ■) upon Drosha knockdown. *PPIA* was used as loading control (shown in Fig. 2C).

henceforth referred to as ALT SPLICE 25 and ALT SPLICE 93 in this article. No such alternatively spliced transcripts corresponding to miRs 106b, 23b, and 27b were observed.

To ascertain whether these novel transcripts included the second 5' exon (i.e., exon 12 of *MCM7*), we performed RT-PCRs with forward primer designed to bind to exon 12 of *MCM7* (labeled as F2 in Supplemental Fig. 2E) and reverse primers designed to bind within the precursor sequences of the miRNAs of interest in both MCF7 and HeLa cells (Fig. 3B). Sequencing of the PCR products confirmed that both ALT SPLICE 25 and ALT SPLICE 93 included the upstream exon of *MCM7* but not the intervening intron. Next, we tried to ascertain whether these novel transcripts represent splice variants of their host gene *MCM7* or whether they represent independent transcription units. We performed RT-PCRs with forward primer designed to bind within exon 7 of *MCM7* and reverse primers designed to bind within the precursor sequences of the miRNAs (Supplemental Fig. 2K). Exon 7 is a conserved exon common to all the protein-coding splice variant mRNAs of *MCM7*. However, we observed no amplicons in the RT-PCR performed with these primers, suggesting that ALT SPLICE 25 and ALT SPLICE 93 do not represent splice variants of *MCM7* and probably arise from independent transcription units similar to PRI 24. Furthermore, both these novel splice variants were observed to be expressed in both MCF7 and HeLa cells, suggesting that they do not represent tissue-specific events. We also observed amplicons corresponding to the partially spliced host gene mRNA as reported by Kim and Kim (2007) (Fig. 3B). These partially spliced host gene mRNAs are reported to be observed in genes which harbor miRNAs in their introns that cause a delay in splicing of these introns as compared with the upstream or downstream introns.

ALT SPLICE 25 has multiple stop codons in all three reading frames, which precludes protein-coding potential (Supplemental Fig. 3B). Though ALT SPLICE 93 had a continuous ORF in one of its three possible reading frames, upon performing a BLAST search (blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins), we did not observe any proteins in the public databases which matched the translated sequence of ALT SPLICE 93 (Supplemental Fig. 3C). There was an accumulation of both ALT SPLICE 25 and ALT SPLICE 93 upon Droscha knockdown in MCF7 cells (Fig. 3C). Therefore, these transcripts can be considered as pri-miRNAs of miRs 25 and 93. Mature miR 93 is hence produced from at least two distinct pri-miRNAs, one from ALT SPLICE 93 and the other from PRI 106b/93, suggesting that miRNAs of intronic origin might mature from multiple pri-miRNA sequences. We also observed an accumulation of the partially spliced host gene mRNA as well, consistent with the report by Kim and Kim (2007) (Fig. 3C). However, we did not observe any significant change in expression of the host gene *MCM7* upon Droscha knockdown (Supplemental Fig. 3D).

In order to rule out the possibility that the novel transcripts identified in our study were not artifacts of PCR amplification,

we designed exon-junction spanning primers corresponding to each of the identified alternatively spliced pri-miRNAs namely ALT SPLICE 25, ALT SPLICE 93, and PRI 24. The forward primers were designed to bind to the regions corresponding to the post-splicing junctions of miR containing exons and upstream exons (Supplemental Fig. 3E). Amplicons of expected sizes, corresponding to the respective pri-miRNAs, were observed in the RT-PCRs performed with these exon-junction spanning primers (Supplemental Fig. 3E). DNA sequencing confirmed that these amplicons were specific to the pri-miRNAs of interest. We also observed an accumulation of the alternatively spliced pri-miRNAs upon Droscha knockdown, when RT-PCRs were performed with exon-junction spanning primers (Supplemental Fig. 3F).

The production of pri-miRNAs PRI 24, ALT SPLICE 25, and ALT SPLICE 93 involves alternative splicing events which were strikingly similar to each other. In all three transcripts, the splice acceptor sites were positioned in such a way that the corresponding pre-miRNA sequences were included while simultaneously excluding the precursor sequences of the other upstream miRNAs belonging to the cluster. We then proceeded to study in silico whether the novel splice sites identified in these pri-miRNAs were recognized by conventional splice-site prediction algorithms by using SplicePort.

SplicePort analysis of pri-miRNAs

SplicePort is a web-based tool which predicts splice sites within a given sequence (<http://spliceport.cbcb.umd.edu/>). We subjected the entire miRNA-harboring intron sequences of *MCM7* and *AP-O* to SplicePort analysis (Materials and Methods). The novel splice acceptor signals observed in the sequences of PRI 24, ALT SPLICE 25, and ALT SPLICE 93 were accurately predicted by SplicePort, suggesting that they were probable genuine splice sites. Interestingly, SplicePort predicted splice acceptor signals upstream of each of the six miRNAs in our study (Table 1). In the case of miRNA cluster 106b–93–25, one splice acceptor site was predicted upstream of each of the three pre-miRNAs (1, 50, and 20 bp upstream of pre-miRs 106b, 93, and 25, respectively) (Supplemental Fig. 4A). In the case of miRNA cluster 23b–27b–24-1, two splice acceptor sites were predicted upstream of each of the three pre-miRNAs (200 and 75 bp upstream of pre-miR 23b; 129 and 105 bp upstream of pre-miR 27b; 91 and 86 bp upstream of pre-miR 24-1) (Supplemental Fig. 4C). The positioning of splice acceptor sites upstream of the pre-miRNAs suggests that they might play a role in uncoupling the expression of clustered miRNAs from each other by selective inclusion of a particular precursor miRNA in the primary transcript (pri-miRNA) as observed in the case of PRI 24, ALT SPLICE 25, and ALT SPLICE 93.

It is known that introns flanking alternatively spliced exons have a higher degree of conservation between humans and mouse as compared with other introns possibly because of the presence of *cis* elements which regulate alternative splicing

TABLE 1. Splice acceptor signals upstream of selected human and mouse intronic miRNAs predicted by SplicePort

S. no.	miRNA	Host gene	Human		Mouse	
			Splice acceptor signal	Distance upstream	Splice acceptor signal	Distance upstream
1	miR 25	<i>MCM7</i>	ctcacaggacag	20 bp ^a	ctcacaggacag	24 bp ^a
2	miR 93	<i>MCM7</i>	GTCCTAGATCCT	50 bp	ACCTTAGTCATG	0 bp
3	miR 106b	<i>MCM7</i>	GCTCCAGCCCTG	1 bp ^a	CTACCAGCCCTG	0 bp ^a
4	miR 23b	<i>AP-O</i>	TCTTCAGTTCCA	200 bp	CTCCCAGTCCCC	146 bp
5	miR 23b	<i>AP-O</i>	ATAACAGGCGTG	75 bp	TATACAGTGCAG	124 bp
6	miR 27b	<i>AP-O</i>	CCTCCAGAAACC	129 bp ^a	CCTGCAGAACCG	132 bp ^a
7	miR 27b	<i>AP-O</i>	ACTGCAGATTGG	105 bp ^a	TCTGCAGTTTGG	107 bp ^a
8	miR 24-1	<i>AP-O</i>	tttgagtcacag	91 bp ^a	tttgagtcacag	86 bp ^a
9	miR 24-1	<i>AP-O</i>	AGTCCAGGCCTT	87 bp ^a	AGTCCAGGTCTC	81 bp ^a
10	miR 330	<i>EML2</i>	ctccaggatcg	32 bp ^a	ctccaggatcg	29 bp ^a
11	miR 26b	<i>CTDSP1</i>	ACCCTAGGCTCT	187 bp ^a	CACCCAGGTTCC	186 bp ^a
12	miR 126	<i>EGFL7</i>	CATCCAGCGCAG	28 bp ^a	CTGGCAGCCTGG	25 bp ^a

The splice acceptor signal sequences depicted in small letters represent those signals whose sequence is entirely conserved between humans and mouse. "Distance upstream" represents the distance between the nucleotides "AG" in the signal sequence and the precursor miRNA sequence. The usage of predicted signal sequences corresponding to human miRs 25, 93, and 24-1 has been validated in our study. EST-based evidence is present for the usage of splice acceptor signals corresponding to human miR 126, mouse miR 27b, and mouse miR 24-1 (Fig. 4).

^aSignals whose "distance upstream" is <10 bp apart between humans and mouse.

(Sorek and Ast 2003). In order to ascertain whether the splice acceptor signals predicted upstream of the six human pre-miRNAs in our study are conserved in the mouse, we subjected the miRNA-harboring introns of mouse *MCM7* and *AP-O* to SplicePort analysis (Supplemental Fig. 4B,D). Interestingly, SplicePort predicted splice acceptor signals upstream of each of the six mouse miRNAs strikingly similar in location to those observed in humans (Table 1). Analogously, splice acceptor signals which are conserved between humans and mouse were observed upstream of three other intronic miRNAs, namely miRs 126, 330, and 26b (Table 1; Supplemental Fig. 5A–C). This suggests that, along with pre-miRNA sequences, their upstream splice acceptor signal sequences might be under evolutionary pressure for conservation.

For the splice sites predicted by SplicePort, we tried to obtain biological evidence for their actual usage by searching for spliced ESTs encoding the precursor sequences of these intronic miRNAs in the human and mouse genome using the UCSC genome browser (<http://genome.ucsc.edu/cgi-bin/hgGateway>). We identified three spliced ESTs, corresponding to human miR 126 (GenBank: BX380440.2), mouse miR 27b (GenBank: BF301119.1), and mouse miR 24-1 (GenBank: AI286629.1) in which splicing occurred at the exact same position as predicted by SplicePort (Supplemental Fig. 6A–C). These ESTs do not have protein-coding potential, suggesting that they are

noncoding RNAs. These spliced ESTs probably represent the pri-miRNAs corresponding to human miR 126, mouse miR 27b, and mouse miR 24-1. The EST BX380440.2 (encoding miR 126) shares TSS with the corresponding host gene *EGFL7* (Fig. 4A). Seven out of eight exons in this EST are common to the host gene *EGFL7* while the eighth exon is unique and produced as a result of alternative splicing within the miR 126-harboring intron of *EGFL7*. The ESTs corresponding to mouse miRs 27b and 24-1 have an exon–intron structure similar to that observed in PRI 24 (Fig. 4B). Notably, the

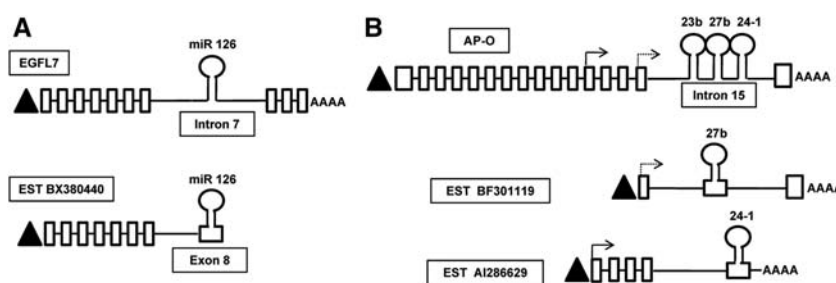


FIGURE 4. Schematic representation of identified ESTs in relation to host genes encoding intronic miRNA. (A) Schematic representation of human EST BX380440 in relation to the host gene *EGFL7*. The upper panel represents the exon–intron structure of the host gene *EGFL7* (Ensembl transcript ID ENST00000308874) which codes for a 273-aa protein. miR 126 is hosted within intron 7 of *EGFL7*. The lower panel represents the gene structure of EST BX380440. Note that the EST shares the TSS and first seven exons with *EGFL7*. However, the precursor sequence of miR 126, which is part of intron 7 of the host gene *EGFL7*, becomes part of the unique exon 8 of BX380440 as a result of alternative splicing. EST BX380440 is a noncoding RNA and probably represents the pri-miRNA of miR 126. (B) Schematic representation of mouse ESTs in relation to the host gene *AP-O*. The upper panel represents the exon–intron structure of the host gene *AP-O* (Ensembl transcript ID ENSMUST00000021911). The arrows represent the location of TSS of the ESTs with respect to the host gene *AP-O*. The lower panels represent the gene structures of mouse ESTs BF301119 and AI286629 corresponding to miRs 27b and 24-1, respectively.

conserved splice acceptor signal upstream of miR 24-1 has biological evidence of actual usage both in humans (PRI 24) and mouse (EST AI286629.1). We propose that a similar strategy of searching for spliced ESTs corresponding to the conserved splice acceptor signals upstream of intronic miRNAs might aid in the discovery of novel pri-miRNAs in the future. Also, the design of exon-junction spanning primers corresponding to predicted splice acceptor signals upstream of miRNAs might be a useful technique to screen for novel pri-miRNAs.

Validation of the novel alternatively spliced transcripts as pri-miRNAs

In order to further validate that the novel alternatively spliced transcripts (ALT SPLICE 25, ALT SPLICE 93, and PRI 24) are genuine pri-miRNAs, we constructed mini-genes in which the sequences corresponding to each of these spliced transcripts were PCR-amplified using exon-junction spanning primers and cloned under CMV promoter of the pTarget Mammalian expression vector (Supplemental Fig. 7A,B). Overexpression of the vector constructs was confirmed by RT-PCR (Supplemental Fig. 7A,B). Overexpression of the alternatively spliced transcripts resulted in increased expression of the corresponding mature miRNAs (Fig. 5A). This conclusively proves that the alternatively spliced transcripts identified in this study are pri-miRNAs.

To ascertain whether splicing influences the production of mature miRNAs, we performed siRNA-mediated knockdown of the key spliceosomal components *PRPF8* and *SNRNP70* (Supplemental Fig. 7C). Knockdown of spliceosome resulted in a significant decrease in expression of miR 25, miR 93, and miR 24-1, suggesting that splicing promotes the maturation of these miRNAs (Fig. 5B).

In order to assess the significance of the splice acceptor signals in generation of alternatively spliced pri-miRNAs as well as mature miRNAs, we constructed artificial mini-genes with normal and mutant splice acceptor sites. Intron 13 of the host gene *MCM7* (harboring the miRNA cluster miR 106b–93–25) along with part of the flanking exons (exons 13 and 14) was PCR-amplified from HeLa genomic DNA and cloned under CMV promoter (referred to as MCM Native vector) (Supplemental Fig. 7D). We constructed another vector, using the MCM Native vector as backbone, in which either the splice acceptor signal situated upstream of ALT SPLICE 25 was mutated (referred to as MCM Mutant vector) (Supplemental Fig. 7D). Transfection of the native MCM vector resulted in increased expression of the spliced *MCM7* mRNA (exon 13–14), as expected (Supplemental Fig. 7D). Interestingly, the expression of ALT SPLICE 25 also increased upon transfection of the native *MCM7* vector. This indicates that both splice acceptor signals (upstream of ALT SPLICE 25 as well as upstream of exon 14 of *MCM7*) were recognized by the spliceosomal complex. Mutation of the splice acceptor signal upstream of miR 25 resulted in decrease in expression

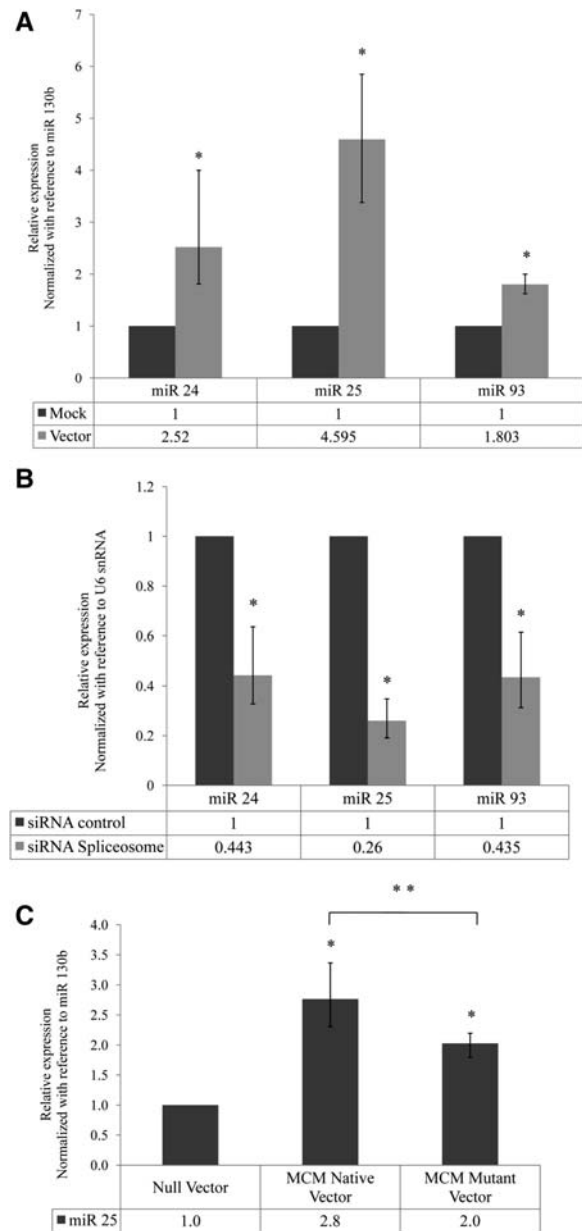


FIGURE 5. Validation of the alternatively spliced transcripts as pri-miRNAs. (A) Expression of miRNAs 24, 25, and 93 from alternatively spliced pri-miRNAs cloned under CMV promoter. The alternatively spliced pri-miRNAs were cloned into mammalian expression vectors, as described in Supplemental Figure 7A. The graph depicts the relative expression of mature miRs 24, 25, and 93 upon transfection of their respective vectors in HeLa cells as compared with mock transfected cells by Real-Time PCR. miR 130b was used as a reference gene for normalization. (*) $P < 0.05$. (B) Expression of miRNAs 24, 25, and 93 after spliceosome knockdown. The graph depicts the relative expression of mature miRs 24, 25, and 93 upon spliceosome knockdown as compared with control siRNA transfected cells by Real-Time PCR. *U6 snRNA* was used as a reference gene for normalization. (*) $P < 0.01$. (C) Effect of splice acceptor site mutation on expression of mature miR 25. The graph depicts the relative expression of mature miR 25 in MCM Native vector and MCM Mutant vector (Supplemental Fig. 7D) transfected HeLa cells as compared with Null vector transfected cells by Real-Time PCR. miR 130b was used as a reference gene for normalization. (*) $P < 0.01$; (**) $P < 0.05$.

of the ALT SPLICE 25 to its basal level (Supplemental Fig. 7D). This conclusively proves that the splice acceptor signal identified upstream of miR 25 is a genuine splice site and is necessary for the production of ALT SPLICE 25.

Transfection of the native MCM vector resulted in increased expression of mature miR 25 (Fig. 5C). The mutation of the splice acceptor signal (MCM Mutant vector) resulted in decreased expression of mature miR 25 as compared with the native MCM vector (Fig. 5C). This is possibly due to the decreased production of pri-miRNA ALT SPLICE 25 as a result of the mutation. However, the expression of mature miR 25 upon transfection of splice acceptor site mutated vector was higher as compared with its expression in null vector transfected cells, suggesting that mature miR 25 can be processed even after blocking the generation of ALT SPLICE 25, presumably from the spliced out intron from the vector (i.e., by the canonical pathway of intronic miRNA biogenesis). Altogether, these results suggest that the splice acceptor signal, upstream of miR 25, plays a significant role in production of the pri-miRNA ALT SPLICE 25 as well as mature miR 25.

DISCUSSION

Though the biological functions of individual miRNAs are being elucidated at a rapid pace, the transcriptional and post-transcriptional regulatory mechanisms governing the expression of these regulatory RNAs are only slowly being unraveled. The identification of full-length pri-miRNA sequences represents the first step toward understanding the transcriptional control of miRNAs. Though pri-miRNA sequences of few intergenic miRNA have been identified in humans, no such information is available for intronic miRNAs (Eis et al. 2005; Saito et al. 2006; Taganov et al. 2006; Woods et al. 2006; Chang et al. 2007; Fukao et al. 2007). The canonical model of intronic miRNA biogenesis predicts that the host gene *hnRNAs* serve as pri-miRNAs for intronic miRNAs and the host gene-intronic miRNA pair is assumed to be under the influence of common transcriptional regulatory mechanisms. As a consequence, pri-miRNA identification studies were confined only to the subgroup of intergenic miRNAs. Our study provides the first direct evidence for the possibility of existence of independently transcribed human intronic miRNAs. We have identified the full-length pri-miRNA sequences of three intronic miRNAs namely miRs 106b, 93, and

24-1 by RLM-RACE, which confirms independent transcription. We demonstrate discordance in expression among the members of clustered intronic miRNAs and also with respect to their host genes by qPCR. We show that novel TSSs within the host genes and alternative splicing events within the miRNA-harboring introns might uncouple the expression of clustered miRNAs from each other and from their host genes by generating unique pri-miRNAs for each member of the miRNA cluster. Based on our results, we propose an alternative model of intronic miRNA biogenesis in addition to the existing canonical model (Fig. 6).

Recent studies by several groups have attempted to characterize the promoters and identify the TSS of miRNA genes, both intronic and intergenic, using bioinformatics approaches (Saini et al. 2007; Fujita and Iba 2008; Marson et al. 2008; Ozsolak et al. 2008; Corcoran et al. 2009; Wang et al. 2009b; Monteys et al. 2010; Chien et al. 2011). Analysis of these studies suggests that ~35% of human intronic miRNAs might possibly use promoters/TSS independent of their host genes (Ozsolak 36%, Corcoran 26%, Monteys 35%, and Chien

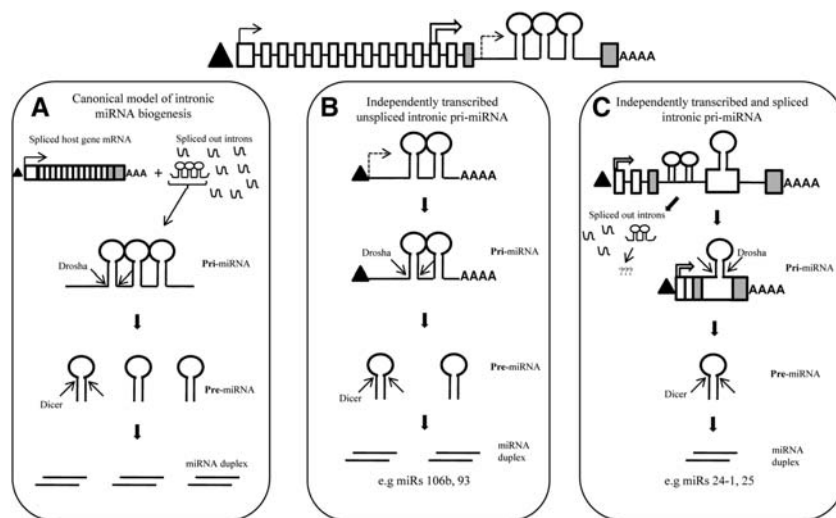


FIGURE 6. Alternative model of intronic miRNA biogenesis. The schematic on top shows the schematic representation of the genomic locus of a hypothetical host gene which harbors a cluster of three miRNAs in one of its introns. The solid arrow represents the TSS of this host gene. Triangle (▲) represents 5' 7-methyl guanosine cap, while "AAAA" represents the 3' poly(A) tail. The exons flanking the miRNA-hosting intron are shaded gray. (A) The canonical model of intronic miRNA biogenesis wherein splicing of the host gene results in release of the miRNA-harboring intron. The pre-miRNAs in this intron are then sequentially processed by Drosha and Dicer (represented by convergent arrows) to release the mature miRNAs. (B,C) The alternative model of biogenesis by independent transcription and alternative splicing as deduced from this study. (B) An independently transcribed unspliced pri-miRNA harboring two intronic miRNAs. This pri-miRNA is transcribed from a unique TSS (dashed arrow) within the host gene intron as observed in the case of PRI 106b/93. The pri-miRNA is then acted upon by Drosha and Dicer to release the mature miRNAs. (C) An independently transcribed spliced pri-miRNA. This pri-miRNA is transcribed from a unique TSS (open arrow) within the host gene as observed in the case of PRI 24. The pre-miRNA of one of the intronic miRNAs is included in this spliced pri-miRNA sequence as a result of alternative splicing within the miRNA-harboring intron as observed in PRI 24, ALT SPLICE 25, and ALT SPLICE 93. Further processing of this pri-miRNA releases the mature miRNA. It is unclear at the present moment whether the spliced out intron harboring the two other members of the miRNA cluster would be processed by Drosha (represented by "???").

40%). However, our study provides the first direct evidence supporting the possibility of independent transcription of human intronic miRNAs. The pri-miRNAs identified in our study also prove useful in validating the TSS predicted by some of these promoter prediction algorithms. For example, the TSS of miR 93 identified in our study lies 305 bp upstream of miR 93. The TSS predicted by Monteys et al. (2010) for this miRNA lies ~457 bp upstream of miR 93. Similarly, the identified TSS for miR 24-1 lies ~25 kb upstream of this miRNA, whereas the predicted TSS lies ~28 kb upstream according to the study by Fujita and Iba (2008) and ~31 kb according to Corcoran et al. (2009).

Since miR 24-1 and miR 106b were observed to be up-regulated under hypoxia, we scanned the 5-kb upstream genomic region of PRI 24 and PRI 106b/93 for the presence of hypoxia responsive elements (Supplemental Fig. 8A,B). Interestingly, we observed several canonical HIF-1 α sites (V\$HIF1_Q6: NRCGTGNGN) in the genomic region upstream of both PRI 24 and PRI 106b/93 which might regulate the transcription of these pri-miRNAs under hypoxic conditions. However, further studies are required to confirm whether these pri-miRNAs are directly regulated by HIF-1 α and the mechanisms by which hypoxia influences miRNA biogenesis.

Several lines of evidence indicate that splicing promotes miRNA biogenesis. Recent studies show that Drosha interacts with the spliceosome (Kataoka et al. 2009) and splicing enhances Drosha processing of pri-miRNAs (Houbaviy et al. 2005). Houbaviy et al. (2005) have shown that spliced pri-miRNA is preferentially cleaved by Drosha and splicing may enhance Drosha processing of pri-miRNAs. Similarly, Kim and Kim (2007) have shown that partially spliced host gene hnRNAs are recognized and cleaved by Drosha more efficiently as compared with the nascent hnRNAs. A recent study by Janas et al. (2011) has shown that knockdown of U1 splicing factors or the mutation of upstream splice donor signals resulted in decreased expression of some intronic as well as intergenic miRNAs. In addition to these reports, we have shown in our study that alternatively spliced transcripts act as pri-miRNAs for the resident intronic miRNAs. The conserved splice acceptor signals observed upstream of human intronic miRNAs and the identification of spliced ESTs corresponding to these predicted splice sites provide further evidence signifying the role of splicing in intronic miRNA biogenesis. Though spliced pri-miRNAs have been identified for several intergenic miRNAs previously, our study is the first to report the existence of spliced pri-miRNAs for human intronic miRNAs (Supplemental Fig. 8C). The existence of spliced pri-miRNAs suggests that splicing might play a significant role in biogenesis of both intronic as well as intergenic miRNAs. Based on these observations, we propose an alternative model of intronic miRNA biogenesis (Fig. 6).

According to the canonical model of intronic miRNA biogenesis, the miRNA-harboring intron of the host gene is considered to be the pri-miRNA for intronic miRNAs, and the completely spliced hnRNA becomes the protein-coding

mRNA (Fig. 6A). Both these events are coupled and not considered to be mutually exclusive. However, our model predicts that at least for some intronic miRNAs, the pri-miRNA might be produced from a unique transcription unit different from that of the host gene *hnRNA* (Fig. 6B). Since the host gene and intronic miRNA in such cases are produced from two different transcripts, the two events are neither coupled nor mutually exclusive. This would help to explain, at least in some instances, the discordance in expression observed between intronic miRNA–host mRNA pairs. Also, clustered miRNAs can be expressed independent of one another according to our model. Independent transcription followed by alternative splicing in the region of clustered miRNAs results in the production of unique pri-miRNAs for each of the cluster members (Fig. 6C). This might allow the cellular machinery to fine-tune the expression of each miRNA of an intronic miRNA cluster without affecting the steady state levels of either its host gene or other miRNAs of the cluster.

Our study provides a novel mechanistic insight into an alternative mechanism of biogenesis of clustered intronic miRNAs. We have shown that a subset of human intronic miRNAs can also be transcribed as nested genes within their hosting transcription units which were hitherto believed to arise only from processing of the introns of their host genes. Our results also suggest that alternative splicing in the vicinity of clustered intronic miRNAs results in the maturation of multiple pri-miRNAs, each of which harbors a single miRNA belonging to the cluster. These findings add to the existing knowledge toward understanding the complexity of regulation of miRNA biogenesis.

MATERIALS AND METHODS

Cell culture and transfections

MCF7 (breast adenocarcinoma) and HeLa (cervical adenocarcinoma) cells were used in the experiments. Cells were maintained in DMEM supplemented with 10% FCS and 3.7 g/dL sodium bicarbonate at 37°C. Cells were cultured in airtight chambers (MART) with exact concentrations of O₂, CO₂, and N₂ maintained using vacuum pumps and gas proportionator (Anoxomat). Cells were maintained at 20% O₂ and 5% CO₂ (normoxia). For the induction of hypoxia, the O₂ concentration was reduced to 0.1%. Drosha siRNA sequences were obtained from Kim and Kim (2007) and siRNA sequences against SNRNP70 and PRPF8 were obtained from Janas et al. (2011) (Supplemental Table 2). siRNA duplexes against Drosha, SNRNP70, PRPF8, or control siRNA (Universal Negative Control siRNA from Invitrogen) were transfected using Lipofectamine 2000 (Invitrogen) as per manufacturer's protocol.

RT-PCR and quantitative PCR

qPCR experiments were performed in Corbett RotorGene 6000 Real-Time PCR instrument (Corbett Research) using SYTO 9 dye (Invitrogen). The sequences of the primers are included in Supplemental Table 1. Random decamer primed cDNA was used

for the qPCR experiments involving host genes, hypoxia responsive genes, and Droscha mRNA quantitation. miRNA quantitation was performed using ABI TAQMAN miRNA quantitation kit (Applied Biosystems) according to the manufacturer's protocol. For the identification and quantitation of alternatively spliced transcripts, RT-PCR was performed using oligo-dT primed cDNA. Appropriate minus RT controls were included in every experiment to ensure absence of genomic DNA contamination.

Data analysis and statistics

Each qPCR reaction was performed in triplicates. The Cq values and amplification efficiencies for each replicate were obtained from the Comparative Quantitation tool of RotorGene software. The fold changes were calculated using Realtime Expression Software Tool (REST 2008) based on the $\Delta\Delta\text{-Ct}$ method. *18s rRNA* and *PPIA* were used as reference genes for normalization. For miRNA quantitation, *U6 snRNA* was used for normalization. *P*-values were calculated using Pair Wise Fixed Reallocation Randomization Test inbuilt in the REST software using 10,000 iterations. Error bars in the figures represent standard error (SEM) of the fold changes corresponding to the triplicates. Each experiment was repeated at least twice. Results corresponding to one representative experiment are shown in the figures.

RLM-RACE

For the identification of 5' TSS and 3' ends of pri-miRNAs, we performed RLM-RACE using the ExactSTART™ Eukaryotic mRNA 5'- and 3'-RACE Kit (Epicentre Biotechnologies) according to the manufacturer's protocol.

SplicePort analysis

In order to predict the splice acceptor signals upstream of the selected intronic miRNAs, we utilized the SplicePort algorithm (<http://spliceport.cs.umd.edu/>). The entire miRNA-harboring intron along with the flanking upstream and downstream exons of the host genes were subjected to SplicePort analysis. The minimum threshold utilized was -0.95 which produces a sensitivity value, $TP/(TP + FN)$, of 99% and a false positive rate, $FP/(FP + TN)$, of 10% for AG (splice acceptor) locations. The splice donor sequences (at the upstream exon-miRNA-harboring intron junction) and the splice acceptor sequences (at the miRNA-harboring intron-downstream exon junction) corresponding to the host gene mRNA were accurately identified by SplicePort for all the genes analyzed and served as a positive control for validating the accuracy of splice sites predicted.

Construction of mini-genes for overexpression of alternatively spliced pri-miRNAs

The alternatively spliced pri-miRNAs were PCR-amplified using HeLa cDNA using exon-junction spanning primers (Supplemental Fig. 7A,B; Supplemental Table 1). The amplicons were inserted into the pTarget Mammalian TA cloning cum Expression Vector System (Promega). The sequences of all the plasmids were confirmed by DNA sequencing. Plasmids were transfected into HeLa cells using Lipofectamine 2000 (Invitrogen) as per manufacturer's recommendations.

Construction of MCM7 mini-gene and site-directed mutagenesis

The MCM7 mini-gene was PCR-amplified as described in Supplemental Figure 7D and inserted into the pTarget Mammalian TA cloning cum Expression Vector System (Promega). Site-directed mutagenesis for the splice acceptor signal was performed using Phusion Site-Directed Mutagenesis Kit (Thermo Scientific) as per manufacturer's recommendations. The primer sequences for mutagenesis are mentioned in Supplemental Table 1. Mutagenesis was confirmed by DNA sequencing. Plasmids were transfected into HeLa cells using Lipofectamine 2000 (Invitrogen) as per manufacturer's recommendations.

SUPPLEMENTAL MATERIAL

Supplemental material is available for this article.

ACKNOWLEDGMENTS

This work was supported by the Department of Biotechnology, Government of India (Sanction Nos. BT/01/COE/05/13, BT/PR11739/Med/30/166/2008, and BT/PR13733/AGR/36/667/2010) and Indian Council of Medical Research (No. 3/2/08-09/PG-the-sis-MPD-3). We thank Dr. Sourav Banerjee (National Brain Research Center, India) for his suggestions and critical comments.

Received August 21, 2013; accepted October 7, 2013.

REFERENCES

- Altuvia Y, Landgraf P, Lithwick G, Elefant N, Pfeffer S, Aravin A, Brownstein MJ, Tuschl T, Margalit H. 2005. Clustering and conservation patterns of human microRNAs. *Nucleic Acids Res* **33**: 2697–2706.
- Bartel DP. 2004. MicroRNAs: Genomics, biogenesis, mechanism, and function. *Cell* **116**: 281–297.
- Baskerville S, Bartel DP. 2005. Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes. *RNA* **11**: 241–247.
- Chang T-C, Wentzel EA, Kent OA, Ramachandran K, Mullendore M, Lee KH, Feldmann G, Yamakuchi M, Ferlito M, Lowenstein CJ. 2007. Transactivation of miR-34a by p53 broadly influences gene expression and promotes apoptosis. *Mol Cell* **26**: 745–752.
- Chien CH, Sun YM, Chang WC, Chiang-Hsieh PY, Lee TY, Tsai WC, Horng JT, Tsou AP, Huang HD. 2011. Identifying transcriptional start sites of human microRNAs based on high-throughput sequencing data. *Nucleic Acids Res* **39**: 9345–9356.
- Corcoran DL, Pandit KV, Gordon B, Bhattacharjee A, Kaminski N, Benos PV. 2009. Features of mammalian microRNA promoters emerge from polymerase II chromatin immunoprecipitation data. *PLoS One* **4**: e5279.
- De Schutter H, Landuyt W, Verbeken E, Goethals L, Hermans R, Nuyts S. 2005. The prognostic value of the hypoxia markers CA IX and GLUT 1 and the cytokines VEGF and IL 6 in head and neck squamous cell carcinoma treated by radiotherapy ± chemotherapy. *BMC Cancer* **5**: 42.
- Eis PS, Tam W, Sun L, Chadburn A, Li Z, Gomez MF, Lund E, Dahlberg JE. 2005. Accumulation of miR-155 and BIC RNA in human B cell lymphomas. *Proc Natl Acad Sci* **102**: 3627–3632.
- Fujita S, Iba H. 2008. Putative promoter regions of miRNA genes involved in evolutionarily conserved regulatory systems among vertebrates. *Bioinformatics* **24**: 303–308.
- Fukao T, Fukuda Y, Kiga K, Sharif J, Hino K, Enomoto Y, Kawamura A, Nakamura K, Takeuchi T, Tanabe M. 2007. An evolutionarily

- conserved mechanism for microRNA-223 expression revealed by microRNA gene profiling. *Cell* **129**: 617–631.
- Gregory RI, Yan K, Amuthan G, Chendrimada T, Doratotaj B, Cooch N, Shiekhattar R. 2004. The microprocessor complex mediates the genesis of microRNAs. *Nature* **432**: 235–240.
- Houbaviy HB, Dennis L, Jaenisch R, Sharp PA. 2005. Characterization of a highly variable eutherian microRNA gene. *RNA* **11**: 1245–1257.
- Hubbi ME, Luo W, Baek JH, Semenza GL. 2011. MCM proteins are negative regulators of hypoxia-inducible factor 1. *Mol Cell* **42**: 700–712.
- Janas MM, Khaled M, Schubert S, Bernstein JG, Golan D, Veguilla RA, Fisher DE, Shomron N, Levy C, Novina CD. 2011. Feed-forward microprocessing and splicing activities at a microRNA-containing intron. *PLoS Genet* **7**: e1002330.
- Kataoka N, Fujita M, Ohno M. 2009. Functional association of the microprocessor complex with the spliceosome. *Mol Cell Biol* **29**: 3243–3254.
- Kim Y-K, Kim VN. 2007. Processing of intronic microRNAs. *EMBO J* **26**: 775–783.
- Kim DH, Sætrom P, Snøve O, Rossi JJ. 2008. MicroRNA-directed transcriptional gene silencing in mammalian cells. *Proc Natl Acad Sci* **105**: 16230–16235.
- Kim YK, Yu J, Han TS, Park SY, Namkoong B, Kim DH, Hur K, Yoo MW, Lee HJ, Yang HK, et al. 2009. Functional links between clustered microRNAs: Suppression of cell-cycle inhibitors by microRNA clusters in gastric cancer. *Nucleic Acids Res* **37**: 1672–1681.
- Kulshreshtha R, Ferracin M, Wojcik SE, Garzon R, Alder H, Agosto-Perez FJ, Davuluri R, Liu C-G, Croce CM, Negrini M, et al. 2006. A microRNA signature of hypoxia. *Mol Cell Biol* **27**: 1859–1867.
- Kulshreshtha R, Davuluri RV, Calin GA, Ivan M. 2008. A microRNA component of the hypoxic response. *Cell Death Differ* **15**: 667–671.
- Lee Y, Jeon K, Lee J-T, Kim S, Kim VN. 2002. MicroRNA maturation: Stepwise processing and subcellular localization. *EMBO J* **21**: 4663–4670.
- Lee Y, Kim M, Han J, Yeom KH, Lee S, Baek SH, Kim VN. 2004. MicroRNA genes are transcribed by RNA polymerase II. *EMBO J* **23**: 4051–4060.
- Marson A, Levine SS, Cole MF, Frampton GM, Brambrink T, Johnstone S, Guenther MG, Johnston WK, Wernig M, Newman J, et al. 2008. Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. *Cell* **134**: 521–533.
- Monteys AM, Spengler RM, Wan J, Tecedor L, Lennox KA, Xing Y, Davidson BL. 2010. Structure and activity of putative intronic miRNA promoters. *RNA* **16**: 495–505.
- Ozsolak F, Poling LL, Wang Z, Liu H, Liu XS, Roeder RG, Zhang X, Song JS, Fisher DE. 2008. Chromatin structure analyses identify miRNA promoters. *Genes Dev* **22**: 3172–3183.
- Pires IM, Bencokova Z, Milani M, Folkes LK, Li J-L, Stratford MR, Harris AL, Hammond EM. 2010. Effects of acute versus chronic hypoxia on DNA damage responses and genomic instability. *Cancer Res* **70**: 925–935.
- Radfar MH, Wong W, Morris Q. 2011. Computational prediction of intronic microRNA targets using host gene expression reveals novel regulatory mechanisms. *PLoS One* **6**: e19312.
- Rodriguez A. 2004. Identification of mammalian microRNA host genes and transcription units. *Genome Res* **14**: 1902–1910.
- Saini HK, Griffiths-Jones S, Enright AJ. 2007. Genomic analysis of human microRNA transcripts. *Proc Natl Acad Sci* **104**: 17719–17724.
- Saito Y, Liang G, Egger G, Friedman JM, Chuang JC, Coetzee GA, Jones PA. 2006. Specific activation of microRNA-127 with downregulation of the proto-oncogene *BCL6* by chromatin-modifying drugs in human cancer cells. *Cancer Cell* **9**: 435–443.
- Schmittgen TD. 2004. A high-throughput method to monitor the expression of microRNA precursors. *Nucleic Acids Res* **32**: e43.
- Sempere LF, Freemantle S, Pitha-Rowe I, Moss E, Dmitrovsky E, 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.
- Sikand K, Slane SD, Shukla GC. 2009. Intrinsic expression of host genes and intronic miRNAs in prostate carcinoma cells. *Cancer Cell Int* **9**: 21.
- Sorek R, Ast G. 2003. Intronic sequences flanking alternatively spliced exons are conserved between human and mouse. *Genome Res* **13**: 1631–1637.
- Taganov KD, Boldin MP, Chang K-J, Baltimore D. 2006. NF- κ B-dependent induction of microRNA miR-146, an inhibitor targeted to signaling proteins of innate immune responses. *Proc Natl Acad Sci* **103**: 12481–12486.
- Ulveling D, Francastel C, Hubé F. 2011. When one is better than two: RNA with dual functions. *Biochimie* **93**: 633–644.
- Wang D, Lu M, Miao J, Li T, Wang E, Cui Q. 2009a. Cepred: Predicting the co-expression patterns of the human intronic microRNAs with their host genes. *PLoS One* **4**: e4421.
- Wang X, Xuan Z, Zhao X, Li Y, Zhang MQ. 2009b. High-resolution human core-promoter prediction with CoreBoost_HM. *Genome Res* **19**: 266–275.
- Woods K, Thomson JM, Hammond SM. 2006. Direct regulation of an oncogenic micro-RNA cluster by E2F transcription factors. *J Biol Chem* **282**: 2130–2134.