Re-evaluation of the roles of *DROSHA*, *Exportin 5*, and *DICER* in microRNA biogenesis

Young-Kook Kim[a,b,c,1], Boseon Kim[a,d], and V. Narry Kim[a,d,1]

[a]Center for RNA Research, Institute for Basic Science (IBS), Seoul 08826, Korea; [b]Department of Biochemistry, Chonnam National University Medical School, Gwangju 61469, Korea; [c]Center for Creative Biomedical Scientists, Chonnam National University Medical School, Gwangju 61469, Korea; and [d]School of Biological Sciences, Seoul National University, Seoul 08826, Korea

Contributed by V. Narry Kim, February 19, 2016 (sent for review November 19, 2015; reviewed by Zissimos Mourelatos and Yukihide Tomari)

Biogenesis of canonical microRNAs (miRNAs) involves multiple steps: nuclear processing of primary miRNA (pri-miRNA) by DROSHA, nuclear export of precursor miRNA (pre-miRNA) by Exportin 5 (XPO5), and cytoplasmic processing of pre-miRNA by DICER. To gain a deeper understanding of the contribution of each of these maturation steps, we deleted *DROSHA*, *XPO5*, and *DICER* in the same human cell line, and analyzed their effects on miRNA biogenesis. Canonical miRNA production was completely abolished in *DROSHA*-deleted cells, whereas we detected a few DROSHA-independent miRNAs including three previously unidentified noncanonical miRNAs (miR-7706, miR-3615, and miR-1254). In contrast to *DROSHA* knockout, many canonical miRNAs were still detected without DICER albeit at markedly reduced levels. In the absence of DICER, pre-miRNAs are loaded directly onto AGO and trimmed at the 3′ end, yielding miRNAs from the 5′ strand (5p miRNAs). Interestingly, in *XPO5* knockout cells, most miRNAs are affected only modestly, suggesting that XPO5 is necessary but not critical for miRNA maturation. Our study demonstrates an essential role of DROSHA and an important contribution of DICER in the canonical miRNA pathway, and reveals that the function of XPO5 can be complemented by alternative mechanisms. Thus, this study allows us to understand differential contributions of key biogenesis factors, and provides with valuable resources for miRNA research.

microRNA | DROSHA | Exportin 5 | DICER | knockout

MicroRNA (miRNA) biogenesis begins with the synthesis of primary miRNA (pri-miRNA) by RNA polymerase II (1). The stem-loop structure embedded in pri-miRNA is cleaved by the Microprocessor complex composed of DROSHA and DGCR8 (2–6). The released hairpin, called precursor miRNA (pre-miRNA), is exported to the cytoplasm by Exportin 5 (XPO5) in a Ran-GTP-dependent manner (7–9). In the cytoplasm, the pre-miRNA is further processed by DICER, producing a duplex RNA of ∼22 nt with its 3′ ends having a two nucleotide overhang (10–13). The duplex is loaded onto the ARGONAUTE (AGO) proteins, and one strand of the duplex remains as mature miRNA, whereas the other strand is discarded from AGO (11, 14). The strand selection is dictated mainly by the relative thermodynamic stability of the two ends of the duplex: the strand whose 5′ terminal nucleotides are less stable is selected as mature miRNA (15, 16). miRNAs originating from the 5′ and 3′ strands of pre-miRNA are referred to as the 5p and 3p miRNAs, respectively. Mammals have four closely related AGO proteins (AGO1-4) that interact with deadenylation factors and translational machinery to induce mRNA degradation and translational repression.

Although the aforementioned canonical pathway accounts for the production of most miRNAs (1), it has also been shown that there exist alternative (noncanonical) pathways for miRNA biogenesis, which bypass a part of the biogenesis steps mentioned above. Mirtrons are one of the first miRNA groups described as noncanonical miRNAs, which do not require DROSHA for their production (17–19). Because mirtrons are located inside short introns of host genes, and their ends often match the splice sites, the spliced-out introns can serve as pre-miRNAs and processed by DICER. A functional miRNA can also be generated from a small nucleolar RNA, ACA45, in a DICER-dependent but DROSHA-independent manner (20). Moreover, endogenous siRNAs (endo-siRNAs) that do not require DROSHA but depend on DICER were identified in somatic tissues (21). Another group called "5′ capped pre-miRNAs" including miR-320 and miR-484 was recently identified (22). The pre-miRNA contains a 7-methylguanosine cap because its 5′ end is generated directly from transcription. The 3′ end of 5′ capped miRNA is thought to be determined by transcriptional termination.

As for DICER-independent maturation, miR-451 is the only known example that can be produced without DICER (23–25). Because of its short length, pre-miR-451 is not cleaved by DICER, and, instead, is directly incorporated into AGO2. Pre-miR-451 is cleaved by AGO2 in the middle of the 3′ strand, and further trimmed by 3′–5′ exoribonuclease PARN to produce the mature form of miR-451 (26).

The nuclear export step is mediated by XPO5, but it has not been investigated as intensively as the other maturation steps. A recent study showed that the miR-320 family requires Exportin 1 (XPO1) instead of XPO5, because these pre-miRNAs have a cap (which is recognized by XPO1 via an adapter molecule PHAX) instead of the typical 5′ monophosphate (22). It remains unknown what fraction of miRNAs are dependent on (or independent of) XPO5 because earlier studies on XPO5 examined individual miRNAs without taking a transcriptomic approach. Moreover, knockout of *XPO5* has not been generated yet. Thus, it is unknown how essential XPO5 is to miRNA biogenesis and whether there are additional alternative pathways for pre-miRNA export.

Significance

MicroRNAs (miRNAs) are noncoding RNAs with diverse roles in development and pathogenesis. Biogenesis of canonical miRNA requires nuclear processing by DROSHA, nuclear export by Exportin 5, and cytoplasmic processing by DICER. To gain a deeper understanding of the maturation processes, we here ablated the *DROSHA*, *Exportin 5*, and *DICER* genes using the same human cell line. Canonical miRNA production was abolished in *DROSHA*-deleted cells, revealing an irreplaceable role of DROSHA. Interestingly, however, some canonical miRNAs were still produced without DICER albeit at markedly reduced levels, and many were detected in *Exportin 5*-deleted cells at only modestly decreased levels. Our study allows us to understand differential contributions of key biogenesis factors, and provides valuable resources for miRNA research.

Author contributions: Y.-K.K. and V.N.K. designed research; Y.-K.K. and B.K. performed research; Y.-K.K. and V.N.K. analyzed data; and Y.-K.K. and V.N.K. wrote the paper.

Reviewers: Z.M., University of Pennsylvania School of Medicine; and Y.T., University of Tokyo.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

Data deposition: The data reported in this paper have been deposited in the Gene Expression Omnibus (GEO) database, www.ncbi.nlm.nih.gov/geo (accession no. GSE77989).

[1]To whom correspondence may be addressed. Email: narrykim@snu.ac.kr or ykk@jnu.ac.kr.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1602532113/-/DCSupplemental.

In this study, we use genome engineering techniques to knockout *DROSHA*, *XPO5*, and *DICER* in the same cell line. By analyzing and comparing the miRNA expression profiles by deep sequencing, we here investigate the essentiality of the key biogenesis factors and discover unexpected alternative mechanisms and previously unidentified noncanonical miRNAs.

Results

Ablation of *DROSHA*, *XPO5*, and *DICER*. To better understand the roles of miRNA maturation factors, we generated knockout cells by transfecting human HCT116 cell line with transcription activator-like effector nuclease (TALEN) or RNA-guided Cas9 endonuclease. We chose the HCT116 cell line because it is near diploid and often used for gene knockout studies. For the knockout of *DROSHA* and *DICER*, we designed guide RNAs complementary to the area near the genomic locus corresponding to the RNase IIIa domain of DROSHA and DICER (Fig. 1*A*). In the case of *XPO5*, we engineered the TALEN to edit the genomic locus matching between the third and fourth HEAT domains of XPO5 (Fig. 1*A*). After single cell cloning, the genomic DNA was analyzed by Sanger sequencing to select the clones that contain frameshift mutations in all alleles. We obtained one clone for *DROSHA*, two clones for *XPO5*, and two clones for *DICER* (Fig. 1*B*). We performed Western blot analysis to confirm that the targeted proteins

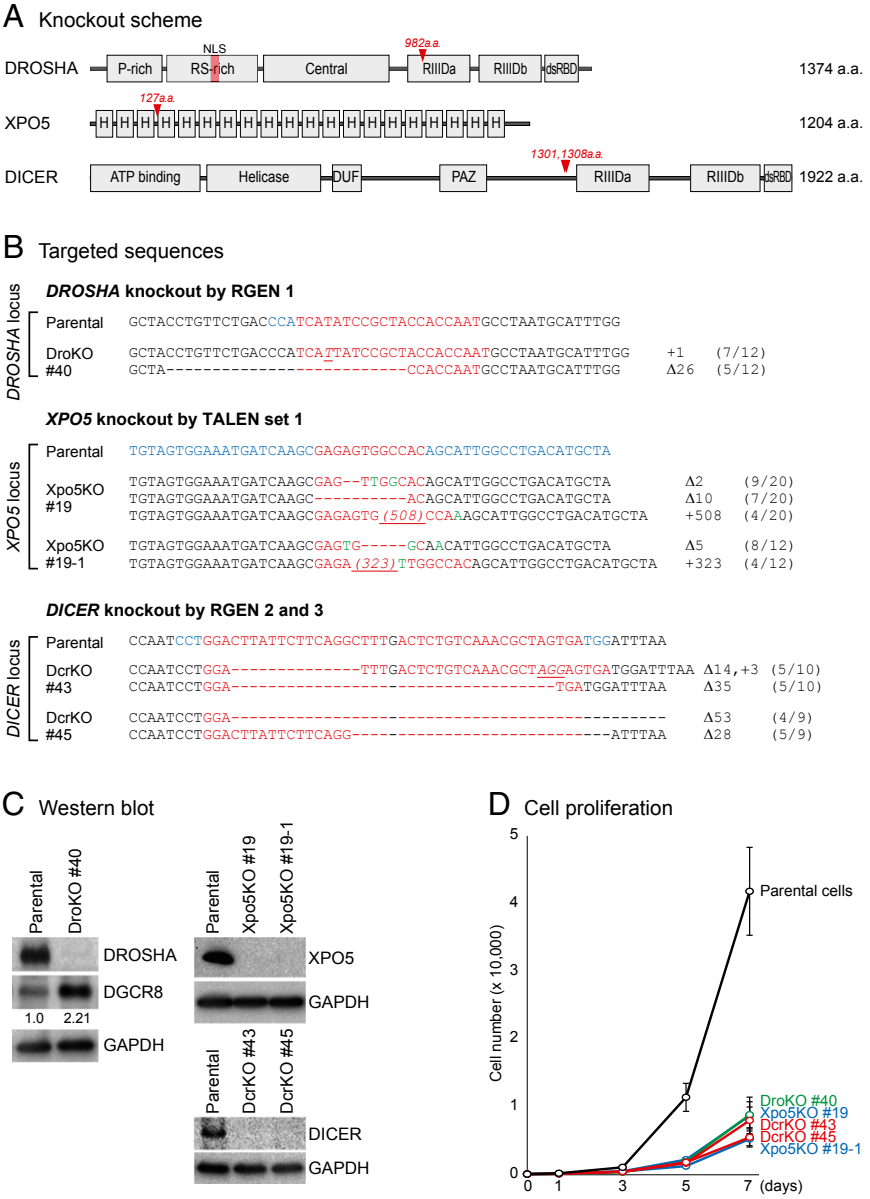


Fig. 1. Nuclease-mediated knockout of miRNA maturation factors. (*A*) Domain structure of DROSHA, XPO5, and DICER proteins. Red triangles indicate the regions corresponding to the genomic DNA sequences that are targeted by nuclease. For the knockout of *DROSHA* and *XPO5*, a single genomic DNA region was selected to be targeted by Cas9 and TALEN, respectively. In the case of *DICER*, two guide RNAs were used to cleave the adjacent genomic regions. (*B*) Targeted genomic sequences. Red letters in *DROSHA* and *DICER* indicate the regions recognized by guide RNAs. Red letters in *XPO5* indicate the region targeted to be cleaved by TALEN. Italic and underlined letters show insertion, whereas green letters stand for substitution. Blue letters in *DROSHA* and *DICER* indicate the Protospacer Adjacent Motif (PAM) recognized by Cas9 protein. Blue letters in *XPO5* indicate the binding region of TALEN pairs. On the right side, the number of mutated nucleotides and the sequencing frequency of the allele in the cell clone are presented. (*C*) Western blot experiments to confirm gene disruption. (*D*) Proliferation of parental and knockout cells measured by cell counting. Error bars show deviation from two independent experiments.

are not produced in the knockout lines (Fig. 1*C* and Fig. S1). For XPO5, we used two different antibodies, one targeting the N-terminal part (Fig. 1*C* and Fig. S1*B*) and another detecting the C-terminal part of XPO5 (Fig. S1*C*), both of which verified the ablation of XPO5 expression. Biogenesis factors other than the targeted ones were not reduced in the knockout cells. Note that in *DROSHA* knockout cells the level of DGCR8 increases as expected from the known activity of DROSHA targeting *DGCR8* mRNA (27). DICER was increased in *DROSHA* and *XPO5* knockout cells (Fig. S1*E*). This result was expected given that *DICER* is subject to feedback control by miRNAs including let-7 (28, 29). The knockout cells displayed lower growth rates compared with the parental cells, indicating impaired cell proliferation presumably due to a deficit in miRNA biogenesis (Fig. 1*D*).

Different Impacts of the Maturation Factors. To investigate miRNA population, we fractionated small RNAs (17–30 nt) and analyzed them by high-throughput sequencing. As expected, the proportion of miRNA reads relative to total reads in the library was markedly decreased in the knockout cells compared with the parental cells (Fig. 2*A*). However, the degree of reduction varied widely depending on the deleted genes.

The reduction was most prominent in the *DROSHA* knockout cell line (Fig. 2*A*, *Left*). The vast majority of miRNAs (193 of 200, 96.5%) were practically eradicated in the knockout cells [median reduction, 0.00067-fold (knockout/parental)] (Fig. 2*B* and Fig. S2*D*). The read counts of each miRNA species were normalized against that of miR-320a-3p whose production is independent of DROSHA (Fig. 2*B* and Fig. S2*D*) (22). After the

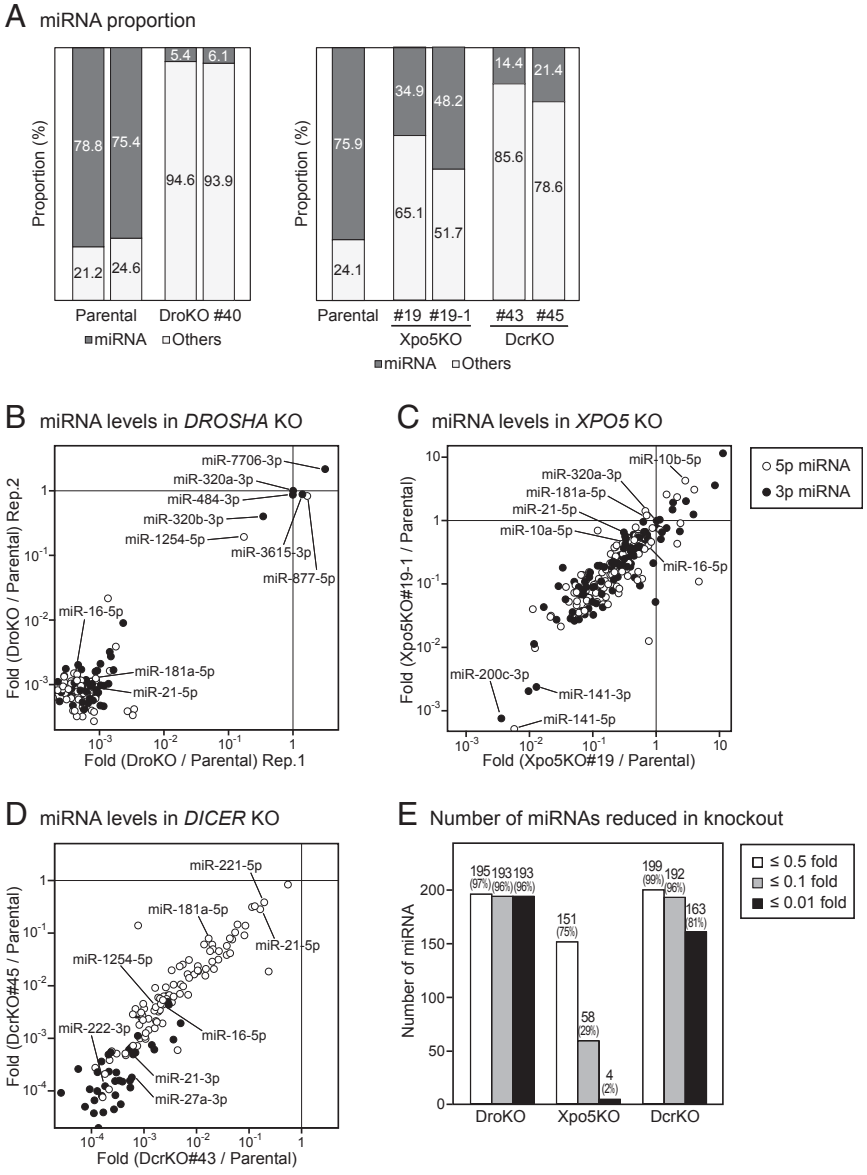


Fig. 2. Global analysis of miRNA expression in knockout cells. (*A*) The proportion of miRNA reads in the small RNA sequencing libraries from knockout cells and their parental cells. Two libraries from independent samples were generated from *DROSHA* knockout cells and its corresponding parental cell line (*Left*). As for *XPO5* and *DICER*, one library was made from each knockout clone (*Right*). (*B-D*) Expression change of miRNAs after the knockout was depicted by scatter plot. The top 200 miRNAs based on their sequencing reads in parental cells were selected for the analysis. For *DROSHA* knockout (*B*), normalized fold changes between two replicates were compared. For *XPO5* (*C*) and *DICER* knockouts (*D*), normalized fold changes between two different knockout clones were compared. The miR-320a-3p level was used for normalization in *DROSHA* and *XPO5* knockout sets. For the *DICER* knockout set, those reads aligned to rRNAs and tRNAs were used for normalization. Outliers and those that are validated by Northern blot are indicated. (*E*) Based on the fold change, the number of miRNAs reduced after the knockout were counted.

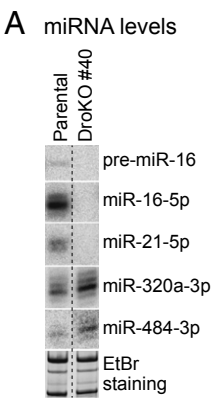
normalization, a minute amount of canonical miRNA reads remained in *DROSHA* knockout library, which may be due to cross-contaminations between libraries, because the abundance ranks of canonical miRNAs in the *DROSHA* knockout are almost identical to those in parental cells (Fig. S2). In contrast, the ranks differ substantially in *XPO5* and *DICER* knockout cells (Fig. S2).

A small subset of miRNAs was apparently independent of *DROSHA*, which include previously reported noncanonical miRNAs: 5′ capped miRNAs (miR-320a-3p, miR-320b-3p, and miR-484-3p) and mirtron (miR-877-5p) (17, 22). We also found some additional miRNAs that are insensitive to *DROSHA* ablation (discussed below).

Unlike in *DROSHA* knockout, we were surprised that a substantial amount of miRNAs were produced in *XPO5* knockout (Fig. 2*A*). For normalization, we divided the sequencing read counts of each miRNA species by that of miR-320a-3p, which depends on XPO1 for its nuclear export instead of XPO5 (Fig. 2*C* and Fig. S2*E*). The majority of miRNAs (151 of 200, 75.5%) were reduced in abundance [≤0.5-fold (knockout/parental)], confirming the role of XPO5 as a miRNA biogenesis factor. However, the degree of reduction (median reduction, 0.23-fold) was distinctively modest compared with that of *DROSHA* knockout.

In the *DICER*-deficient cells, almost all miRNAs were markedly reduced in abundance (median reduction, 0.00058-fold) (Fig. 2*D* and Fig. S2*F*). In this analysis, we had to use the combined reads from rRNAs and tRNAs for normalization because a known DICER-independent miRNA, miR-451, is not expressed in HCT116 cells. So it should be noted that the fold changes may globally shift depending on normalization method. Intriguingly, the reads from the 3′ strand (3p miRNAs) decreased more severely in the knockout cells (median reduction, 0.00009-fold) than those from the 5′ strand (5p miRNAs) (median reduction, 0.0041-fold) (Fig. 2*D* and Fig. S2*F*).

In summary, although DROSHA, XPO5, and DICER are indeed required for miRNA biogenesis, each factor contributes



B miRNAs produced in *DROSHA* KO

miRNA	Fold change in KO			Comments
	Dro	Xpo5	Dcr	
miR-7706-3p	**2.66**	**5.98**	**0.00**	**5′ capped**
miR-877-5p	1.20	0.06	0.00	mirtron
miR-3615-3p	**1.13**	**2.54**	**0.00**	**5′ capped**
miR-320a-3p	1.00	1.00	0.00	5′ capped
miR-484-3p	0.95	1.64	0.00	5′ capped
miR-320b-3p	0.47	0.51	0.00	5′ capped
miR-1254-5p	**0.19**	**0.11**	**0.00**	**Endo-siRNA?**
miR-16-5p	0.00	0.54	0.00	Canonical
miR-27a-3p	0.00	0.62	0.00	Canonical

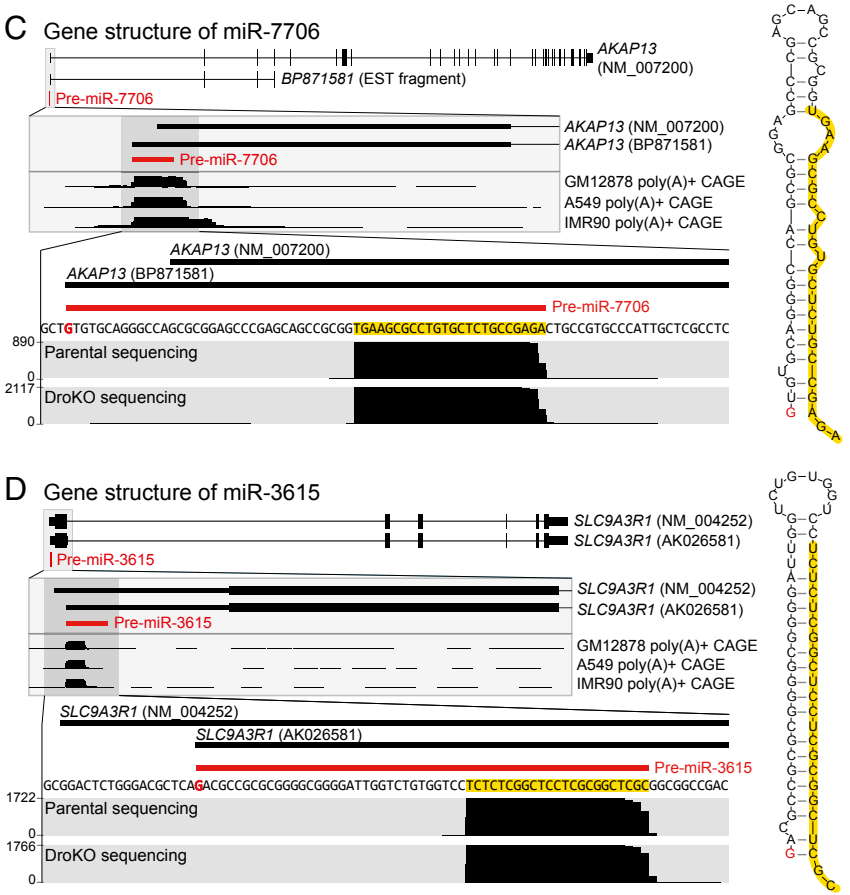


Fig. 3. miRNA expression in the *DROSHA* knockout cells. (*A*) Northern blot analysis to validate the changes in miRNA levels in the knockout. Canonical miRNAs such as miR-16-5p and miR-21-5p are not detectable, whereas noncanonical miRNAs including miR-320a-3p and miR-484–3p were readily observed. The dashed lines indicate discontinuous lanes from a single gel, which is true for all of the figures with dashed lines throughout this paper. (*B*) List of miRNAs produced in the *DROSHA* knockout cells. Bold letters indicate those that are newly identified DROSHA-independent miRNAs in this study. Those miRNAs with a fold change larger than 0.01 (*DROSHA* knockout/parental) are shown. A table for top 200 miRNAs based on the sequencing reads in parental cells is provided in Dataset S1. According to the sequencing reads and gene structure, miRNAs were classified into 5′ capped miRNA, mirtron, or endo-siRNA. (*C*) Predicted secondary structure of pre-miR-7706 and the sequencing data in comparison with the *AKAP13* gene structure. The CAGE data from ENCODE project were obtained from UCSC genome browser (genome.ucsc.edu/). The sequencing results combined from two libraries were used to depict the graph showing the reads at each nucleotide position. The *y* axis of the graph was normalized based on the level of miR-320a-3p, a DROSHA-independent miRNA. The miRNA sequence produced from 3p strand is indicated with yellow shade. Based on the sequencing data and the structure of the host gene, the secondary structure of pre-miR-7706 was predicted and shown on the right. Note that the 5′ end nucleotide (indicated with a red letter) of pre-miR-7706 coincides with the transcription start site of *AKAP13* isoform (BP871581), a probable host gene of miR-7706. Because there is no CAGE data available for colorectal cancer cell lines including HCT116, we analyzed the CAGE data from other randomly selected cell lines, GM12878, A549, and IMP90. All of the cell lines included in the ENCODE project showed similar CAGE patterns. The same analysis as in *C* was applied to *D* miR-3615. In addition to the RefSeq mRNA *SLC9A3R1* (NM_004252), an mRNA from GenBank, AK026581, was also shown in *D* because the latter is expected to be a more probable host gene of miR-3615 based on its 5′ end sequence.

differently to miRNA production (Fig. 2*E*). Canonical miRNA biogenesis was abolished in the *DROSHA* knockout cells. In contrast, many miRNAs were produced in the absence of XPO5. DICER is critical for most miRNAs, but the 5p miRNAs appear to be produced to some extent even without DICER. Thus, DROSHA may be the only essential factor for the canonical miRNA pathway, at least in this cell type.

miRNA Expression in the *DROSHA* Knockout Cells. To validate the sequencing data, we performed Northern blot analysis on two canonical miRNAs, miR-16-5p and miR-21-5p, that are abundant in parental cells (Fig. 3*A*). These miRNAs were undetectable in the *DROSHA* knockout cells. Thus, canonical miRNAs are strictly dependent on DROSHA.

Unlike the canonical ones, miR-320a-3p and miR-484-3p were detected readily by Northern blotting in the *DROSHA* knockout cells (Fig. 3*A*), consistent with the previous finding (22). Notably, the levels of miR-320a-3p and miR-484-3p increased reproducibly in the absence of canonical miRNAs. This result suggests that small RNAs may compete against each other for biogenesis factors and/or AGOs.

In addition to miR-320 and miR-484, we found several other miRNAs that are independent of DROSHA: miR-7706, miR-3615, and miR-1254 (Figs. 2*B* and 3*B* and Fig. S2*D*). To understand how these miRNAs are generated without DROSHA, we examined the genomic distribution of sequencing reads together with transcriptomic data from public databases. At the *mir-7706* locus (Fig. 3*C*), we did not find any small RNA read corresponding to the 5p arm. Assuming that the precursor miRNA has a 2-nt overhang at its 3′ end, we predicted the structure of pre-miR-7706 (Fig. 3*C*, *Right*). When we mapped this precursor sequence to the genome, the expected 5′ end of pre-miR-7706 matched exactly the start site of the *AKAP13* isoform BP871581. Although the reference sequence (RefSeq, NM_007200) starts 13 nt downstream, the 5′ end of the isoform BP871581 is supported by the transcription initiation site determined by cap analysis gene expression (CAGE) (30), indicating that transcription is indeed initiated at the 5′ end of pre-miR-7706 (Fig. 3*C*). These observations suggest that pre-miR-7706 is generated directly by transcription and does not require DROSHA. Thus, it is likely that miR-7706 is a previously unidentified member of the 5′ capped pre-miRNA family.

Similarly, we did not detect any 5p miRNA reads from the *mir-3615* locus. The 5′ end of the predicted precursor maps to the 5′ end of *SLC9A3R1* (AK026581) and that of the CAGE signals, indicating that the precursor is generated by transcription rather than from DROSHA-mediated processing (Fig. 3*D*). Thus, miR-3615 may also belong to the 5′ capped pre-miRNA family. It is intriguing that the same promoter is used for both miRNA and protein genes. These two uses must be mutually exclusive such that transcriptional termination will result in miRNA production, whereas transcriptional elongation will generate a hairpin RNA with an ORF in the 3′ tail, which is too long for DICER to process. It will be interesting to investigate the functional relationship between the miRNAs and their overlapping protein-coding genes.

As for miR-484, the CAGE signal was observed downstream of its host gene *NDE1* (NM_001143979) and exactly matched the 5′ end of another isoform of *NDE1* (NM_017668) (Fig. S3*A*). This data implies that two independent promoters are used to produce pre-miR-484 and the *NDE1* mRNA separately. The miR-320 family miRNAs do not seem to overlap with the downstream protein coding gene.

miR-1254-5p is another noncanonical miRNA that is not affected substantially by DROSHA knockout but is dependent strictly on DICER. miR-1254 is produced from an intron of protein coding gene, *CCAR1*. The genomic location of *mir-1254-1* overlaps with that of an Alu, a type of short interspersed nuclear elements (SINEs), and belongs to the AluJr subgroup (Fig. S3*B*). Unlike the other Alu subtypes, this Alu sequence can form a stable hairpin. Our results suggest that the hairpin may be cleaved by DICER, yielding a kind of a DROSHA-independent, DICER-dependent endogenous siRNA (31).

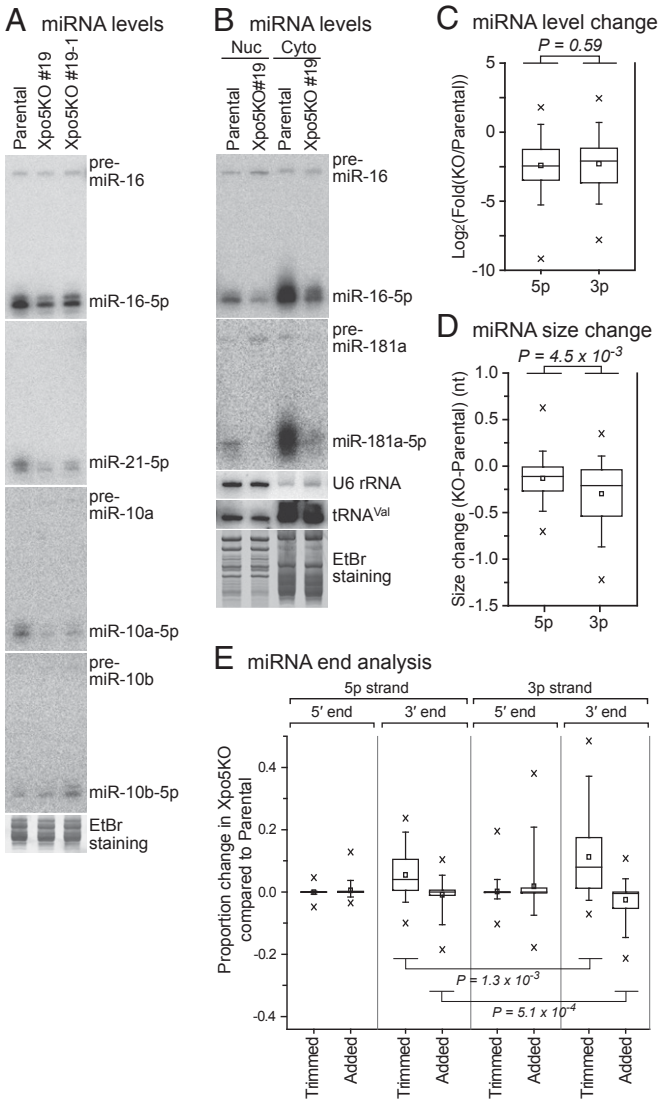


Fig. 4. miRNA expression in the *XPO5* knockout cells. (*A*) The expression of miRNAs from parental cells and *XPO5* knockout cells was measured. (*B*) The expression level of miRNAs in the nucleus and cytoplasm of parental cells and *XPO5* knockout cells, respectively, was compared. (*C*) The change in expression level after *XPO5* knockout was compared between the miRNAs from 5p and 3p strands. Only the reads with perfect match to miRNA sequences were selected for analysis. *P* value was calculated by two-sided Wilcoxon rank-sum test. (*D*) The change in length after *XPO5* knockout was compared between the miRNAs from 5p and 3p strands. *P* value was calculated by two-sided Wilcoxon rank-sum test. (*E*) The proportion change in miRNAs that are trimmed or added after *XPO5* knockout was analyzed at each end of the miRNA sequences. The proportions of miRNAs with sizes less than those of reference sequences were analyzed using the sequencing data from parental cells and *XPO5* knockout cells, respectively, and their differences were calculated to be used as the change in "Trimmed" proportion. For the analysis of 5′ end, only the miRNA reads whose 3′ end coincides with the 3′ end of reference sequence were used, and vice versa. The same analysis was applied to calculate the proportion change of "Added" miRNA except that longer sequences than reference sequence were collected for the analysis. *P* value was calculated by two-sided Wilcoxon rank-sum test.

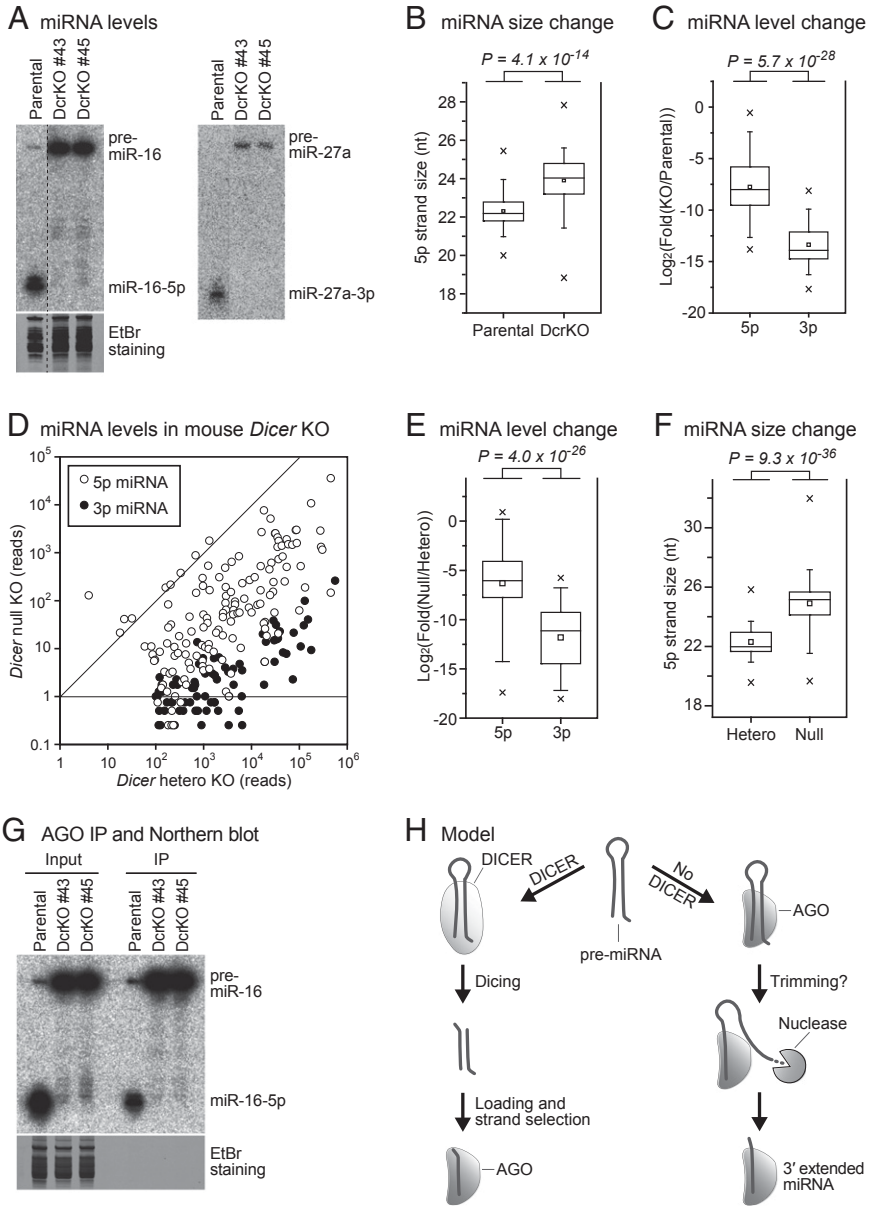


Fig. 5. miRNA expression in the *DICER* knockout cells. (*A*) The expression of miRNAs in parental cells and *DICER* knockout cells was measured. Note the bands for miRNA fragments between pre-miRNA and mature miRNA. (*B*) Change in the length of 5p strand miRNAs was compared between parental and *DICER* knockout cells. *P* value was calculated by two-sided Wilcoxon rank-sum test. (*C*) Change in the expression level of miRNAs was compared between 5p and 3p strand miRNAs. *P* value was calculated by two-sided Wilcoxon rank-sum test. (*D*) The expression of individual miRNA was compared between the sequencing libraries made with *Dicer* heterozygous knockout and null knockout mouse cells, respectively (35). The miRNAs produced from 5p and 3p strands are shown with different colors. (*E*) The change in expression level of miRNAs produced from 5p and 3p strands were compared. (*F*) The length change of 5p strand miRNAs were compared between *Dicer* heterozygous knockout and null knockout cells. The same data in *D* were used for the analysis in *E* and *F*. *P* value was calculated by two-sided Wilcoxon rank-sum test. (*G*) Association of 3′ extended-5p miRNAs with AGO proteins. Input RNA was prepared from 200 μg of protein extracts. For the immunoprecipitation of AGO proteins, 2.4 mg of protein extract was used. The associated miRNAs with AGO proteins were measured by Northern blot. (*H*) Model for the generation of 3′ extended-5p miRNAs in the absence of DICER. Although pre-miRNAs are loaded into AGO proteins, their 3′ ends may be more vulnerable to nuclease attack if they are not processed by DICER rapidly. The nuclease may trim the pre-miRNA until most of the terminal loop is removed, but further trimming might be hindered by Ago proteins making 3′ extended-5p miRNAs.

Generation of miRNAs in the *XPO5* Knockout Cells. We next examined the *XPO5* knockout cells by Northern blot analysis. Validating the sequencing data (Fig. 2*C* and Fig. S2*E*), miRNAs were reduced in abundance but still readily detectable in the knockout (Fig. 4*A*). To examine the intracellular distribution of pre-miRNAs, we performed subcellular fractionation and Northern blotting (Fig. 4*B*). The nuclear pre-miRNA level increased in the knockout cells, as expected for the established function of XPO5 (Fig. 4*B*). However, pre-miRNAs, particularly pre-miR-16, were readily detected in the cytoplasm of the knockout cells (Fig. 4*B*). This result suggests that although XPO5 indeed mediates pre-miRNA export, alternative pathway(s) may exist to translocate some pre-miRNAs. Note that the mature miRNAs are markedly decreased in the cytoplasm of *XPO5* knockout cells, whereas the nuclear accumulation of pre-miRNAs is less prominent, as previously observed (7) (Fig. 4*B*). It is plausible that pre-miRNA is less stable than mature miRNA which is loaded and protected by AGO proteins.

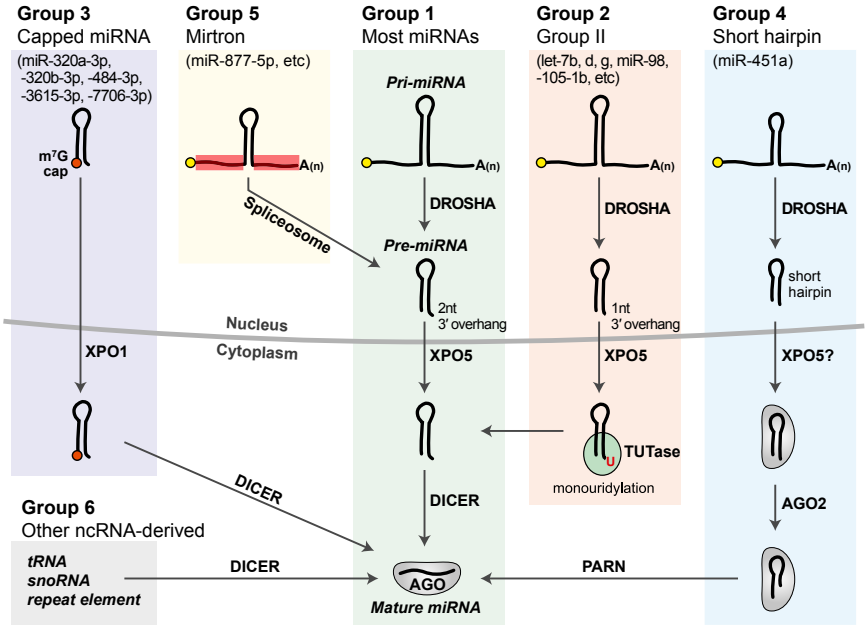


Fig. 6. Biogenesis pathways of miRNAs. The miRNAs are categorized into six groups based on the requirement for each biogenesis factor. Nuclear export of pre-miRNAs can be mediated by the factor other than XPO5 or XPO1, although not indicated in this figure. This figure was modified from ref. 1. Please refer to the text for a detailed description.

A possibility to consider is that pre-miRNAs simply diffuse during mitosis when nuclear membrane disintegrates. Arguing against this possibility, however, the knockout effect varied widely among miRNAs: miR-16-5p and miR-10b-5p were affected more modestly than miR-10a-5p, miR-21-5p, and miR-181a-5p. Thus, simple diffusion is unlikely to explain the effective maturation of selective miRNAs in the *XPO5* knockout cells. We do not exclude the possibility that some of the nuclear retained pre-miRNAs are slowly processed by nuclear DICER, given that DICER is detected in the nucleus albeit at a lower level than that in the cytoplasm (32, 33). Moreover, some pri-miRNAs may escape nuclear processing and get processed by the cytoplasmic Microprocessor. Although all of these alternative scenarios are expected to be of low efficiency, their additive actions may complement the loss of XPO5.

Intriguingly, the 5p and 3p miRNAs are affected differently by the XPO5 deletion. Their abundance was reduced to a comparable extent (Figs. 2*C* and 4*C* and Fig. S2*E*). However, we noticed that the 3p miRNAs were selectively shortened in the *XPO5* knockout cells (Fig. 4*D*). The shortening occurs mainly at the 3′ end of 3p miRNAs (Fig. 4*E*), suggesting that the 3′ ends of pre-miRNAs may be trimmed by nuclear 3′–5′ exoribonuclease (s) when pre-miRNA export is delayed. It is also possible that the XPO5 binds and protects the 3′ ends of 3p miRNAs from the degradation (7). Despite the 3′ shortening, the DICER cleavage site (which is reflected by the 5′ end of the 3p miRNA) remains largely unaltered in the knockout (Fig. 4*E*). This data is consistent with our previous finding that human DICER measures mainly from the 5′ end rather than the 3′ end of pre-miRNA to determine the cleavage site (34). Given that the targeting specificity of miRNA is determined by the sequences at nucleotides 2–8 relative to the 5′ end of miRNA, it is unlikely that the shortening of the 3′ end of mature 3p miRNA changes their targeting specificity.

Direct AGO Loading of pre-miRNA in *DICER*-Deleted Cells. Northern blot analysis of miR-16-5p and miR-27a-3p confirmed a strong reduction of mature miRNAs and a concomitant accumulation of pre-miRNAs in *DICER* knockout cells (Fig. 5*A*). Thus, DICER is indeed an important enzyme in miRNA biogenesis.

Interestingly, however, additional bands of ∼23–35 nt were visible when we probed the blot for miR-16-5p, but not in the case of miR-27a-3p (Fig. 5*A*). This result is consistent with the sequencing data: the 5p miRNA species with a 3′ extension was often observed (Fig. S4*A*). These unusually long miRNA species originate selectively from the 5′ strand, and have an extension at the 3′ end (Fig. 5*B*). The extended sequences match the genomic sequences, indicating that these elongated miRNAs are generated from 3′–5′ trimming of pre-miRNAs rather than from nontemplated nucleotide addition (tailing) of mature miRNAs. When pre-miRNAs are trimmed in a 3′–5′ direction, only the 5p miRNAs can be produced, whereas the 3p miRNA sequences are chipped away. Consistent with this model, the abundance of the 3p miRNAs were more severely reduced than that of the 5p miRNAs (Fig. 5*C*), and we could not detect any 3p miRNAs by Northern blotting (Fig. 5*A*).

In addition, when we reanalyzed previously published sequencing data from *Dicer* knockout mouse sarcoma cell line (35), we made a similar observation: the 3p miRNAs were reduced more strongly in abundance than the 5p miRNAs were (Fig. 5 *D* and *E*). Furthermore, the 5p miRNAs increased in size due to the 3′ extension in the null mutant cells (Fig. 5*F*). We also observed a similar pattern in another published sequencing data from *dicer* knockout zebrafish embryo (24) (Fig. S4 *B–D*). These data support our model and demonstrate the evolutionary conservation of the mechanism.

Previous studies have shown that some pre-miRNAs can be directly loaded onto AGO (36, 37). In the case of miR-451, pre-miRNA is cleaved endonucleolytically by AGO and trimmed by PARN to yield miRNA of variable size (20–30 nt) (23–26). To examine if pre-miRNAs are directly loaded onto AGO in the knockout cells, we performed immunoprecipitation and Northern blotting (Fig. 5*G*). Both pre-miR-16 and its fragments (23–30 nt) are associated with AGO (Fig. 5*G*). This result confirms and extends the previous observations: when DICER processing is compromised, pre-miRNA can be directly loaded onto AGO, allowing selective maturation of the 5p miRNAs (Fig. 5*H*). Given that most 5p miRNA species are detected in DICER-deficient

cells, direct AGO loading may take place more widely than previously appreciated.

Discussion

Our current study provides valuable resources for miRNA research. By analyzing small RNA population in these knockout cells, we validate the current dogmatic model of canonical miRNA pathway. DROSHA and DICER ablation resulted in a depletion of 96.5% and 96% of detected miRNA species by ≤0.1-fold (knockout/parental), respectively (Fig. 2*E*). We also confirmed the contribution of XPO5 to the majority of miRNAs: 75.5% of detected miRNA species were decreased by ≤0.5-fold.

It was surprising, however, that in *XPO5* knockout the reduction was much more modest than expected: only 29% of miRNAs were decreased by ≤0.1-fold. Therefore, DROSHA and DICER are indeed critical for canonical miRNA biogenesis, whereas XPO5 is not an indispensable factor and may be replaced by other potentially multiple mechanisms. Thus, the requirement for XPO5 cannot be an effective classifier for miRNA.

To our knowledge, this is the first knockout study on XPO5. In a previous study where XPO5 was identified as a nuclear export factor of pre-miRNA, the conclusion was based on binding assays, knockdown experiments, and detection of let-7a-5p (9). We found that let-7a-5p is one of the most sensitive miRNAs to *XPO5* ablation [0.027-fold reduction (knockout/parental) (Dataset S1)]. In another study, the function of XPO5 was examined by measuring the level of ectopically expressed miRNA in the XPO5-knockdown cells (7). Our current global analysis using *XPO5* knockout cells confirms the involvement of XPO5 in the miRNA pathway, but at the same time, reveals that XPO5 is not an essential factor for miRNA biogenesis, at least in HCT116 cells. It was recently reported that a genetic defect in *XPO5* reduces the production of mature miRNAs (38). A heterozygotic mutation results in a C-terminal truncation of XPO5 protein in a subset of cancer cell lines. By reanalyzing the microarray data from that study, however, we found that considerable amounts of miRNAs were still expressed in cells with the XPO5 truncation (38). Thus, the truncating mutation may have a limited effect on miRNAs, although the modest reduction may still be sufficient to influence cancer physiology.

Based on the results from this and previous studies, we can categorize miRNAs into six groups (Fig. 6, modified from ref. 1). Most miRNAs are classified as group 1 miRNAs, which require both DROSHA and DICER for their biogenesis. On the other hand, additional processing steps or a modified biogenesis pathway is used for other noncanonical miRNAs (groups 2 through 6). The production of group 2 miRNAs requires monouridylation of pre-miRNAs for their efficient processing by DICER, because these pre-miRNAs have a short (1 nt) 3′ overhang (39). Most vertebrate let-7 members and miR-105 belong to group 2. Group 3 miRNAs are derived from the 5′ capped pre-miRNAs that do not require DROSHA for their production (22). In the current study, we identify two additional noncanonical miRNAs that may belong to group 3 (miR-3615 and miR-7706). Their sequences are conserved only in primates, suggesting that these miRNAs may have evolved recently. Group 4 includes miR-451a, a short hairpin miRNA that does not require a DICER cleavage step during its biogenesis (23–25). Group 5 is composed of mirtrons that are produced from spliced-out intron, instead of DROSHA-mediated processing (17–19). Finally, miRNAs that are processed from structured noncoding RNAs by DICER can be categorized as group 6 miRNAs (1). Our data suggest that miR-1254-5p may belong to group 6, as it appears to be produced from Alu-derived long stem-loop and directly cleaved by DICER. Consistent with previous reports, the noncanonical miRNAs of groups 3–6 found in this study are generally low in abundance and poorly conserved, except for the miR-320 family (conserved in vertebrates) and miR-484 (conserved in mammals).

Materials and Methods

Knockout Procedure. The colorectal cancer cell line, HCT116, was maintained with McCoy's 5A media supplemented with 10% FBS (WelGene). The TALEN and Cas9 constructs were synthesized by ToolGen Inc. as described previously (40). The DNA binding sequences of TALEN constructs and guide RNA sequences are shown in Fig. 1*B*. The knockout and screening were performed by ToolGen, and overall procedures were described (41). We used a reporter construct harboring the recognition sequence of TALEN or guide RNA to enrich the cells with DNA mutation as shown in a previous report (42).

Small RNA Sequencing and Analysis. TruSeq Small RNA Sample Prep Kit (Illumina) was used for the preparation of the small RNA sequencing library. In brief, 10 μg of total RNA from parental and knockout cells was extracted with TRIzol reagent (Life Technologies) and size-separated on a 15% urea-polyacrylamide gel. The region of the gel containing RNA with the size from 17 to 30 nucleotides was excised and eluted for adaptor ligation at both 5′ and 3′ ends. The ligated RNAs were reverse-transcribed with SuperScript III (Life Technologies) reverse transcriptase and amplified using Phusion High-Fidelity DNA polymerase (Thermo Scientific). The final products were sequenced by MiSeq system (Illumina). FASTQ sequences produced from the sequencer were aligned to the human reference genome (GRCh38) by Bowtie2 (43). By comparing with genomic coordinates of miRNAs (obtained from mirbase.org), miRNA reads were chosen. For those miRNAs whose strand information (that is, 5p or 3p) is not annotated into their mature miRNA names, we manually examined the secondary structure of pre-miRNAs, and annotated their strand information. Only the reads that matched perfectly with miRNA sequences were selected for further analyses. To normalize miRNA reads, the read of miR-320a-3p was used for *DROSHA* and *XPO5* knockouts. In the case of *DICER* knockout, the combined reads from rRNAs and tRNAs were used for the normalization. The proportions of combined reads of rRNAs and tRNAs among whole sequencing reads were 27.6% for no. 43 clone and 27.2% for no. 45 clone, respectively, suggesting the degree of spontaneous contamination was reproducible.

To calculate the fold change of miRNA levels in Dataset S1, Figs. 4*C*, and 5*C*, the top 200 miRNAs based on their sequencing reads in parental cells were selected. For the analysis of *XPO5* or *DICER* knockout cells, the miRNA reads between two knockout libraries were averaged, respectively, and the fold change of miRNAs between parental and knockout cells was calculated. To analyze the size change of miRNAs in Figs. 4*D* and 5*B*, we filtered out the miRNAs with read numbers in the parental library less than 10, and calculated average size of the reads aligned into the same miRNA locus in each of parental or knockout library. The miRNAs with a size difference between the two knockout libraries greater than 0.5 nucleotides were discarded. After averaging the size of each miRNA between two knockout libraries, the size change between parental and knockout library were calculated. To analyze the change in proportion of trimmed or added reads at each ends of miRNAs in Fig. 4*E*, we calculated the ratios of miRNAs with the size shorter and longer than the reference sequence and designated them as 'trimmed' and 'added' ratios, respectively. The difference in the ratios between parental and each of two knockout libraries was calculated and averaged. Note that miR-7974-3p was not included in the analysis of top 200 miRNAs because it was not affected substantially in any of the knockout cell lines. Given that miR-7974-3p is not conserved beyond primates and that the hairpin structure is not stable, this RNA is unlikely to be a miRNA and need to be excluded from the miRNA database.

Northern Blot. Total RNA was separated on a 15% urea-polyacrylamide gel, and then transferred to a Hybond-NX membrane (Amersham). The membrane was cross-linked chemically with 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (44) and hybridized with a 5′ end-labeled-oligonucleotide probe that has a complementary sequence against each miRNA. The radioactive signals were analyzed using a BAS-2500 (Fujifilm).

Nuclear-Cytoplasmic Fractionation. To fractionate cell lysate into nuclear and cytoplasmic pools, we collected the cells and treated them with hypotonic buffer [Buffer A; 10 mM Hepes (pH 7.9), 10 mM KCl, 0.1 mM EDTA, 1 mM DTT]. After 25 min of incubation in ice, Nonidet P-40 was added to a final concentration of 0.25% and incubated for an additional 2 min. After spin-down, the pellet was used for the nuclear fraction whereas the supernatant was used for the cytoplasmic fraction.

AGO Immunoprecipitation. After sonication of parental and *DICER* knockout cells in a buffer solution with 150 mM KCl, 20 mM Tris·HCl at pH 8.0, and 0.2 mM EDTA, the supernatant was collected by centrifugation at full speed. After measuring the concentration of protein, 200 μg of protein extract was used for RNA extraction and designated as input. To immunoprecipitate AGO-miRNA

complexes, 2.4 mg of protein was incubated with pan-AGO antibody (2A8, a kind gift from Dr. Z. Mourelatos, University of Pennsylvania School of Medicine, Philadelphia) for 3 h, and then protein G beads were added for additional 1 h of incubation. The beads were washed four times and TRIzol was directly added to the beads for RNA extraction from AGO proteins.

ACKNOWLEDGMENTS. We thank members of the laboratories, particularly Haedong Kim, for helpful discussion and critical reading of this manuscript. This study was financially supported by IBS-R008-D1 of the Institute for Basic Science from the Ministry of Science, ICT, and Future Planning of Korea (to Y.-K.K., B.K, and V.N.K.), and by Chonnam National University, 2014 (to Y.-K.K.).

1. Ha M, Kim VN (2014) Regulation of microRNA biogenesis. *Nat Rev Mol Cell Biol* 15(8): 509–524.
2. Lee Y, et al. (2003) The nuclear RNase III Drosha initiates microRNA processing. *Nature* 425(6956):415–419.
3. Denli AM, Tops BB, Plasterk RH, Ketting RF, Hannon GJ (2004) Processing of primary microRNAs by the Microprocessor complex. *Nature* 432(7014):231–235.
4. Gregory RI, et al. (2004) The Microprocessor complex mediates the genesis of microRNAs. *Nature* 432(7014):235–240.
5. Han J, et al. (2004) The Drosha-DGCR8 complex in primary microRNA processing. *Genes Dev* 18(24):3016–3027.
6. Landthaler M, Yalcin A, Tuschl T (2004) The human DiGeorge syndrome critical region gene 8 and Its D. melanogaster homolog are required for miRNA biogenesis. *Curr Biol* 14(23):2162–2167.
7. Yi R, Qin Y, Macara IG, Cullen BR (2003) Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev* 17(24):3011–3016.
8. Bohnsack MT, Czaplinski K, Gorlich D (2004) Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. *RNA* 10(2): 185–191.
9. Lund E, Güttinger S, Calado A, Dahlberg JE, Kutay U (2004) Nuclear export of microRNA precursors. *Science* 303(5654):95–98.
10. Bernstein E, Caudy AA, Hammond SM, Hannon GJ (2001) Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* 409(6818):363–366.
11. Grishok A, et al. (2001) Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control C. elegans developmental timing. *Cell* 106(1):23–34.
12. Hutvágner G, et al. (2001) A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science* 293(5531):834–838.
13. Knight SW, Bass BL (2001) A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in Caenorhabditis elegans. *Science* 293(5538):2269–2271.
14. Mourelatos Z, et al. (2002) miRNPs: A novel class of ribonucleoproteins containing numerous microRNAs. *Genes Dev* 16(6):720–728.
15. Khvorova A, Reynolds A, Jayasena SD (2003) Functional siRNAs and miRNAs exhibit strand bias. *Cell* 115(2):209–216.
16. Schwarz DS, et al. (2003) Asymmetry in the assembly of the RNAi enzyme complex. *Cell* 115(2):199–208.
17. Berezikov E, Chung WJ, Willis J, Cuppen E, Lai EC (2007) Mammalian mirtron genes. *Mol Cell* 28(2):328–336.
18. Okamura K, Hagen JW, Duan H, Tyler DM, Lai EC (2007) The mirtron pathway generates microRNA-class regulatory RNAs in Drosophila. *Cell* 130(1):89–100.
19. Ruby JG, Jan CH, Bartel DP (2007) Intronic microRNA precursors that bypass Drosha processing. *Nature* 448(7149):83–86.
20. Ender C, et al. (2008) A human snoRNA with microRNA-like functions. *Mol Cell* 32(4): 519–528.
21. Castellano L, Stebbing J (2013) Deep sequencing of small RNAs identifies canonical and non-canonical miRNA and endogenous siRNAs in mammalian somatic tissues. *Nucleic Acids Res* 41(5):3339–3351.
22. Xie M, et al. (2013) Mammalian 5′-capped microRNA precursors that generate a single microRNA. *Cell* 155(7):1568–1580.
23. Cheloufi S, Dos Santos CO, Chong MM, Hannon GJ (2010) A dicer-independent miRNA biogenesis pathway that requires Ago catalysis. *Nature* 465(7298):584–589.
24. Cifuentes D, et al. (2010) A novel miRNA processing pathway independent of Dicer requires Argonaute2 catalytic activity. *Science* 328(5986):1694–1698.
25. Yang JS, et al. (2010) Conserved vertebrate mir-451 provides a platform for Dicer-independent, Ago2-mediated microRNA biogenesis. *Proc Natl Acad Sci USA* 107(34): 15163–15168.
26. Yoda M, et al. (2013) Poly(A)-specific ribonuclease mediates 3′-end trimming of Argonaute2-cleaved precursor microRNAs. *Cell Reports* 5(3):715–726.
27. Han J, et al. (2009) Posttranscriptional crossregulation between Drosha and DGCR8. *Cell* 136(1):75–84.
28. Martello G, et al. (2010) A MicroRNA targeting dicer for metastasis control. *Cell* 141(7):1195–1207.
29. Tokumaru S, Suzuki M, Yamada H, Nagino M, Takahashi T (2008) let-7 regulates Dicer expression and constitutes a negative feedback loop. *Carcinogenesis* 29(11): 2073–2077.
30. Consortium EP; ENCODE Project Consortium (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature* 489(7414):57–74.
31. Golden DE, Gerbasi VR, Sontheimer EJ (2008) An inside job for siRNAs. *Mol Cell* 31(3): 309–312.
32. Tan GS, et al. (2009) Expanded RNA-binding activities of mammalian Argonaute 2. *Nucleic Acids Res* 37(22):7533–7545.
33. Doyle M, et al. (2013) The double-stranded RNA binding domain of human Dicer functions as a nuclear localization signal. *RNA* 19(9):1238–1252.
34. Park JE, et al. (2011) Dicer recognizes the 5′ end of RNA for efficient and accurate processing. *Nature* 475(7355):201–205.
35. Ravi A, et al. (2012) Proliferation and tumorigenesis of a murine sarcoma cell line in the absence of DICER1. *Cancer Cell* 21(6):848–855.
36. Diederichs S, Haber DA (2007) Dual role for argonautes in microRNA processing and posttranscriptional regulation of microRNA expression. *Cell* 131(6):1097–1108.
37. Liu X, Jin DY, McManus MT, Mourelatos Z (2012) Precursor microRNA-programmed silencing complex assembly pathways in mammals. *Mol Cell* 46(4):507–517.
38. Melo SA, et al. (2010) A genetic defect in exportin-5 traps precursor microRNAs in the nucleus of cancer cells. *Cancer Cell* 18(4):303–315.
39. Heo I, et al. (2012) Mono-uridylation of pre-microRNA as a key step in the biogenesis of group II let-7 microRNAs. *Cell* 151(3):521–532.
40. Kim Y, et al. (2013) A library of TAL effector nucleases spanning the human genome. *Nat Biotechnol* 31(3):251–258.
41. Kim YK, et al. (2013) TALEN-based knockout library for human microRNAs. *Nat Struct Mol Biol* 20(12):1458–1464.
42. Kim H, et al. (2013) Magnetic separation and antibiotics selection enable enrichment of cells with ZFN/TALEN-induced mutations. *PLoS One* 8(2):e56476.
43. Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9(4):357–359.
44. Pall GS, Codony-Servat C, Byrne J, Ritchie L, Hamilton A (2007) Carbodiimide-mediated cross-linking of RNA to nylon membranes improves the detection of siRNA, miRNA and piRNA by northern blot. *Nucleic Acids Res* 35(8):e60.