Biogenesis of mammalian microRNAs by a non-canonical processing pathway

Mallory A. Havens[1], Ashley A. Reich[2], Dominik M. Duelli[3] and Michelle L. Hastings[1],*

[1]Department of Cell Biology and Anatomy, Chicago Medical School, Rosalind Franklin University of Medicine and Science, North Chicago, IL 60064, [2]Department of Biology, Lake Forest College, Lake Forest, IL 60045 and [3]Department of Cellular and Molecular Pharmacology, Chicago Medical School, Rosalind Franklin University of Medicine and Science, North Chicago, IL 60064, USA

ABSTRACT

Canonical microRNA biogenesis requires the Microprocessor components, Drosha and DGCR8, to generate precursor-miRNA, and Dicer to form mature miRNA. The Microprocessor is not required for processing of some miRNAs, including mirtrons, in which spliceosome-excised introns are direct Dicer substrates. In this study, we examine the processing of putative human mirtrons and demonstrate that although some are splicing-dependent, as expected, the predicted mirtrons, miR-1225 and miR-1228, are produced in the absence of splicing. Remarkably, knockout cell lines and knockdown experiments demonstrated that biogenesis of these splicing-independent mirtron-like miRNAs, termed 'simtrons', does not require the canonical miRNA biogenesis components, DGCR8, Dicer, Exportin-5 or Argonaute 2. However, simtron biogenesis was reduced by expression of a dominant negative form of Drosha. Simtrons are bound by Drosha and processed *in vitro* in a Drosha-dependent manner. Both simtrons and mirtrons function in silencing of target transcripts and are found in the RISC complex as demonstrated by their interaction with Argonaute proteins. These findings reveal a non-canonical miRNA biogenesis pathway that can produce functional regulatory RNAs.

INTRODUCTION

MicroRNAs (miRNAs) are small non-coding RNAs that regulate gene expression by direct base-pairing with target mRNAs (1–3). Canonical miRNAs in animals are transcribed as primary miRNAs (pri-miRNAs) and subsequently cleaved by the Microprocessor complex, comprised of the RNase III enzyme Drosha (4–6) and the double-stranded RNA (dsRNA)-binding protein, DGCR8/Pasha (4,5,7–9) to yield a pre-miRNA that is then exported to the cytoplasm by Exportin-5 (XPO5) (10–12). In the cytoplasm, pre-miRNA is processed into a ~21–23-nt mature miRNA duplex by the RNAse III enzyme, Dicer (13–17). One strand of the mature miRNA duplex is preferentially loaded into the RNA-induced silencing complex (RISC) with members of the Argonaute family of proteins, producing a functional complex for targeting mRNA via direct base pairing (18–20). The resulting miRNA/mRNA hybrids can alter protein expression of the targeted mRNA by different mechanisms, such as translational repression and mRNA degradation (21–24).

A number of non-canonical pathways for miRNA biogenesis have also been described (25–33). However, a common feature of all other pathways is the cleavage of the intermediate precursor by Dicer. One exception is the processing of miR-451, which has been shown to bypass Dicer cleavage and instead is cleaved by Argonaute-2 (Ago2) (34–37).

Mirtrons are a type of miRNA that are processed by a non-canonical miRNA pathway. Mirtrons have a pre-miRNA that is defined by the entire length of the intron in which it is located, and require pre-mRNA splicing rather than the Microprocessor for the first step in their biogenesis. The biogenesis pathway for a number of mirtrons has been characterized in *Drosophila melanogaster* and *Caenorhabditis elegans* (38,39). The pre-miRNA excised by splicing is initially in the form of an intron lariat which is subsequently linearized by the debranching enzyme, Ldbr (DBR1 in humans), allowing the intron to form a structure that is exported to the cytoplasm by XPO5 and recognized and cleaved by the Dicer complex to form a mature miRNA (38,39). In all cases, mirtronic miRNAs contain all or a portion of either the 5′ splice site, in the case of 5′ miRNAs, or the 3′splice site, if a 3′ miRNA is formed. Thus, the only known way that both the mRNA and miRNA can be generated from the same primary transcript would be for splicing to occur first and the miRNA to be generated

*To whom correspondence should be addressed. Tel: +1 847 578 8517; Fax: +1 847 578 3253; Email: michelle.hastings@rosalindfranklin.edu

from the excised intron, as has been demonstrated for mirtrons (38,39).

Mirtrons have been documented in mammals, avians and plants by deep-sequencing approaches (40–42). For humans, 13 mirtrons were predicted based on their structure, conservation, location within small introns and cloning evidence (40). Two mirtrons, miR-877 and miR-1224, were shown to be insensitive to changes in cellular DGCR8 or Drosha levels, as expected for a mirtron (25,40,43). Mammalian mirtrons, miR-877, 1226 and 1224, have been shown to be splicing-dependent, based on a GFP splicing reporter (44).

In this study, we investigated the biogenesis of predicted human mirtrons in the context of their natural flanking exons and made the unexpected discovery that, while some of the predicted mirtrons are splicing-dependent, a subset we term 'simtrons' are not processed by the canonical miRNA pathway or by the mirtron-processing pathway. Rather, simtron biogenesis occurs by a novel pathway that involves Drosha but does not require Drosha's binding partner DGCR8 or the endonuclease, Dicer. We further demonstrate that this novel class of non-canonical miRNAs is capable of gene silencing and associates with all four of the human Argonaute proteins. The identification of simtrons expands the mechanisms by which small RNAs can be generated to produce regulatory molecules.

MATERIALS AND METHODS

Primers

Primer and siRNA sequences are provided in Supplementary Table S1.

Construction of plasmids

PKD1 exons 43–46, *LRP1* exons 48–50 and *DHX30* exons 20–22 were amplified from human genomic DNA by RT–PCR using the Phire PCR kit (NEB). Primers for amplification had restriction sites incorporated on the termini. PCR products were digested with BamHI and HindIII and inserted into the similarly digested pTT3 plasmid (45) using T4 DNA ligase. *ABCF1* exons 12–15 were cloned into the pCI vector (Promega) using the restriction sites for Xho and NotI. Splicing mutations were made using the Quik Change Lightening kit (Agilent) per manufacturer's instructions.

For construction of the miRNAs expressed in an intergenic context, pcDNA-1225 and -877, the mirtonic introns were PCR amplified with Phire polymerase (NEB) using the previously constructed minigenes as templates and primers with restriction sites incorporated on the termini. PCR products were digested with BamHI and EcoR1 and inserted into the similarly digested pcDNA3.1+ plasmid (Invitrogen) using T4 DNA ligase.

For pmiRGLO-luciferase reporter plasmid construction, oligonucleotides miR-1225-5p Target 3′-UTR and miR-1225-5p Target 3′-UTR R were annealed and inserted into an XbaI- and XhoI-digested pmiRGLO plasmid (Promega) as per manufacturer's instructions. Oligonucleotides miR-877 Target 3′-UTR and miR-877 Target 3′-UTR R were annealed and inserted into a

XbaI and XhoI digested pmiRGLO plasmid (Promega) as per manufacturer's instructions.

FLAG-pcK-Drosha (pFLAG-Drosha) and FLAG-pcK-Dicer (pFLAG-Dicer) were generated by quick-change mutagenesis of the pFLAG-TN-Drosha and pFLAG-TN-Dicer (a generous gift of V.N. Kim) plasmids, respectively, using the Quik Change Lightening kit (Agilent) per manufacturer's instructions.

All plasmid constructions were confirmed by sequence analysis.

Cell culture and transfection

HEK-293T, HeLa, NIH/3T3 and mouse embryonic fibroblast *Ago2* knockout cells ($Ago2^{-/-}$) (46) (a generous gift from G.J. Hannon) were cultured in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum. Optifect (Invitrogen) or Lipofectamine 2000 (Invitrogen) was used for transfection of HEK-293T with plasmids as per manufacturer's instructions. Lipofectamine 2000 (Invitrogen) was used for transfection of HeLa, NIH/3T3 and $Ago2^{-/-}$ cells. RNA was isolated after 48 h.

For *DGCR8* RNAi experiments, HeLa cells were transfected with siRNA using Lipofectamine 2000 (Invitrogen) as per the manufacturer's instructions. Cells were transfected with 32 nM siRNA on the first day and again after 48 and 96 h. RNA was collected 72 h after the final siRNA treatment.

For TN-Drosha experiments HEK-293T cells were transfected with 3 μg of pFLAG-TN-Drosha (a generous gift from V.N. Kim) (47) with or without 3 μg of minigene plasmid or mock transfection control using Lipofectamine 2000 (Invitrogen). Cells were split 1:2, 24-h post-transfection and RNA was collected 24 h later.

For *XPO5* RNAi experiments, HeLa cells were transfected with 40 nM siRNA using Lipofectamine 2000 (Invitrogen). After 24 h, cells were transfected again with 40 nM siRNA and 3 μg of minigene plasmid. After 48 h, RNA was collected.

Mouse embryonic stem cell lines, Control (*Dicer* conditional knockout) and *Dicer* knockout ($Dicer^{-/-}$) cells (from G.J. Hannon) (48) and *DGCR8* knockout ($DGCR8^{-/-}$) cells (Novus Biologicals) were grown on a gelatin layer in Knockout Dulbecco's modified Eagle's medium (Gibco) supplemented with 15% ES cell FBS (Gibco), 1% non-essential amino acids, 1% L-glutamine, 1% penicillin/streptomycin/Amphotericin B, 0.1% ESGRO-LIF and 0.008% beta-mercaptoethanol. Cells were transfected with 3 μg of minigene plasmid using Lipofectamine 2000 (Invitrogen) as per the manufacturer's instructions and RNA was collected after 48 h.

RNA was collected for all experiments using Trizol reagent (Invitrogen).

RT–PCR

Reverse transcription was performed using a High Capacity Reverse Transcription kit (Applied Biosystems), Superscript III (Invitrogen) or GoScript (Promega) as per manufacturer's instructions. The same kits were used to reverse transcribe miRNAs with gene-specific stemloop

primers as previously described (49) or linker-specific primers. PCR was performed with GoTaq (Promega) and ^{32}P-dCTP. Various PCR cycle numbers were tested to insure amplification in the linear range. Endogenous simtrons and mirtrons were analysed following 35 PCR cycles and minigene-derived miRNAs and snoRNA65 were analysed following 25 PCR cycles. miR-16 was analysed after 20 cycles. All RNAs from immunoprecipitation experiments were subjected to 35 PCR cycles. Amplification products from stemloop PCR of miRNAs were separated on 12% native or 15% denaturing PAGE gels and mRNA products on 5% native PAGE gels. Products were quantitated using a Typhoon Phosphorimager (GE Healthcare).

Western blotting

Proteins were separated by sodium dodecyl sulphate (SDS)–PAGE and transferred to Immobilon-FL membrane (Millipore). Blots were probed with rabbit polyclonal antibodies specific to DGCR8 (Novus Biologicals) and mouse monoclonal antibodies specific to β-actin (Sigma) followed by horseradish peroxidase (HRP)-conjugated goat anti-mouse or anti-rabbit secondary antibody. Detection was preformed with Luminata Forte Western HRP Substrate (Millipore). Quantitation was preformed with NIH-ImageJ software.

Immunoprecipitation

HEK-293T cells were transfected with 3 μg of minigene plasmid and 18 μg of pFLAG-GFP, Argonaute-1, (50) 2, 3, 4 (51) (Addgene), -Dicer or -Drosha. Cells were harvested after 48 h and immunoprecipitation was preformed using 20 μl M2-FLAG beads (Sigma) in RSB-250. Portions of the input (In), unbound (Un) and immunoprecipitated (IP) fractions were proteinase K (Promega) treated and RNA was isolated by phenol/chloroform extraction followed by ethanol precipitation. Isolated, immunoprecipitated RNA was analysed using radiolabelled stemloop RT–PCR as described earlier.

In vitro transcription of RNAs

T7-DNA templates were generated via RT–PCR. RNA substrates were transcribed *in vitro* by T7 RNA polymerase (Promega) in the presence of α-^{32}P UTP with 5× transcription buffer (Promega), 100 mM DTT, 2.5 mM A, C and G and 0.1 mM U, RNase inhibitor (Promega) and T7 RNA polymerase (Promega). The reaction was incubated for 1 h at 37°C then treated with RQ1 DNase (Promega) for 30 min at 37°C. Reaction products were separated on a 5% denaturing PAGE gel, excised, eluted, ethanol precipitated and resuspended in water prior to use.

In vitro processing of RNA

FLAG-Drosha and FLAG-DGCR8 (AddGene), FLAG-Drosha alone, FLAG-TN-Drosha or FLAG-GFP were overexpressed in HEK-293T cells and FLAG-tagged proteins were isolated on M2-FLAG beads (Sigma). Briefly, 6 μg of plasmid was transfected into cells using Lipofectamine 2000 (Invitrogen) as per manufacturer's instructions. Approximately 48 h after transfection, cells were washed with phosphate buffered saline, harvested and lysed via sonication in lysis buffer (20 mM Tris–HCl pH 8.0, 100 mM KCl and 0.2 mM EDTA). Lysate was incubated with M2-FLAG beads (Sigma) for 1 h at 4°C. Beads were washed five times with lysis buffer and 15 μl of beads/buffer were combined with 100 000 cpm RNA, 10× reaction buffer (64 mM MgCl$_2$) and RNase inhibitor (Promega) for 90 min at 37°C. Alternatively, cell lysates were directly incubated (referred to as whole-cell extracts, WCE) with *in vitro* transcribed RNA. Products were phenol/chloroform extracted and ethanol precipitated and analysed on 8% denaturing PAGE gels.

Luciferase assay

HEK-293T cells were transfected with 0.4 or 1 μg of minigene plasmid and 10 ng of pmiRGLO (Promega) using Optifect (Invitrogen). After 24 h, cells were split in triplicate into 96-well black bottom plates. After 24 h, the media was changed to Dulbecco's modified Eagle's medium High Glucose without phenol red and luciferase and renilla expression from the pmiRGLO plasmid was quantitated using the Dual-Glo Luciferase assay system (Promega) and a Synergy HT luminometer (BioTek) as per the manufacturer's instructions. For each experiment, all wells were analysed in triplicate. Values greater than ±1 SD within the triplicate measurements were not considered. For the mismatch controls, miR-877 was transfected with the luciferase reporter plasmid for miR-1225 and miR-1225 was transfected with the luciferase reporter plasmid for miR-877. The seed sequence for miR-877 and miR-1225 are not matched and thus serve effectively as mismatch controls.

Statistics

The Student's *t*-test was used to analyse all experimental results with the exception that the Wilcoxon matched pairs signed-rank tests was used for sample sets that did not have Gaussian distributions. Data were considered significant when $P \leq 0.05$.

RESULTS

Identification of simtrons, splicing-independent mirtron-like miRNAs

Mammalian mirtrons have been previously predicted by computational and sequencing approaches (40). In order to be definitively classified as a mirtron, miRNA production must require splicing. To determine whether predicted mammalian mirtrons are splicing-dependent in the context of their natural host gene exons, we constructed minigene expression plasmids comprised of the mirtron and flanking exons from *ABCF1* (miR-877), *DHX30* (miR-1226), *PKD1* (miR-1225) and *LRP1* (miR-1228) (Figure 1A). In addition, a version of each minigene was made in which single nucleotide changes were introduced into the 5′ splice site and the 3′ splice site to inactivate splicing of the mirtronic intron (Δss, Figure 1A). Compensatory mutations at the 5′ and 3′ splice sites
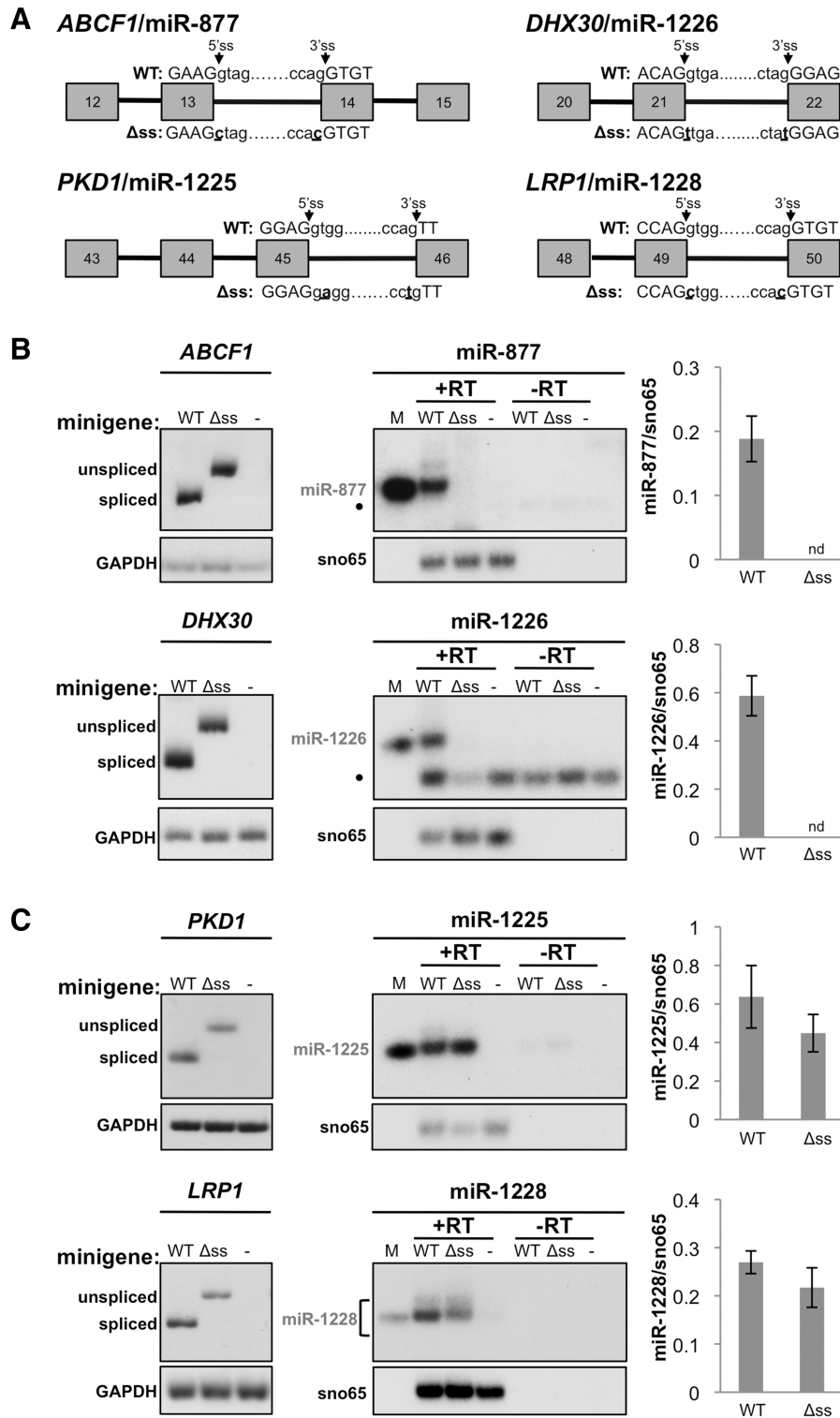
Figure 1. Identification of splicing-independent mirtron-like miRNAs (simtrons). (**A**) Minigene structures. Boxes and lines indicate exons and introns, respectively. Splice site mutations are in bold and underlined. Arrows refer to the 5′ and 3′ splice sites (ss). (**B**) Splicing-dependent mirtrons. Left panels: splicing analysis of host gene transcripts. wt, splicing-deficient (Δss) minigene transcripts or empty vector control (−) were transfected into HEK-293T cells and mRNA splicing was analysed by radiolabelled RT–PCR. *GAPDH* was used as a loading control. Unspliced indicates the mRNA product retaining only the mirtronic intron, all other introns were excised. Middle panels: miRNA expression in minigene-transfected HEK-293T cells was analysed by radiolabelled, stemloop RT–PCR with snoRNA65 (sno65) used as a loading control. +RT indicates RNA that was reverse transcribed with stemloop primers and −RT indicates RNA that was not reverse transcribed but had stemloop primers added as a control. M indicates a synthetic size marker. Filled circle indicates a non-specific primer dimer. Right panels: Quantitation of miRNA expression normalized to sno65. nd indicates that the miRNA was not detected. Bars represent the average values ± SEM; *ABCF1*/miR-877 *n* = 12, *DHX30*/miR-1226 *n* = 11. (**C**) Splicing-independent mirtron-like miRNAs analysed as in B. *PKD1*/miR-1225, *n* = 8; *LRP1*/miR-1228, *n* = 9.

were designed to maintain the predicted secondary structure of the stem of each mirtron.

These minigene reporters were transiently transfected into HEK-293T cells, RNA was isolated and mRNA splicing of minigene transcripts and miRNA production from the minigene-derived RNA was analysed. Given that stemloop RT–PCR (49) is the basis for detection in pre-designed Taqman miRNA assays (Applied Biosystems), we used a modified version of this method, radiolabelled stemloop RT–PCR, for analysis of miRNAs generated from the minigene transcripts. This method was also chosen because of its detection sensitivity and track-record of specificity (52,53). The specificity of the stemloop primers for our substrates was confirmed through a number of control experiments. First, we demonstrate that DNA is not being amplified, as reverse transcription is required for product detection (Figure 1B and C). Second, the stemloops are specific for amplifying the targeted terminal RNA sequence, as a synthetic pre-mRNA comprised of the intron and flanking exons is not detected in the stemloop RT–PCR reaction (Supplementary Figure S1A and B), whereas a synthetic RNA substrate with the 3′ termini corresponding to the stemloop is amplified (Supplementary Figure S1A and C). These controls confirm the effectiveness and specificity of the radiolabelled stemloop RT–PCR methodology.

Each of the wild-type (wt) minigene-derived RNA transcripts gave rise to mature miRNA, as well as spliced mRNA (Figure 1B and C). miRNA species from both the 5′ and 3′ arm of these miRNAs have been identified in small RNA libraries from deep sequencing (54,55). Sequencing of the products confirmed that the miRNAs correspond to those previously sequenced from mammals and predicted to be mirtrons (40,54,55) (Supplementary Figure S2). The minigene transcripts with splice site mutations (Δss) were not spliced and yielded RNA transcripts that retained the mirtronic intron (Figure 1B and C). As expected for mirtrons, the *ABCF1*Δss and *DHX30*Δss minigene transcripts did not produce miR-877 or miR-1226, respectively (Figure 1B). This result demonstrates the existence of splicing-dependent mirtrons in mammalian cells in the context of their host gene.

Unexpectedly, *PKD1*Δss and *LRP1*Δss minigene transcripts, though not spliced, produced miR-1225-5p (referred to as miR-1225) and miR-1228, respectively, at similar amounts as their wt counterparts (Figure 1C). These results were confirmed using Taqman qPCR (Supplementary Figure S3), demonstrating that miR-1225 and miR-1228 are produced in the absence of splicing. Two species of mature miR-1228 were produced from both the wt and Δss minigenes one of which had two additional 3′ nt as confirmed by sequencing (Supplementary Figure S2), indicating 3′ heterogeneity. Stemloops specific for all previously reported 3′-ends were included in the reaction in order to detect the different 3′-ends of the miRNA. This extended 3′-end has been reported previously for miR-1228 (54,55). These results demonstrate that miR-1225 and miR-1228 are not strictly splicing-dependent mirtrons, as they can be generated in the absence of splicing. We propose to call this new type of miRNAs splicing independent mirtron-like miRNAs or 'simtrons'.

Simtron biogenesis involves Drosha but does not require DGCR8

The Microprocessor, composed of Drosha and DGCR8 is required for the first cleavage step to generate canonical pre-miRNAs, but is not required to generate mirtrons, which are excised from the primary RNA transcript during pre-mRNA splicing. To test whether biogenesis of endogenous miR-1225 and miR-1228, requires the Microprocessor, we reduced *DGCR8* expression in HeLa cells using *DGCR8*-specific siRNAs, and examined endogenous miRNA levels. *DGCR8* mRNA was reduced by 63% (Figure 2A) and protein by 46% (Figure 2B), and resulted in a 40% reduction in miR-16 ($P = 0.0313$) (Figure 2C), a canonical, DGCR8-dependent miRNA (8,56,57). This result confirms that DGCR8 was effectively reduced in the cells.

The abundance of the splicing-dependent mirtrons, miR-1226 and miR-877, did not change significantly following *DGCR8*-knockdown (Figure 2C), confirming that miR-1226 and miR-877 are bona fide mirtrons. The splicing-independent simtrons, miR-1225 and miR-1228, were also not significantly affected by *DGCR8*-knockdown (Figure 2C), suggesting that endogenous miR-1225 and miR-1228 are not dependent on the canonical miRNA biogenesis pathway.

We next tested whether simtron processing requires Drosha, the endonuclease component of the Microprocessor. Knockdown of *Drosha* using siRNAs reduced *Drosha* by >50% but did not significantly reduce miR-16 production, suggesting that the enzyme was not sufficiently reduced to affect overall miRNA production. Efforts to further reduce *Drosha* have been unsuccessful largely due to cellular toxicity. As an alternative approach to test the requirement of Drosha in simtron processing, we expressed a transdominant negative Drosha (TN-Drosha) (47) in HEK-293T cells (Figure 2D–G). The endonuclease domains in TN-Drosha contain the mutations E1045Q and E1222Q, which prevent nuclease activity at the 3′-end and 5′-end of the pre-miRNA hairpin, respectively, while Drosha binding to the pre-miRNA is not affected (4). Endogenous miR-16 was reduced by 38% ($P = 0.0213$) whereas endogenous miR-877, miR-1226, miR-1225 and miR-1228 were not affected (Figure 2E). These results suggest that endogenous miR-1225 and miR-1228 are not dependent on Drosha for processing.

It is possible that miR-1225 and miR-1228 production is only dependent on Drosha and DGCR8 when pre-miR processing cannot occur via splicing. To test this idea, we expressed TN-Drosha in HEK-293T cells co-transfected with wt and Δss minigenes. miR-1225 and miR-1228 derived from the wt minigenes were not affected (miR-1225) or increased (miR-1228) by 27% ($P = 0.0239$) when expressed with the dominant negative form of Drosha (Figure 2F and G). In contrast miR-1225 and miR-1228 derived from the Δss minigenes decreased by 40% ($P = 0.0456$) and 24% ($P = 0.0499$), respectively,
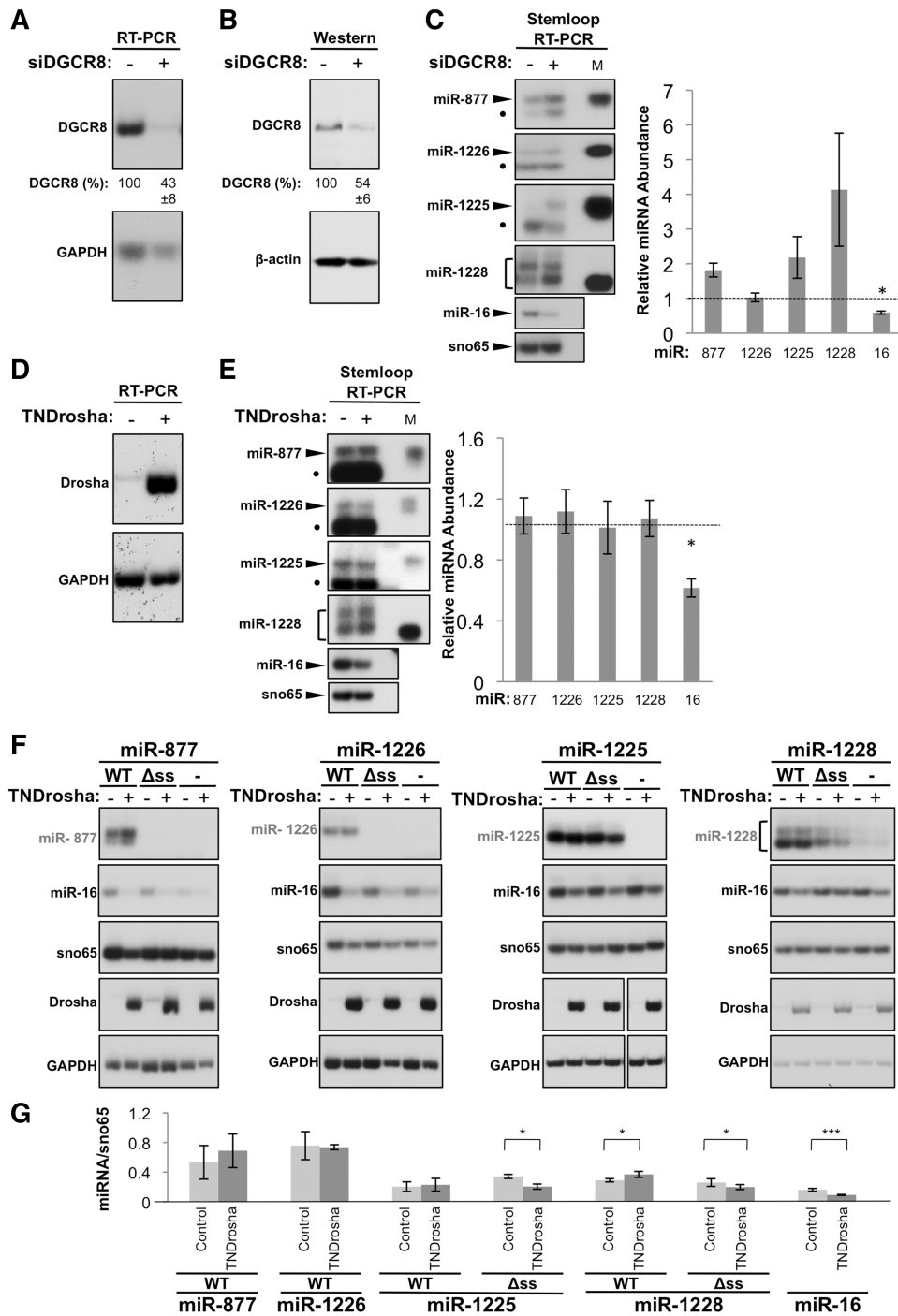
Figure 2. Simtron biogenesis involves Drosha but not DGCR8. Knockdown of DGCR8 in HeLa cells using siRNA was quantitated by (**A**) RT–PCR analysis of *DGCR8* mRNA and (**B**) western blot analysis of DGCR8 protein expression. The percentage of knockdown of DGCR8 was quantitated for *DGCR8* mRNA using the equation $100 - [((DGCR8_{knockdown}/GAPDH)/(DGCR8_{control}/GAPDH)) \times 100]$, $n = 5$ and for DGCR8 protein using the equation $100 - [((DGCR8_{knockdown}/\beta\text{-}actin)/(DGCR8_{control}/\beta\text{-}actin)) \times 100]$. (**C**) Changes in endogenous miRNA levels following DGCR8 knockdown were analysed by stemloop RT–PCR analysis. miR-16 is a canonical miRNA control and sno65 is a loading control. Graph shows quantitation of miRNA abundance using the equation: $(miRNA_{experimental\ condition}/sno65)/(miRNA_{control}/sno65)$. $n = 4$ for all miRNAs except for miR-16, $n = 5$; asterisk indicates $P \le 0.05$ (Wilcoxon matched pairs signed-rank test). M indicates a synthetic size marker and filled circle indicates a non-specific primer dimer. (**D**) RT–PCR analysis of *Drosha* mRNA following expression of TN-Drosha in HEK-293T cells. (**E**) The effect of TN-Drosha expression on endogenous miRNA abundance was analysed by stemloop RT–PCR. Graph shows quantitation of miRNA abundance using the same equation as in C, $n = 6$; asterisk indicates $P \le 0.05$ (Student's *t*-test). (**F**) Stemloop RT–PCR analysis of minigene-derived miR-877, 1226, 1225, 1228 and endogenous miR-16 isolated from HEK-293T cells transiently transfected with TN-Drosha. sno65 was used as a control. *TN-Drosha* mRNA expression in HEK-293T cells was analysed by radiolabelled RT–PCR. *GAPDH* was used as a control. (**G**) Quantitation of miRNA abundance relative to sno65 using the equation: miRNA/sno65. $n = 3$ for miR-877, 1226 and 1225, $n = 5$ for miR-1228 and $n = 14$ for miR-16; *$P \le 0.05$, ***$P \le 0.0001$ (). Data sets were analysed using the Student's *t*-test with the exception of miR-16, which was analysed using the Wilcoxon matched pairs signed-rank test. In all panels, bars represent the average $\pm$ SEM. The horizontal dotted line indicates normalized control levels.

following pTN-Drosha treatment (Figure 2F and G). This decrease is comparable to the 40% decrease ($P = 0.0001$) of the canonical miRNA miR-16. These results suggest that simtrons are sensitive to Drosha activity only when splicing is inactive, perhaps indicating that simtrons can be excised by either the mirtron processing pathway or by a pathway involving Drosha.

To further test whether the canonical miRNA pathway is required for miR-1225 and miR-1228 processing when splicing is inactive, we analysed miR-1225 and miR-1228 processing in a *DGCR8* knockout mouse embryonic stem cell line (DGCR8$^{-/-}$) (58). wt and Δss minigenes for *PKD1* (miR-1225) and *LRP1* (miR-1228) were transfected into the DGCR8$^{-/-}$ cells and RNA was subsequently collected and analysed for miRNA production. Endogenous, mature miR-16 was not detected in these cells, as expected for a canonical miRNA (Figure 3A). In contrast, mature miR-1225 and miR-1228 abundance did not change significantly in DGCR8$^{-/-}$ cells compared to control cells (Figure 3A). These results indicate that DGCR8 is not required for miR-1225 or miR-1228 processing. Taken together, our results demonstrate that miR-1225 and miR-1228 are neither mirtrons (dependent on splicing for biogenesis) nor canonical miRNAs (dependent on Drosha and DGCR8). Rather, it appears that splicing may be one, but not the only mechanism for production of these miRNAs and that simtronic miRNAs can be generated by an alternative processing pathway that does not require splicing or DGCR8.

Simtron biogenesis does not require Dicer, Ago2 or XPO5

To further elucidate the requirements for simtron biogenesis, we tested the requirement of other factors involved in the maturation process of miRNAs. In the canonical miRNA biogenesis pathway Dicer, or in one case, Ago2, processes pre-miRNA into mature miRNA. Biogenesis of mirtrons in *D. melanogaster* and *C. elegans* has also been shown to require Dicer and XPO5 (12,38,39). To determine whether the simtrons, miR-1225 and miR-1228, require XPO5, Dicer or Ago2, we transiently transfected wt and Δss minigenes into *Dicer* knockout mouse embryonic stem cells (Dicer$^{-/-}$), *Ago2* knockout mouse embryonic fibroblasts (Ago2$^{-/-}$) or control cells (NIH-3T3 or Dicer conditional mouse embryonic stem cells). miR-16 was not detected in the Dicer$^{-/-}$ cells and was reduced dramatically in Ago2$^{-/-}$ cell (Figure 3A and C) as expected for a canonical miRNA. In contrast, there was no significant change in miR-1225 or miR-1228 production from either the wt or Δss minigenes in either knockout cell line relative to control cells, though miR-1225 increased slightly from both wt and Δss minigenes in the Ago2$^{-/-}$ cells and the Dicer$^{-/-}$ cells (Figure 3A and C). The production of both miR-1225 and miR-1228 in the absence of Dicer and Ago2 indicates that maturation of these simtrons does not require these enzymes for processing.

The processing of miR-1225 in the absence of Dicer was unexpected, as most miRNAs require Dicer cleavage to produce a mature miRNA. To further test the Dicer requirement, pFLAG-Dicer constructs were expressed in HEK-293T cells. FLAG-Dicer was immunoprecipitated from the cell lysates and co-immunoprecipitated miRNAs were analysed. Consistent with our results in the Dicer$^{-/-}$ cells, miR-1225 and miR-1228 were not detected above background levels in immunoprecipitates from cells co-transfected with the miR-1225 or miR-1228 Δss minigenes (Figure 3B). In contrast, miR-1225, miR-1228 and miR-877 from the wt minigenes were efficiently co-immunoprecipitated with FLAG-Dicer (Figure 3B). These results suggest that miR-1225, miR-1228 and miR-877 generated from the mirtronic processing pathway interact with Dicer, and the simtrons generated by a splicing-independent pathway, do not interact with Dicer.

In the canonical miRNA biogenesis pathway XPO5 exports pre-miRNA to the cytoplasm. To test the requirement for XPO5 in the simtron biogenesis pathway, HeLa cells were transiently transfected with wt or Δss minigenes and *XPO5* siRNA. *XPO5* expression was reduced by 87% relative to cells treated with a control siRNA and resulted in a 73% reduction in the mirtron miR-877 and a 415% increase in pre-miR-877 (Figure 3D), suggesting that in the absence of XPO5, pre-miRNA is trapped in the nucleus, unable to enter the cytoplasm for processing to a mature miRNA. In contrast, the simtrons, miR-1225 and miR-1228, were not affected by *XPO5* knockdown (Figure 3D), suggesting that simtrons do not require XPO5 for production of a mature miRNA.

Processing of simtrons by Drosha

Experiments with a dominant negative form of Drosha (Figure 2F and G), suggest that Drosha is involved in simtron biogenesis. Because Drosha was the only canonical miRNA processing factor analysed that had any influence on simtron biogenesis, we further tested the role of Drosha in the cleavage of the simtron pre-miRNA from the primary transcript. First, interactions between simtrons and Drosha were examined. For this, FLAG-Drosha was co-expressed with 1225 wt or Δss minigenes or 877 wt in HEK-293T cells. FLAG-Drosha was immunoprecipitated and the co-immunoprecipitated pre-miRNAs were analysed. Pre-miR-1225 derived from both wt and Δss transcripts immunoprecipitated with FLAG-Drosha, whereas pre-miR-1225 was not immunoprecipitated from lysates of mock transfected cells (Figure 4A). miR-877 did not immunoprecipitate with FLAG-Drosha (Figure 4A) further verifying it as a mirtron. These results demonstrate that Drosha interacts with the simtron miR-1225.

Next, an *in vitro* Drosha processing assay was used to test more directly whether Drosha promotes pre-miRNA processing from a primary RNA transcript. Plasmids expressing FLAG-Drosha, FLAG-DGCR8 or FLAG-TN-Drosha were expressed in HEK-293T cells and immunoprecipitated with anti-FLAG M2 beads. Immunoprecipitated FLAG-tagged proteins were incubated with a radiolabelled, *in vitro* transcribed and purified pri-miRNA consisting of the miRNA containing intron and flanking exons, or the flanking ~120 nt in the case of miR-16. The purified Drosha/DGCR8 complex
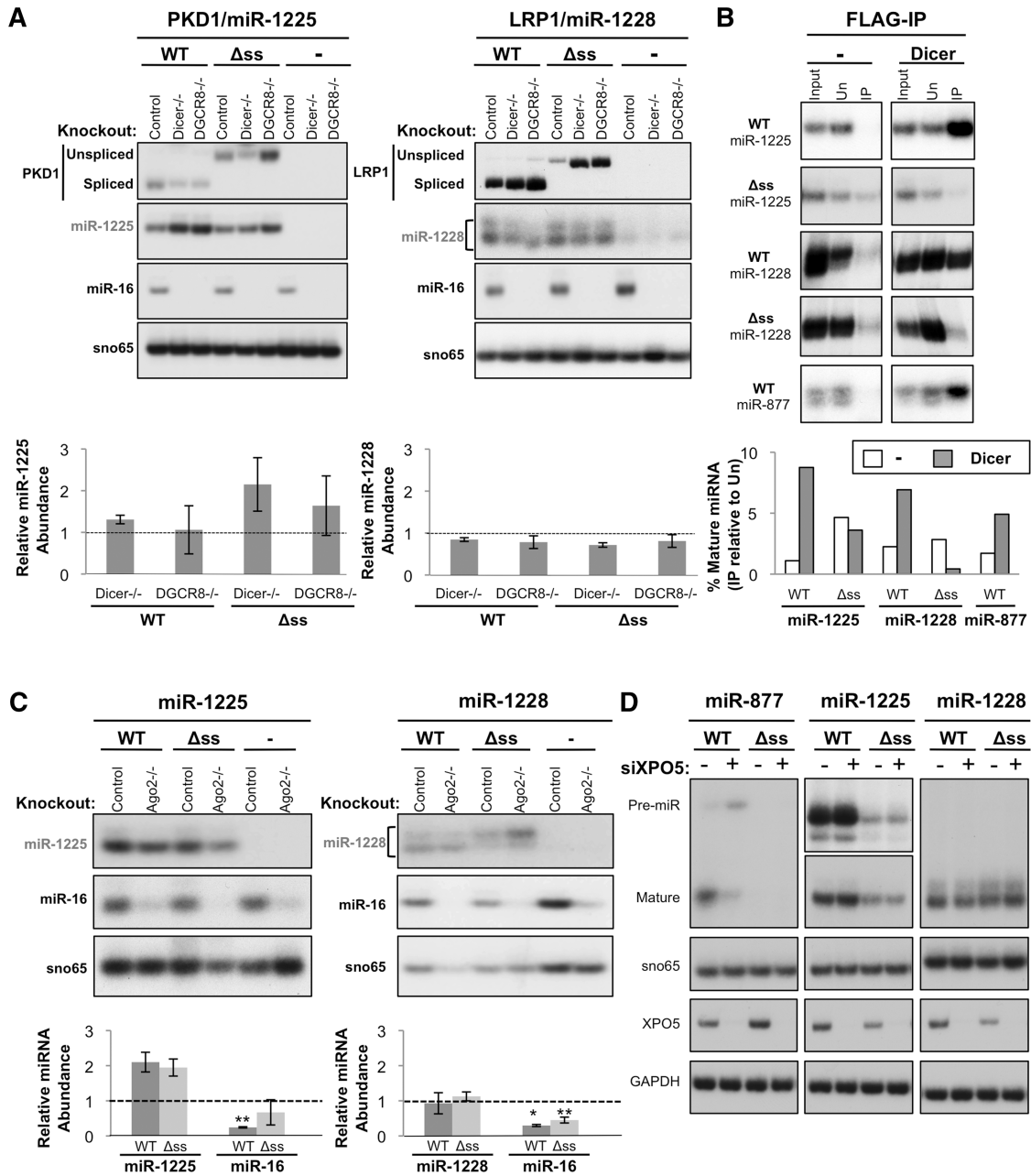
Figure 3. Simtron biogenesis does not require DGCR8, Dicer, Ago2 or XPO5. (**A**) RT–PCR analysis of minigene-derived host gene mRNA and stemloop RT–PCR analysis of minigene-derived miRNA and endogenous miR-16 in *Dicer* and *DGCR8* knockout mouse embryonic stem cells transfected with the wt or splicing-deficient minigene (Δss) or empty vector control (−). sno65 was used as a loading control. Graphs show quantitation of miRNA using the equation: (miRNA$_{experimental\ condition}$/sno65)/(miRNA$_{control}$/sno65). Bars represent the average ± SEM, $n = 3$. The horizontal dotted lines indicate normalized control levels. (**B**) Stemloop RT–PCR analysis of miR-1225 and miR-1228 immunoprecipitated from HEK-293T cell lysates that were transiently transfected with wt or Δss minigenes, or miR-877 from wt minigene along with pFLAG-Dicer (Dicer) or without (−) and immunoprecipitated with an antibody against the FLAG epitope. Input refers to cell lysates before FLAG immunoprecipitation; Un is the unbound fraction and IP is the immunoprecipitated fraction. Un is 1/20 IP and Input is 1/5 IP. The graph represents the percent of the mature miRNA found in the IP fraction versus the amount that remained in the Un fraction using the equation: (IP/(IP + (Un × 20)) × 100). (**C**) Stemloop RT–PCR analysis of minigene-derived miR-1225, miR-1228 and endogenous miR-16 from *Ago2* knockout mouse embryonic fibroblasts. sno65 was used as a loading control. Cells were transiently transfected with wt or Δss minigenes or empty vector control (−). Graph shows quantitation of miRNA abundance using the same equation as in A. Bars represent the average ± SEM, $n = 3$ and *$P \le 0.05$ or **$P \le 0.01$, Student's *t*-test. The horizontal dotted lines indicate normalized control levels. (**D**) Stemloop RT–PCR and RT–PCR analysis of miR-877 (left panel), miR-1225 (middle panel) and miR-1228 (right panel) minigene-expression in HeLa cells following siRNA-directed knockdown of *XPO5*. sno65 is a loading control for miRNA using stemloop RT–PCR and *GAPDH* is a loading control for RT–PCR of *XPO5* mRNA.
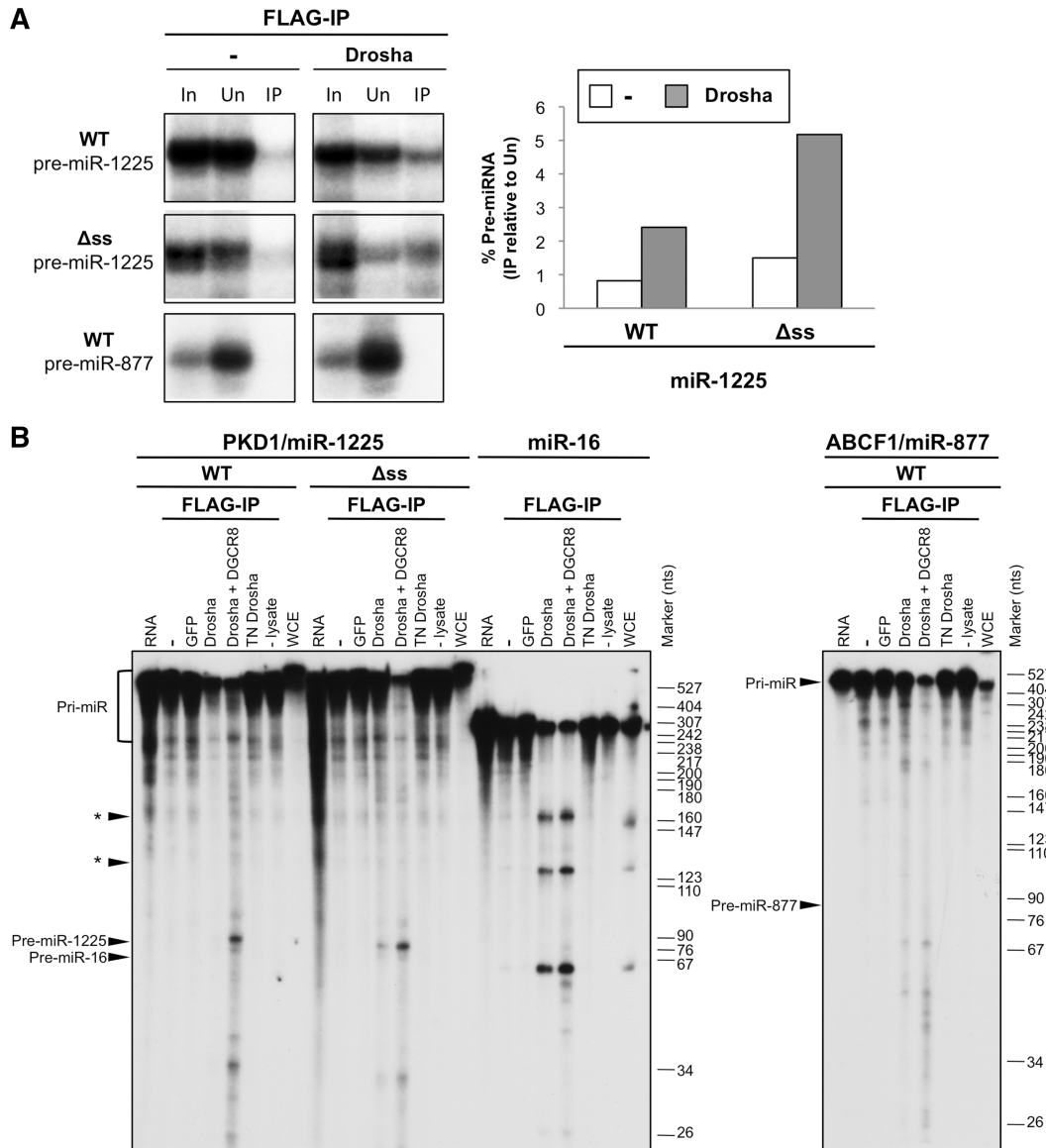
Figure 4. Immunoprecipitation and *in vitro* processing of simtrons with Drosha. (**A**) Pre-miR-1225 co-immunoprecipitates with Drosha. Pre-miR-1225 derived from wt and Δss minigenes and pre-miR-877 from wt minigene were transiently transfected into HEK-293T cells with pFLAG-Drosha (Drosha) or without (−), and immunoprecipitated with an antibody against the FLAG epitope. Isolated pre-miRNAs were analysed by radiolabelled stemloop RT–PCR and products were separated by 12% native PAGE. Input (In) refers to cell lysates before FLAG immunoprecipitation; Un is the unbound fraction and IP is the immunoprecipitated fraction. Un is 1/20 IP and Input is 1/5 IP. The graph represents the percent of the pre-miRNA found in the IP fraction versus the amount that remained in the Un fraction using the equation: (IP/ (IP + (Un × 20)) × 100). (**B**) Drosha-dependent *in vitro* simtron processing. Radiolabelled RNA transcribed from a *PKD1* wt or Δss, *ABCF1* wt or pri-miR-16-1 DNA template was incubated with the FLAG-immunoprecipitates from HEK-293T cells, or with HEK-293T WCEs from cells that were not transfected. FLAG-immunoprecipitates were derived from cells transfected with mock transfection (−), pFLAG-GFP (GFP), pFLAG-Drosha (Drosha), pFLAG-Drosha and pFLAG-DGCR8 (Drosha + DGCR8), pFLAG-TN-Drosha (TN Drosha), or FLAG-M2-beads that were incubated with lysis buffer but no cell lysate (−lysate). Template RNA was included as a control (RNA). Reaction products were separated by 8% denaturing PAGE. The sizes of pre-miRNAs are indicated. Asterisk indicates uncharacterized miR-16 cleavage fragments (59).

efficiently processed the canonical pri-miR-16 (~60 nt pre-miRNA) substrate as well as the wt and Δss versions of the simtron, miR-1225 (~90 nt pre-miRNA), but not the mirtron, miR-877 as expected (~86 nt pre-miRNA) (Figure 4B). wt Pre-miR-1225 (90 nt) was produced when *Drosha* was overexpressed with its binding partner *DGCR8*, whereas Δss pre-miRNA-1225

was generated from incubation with Drosha immunoprecipitates alone or in combination with DGCR8. Simtron miR-1225, miR-16 and miR-877 were not processed by immunoprecipitates from mock transfected cell lysates (−), FLAG-GFP or FLAG-TN-Drosha lysates. These *in vitro* data support our *in vivo* data that simtrons, but not mirtrons, can be processed by Drosha.

Simtron processing is not dependent on intronic location

Although processing of simtrons does not require splicing, simtron processing may depend on its location within the intron, perhaps involving the recruitment of splicing factors via *cis*-acting sequences within the flanking exons. To test this idea, we removed the intron from the host gene and inserted it into the pcDNA3.1+ plasmid under the control of a CMV promoter, effectively changing the intronic location to an intergenic one (Figure 5A). We transiently transfected the plasmids into control, DGCR8$^{-/-}$ or Dicer$^{-/-}$ cells and analysed RNA abundance. As expected, miR-877, which is splicing-dependent, was not produced in the absence of the flanking exons in either control, DGCR8$^{-/-}$ or Dicer$^{-/-}$ cells (Figure 5B). However, intergenic miR-1225 derived from wt and Δss minigenes were produced in the absence of their flanking exons and in the absence of Dicer and DGCR8 (Figure 5B and C), indicating that simtron bio-genesis is not context-dependent and that all the sequence requirements necessary for processing of the miRNAs are contained within the intronic sequence. Similar results were obtained for miR-1228 (Supplementary Figure S4A and B).

Simtrons function in gene silencing and bind Argonaute proteins

To test whether small RNAs generated from mammalian simtrons and mirtrons are functional miRNAs, we assessed the gene-silencing capabilities of miR-1225, miR-1228 and miR-877 in HEK-293T cells using a luciferase reporter assay. Two exact complement target sequences for miR-1225, miR-1228 or miR-877 were cloned into the 3′-UTR of the luciferase gene of the pmiRGLO plasmid, which also has a renilla reporter for signal normalization. The wt or Δss minigenes were transiently transfected with a match or mismatched luciferase reporter plasmid. The wt miR-877 minigene reduced luciferase activity of its pmiRGlo-derived target by up to 30% ± 6 as compared to mismatch control (Figure 6). The *ABCF1*Δss minigene does not produce miRNA, and, as expected, did not reduce luciferase activity (Figure 6). These results demonstrate that human mirtrons are functional in targeted gene silencing.

The simtron, miR-1225, produced from either the wt or Δss plasmid, induced a dose-responsive reduction in luciferase activity by up to 52% ± 11 and 73% ± 5, respectively, compared to the mismatch control (Figure 6). Similar to miR-1225, miR-1228 also induced a dose-responsive reduction in luciferase activity by 23% ± 8 and 29% ± 11, respectively, compared to a mismatch control (Supplementary Figure S4C). These results demonstrate that miR-1225 and miR-1228 are capable of silencing a target transcript in a dose-dependent manner and thus indicate that miRNAs derived from the simtron processing pathway can function as miRNAs.

Canonical miRNAs are loaded into the RISC complex with Ago proteins to produce a functional complex for silencing. To test whether simtrons associate with the RISC complex, we performed immunoprecipitation assays on all four human Ago proteins. We found that
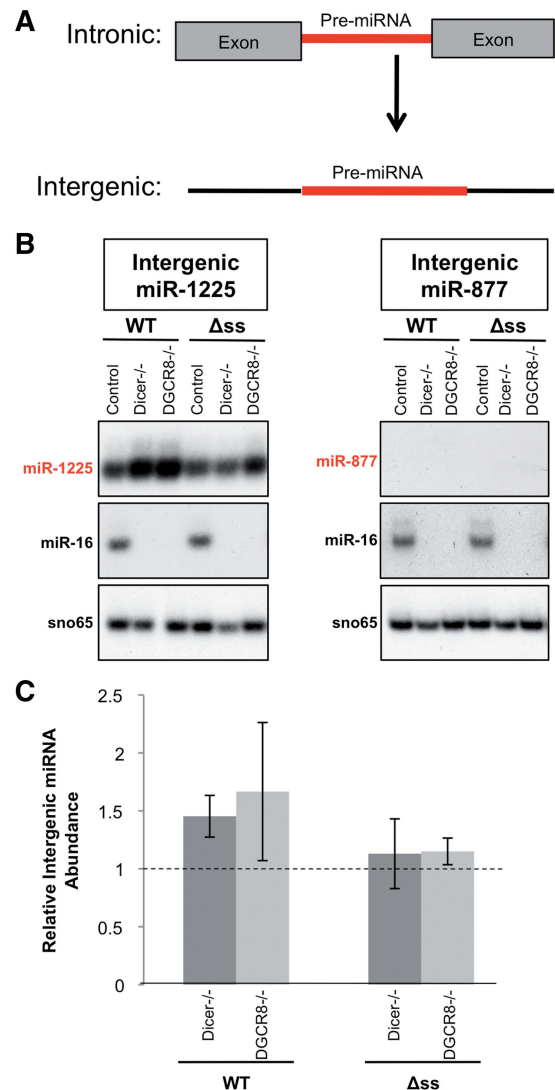

Figure 5. Simtron processing is context independent. (**A**) Diagram comparing intronic and intergenic pre-miRNA expression. (**B**) Control, *Dicer* (Dicer$^{-/-}$) or *DGCR8* (DGCR8$^{-/-}$) knockout mouse embryonic stem cells were transiently transfected with the intergenic wt minigene, or intergenic splicing-deficient minigene (Δss). Minigene-derived miRNAs and endogenous miR-16 were analysed by stemloop RT–PCR. Left panel: simtron miR-1225. Right panel: mirtron miR-877. sno65 was analysed as a loading control. (**C**) Graph shows quantitation of miR-1225 abundance using the equation: (miRNA$_{Dicer-/-\ or\ DGCR8-/-}$/sno65)/(miRNA$_{control}$/sno65). Bars represent the average values ± SEM, $n = 4$ for Dicer$^{-/-}$ and $n = 3$ for DGCR8$^{-/-}$. The horizontal dotted line indicates normalized control cell levels.

all four Ago proteins pulled down wt and Δss versions of miR-1225 (Figure 7). All four Ago proteins also pulled down wt and Δss-derived miR-1228, with the exception of Ago4 which did not pull down Δss-derived miR-1228 (Supplementary Figure S4D and E). Canonical miR-16 also interacted with all four Ago proteins. Although a detectable amount of snoRNA65 immunoprecipitated with the Ago proteins, it was signifi-cantly less than the miRNAs and therefore substantiates the specific interaction of the Ago proteins with the
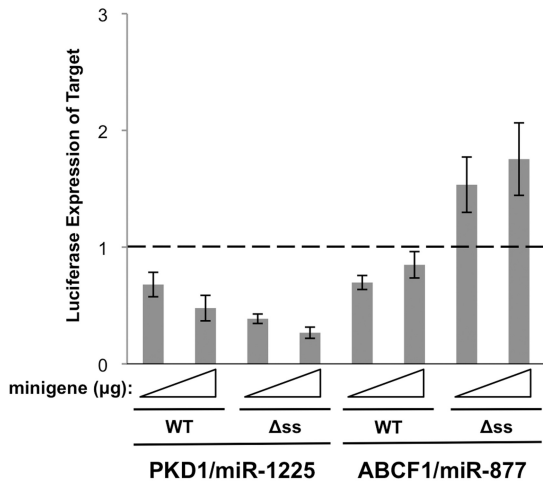
Figure 6. Mammalian mirtrons and simtrons function in silencing. Luciferase expression in HEK-293T cells transiently co-transfected with pmiRGLO with matching miRNA target sequences or mis-match miRNA target sequences and two different concentrations (0.4 or 1 μg) of the wt or splicing-deficient (Δss) minigene. Horizontal dotted line represents the normalized mismatch control value. The following equation was used to analyse the data: $[(\text{Luciferase Target}/ \text{Renilla})_{\text{miR\_match}}]_{\text{avg}} / [(\text{Luciferase Target/Renilla})_{\text{miR\_mismatch}}]_{\text{avg}}$. Each sample was analysed in triplicate. Bars represent the average ± SEM, $n = 5$ independent experiments.

miRNAs. Taken together, these results demonstrate the existence of a novel processing-pathway that generates small RNAs that can function to silence gene expression.

DISCUSSION

In this study, we identify small regulatory RNAs that are processed by a non-canonical biogenesis pathway that involves Drosha, but does not require DGCR8 or the other components of the canonical biogenesis pathway. These miRNAs were originally predicted to be mirtrons, due to their predicted hairpin structure, location within short introns, and identification from deep-sequencing analysis (40). Although we confirm that the predicted mammalian mirtrons, miR-877 and miR-1226 are, indeed, splicing-dependent in the context of their natural host gene, two other predicted mirtrons, miR-1225 and miR-1228 do not require splicing for their biogenesis. We classify these miRNAs as 'simtrons' (splicing-independent mirtron-like miRNAs), for their resemblance to mirtrons in their structure and genomic context spanning introns. To our knowledge, simtrons are the only class of miRNAs characterized to date that are processed in a manner that does not require splicing, the Microprocessor, Dicer or Ago2.

A number of non-canonical miRNAs have been described. However, they all require Dicer to produce the mature miRNA. Flynt *et al.* (60) discovered a miRNA in *Drosophila,* whose 5′ pre-miRNA end is generated as a result of splicing and whose 3′ pre-miRNA end is generated by exosome-mediated trimming. These 3′ tailed mirtrons are distinct from simtrons because they require splicing. Endogenous
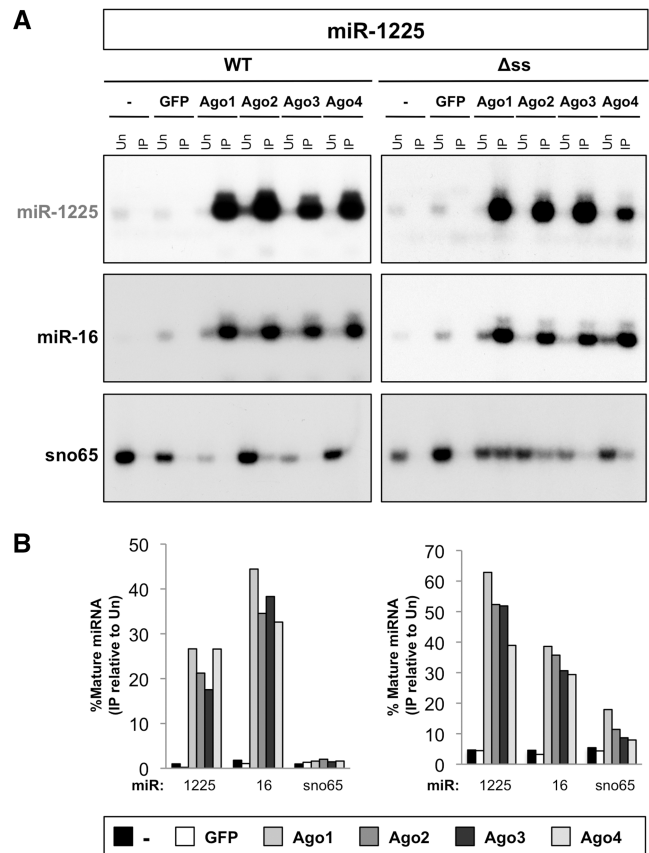

Figure 7. Simtrons associate with Argonaute proteins. (**A**) miR-1225 derived from wt and Δss minigenes were co-transfected with pFLAG-GFP, pFLAG-Ago1, pFLAG-Ago2, pFLAG-Ago3, pFLAG-Ago4 or mock (−) into HEK-293T cells. FLAG-tagged proteins were immunoprecipitated from cell lysates and associated miRNAs were analysed by radiolabelled stemloop RT–PCR. Un indicates the unbound fraction, and IP is the immunoprecipitated miRNA. For wt minigenes, unbound is 1/20 of IP. For Δss minigenes, unbound is 1/5 of IP. Sno65 is a control non-coding RNA. miR-16 is a canonical miRNA control. (**B**) The graphs represent the percent of miRNA in the IP fraction as determined using the equation: $(\text{IP}/(\text{IP} + (\text{Un} \times 20 \text{ or } 5)) \times 100$. Left panel: wt miR-1225, Right panel: Δss miR-1225.

shRNAs are non-canonical miRNAs that do not require Drosha or splicing but are cleaved by Dicer to generate a mature miRNA (25). Another class of non-canonical miRNAs is encoded by a murine herpesvirus and processed by tRNAseZ rather than by Drosha (26). Likewise, endogenous siRNAs are generated from the sequential processing of long hairpins by Dicer but are not considered miRNAs because they produce many different small RNAs (1). Finally, miRNAs have been identified that are derived from snoRNAs and tRNAs by non-canonical pathways that do not require the Microprocessor but require Dicer (25,27–31,33). These examples highlight the diversity of small RNAs and their processing pathways in the cells.

We demonstrated that simtrons are distinct from other classes of small RNAs in that they are not processed by Dicer or Ago2 (Figure 3), and yet are functional in gene silencing (Figure 6 and Supplementary Figure S4C).

Although we identified simtrons during our examination of mammalian mirtrons, it is possible, and perhaps likely, that the processing pathway responsible for the biogenesis of this non-canonical type of miRNA has a wider range of substrates. Indeed, several studies have observed miRNAs whose abundance, as determined by large-scale sequencing analysis, is not altered in *Dicer* knockout, *Drosha* knockout or *DGCR8* knockout cells (25,61), suggesting that additional processing pathways exist.

Our data indicate that Drosha is involved in the processing of simtrons, though this may not be a strict requirement. We also show that DGCR8 is not required for simtron processing, though we cannot exclude the possibility that DGCR8 may be involved in simtron biogenesis at some level (Figures 2–4). These findings distinguish simtrons from other miRNAs. A Drosha complex that does not contain DGCR8 has been identified but was reported to be less efficient at miRNA processing than the complex containing both DGCR8 and Drosha (8). It is possible that this alternative Drosha complex, that lacks DGCR8, processes simtrons. Although Drosha contains a double-stranded RNA binding domain (DSB), it appears to require an adaptor, such as DGCR8, for proper positioning and precise cleavage (62). DGCR8 functions as an anchor that recognizes both dsRNA and ssRNA and directs Drosha to cleave the stemloop ~11 bp from the dsRNA/ssRNA junction (5). Drosha is capable of cleaving hairpin-structured mRNA without producing miRNAs (61), however, this is the first evidence that Drosha can process a subclass of miRNAs in the absence of DGCR8. It is possible that another RNA binding protein plays a role similar to DGCR8 in the processing of some miRNAs such as simtrons. A number of splicing factors, for example have been shown to promote Drosha cleavage and silencing (63–66). Such a factor, which binds in the intron, could potentially act as a cofactor for Drosha cleavage of simtrons. Indeed, though DGCR8 is not required for simtron processing (Figure 3A), it appears to enhance processing *in vitro* (Figure 4B) possibly by co-immunoprecipitating an important co-factor or by stabilizing Drosha in the *in vitro* reaction.

Interestingly, simtrons generated from both the wt and splicing-deficient minigene transcripts immunoprecipitated with Drosha (Figure 4A), suggesting that Drosha processing by the simtron pathway may occur even when splicing is active. In this case, splicing and Drosha processing may be in competition. Competition between splicing and Drosha processing is further supported by results from experiments with the dominant negative form of Drosha. Only miR-1225 derived from the splicing deficient transcripts (Δss) was reduced as a result of TN-Drosha expression. Reduction of Drosha activity is not expected to affect miR-1225 levels from the wt transcript because splicing can generate the pre-miR-1225 via the mirtron pathway. Processing of the miR-1225 by the mirtron pathway is also suggested by the interaction of mature miR-1225 with Dicer. The simtron pathway, on the other hand, does not require Dicer (Figure 3A), and miR-1225 generated by this pathway does not interact with Dicer (Δss, Figure 3B). Taken

together, these results suggest that miR-1225 can be excised from the host gene RNA transcript by either the mirtron pathway (Drosha/DGCR8-independent, Dicer-dependent) or the simtron pathway (involving Drosha but DGCR8 and Dicer-independent) (Figure 8).

Our results indicate that Drosha is involved in the processing of simtrons, however, it seems unlikely that Drosha alone completes all steps of simtron processing. Known canonical and non-canonical miRNA processing pathways require two endonucleolytic cleavage events, carried out by distinct enzymes, to generate a mature miRNA. There are a number of human proteins that are predicted to have RNase activity (67). For example, human RNase P and tRNAse Z generate the 5′ and 3′-ends of tRNAs, respectively (68,69), and could feasibly recognize the highly structured simtronic introns and cleave the intron to generate the pre-miRNA. Additionally, the human tRNA splicing endonuclease complex cleaves both the 5′ and 3′ splice sites of a
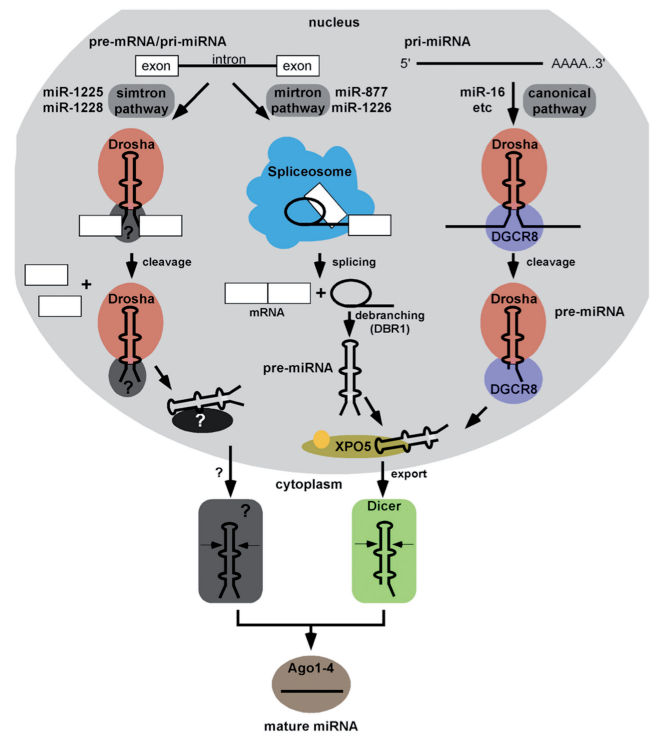

Figure 8. Proposed model of simtron biogenesis compared to other miRNA processing pathways. The pathways shown begin with the primary transcript and end with the mature product. Left: simtron pathway, Middle: mirtron pathway, Right: canonical miRNA pathway. Exons are depicted as boxes and introns and miRNAs as lines. Each protein or protein complex is labelled. Proteins labelled with question marks are proposed but not known. Simtrons (such as miR-1225 and miR-1228) processing from the intron involves Drosha and possibly an unknown binding partner. Simtrons are further processed by unknown factors and enter the RISC complex with any of the four human Argonaute proteins. Mirtrons (such as miR-877 and miR-1226) are excised from the host gene by the spliceosome, are debranched, exported from the nucleus by exportin5 (XPO5), cleaved by Dicer and enter the RISC complex. Canonical miRNAs (such as miR-16) are processed by Drosha and DGCR8, exported from the nucleus by XPO5, cleaved by Dicer and enter the RISC complex. All three pathways result in functional miRNAs.

pre-tRNA intron to release the intron from the transcript (70,71). It is therefore possible, that one of these enzymes also plays a role in simtron biogenesis.

The simtron biogenesis pathway appears to follow traditional miRNA assembly into the RISC complex with Ago proteins (Figure 7 and Supplementary Figure S4D and E) and targeting (Figure 6 and Supplementary Figure S4C). Even though miR-1225 and miR-1228 abundance was not affected by the absence of Argonaute-2 in the knockout cell lines (Figure 3C), they do associate with Ago2 when both are present (Figure 7 and Supplementary Figure S4D and E) suggesting that simtronic miRNAs in the Ago2$^{-/-}$ cells sort into Agos-1,3 and 4, which can compensate for the lack of Ago2.

One implication of our results is that, in certain situations, splicing and miRNA biogenesis may be in competition with one another. Competition between the splicing of pre-mRNA and the production of the miRNA would impact not only miRNA target gene expression, but also miRNA host gene expression. The generation of simtrons, by virtue of their intron-spanning location and processing pathway, could have an effect on the expression of the mRNA from which it is removed. Although excision of miRNAs from introns has been shown to be compatible with splicing of the intron, yielding both a pre-miRNA and spliced mRNA, with little if any impact on expression of either RNA (72,73), these studies examined miRNAs which are housed in larger introns and spatially removed from the essential 5′ splice site, 3′ splice site, branchpoint sequence and polypyrimidine tracts. Simtrons, on the other hand, encompass and excise all of these splice site sequences during processing. Thus, simtron biogenesis and splicing would appear to be mutually exclusive if splicing does not occur first (Figure 8).

Considering that many small RNAs are processed from the introns and non-coding sequences of pre-mRNA transcripts, a mechanism by which excision and expression of these RNAs is linked to mRNA expression is important to consider when investigating such phenomenon as genetic modifiers to disease. For example, *PKD1*, the host gene for miR-1225, is the most frequently mutated gene in autosomal dominant polycystic kidney disease (85% of cases) (74,75). Polycystic kidney disease, for reasons that are not clear, has variable disease severity (74,75). One intriguing possibility is that miR-1225 is a disease modifier that, depending on the type of mutation in the gene, could influence the penetrance of the disease (74). Furthermore, mutations found only within the simtron containing intron are associated with disease (76–78). Likewise, miR-1228 is housed within the *LRP1* gene, which has diverse functions in the cell and has been implicated to play a role in atherosclerosis and Alzheimer's disease (79). The idea that two separate biogenesis pathways (mirtronic and simtronic) can generate the same miRNA may indicate the importance of these miRNAs in the maintenance of homeostasis.

The discovery of the simtron biogenesis pathway demonstrates the complexity of RNA processing and uncovers another mechanism by which small non-coding RNAs can be produced from an RNA transcript. Our results also caution against the assumption of a biogenesis mechanism based on sequence, location and predicted structure alone. It is possible that other predicted mirtrons are not splicing-dependent, but are instead processed by the simtron pathway.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Table 1, Supplementary Figures 1–4, Supplementary Methods and Supplementary References [40,80,81].

REFERENCES

1. Bartel,D.P. (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, **116**, 281–297.
2. Kim,V.N. (2005) MicroRNA biogenesis: coordinated cropping and dicing. *Nat. Rev. Mol. Cell. Biol.*, **6**, 376–385.
3. Lai,E.C. (2003) microRNAs: runts of the genome assert themselves. *Curr. Biol.*, **13**, R925–R936.
4. Han,J., Lee,Y., Yeom,K.H., Kim,Y.K., Jin,H. and Kim,V.N. (2004) The Drosha-DGCR8 complex in primary microRNA processing. *Genes Dev.*, **18**, 3016–3027.
5. Han,J., Lee,Y., Yeom,K.H., Nam,J.W., Heo,I., Rhee,J.K., Sohn,S.Y., Cho,Y., Zhang,B.T. and Kim,V.N. (2006) Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. *Cell*, **125**, 887–901.
6. Lee,Y., Ahn,C., Han,J., Choi,H., Kim,J., Yim,J., Lee,J., Provost,P., Radmark,O., Kim,S. *et al.* (2003) The nuclear RNase III Drosha initiates microRNA processing. *Nature*, **425**, 415–419.
7. Denli,A.M., Tops,B.B., Plasterk,R.H., Ketting,R.F. and Hannon,G.J. (2004) Processing of primary microRNAs by the Microprocessor complex. *Nature*, **432**, 231–235.
8. Gregory,R.I., Yan,K.P., Amuthan,G., Chendrimada,T., Doratotaj,B., Cooch,N. and Shiekhattar,R. (2004) The Microprocessor complex mediates the genesis of microRNAs. *Nature*, **432**, 235–240.
9. Landthaler,M., Yalcin,A. and Tuschl,T. (2004) The human DiGeorge syndrome critical region gene 8 and Its D. melanogaster homolog are required for miRNA biogenesis. *Curr. Biol.*, **14**, 2162–2167.
10. Bohnsack,M.T., Czaplinski,K. and Gorlich,D. (2004) Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. *RNA*, **10**, 185–191.
11. Lund,E., Guttinger,S., Calado,A., Dahlberg,J.E. and Kutay,U. (2004) Nuclear export of microRNA precursors. *Science*, **303**, 95–98.
12. Yi,R., Qin,Y., Macara,I.G. and Cullen,B.R. (2003) Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev.*, **17**, 3011–3016.

13. Bernstein,E., Caudy,A.A., Hammond,S.M. and Hannon,G.J. (2001) Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature*, **409**, 363–366.

14. Grishok,A., Pasquinelli,A.E., Conte,D., Li,N., Parrish,S., Ha,I., Baillie,D.L., Fire,A., Ruvkun,G. and Mello,C.C. (2001) Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control C. elegans developmental timing. *Cell*, **106**, 23–34.

15. Hutvagner,G., McLachlan,J., Pasquinelli,A.E., Balint,E., Tuschl,T. and Zamore,P.D. (2001) A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science*, **293**, 834–838.

16. Ketting,R.F., Fischer,S.E., Bernstein,E., Sijen,T., Hannon,G.J. and Plasterk,R.H. (2001) Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans. *Genes Dev.*, **15**, 2654–2659.

17. Knight,S.W. and Bass,B.L. (2001) A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in Caenorhabditis elegans. *Science*, **293**, 2269–2271.

18. Du,T. and Zamore,P.D. (2005) microPrimer: the biogenesis and function of microRNA. *Development*, **132**, 4645–4652.

19. Khvorova,A., Reynolds,A. and Jayasena,S.D. (2003) Functional siRNAs and miRNAs exhibit strand bias. *Cell*, **115**, 209–216.

20. Schwarz,D.S., Hutvagner,G., Du,T., Xu,Z., Aronin,N. and Zamore,P.D. (2003) Asymmetry in the assembly of the RNAi enzyme complex. *Cell*, **115**, 199–208.

21. Chekulaeva,M. and Filipowicz,W. (2009) Mechanisms of miRNA-mediated post-transcriptional regulation in animal cells. *Curr. Opin. Cell Biol.*, **21**, 452–460.

22. Filipowicz,W., Bhattacharyya,S.N. and Sonenberg,N. (2008) Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nat. Rev. Genet.*, **9**, 102–114.

23. Guo,H., Ingolia,N.T., Weissman,J.S. and Bartel,D.P. (2010) Mammalian microRNAs predominantly act to decrease target mRNA levels. *Nature*, **466**, 835–840.

24. Valencia-Sanchez,M.A., Liu,J., Hannon,G.J. and Parker,R. (2006) Control of translation and mRNA degradation by miRNAs and siRNAs. *Genes Dev.*, **20**, 515–524.

25. Babiarz,J.E., Ruby,J.G., Wang,Y., Bartel,D.P. and Blelloch,R. (2008) Mouse ES cells express endogenous shRNAs, siRNAs, and other Microprocessor-independent, Dicer-dependent small RNAs. *Genes Dev.*, **22**, 2773–2785.

26. Bogerd,H.P., Karnowski,H.W., Cai,X., Shin,J., Pohlers,M. and Cullen,B.R. (2010) A mammalian herpesvirus uses noncanonical expression and processing mechanisms to generate viral MicroRNAs. *Mol. Cell*, **37**, 135–142.

27. Brameier,M., Herwig,A., Reinhardt,R., Walter,L. and Gruber,J. (2011) Human box C/D snoRNAs with miRNA like functions: expanding the range of regulatory RNAs. *Nucleic Acids Res.*, **39**, 675–686.

28. Cole,C., Sobala,A., Lu,C., Thatcher,S.R., Bowman,A., Brown,J.W., Green,P.J., Barton,G.J. and Hutvagner,G. (2009) Filtering of deep sequencing data reveals the existence of abundant Dicer-dependent small RNAs derived from tRNAs. *RNA*, **15**, 2147–2160.

29. Ender,C., Krek,A., Friedlander,M.R., Beitzinger,M., Weinmann,L., Chen,W., Pfeffer,S., Rajewsky,N. and Meister,G. (2008) A human snoRNA with microRNA-like functions. *Mol. Cell*, **32**, 519–528.

30. Ono,M., Scott,M.S., Yamada,K., Avolio,F., Barton,G.J. and Lamond,A.I. (2011) Identification of human miRNA precursors that resemble box C/D snoRNAs. *Nucleic Acids Res.*, **39**, 3879–3891.

31. Scott,M.S., Avolio,F., Ono,M., Lamond,A.I. and Barton,G.J. (2009) Human miRNA precursors with box H/ACA snoRNA features. *PLoS Comput. Biol.*, **5**, e1000507.

32. Shapiro,J.S., Varble,A., Pham,A.M. and Tenoever,B.R. (2010) Noncanonical cytoplasmic processing of viral microRNAs. *RNA*, **16**, 2068–2074.

33. Taft,R.J., Glazov,E.A., Lassmann,T., Hayashizaki,Y., Carninci,P. and Mattick,J.S. (2009) Small RNAs derived from snoRNAs. *RNA*, **15**, 1233–1240.

34. Cheloufi,S., Dos Santos,C.O., Chong,M.M. and Hannon,G.J. (2010) A dicer-independent miRNA biogenesis pathway that requires Ago catalysis. *Nature*, **465**, 584–589.

35. Cifuentes,D., Xue,H., Taylor,D.W., Patnode,H., Mishima,Y., Cheloufi,S., Ma,E., Mane,S., Hannon,G.J., Lawson,N.D. et al. (2010) A novel miRNA processing pathway independent of Dicer requires Argonaute2 catalytic activity. *Science*, **328**, 1694–1698.

36. Yang,J.S. and Lai,E.C. (2010) Dicer-independent, Ago2-mediated microRNA biogenesis in vertebrates. *Cell Cycle*, **9**, 4455–4460.

37. Yang,J.S., Maurin,T., Robine,N., Rasmussen,K.D., Jeffrey,K.L., Chandwani,R., Papapetrou,E.P., Sadelain,M., O'Carroll,D. and Lai,E.C. (2010) Conserved vertebrate mir-451 provides a platform for Dicer-independent, Ago2-mediated microRNA biogenesis. *Proc. Natl Acad. Sci. USA*, **107**, 15163–15168.

38. Okamura,K., Hagen,J.W., Duan,H., Tyler,D.M. and Lai,E.C. (2007) The mirtron pathway generates microRNA-class regulatory RNAs in Drosophila. *Cell*, **130**, 89–100.

39. Ruby,J.G., Jan,C.H. and Bartel,D.P. (2007) Intronic microRNA precursors that bypass Drosha processing. *Nature*, **448**, 83–86.

40. Berezikov,E., Chung,W.J., Willis,J., Cuppen,E. and Lai,E.C. (2007) Mammalian mirtron genes. *Mol. Cell*, **28**, 328–336.

41. Glazov,E.A., Cottee,P.A., Barris,W.C., Moore,R.J., Dalrymple,B.P. and Tizard,M.L. (2008) A microRNA catalog of the developing chicken embryo identified by a deep sequencing approach. *Genome Res.*, **18**, 957–964.

42. Zhu,Q.H., Spriggs,A., Matthew,L., Fan,L., Kennedy,G., Gubler,F. and Helliwell,C. (2008) A diverse set of microRNAs and microRNA-like small RNAs in developing rice grains. *Genome Res.*, **18**, 1456–1465.

43. Chiang,H.R., Schoenfeld,L.W., Ruby,J.G., Auyeung,V.C., Spies,N., Baek,D., Johnston,W.K., Russ,C., Luo,S., Babiarz,J.E. et al. (2010) Mammalian microRNAs: experimental evaluation of novel and previously annotated genes. *Genes Dev.*, **24**, 992–1009.

44. Sibley,C.R., Seow,Y., Saayman,S., Dijkstra,K.K., El Andaloussi,S., Weinberg,M.S. and Wood,M.J. (2011) The biogenesis and characterization of mammalian microRNAs of mirtron origin. *Nucleic Acids Res.*, **40**, 438–448.

45. Durocher,Y., Perret,S. and Kamen,A. (2002) High-level and high-throughput recombinant protein production by transient transfection of suspension-growing human 293-EBNA1 cells. *Nucleic Acids Res.*, **30**, E9.

46. Liu,J., Carmell,M.A., Rivas,F.V., Marsden,C.G., Thomson,J.M., Song,J.J., Hammond,S.M., Joshua-Tor,L. and Hannon,G.J. (2004) Argonaute2 is the catalytic engine of mammalian RNAi. *Science*, **305**, 1437–1441.

47. Heo,I., Joo,C., Cho,J., Ha,M., Han,J. and Kim,V.N. (2008) Lin28 mediates the terminal uridylation of let-7 precursor MicroRNA. *Mol. Cell*, **32**, 276–284.

48. Murchison,E.P., Partridge,J.F., Tam,O.H., Cheloufi,S. and Hannon,G.J. (2005) Characterization of Dicer-deficient murine embryonic stem cells. *Proc. Natl Acad. Sci. USA*, **102**, 12135–12140.

49. Chen,C., Ridzon,D.A., Broomer,A.J., Zhou,Z., Lee,D.H., Nguyen,J.T., Barbisin,M., Xu,N.L., Mahuvakar,V.R., Andersen,M.R. et al. (2005) Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucleic Acids Res.*, **33**, e179.

50. Lian,S.L., Li,S., Abadal,G.X., Pauley,B.A., Fritzler,M.J. and Chan,E.K. (2009) The C-terminal half of human Ago2 binds to multiple GW-rich regions of GW182 and requires GW182 to mediate silencing. *RNA*, **15**, 804–813.

51. Meister,G., Landthaler,M., Patkaniowska,A., Dorsett,Y., Teng,G. and Tuschl,T. (2004) Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. *Mol. Cell*, **15**, 185–197.

52. Biggar,K.K., Kornfeld,S.F. and Storey,K.B. (2011) Amplification and sequencing of mature microRNAs in uncharacterized animal models using stem-loop reverse transcription-polymerase chain reaction. *Anal. Biochem.*, **416**, 231–233.

53. Varkonyi-Gasic,E. and Hellens,R.P. (2011) Quantitative stem-loop RT-PCR for detection of microRNAs. *Methods Mol. Biol.*, **744**, 145–157.

54. Bar,M., Wyman,S.K., Fritz,B.R., Qi,J., Garg,K.S., Parkin,R.K., Kroh,E.M., Bendoraite,A., Mitchell,P.S., Nelson,A.M. et al. (2008) MicroRNA discovery and profiling in human embryonic

stem cells by deep sequencing of small RNA libraries. *Stem Cells*, **26**, 2496–2505.

55. Stark,M.S., Tyagi,S., Nancarrow,D.J., Boyle,G.M., Cook,A.L., Whiteman,D.C., Parsons,P.G., Schmidt,C., Sturm,R.A. and Hayward,N.K. (2010) Characterization of the Melanoma miRNAome by Deep Sequencing. *PLoS One*, **5**, e9685.

56. Krutzfeldt,J., Rajewsky,N., Braich,R., Rajeev,K.G., Tuschl,T., Manoharan,M. and Stoffel,M. (2005) Silencing of microRNAs in vivo with 'antagomirs'. *Nature*, **438**, 685–689.

57. Tang,F., Hajkova,P., O'Carroll,D., Lee,C., Tarakhovsky,A., Lao,K. and Surani,M.A. (2008) MicroRNAs are tightly associated with RNA-induced gene silencing complexes in vivo. *Biochem. Biophys. Res. Commun.*, **372**, 24–29.

58. Wang,Y., Medvid,R., Melton,C., Jaenisch,R. and Blelloch,R. (2007) DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell self-renewal. *Nat. Genet.*, **39**, 380–385.

59. Lee,Y., Jeon,K., Lee,J.T., Kim,S. and Kim,V.N. (2002) MicroRNA maturation: stepwise processing and subcellular localization. *EMBO J.*, **21**, 4663–4670.

60. Flynt,A.S., Greimann,J.C., Chung,W.J., Lima,C.D. and Lai,E.C. (2010) MicroRNA biogenesis via splicing and exosome-mediated trimming in Drosophila. *Mol. Cell*, **38**, 900–907.

61. Chong,M.M., Zhang,G., Cheloufi,S., Neubert,T.A., Hannon,G.J. and Littman,D.R. (2010) Canonical and alternate functions of the microRNA biogenesis machinery. *Genes Dev.*, **24**, 1951–1960.

62. Kim,V.N., Han,J. and Siomi,M.C. (2009) Biogenesis of small RNAs in animals. *Nat. Rev. Mol. Cell. Biol.*, **10**, 126–139.

63. Bayne,E.H., Portoso,M., Kagansky,A., Kos-Braun,I.C., Urano,T., Ekwall,K., Alves,F., Rappsilber,J. and Allshire,R.C. (2008) Splicing factors facilitate RNAi-directed silencing in fission yeast. *Science*, **322**, 602–606.

64. Trabucchi,M., Briata,P., Garcia-Mayoral,M., Haase,A.D., Filipowicz,W., Ramos,A., Gherzi,R. and Rosenfeld,M.G. (2009) The RNA-binding protein KSRP promotes the biogenesis of a subset of microRNAs. *Nature*, **459**, 1010–1014.

65. Wu,H., Sun,S., Tu,K., Gao,Y., Xie,B., Krainer,A.R. and Zhu,J. (2010) A splicing-independent function of SF2/ASF in microRNA processing. *Mol. Cell*, **38**, 67–77.

66. Guil,S. and Caceres,J.F. (2007) The multifunctional RNA-binding protein hnRNP A1 is required for processing of miR-18a. *Nat. Struct. Mol. Biol.*, **14**, 591–596.

67. Tomecki,R. and Dziembowski,A. (2010) Novel endoribonucleases as central players in various pathways of eukaryotic RNA metabolism. *RNA*, **16**, 1692–1724.

68. Kikovska,E., Svard,S.G. and Kirsebom,L.A. (2007) Eukaryotic RNase P RNA mediates cleavage in the absence of protein. *Proc. Natl Acad. Sci. USA*, **104**, 2062–2067.

69. Vogel,A., Schilling,O., Spath,B. and Marchfelder,A. (2005) The tRNase Z family of proteins: physiological functions, substrate specificity and structural properties. *Biol. Chem.*, **386**, 1253–1264.

70. Paushkin,S.V., Patel,M., Furia,B.S., Peltz,S.W. and Trotta,C.R. (2004) Identification of a human endonuclease complex reveals a link between tRNA splicing and pre-mRNA 3' end formation. *Cell*, **117**, 311–321.

71. Trotta,C.R., Miao,F., Arn,E.A., Stevens,S.W., Ho,C.K., Rauhut,R. and Abelson,J.N. (1997) The yeast tRNA splicing endonuclease: a tetrameric enzyme with two active site subunits homologous to the archaeal tRNA endonucleases. *Cell*, **89**, 849–858.

72. Kim,Y.K. and Kim,V.N. (2007) Processing of intronic microRNAs. *EMBO J.*, **26**, 775–783.

73. Morlando,M., Ballarino,M., Gromak,N., Pagano,F., Bozzoni,I. and Proudfoot,N.J. (2008) Primary microRNA transcripts are processed co-transcriptionally. *Nat. Struct. Mol. Bio.l*, **15**, 902–909.

74. Tan,Y.C., Blumenfeld,J. and Rennert,H. (2011) Autosomal dominant polycystic kidney disease: genetics, mutations and microRNAs. *Biochim. Biophys. Acta*, **1812**, 1202–1212.

75. van Gulick,J.J., Gevers,T.J., van Keimpema,L. and Drenth,J.P. (2011) Hepatic and renal manifestations in autosomal dominant polycystic kidney disease: a dichotomy of two ends of a spectrum. *Neth. J. Med.*, **69**, 367–371.

76. Aguiari,G., Savelli,S., Garbo,M., Bozza,A., Augello,G., Penolazzi,L., De Paoli Vitali,E., La Torre,C., Cappelli,G., Piva,R. *et al.* (2000) Novel splicing and missense mutations in autosomal dominant polycystic kidney disease 1 (PKD1) gene: expression of mutated genes. *Hum. Mutat.*, **16**, 444–445.

77. Badenas,C., Torra,R., San Millan,J.L., Lucero,L., Mila,M., Estivill,X. and Darnell,A. (1999) Mutational analysis within the 3' region of the PKD1 gene. *Kidney Int.*, **55**, 1225–1233.

78. Burtey,S., Lossi,A.M., Bayle,J., Berland,Y. and Fontes,M. (2002) Mutation screening of the PKD1 transcript by RT-PCR. *J. Med. Genet.*, **39**, 422–429.

79. Lillis,A.P., Van Duyn,L.B., Murphy-Ullrich,J.E. and Strickland,D.K. (2008) LDL receptor-related protein 1: unique tissue-specific functions revealed by selective gene knockout studies. *Physiol. Rev.*, **88**, 887–918.

80. Jodelka,F.M., Ebert,A.D., Duelli,D.M. and Hastings,M.L. (2010) A feedback loop regulates splicing of the spinal muscular atrophy-modifying gene, SMN2. *Hum. Mol. Genet.*, **19**, 4906–4917.

81. Cikos,S., Bukovska,A. and Koppel,J. (2007) Relative quantification of mRNA: comparison of methods currently used for real-time PCR data analysis. *BMC Mol. Biol.*, **8**, 113.