이학박사학위논문

마이크로RNA 전구체의 모노유리딘화에 대한 연구

Studies on Mono-Uridylation of Pre-MicroRNAs

2013년 2월

서울대학교 대학원

생명과학부

하 민 주

Studies on Mono-Uridylation of Pre-MicroRNAs

by

Minju Ha

Advisor

Associate Professor V. Narry Kim, D. Phil.

A Thesis for the Degree of Doctor of Philosophy

February, 2013

School of Biological Sciences

Seoul National University

ABSTRACT

MicroRNAs (miRNAs) are single-stranded RNAs of ~22 nucleotide (nt) in length that function as guide molecules of gene silencing. Biogenesis of miRNA involves multiple maturation steps. RNase III Drosha initiates miRNA maturation by cleaving a primary miRNA (pri-miRNA) transcript and releasing a hairpin shaped precursor miRNA (pre-miRNA) with a 2-nt 3′ overhang. The downstream maturation factor, Dicer, recognizes the 2-nt 3′ overhang structure to selectively process pre-miRNAs.

Biogenesis of the let-7 miRNA family, one of the most ancient and highly conserved miRNAs, is suppressed in embryonic stage and in certain cancer cells. It was previously reported that Terminal uridylyl transferase 4 (TUT4) down-regulates let-7 miRNA biogenesis by oligo-uridylating let-7 precursor (pre-let-7) in mouse embryonic stem cells and that a pluripotency marker Lin28 acts as a processivity factor of TUT4.

Here I find that TUTs positively controls let-7 biogenesis via mono-uridylation of pre-let-7 in a Lin28-independent manner. I identify TUT7/ZCCHC6, TUT4/ZCCHC11, and TUT2/PAPD4/GLD2 as the terminal uridylyl transferases responsible for pre-miRNA mono-uridylation. Such mono-uridylation enhances Dicer processing by generating an optimal 3′ 2nt overhnag end structure of pre-let-7 for Dicer recognition.

I further suggest that miRNAs is subdivided into two groups, group I (which has 2nt 3′ overhang) and group II (which has 1nt 3′ overhang) based on the end structure of the precursor and dependency for mono-uridylation. Unlike prototypic pre-miRNAs (group I), group II pre-miRNAs acquire a shorter (1-nt) 3′ overhang from Drosha processing and therefore require a 3′-end mono-uridylation for Dicer processing. The majority of let-7 miRNAs and miR-105 belong to group II. The TUTs act specifically on dsRNAs with a 1-nt 3′ overhang, thereby creating a 2-nt 3′ overhang for optimal Dicer processing. Depletion of TUTs reduces let-7 levels and disrupts let-7 function in cells.

While the let-7 suppressor, Lin28, induces inhibitory oligo-uridylation in embryonic stem cells, mono-uridylation occurs in somatic cells lacking Lin28 to promote let-7 biogenesis. My study reveals functional duality of uridylation and introduces TUT7/4/2 as novel components of the miRNA biogenesis pathway.

Key words

microRNA (miRNA); miRNA biogenesis; post-transcriptional regulation; pre-microRNA (pre-miRNA); let-7; uridylation; TUT; TUTase; terminal uridylyl transferse; TUT7; ZCCHC6; TUT4; ZCCHC11; TUT2; PAPD4; GLD2; Dicer

CONTENTS

ABSTRACT i

CONTENTS iii

LIST OF FIGURES AND TABLES v

ABBREVIATIONS vii

CHAPTER I. Overview 1

I-1. What is MicroRNA? 2

I-2. Biogenesis of MicroRNA 2

I-3. Regulation of MicroRNA Biogenesis 5

CHAPTER II. Mono-Uridylation of Pre-MicroRNA in the Regulation of MicroRNA Biogenesis 9

II-1. Introduction 10

II-2. Results 17

II-3. Discussion 55

II-4. Materials and Methods 60

CHAPTER III. Conclusion 92

REFERENCES **99**

PUBLICATIONS **115**

국문 초록 / ABSTRACT IN KOREAN **116**

감사의 글 / ACKNOWLEDGEMENTS **120**

LIST OF FIGURES AND TABLES

Figure I-1. miRNA biogenesis pathway in animals.·····3
Figure II-1. Lin28 suppresses let-7 biogenesis in Huh7 and ES cells·····11
Figure II-2. Knockdown of TUT4 and Lin28 increases let-7.·····13
Figure II-3. Human TUTases.·····14
Figure II-4. Let-7 biogenesis is suppressed through oligo-uridylation by Lin28 and TUT4.·····15
Figure II-5. Scheme of pre-let-7 sequencing.·····18
Figure II-6. Pre-let-7 is mono-uridylated in the absence of Lin28.·····19
Figure II-7. TUT7, TUT4, and TUT2 mono-uridylate pre-let-7 in the absence of Lin28.·····21
Figure II-8. TUT4 and TUT2 catalyze mono-uridylation of pre-let-7.·····23
Figure II-9. TUT7, TUT4, and TUT2 redundantly promote biogenesis of let-7.·····25
Figure II-10. TUT7/4/2 function redundantly in let-7 biogenesis.·····28
Figure II-11. TUT7, TUT4, and TUT2 are required for mono-uridylation of pre-let-7.·····31
Figure II-12. Structure of pre-let-7.·····32
Figure II-13. Mono-uridylation of pre-let-7 enhances Dicer processing.··34
Figure II-14. The let-7 family is subdivided into two groups.·····35
Figure II-15. Mono-uridylated sequences in let-7 3p reads.·····39
Figure II-16. TUT7, TUT4, and TUT2 are required for the biogenesis of group II let-7.·····41
Figure II-17. Group II let-7 is specifically regulated through mono-uridylation by TUT7/4/2.·····42
Figure II-18. TUT7 acts specifically on group II pre-miRNAs.·····44

Figure II-19. TUT4 and TUT2 act specifically on pre-let-7a-1 with a 1-nt 3′ overhang.·····47
Figure II-20. TUTs act specifically on dsRNAs with a 1-nt 3′ overhang.··48
Figure II-21. Sequencing of group I pre-let-7 following TUT knock-down·····49
Figure II-22. Mono-uridylation is critical for functionality of let-7.·····51
Figure II-23. Phylogenetic distribution of the let-7 family.·····53
Figure II-24. Other group II miRNAs.·····57
Table II-1. Results of pre-let-7 sequencing following knock-down of TUTs in HeLa·····71
Table II-2. List of small RNA sequencing data used in miRNA analysis.·75
Table II-3. Results of small RNA deep sequencing following knock-down of TUTs in HeLa.·····83
Table II-4. Results of group I pre-let-7 sequencing following knock-down of TUTs in HeLa.·····85
Table II-5. Oligonucleotides Used in This Study.·····88
Figure III-1. Mono-uridylation is required for group II miRNA.·····94
Figure III-2. Dual effects of uridylation on the let-7 biogenesis.·····96
Figure III-3. Post-transcriptional control of miRNA precursors.·····97

ABBREVIATIONS

Ago,	Argonaute
ATP,	adenosine triphosphate
bp,	base pair
CTP,	cytidine triphosphate
DGCR8,	DiGeorge syndrome critical region gene 8
DNA,	deoxyribonucleic acid
dsRNA,	double stranded RNA
ES cells,	embryonic stem cells
EXP5,	Exportin 5
GEO,	gene expression omnibus
GTP,	guanosine triphosphate
HCC cells,	hepatocellular carcinoma cells
miRNA,	microRNA
mRNA,	messenger RNA
nt,	nucleotide
NTP,	nucleoside triphosphate
PAP,	poly (A) polymerase
pre-miRNA,	precursor microRNA
pri-miRNA,	primary microRNA

PUP,	poly (U) polymerase
qRT-PCR,	quantitative real-time polymerase chain reaction
RISC,	RNA induced silencing complex
RNA,	ribonucleic acid
RNase III,	ribonuclease III
RPM,	read per million
RT,	reverse transcription
siRNA,	small interfering RNA
TUTase,	terminal uridylyl transferase
U-tail,	uridine tail
UTP,	uridine triphosphate
UTR,	untranslated region
ZCCHC11,	zinc finger CCHC domain containing 11

CHAPTER I

Overview

I-1. What is MicroRNA ?

MicroRNAs (miRNAs) are single-stranded RNAs of ~22 nucleotide (nt) in length that are generated from endogenous hairpin-shaped transcripts (Kim et al., 2009). MiRNAs function as guide molecules by base-pairing with target mRNAs that have complementary sequence in their 3' untranslated region (UTR) (Bartel, 2009), which lead to mRNA decay and/or translational repression (Huntzinger and Izaurralde, 2011). Since the discovery of first miRNA lin-4 in 1993 (Lee et al., 1993), vast numbers of miRNAs have been identified through sequencing and computational prediction methods. MiRNAs are found in animals, plants, virus, and unicellular eukaryotes (Griffiths-Jones et al., 2008). Mammals have at least 196 conserved miRNA gene families and among them, 89 miRNAs are more broadly conserved in chickens, lizards, frogs, and fish (Chiang et al., 2010). MiRNAs are predicted to regulate over one third of all human genes (Lewis et al., 2005). These tiny regulators are involved in various cellular pathways including development, cell differentiation, apoptosis, cell proliferation and tumorigenesis.

I-2. Biogenesis of MicroRNA

Biogenesis of microRNA (miRNA) involves multiple maturation steps (**Figure I-1**) (Kim et al., 2009). As miRNA sequences are embedded in the

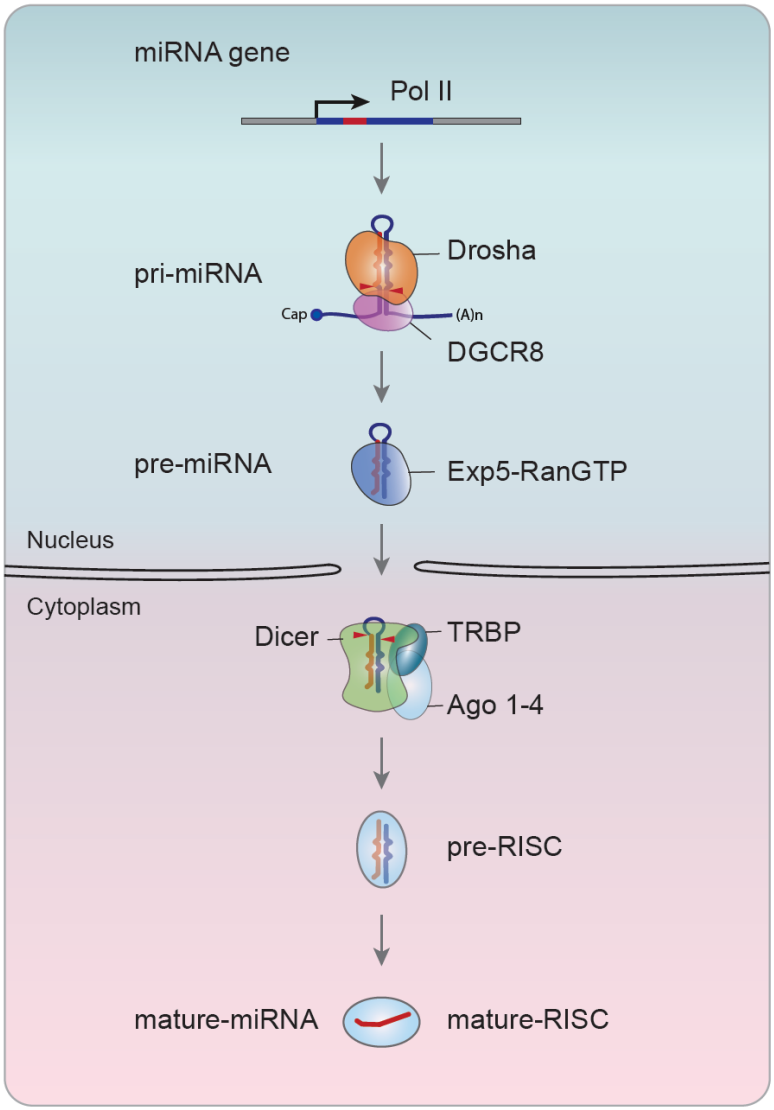


Figure I-1. Biogenesis of canonical miRNA

stem of a local hairpin in a nascent transcript (primary miRNA or pri-miRNA), a couple of endonucleolytic reactions are needed to yield a functional miRNA. The nuclear RNase III Drosha initiates the maturation process by cleaving a pri-miRNA to release a ~70-nt hairpin-shaped RNA (pre-miRNA) (Lee et al., 2003). Together with its cofactor DGCR8 (also known as Pasha), Drosha cuts the hairpin at 11 bp away from the base of the hairpin (Denli et al., 2004; Gregory et al., 2004; Han et al., 2004; Han et al., 2006; Landthaler et al., 2004). Like other RNase III type endonucleases, Drosha introduces a staggered cut such that the product acquires a characteristic 2-nt overhang at the 3′ terminus. After cleavage, the pre-miRNA is exported to the cytoplasm by exportin 5 in a complex with Ran-GTP (Bohnsack et al., 2004; Lund et al., 2004; Yi et al., 2003).

The cytoplasmic RNase III Dicer processes the pre-miRNA further to liberate a small RNA duplex (Bernstein et al., 2001; Grishok et al., 2001; Hutvagner et al., 2001; Ketting et al., 2001; Knight and Bass, 2001). Human Dicer binds to the pre-miRNA with a preference for the 2-nt 3′ overhang (Zhang et al., 2004). The 5′ and 3′ ends of pre-miRNA are accommodated in two basic pockets (5′ and 3′ pockets, respectively) located in the PAZ domain of Dicer (Park et al., 2011). Dicer measures 22-nt from the 5′ phosphorylated end of pre-miRNA and cleaves near the terminal loop (Park et al., 2011; Vermeulen et al., 2005; Zhang et al., 2002; Zhang et al., 2004). The resulting small RNA duplex is loaded on to Argonaute and one of the strands is

selected to form an active RNA-induced silencing complex (RISC) (Hammond et al., 2001; Mourelatos et al., 2002; Tabara et al., 1999).

I-3. Regulation of MicroRNA Biogenesis

It has become apparent that miRNA regulators themselves are under control of intensive regulation (Kim et al., 2009; Krol et al., 2010b; Siomi and Siomi, 2010). Precise control of miRNA level is essential to maintain normal cellular functions, and its dysregulation is often associated with diseases. Regulation of miRNA biogenesis can be achieved at transcriptional level or post-transcriptional level, which control miRNA processing and stability of miRNA.

Most miRNAs are transcribed by RNA polymerase II and they are regulated in a similar manner with protein coding genes. Developmental stage- or tissue- specific miRNAs are controlled by signaling (Saj and Lai, 2011). Transcription factors like Myc, p53, and MYOD1 regulate miRNA expression positively or negatively (Kim et al., 2009; Krol et al., 2010b). In addition, epigenetic control such as DNA methylation and modifications of histones also contributes to miRNA gene regulation (Davis-Dusenbery and Hata, 2010).

Drosha processing can be regulated by accessory proteins that bind to biogenesis factors or miRNA precursors. SMAD proteins are activated by

bone morphogenetic protein (BMP)/ transforming growth factor- β (TGF β) and interact with Drosha and DDX5 (also known as p68) to stimulate Drosha processing of miR-21and miR-199a (Davis et al., 2008). Heterogeneous ribonucleoprotein particle A1 (hnRNPA1) binds loop regions of pri-miR-18a and facilitates its Drosha processing (Guil and Caceres, 2007; Michlewski et al., 2008).

Some RNA binding proteins are reported to affect both on Drosha and Dicer processing. KSRP also serve as auxiliary protein in biogenesis of a subset of miRNAs by binding their terminal loop and stimulating cleavage by Drosha or Dicer (Trabucchi et al., 2009). RNA binding protein Lin28 is responsible for the suppression of let-7 maturation (Heo et al., 2008; Newman et al., 2008; Rybak et al., 2008; Viswanathan et al., 2008). The Lin28 proteins bind to terminal loop of pre-let-7 to suppress Drosha and Dicer processing (Heo et al., 2009; Loughlin et al., 2012; Nam et al., 2011). Lin28 also recruits terminal uridylyl transferase 4 (TUT4) to induce oligo-uridylation of pre-let-7 by TUT4, which interferes Dicer processing and facilitates decay (Hagan et al., 2009; Heo et al., 2009).

Other than 3′ uridylation, several types of RNA modification is reported to impair Drosha or Dicer processing. RNA editing, the alteration of adenines to inosines that is mediated by adenine deaminases (ADARs), has been observed in pri-miR-142 (Yang et al., 2006) and pre-miR-151 (Kawahara et al., 2007). The editing modification makes pri-miRNAs or pre-

miRNAs poor substrate of Drosha or Dicer. Recently, human RNA methyltransferase BCDIN3D was shown to O-methylate 5' monophosphate of pre-miR-145 and pre-miR-23b, which impairs Dicer processing (Xhemalce et al., 2012).

MiRNA biogenesis can be also regulated by modulating the stability of precursor or mature miRNA level (Ruegger and Grosshans, 2012). Some reports showed that precursor miRNAs are cleaved by specific nucleases in human. Mammalian immune regulator MCPIP1 (also known as Zc3h12a) ribonuclease was identified as suppressing miRNA biogenesis via cleavage of the terminal loops of pre-miRNAs (Suzuki et al., 2011). More recently, endoplasmic reticulum (ER) transmembrane kinase-endoribonuclease, IRE1α was shown to be activated upon ER stress, and cleaves select pre-miRNAs (miR-17, 34a, 96, 125b) to derepress translation of proapoptotic Caspase-2 (Upton et al., 2012).

Turnover of mature miRNA is a largely unexplored area, especially in terms of substrate specificity and phylogenetic conservation of miRNA-degrading enzymes. Active degradation of miRNAs was initially reported in Arabidopsis thaliana, where it is mediated by a group of exoribonucleases named small RNA degrading nuclease (SDN) (Ramachandran and Chen, 2008). Although SDN uses 3’-to-5’ exonucleolytic mechanism, the 5’-to-3’ exoribonucleases XRN-1 and XRN-2 were shown to degrade mature miRNA in Caenorhabditis elegnas (Chatterjee et al., 2011; Chatterjee and Grosshans,

2009). In humans, Krol and colleagues found that miRNA-cluster miR-183/96/192, as well as miR-204 and miR-211 are rapidly decreased upon dark adaptation (Krol et al., 2010a). In addition, human polynucleotide phosphorylase (PNPasold-35), which is an interferon (IFN)-inducible 3’-to-5’ exonuclease degrades certain mature miRNAs (including miR-221, miR-222, and miR-106) in melanoma cells (Das et al., 2010). Another interesting example is that viral encoded poly (A) polymerase of insect and mammalian poxviruses induces the 3' polyadenylation of host miRNAs which leads to degradation by host machinery (Backes et al., 2012).

CHAPTER II

Mono-Uridylation of Pre-MicroRNA

in the Regulation of MicroRNA Biogenesis

II-1. Introduction

The let-7 miRNA family is highly conserved throughout bilaterian animals (Pasquinelli et al., 2000; Reinhart et al., 2000; Roush and Slack, 2008). Let-7 miRNAs suppress cell proliferation and promote cell differentiation by targeting multiple genes including HMGA2, RAS, and Lin28 (Bussing et al., 2008). At the organismal level, let-7 has been implicated in multiple processes such as larval development in C. elegans and growth and glucose metabolism in mammals (Grosshans et al., 2005; Meneely and Herman, 1979; Pasquinelli et al., 2000; Reinhart et al., 2000; Zhu et al., 2010; Zhu et al., 2011).

Biogenesis of let-7 is suppressed in embryonic stage and in certain cancer cells (Bussing et al., 2008). I and other groups have previously shown that let-7 maturation is inhibited by an RNA-binding protein Lin28 (**Figure II-1**) (Heo et al., 2008; Newman et al., 2008; Rybak et al., 2008; Viswanathan et al., 2008). There are two paralogues of Lin28 (Lin28A and Lin28B) in mammals that are biochemically similar but are distinct in expression patterns and subcellular localization (Balzer and Moss, 2007; Guo et al., 2006; Piskounova et al., 2011; Polesskaya et al., 2007; Richards et al., 2004; Yang and Moss, 2003). Lin28A is expressed in embryonic cells and cancer cells and localized mainly in the cytoplasm whereas Lin28B is induced in cancer cells which do not express Lin28A and is predominantly located in the nucleolus.

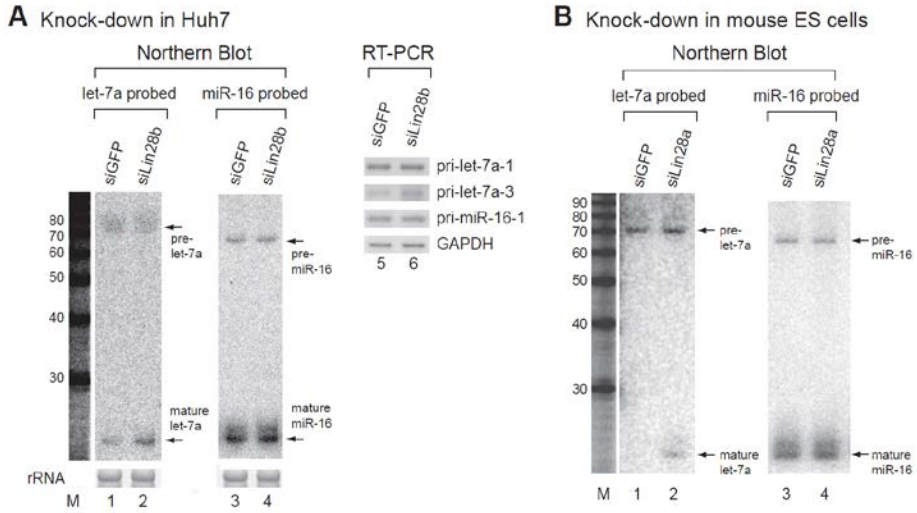


Figure II-1. Lin28 suppresses let-7 biogenesis in Huh7 and ES cells.[1]

(A) Lin28b was knocked down in Huh7 cells. MiRNA was analyzed with northern blotting and RT-PCR analyses. Pri-let-7a-2 was not detected even after 43 cycles (data not shown). **(B)** Lin28a was knocked down in mouse ES cells. MiRNA was analyzed with northern blotting.

[1] 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. Molecular cell 32, 276-284.

How Lin28 suppresses let-7 biogenesis? The Lin28 proteins bind to the terminal loop of let-7 precursors through a conserved motif GGAG to interfere with Drosha and Dicer processing (Heo et al., 2009; Loughlin et al., 2012; Nam et al., 2011). Lin28 also interacts with terminal uridylyl transferase 4 (TUT4, also known as ZCCHC11, PAPD3, and Hs3) to induce oligo-uridylation (10-30 nt) of pre-let-7 by TUT4 (Hagan et al., 2009; Heo et al., 2009). I have previously shown that biogenesis of let-7 is suppressed by TUT4 in mouse embryonic stem cells (**Figure II-2**) (Heo et al., 2009). There are seven terminal uridylyl transferases (TUTases or TUTs, also called poly (U) polymerases) in human (**Figure II-3**) (Martin and Keller, 2007; Stevenson and Norbury, 2006; Wilusz and Wilusz, 2008). Among them, my group concluded that TUT4 is the major enzyme responsible for pre-let-7 uridylation and TUT7 (also known as ZCCHC6, PAPD6, and Hs2) has a similar but weaker activity compared to that of TUT4 (Heo et al., 2009). Because Dicer disfavors a substrate with such a long single-stranded RNA tail, oligo-uridylation provides an effective way of suppressing let-7 maturation (**Figure II-4**). Uridylation has also been associated with RNA degradation and the U-tails are thought to facilitate the recruitment of exonucleases (Ji and Chen, 2012; Kim et al., 2010; Ren et al., 2012; Wickens and Kwak, 2008; Wilusz and Wilusz, 2008; Zhao et al., 2012).

Recent high-throughput studies of miRNA population in various cell types suggested that miRNAs and their precursors may undergo multiple

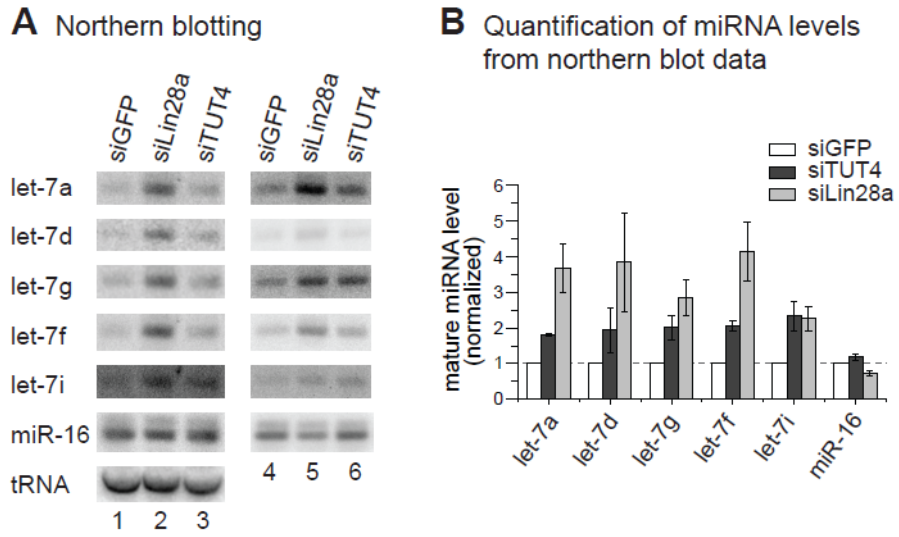


Figure II-2. Knockdown of TUT4 and Lin28 increases let-7.[2]

(A) Northern blotting analysis was carried out following the knockdown of TUT4 and Lin28a in mES cells (A3-1). **(B)** For quantification of the northern blotting, the band intensity was measured by phosphoimager and normalized against signals from siGFP-treated sample. The standard errors are from two independent experiments that include the data shown in **[A]**.

[2] Heo, I., Joo, C., Kim, Y.K., Ha, M., Yoon, M.J., Cho, J., Yeom, K.H., Han, J., and Kim, V.N. (2009). TUT4 in concert with Lin28 suppresses microRNA biogenesis through pre-microRNA uridylation. Cell 138, 696-708.

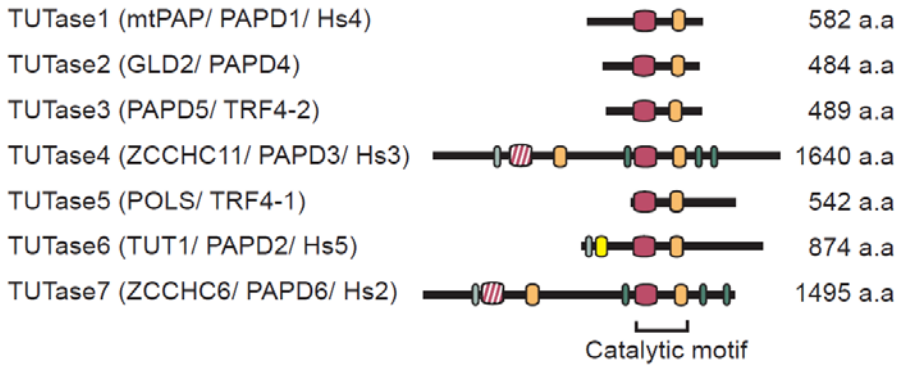


Figure II-3. Human TUTases.

Domain organization of human TUTs. (red: nucleotidyl transferase domain, orange: PAP-associated domain, hatched red: inactive nucleotidyl transferase domain due to sequence variations, light green: C2H2 zinc finger domain, green: CCHC zinc finger domain, yellow: RNA recognition motif)

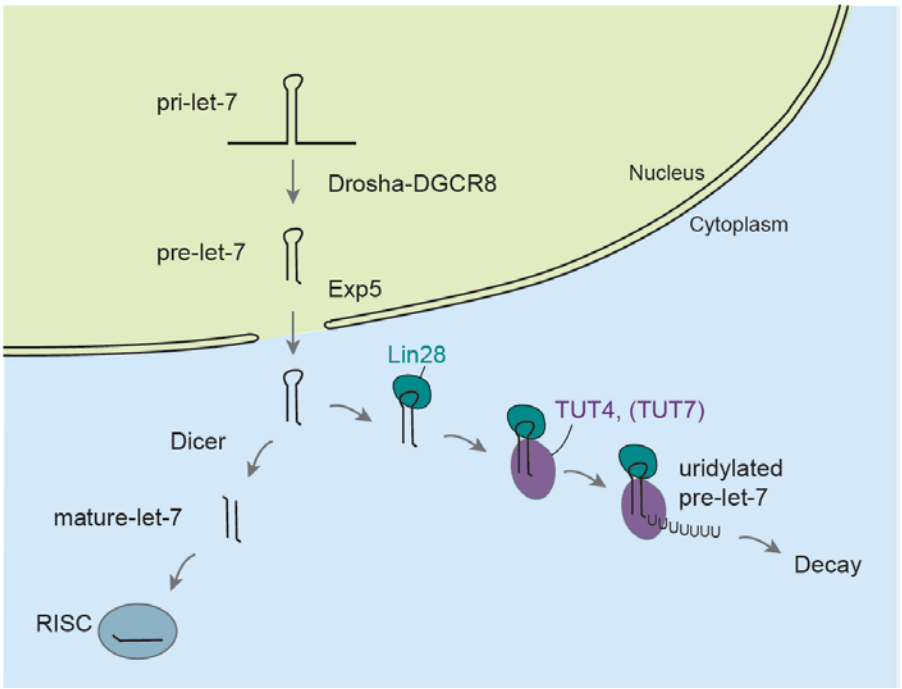


Figure II-4. Let-7 biogenesis is suppressed through oligo-uridylation by Lin28 and TUT4.[3]

Pri-miRNA is cropped into pre-miRNA by Microprocessor consisting of Drosha and DGCR8. The pre-miRNA is exported to the cytoplasm by exportin 5 (EXP5). After the export, the pre-miRNA is bound to Lin28, which interferes with Dicer processing. Lin28 recruits TUT4, which uridylates the pre-miRNA. The uridylated pre-miRNA (up-miRNA) fails to be processed by Dicer and gets degraded by nuclease(s).

[3] This figure is modified from :
Heo, I., Joo, C., Kim, Y.K., Ha, M., Yoon, M.J., Cho, J., Yeom, K.H., Han, J., and Kim, V.N. (2009). TUT4 in concert with Lin28 suppresses microRNA biogenesis through pre-microRNA uridylation. Cell 138, 696-708.

types of posttranscriptional modifications (Burroughs et al., 2010; Chiang et al., 2010; Jones et al., 2009; Newman et al., 2011; Wu et al., 2009; Wyman et al., 2011). However, most modification events appear to be rare and their functional relevance remains unclear. In the current study, I find that certain pre-miRNAs are frequently mono-uridylated and that the uridylating enzymes play critical roles in the maturation and function of miRNAs.

II-2. Results

Mono-uridylation of pre-let-7 in the absence of Lin28

To investigate post-transcriptional regulation of miRNAs at the pre-miRNA level, I cloned and sequenced pre-let-7 from HeLa cells (**Figure II-5**). Surprisingly, a considerable portion of the pre-let-7 clones (20%) carried one untemplated uridyl residue at the 3' end (**Figure II-6A**). Because HeLa cells do not express Lin28A or Lin28B (**Figure II-6B**), this result indicates that pre-let-7 may undergo uridylation even in the absence of Lin28. Other modifications appear to be rare events: only 1% carried an untemplated adenosine and there was no clone with extra guanosine or cytidine (**Figure II-6A** and **Table II-1**). Most of the remaining clones were shorter than pre-let-7, likely representing degradation intermediates.

I refer to the non-templated addition of a single uridine as “mono-uridylation” because the length of the U tail is clearly distinct from that in “oligo-uridylation” (10-30 nt) observed in Lin28-expressing cells (Hagan et al., 2009; Heo et al., 2008; Heo et al., 2009). In a recent report, Hammond and colleagues made a similar observation of mono-uridylation in various cell types (Newman et al., 2011). A caveat of my and Hammond group’s findings, however, is that only the steady-state levels were examined. Thus, it is unclear from these data whether the mono-uridylated pre-miRNAs are active precursors to be processed or non-functional dead-end products facing decay.

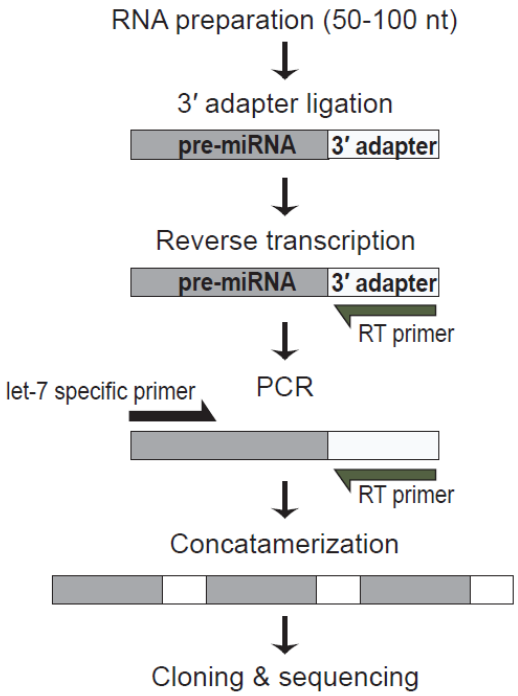


Figure II-5. Scheme of pre-let-7 sequencing.

RNAs of 50-100 nt were gel purified and ligated to 3′ adapter. The ligated RNAs were reverse transcribed with a RT primer which is complementary to the 3′ adapter. The RT products were amplified by PCR with a let-7 specific forward primer and the RT primer. The PCR products were cleaved by BanI restriction enzyme and then concatamerized using DNA ligase. The concatamers were cloned for Sanger sequencing.

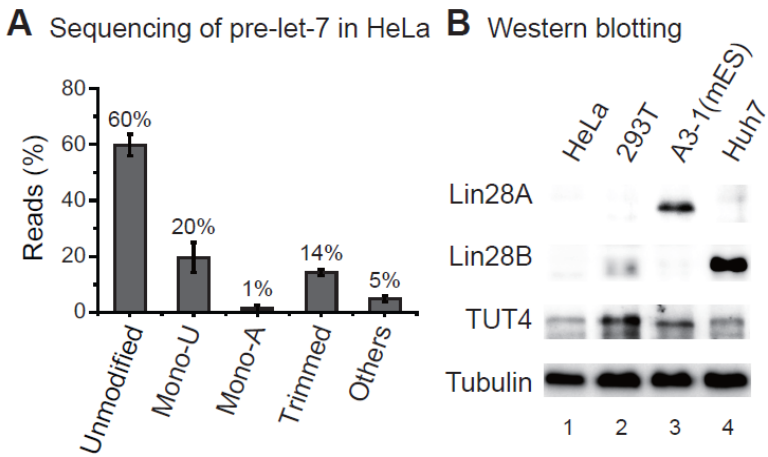


Figure II-6. Pre-let-7 is mono-uridylated in the absence of Lin28.[4]

(A) Significant amount of pre-let-7 carries an untemplated uridine at its 3′ end (Mono-U). A total of 145 pre-let-7 clones were obtained from two independent experiments using HeLa cells. "Trimmed" reads are shorter than pre-let-7 and "others" reads do not belong to any other categories (see **Table II-1**). Error bars indicate standard deviations (SD). **(B)** Expression pattern of Lin28A, Lin28B, and TUT4 in HeLa (human cervical adenocarcinoma), 293T (human embryonic kidney), A3-1 (mouse embryonic stem cell), and Huh7 (human hepatocellular carcinoma). Tubulin was detected as a loading control.

[4] These experiments were performed in collaboration with Inha Heo and Jaechul Lim.

TUT7, TUT4, and TUT2 mono-uridylate pre-let-7 in vitro

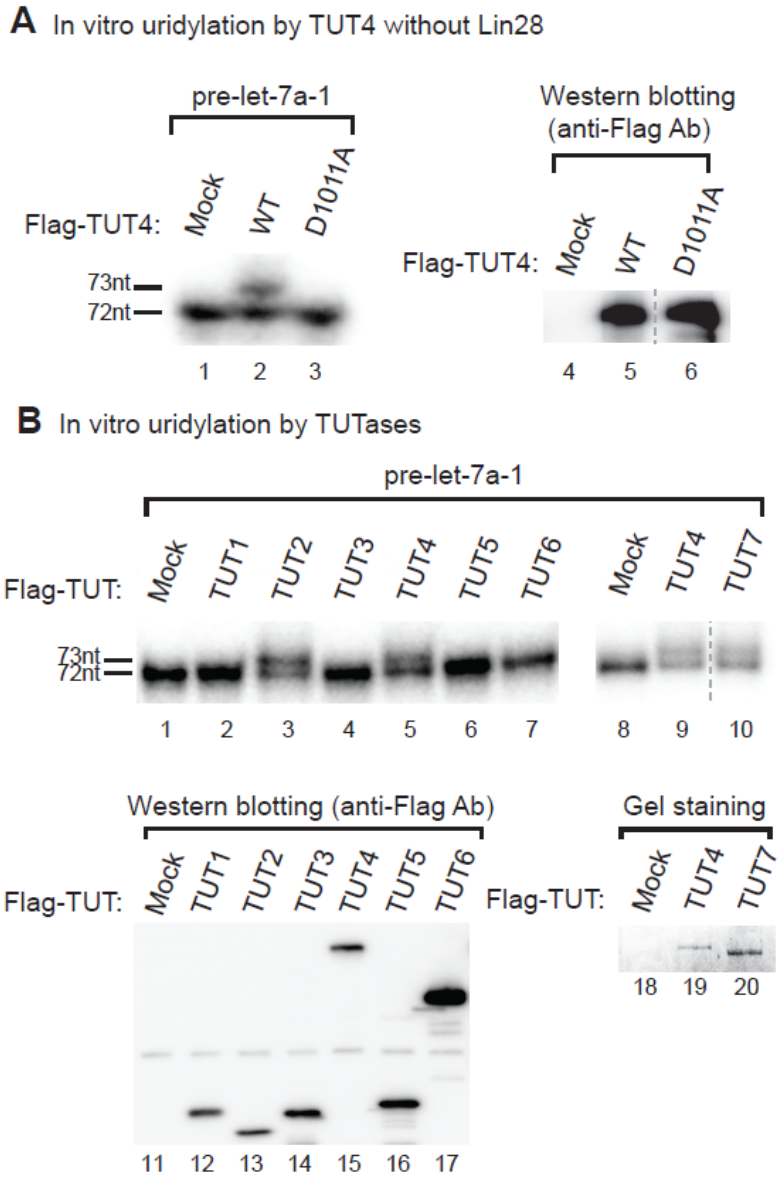
To understand the functional significance of mono-uridylation, it was critical to identify the factors responsible for the modification. Interestingly, from in vitro uridylation assay, TUT4 has a mono-uridylating activity in the absence of Lin28 (**Figure II-7A**) while the same enzyme catalyzes oligo-uridylation when Lin28 is bound (Heo et al., 2009). Pre-let-7a-1 was extended by 1-nt when 5′ end-labeled pre-let-7a-1 was incubated with wild type TUT4, but not with a catalytically dead mutant (D1011A) (**Figure II-7A**). Consistent with this result, TUT4 interacts with pre-let-7 even in the absence of Lin28, albeit transiently, as recently shown by single molecule detection technique called SIMPlex (Yeom et al., 2011). The duration of the interaction between TUT4 and pre-let-7 is very short (1.1 ± 0.2 sec), explaining why the interaction was not detected in the previous study by bulk binding assays (Heo et al., 2009).

Other terminal uridylyl transferases (TUTases or TUTs) were further examined for their activity towards pre-let-7 (Martin and Keller, 2007; Stevenson and Norbury, 2006; Wilusz and Wilusz, 2008). Seven TUTases (TUT1, TUT2, TUT3, TUT4, TUT5, TUT6, and TUT7) were immunoprecipitated for in vitro uridylation experiments (**Figure II-7B**). Interestingly, TUT7 and TUT2 as well as TUT4 can catalyze mono-uridylation of pre-let-7a-1 (**Figure II-7B**). Such mono-uridylation activity was not detected when point mutants of the conserved catalytic site in TUT7 (D1060A) or TUT2 (D215A) were used (**Figure II-8**).

Figure II-7. TUT7, TUT4, and TUT2 mono-uridylate pre-let-7 in the absence of Lin28.[5]

(A) TUT4 catalyzed mono-uridylation of pre-let-7a-1 in the absence of Lin28 (left). Immunopurified wild type or catalytically dead mutant (D1011A) TUT4 was incubated with 5′ end-labeled pre-let-7a-1 and 0.25 mM UTP. (right) Comparable amounts of proteins were used in the reaction. **(B)** Not only TUT4 but also TUT7 and TUT2 mono-uridylate pre-let-7a-1 (top). In vitro uridylation was performed using immunopurified human TUTases (TUTs). (bottom) The immunopurified proteins were visualized by either western blotting with anti-Flag-antibody or coomassie blue gel staining. Dashed line indicates discontinuous lanes from the same gel.

[5] These experiments were performed in collaboration with Mi-Jeong Yoon.


A In vitro uridylation by TUT4 without Lin28
pre-let-7a-1
Western blotting (anti-Flag Ab)
Flag-TUT4:
Mock
WT
D1011A
73nt
72nt
1 2 3
4 5 6
B In vitro uridylation by TUTases
pre-let-7a-1
Flag-TUT:
Mock
TUT1
TUT2
TUT3
TUT4
TUT5
TUT6
TUT7
1 2 3 4 5 6 7 8 9 10
Western blotting (anti-Flag Ab)
Gel staining
11 12 13 14 15 16 17
18 19 20

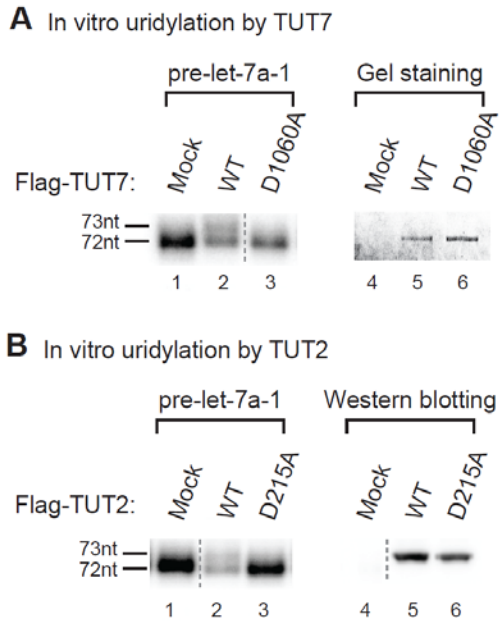


Figure II-8. TUT4 and TUT2 catalyze mono-uridylation of pre-let-7.

(A) (left) In vitro uridylation was performed by incubating 5′ end-labeled pre-let-7a-1 and 0.25 mM UTP with either wild type or catalytically dead mutant (D1060A) of TUT7. The RNA was extracted and analyzed on 6% denaturing polyacrylamide sequencing gel. (right) The immunopurified wild type and mutant TUT7 proteins were visualized by coomassie blue gel staining. Dashed line indicates discontinuous lanes from the same gel. **(B)** (left) In vitro uridylation was carried out using immunopurified wild type or catalytically dead mutant (D215A) of TUT2. Mono-uridylated pre-let-7a-1 by wild type TUT2 was demonstrated with 6% denaturing polyacrylamide sequencing gel. (right) Wild type and mutant TUT2 were immunopurified and shown by western blotting with anti-Flag antibody.

TUT7 is closely related to TUT4 (**Figure II-3**), but its oligo-uridylation activity for pre-let-7 is lower compared to that of TUT4 (Heo et al., 2009). It was unexpected that TUT2 (also known as GLD2 and PAPD4) has uridylating activity and acts on a pre-miRNA. TUT2 is known to induce translation by poly-adenylating mRNAs at synapses (Rouhana et al., 2005) and during oogenesis (Nakanishi et al., 2006). In addition, TUT2 mono-adenylates and stabilizes mature miR-122 in mammalian liver cells and fibroblasts (Burns et al., 2011; Katoh et al., 2009).

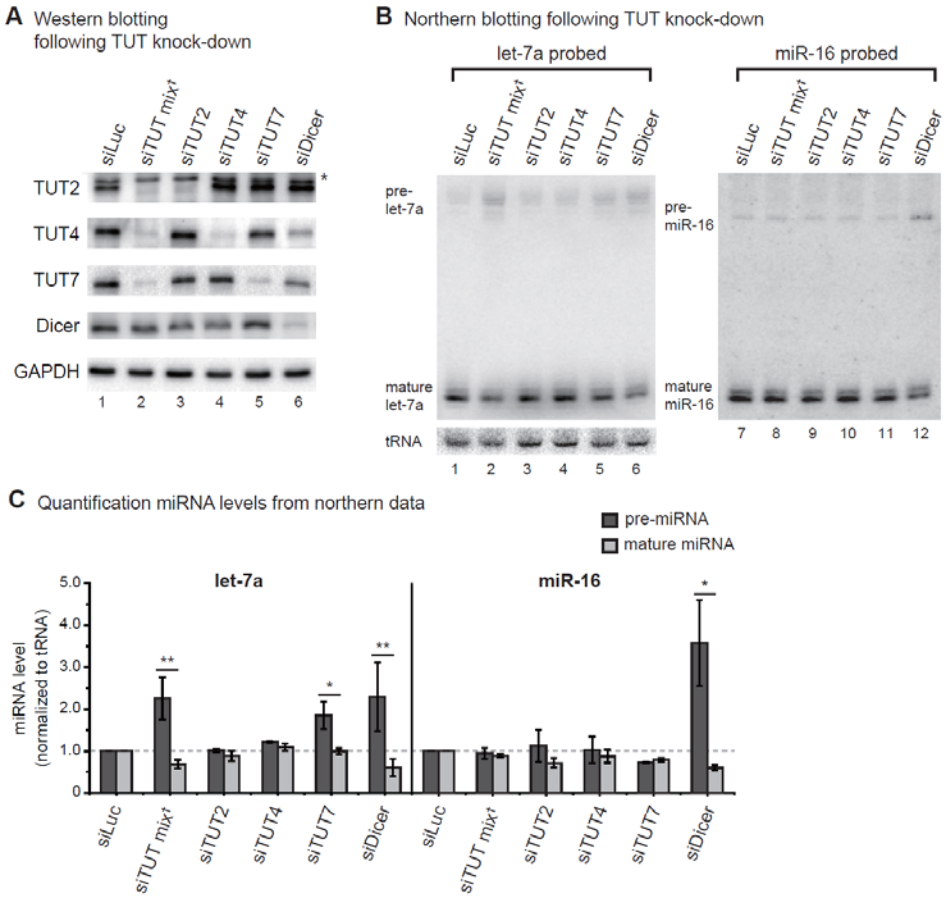
TUT7, TUT4, and TUT2 promote let-7 biogenesis in cells

I next asked whether TUT7, TUT4, and TUT2 (TUT7/4/2) have any effect on let-7 biogenesis by transfecting siRNAs against the TUTs into HeLa cells. Western blotting confirmed the depletion of each protein (**Figure II-9A**). An individual knock-down of TUTs induced marginal alteration at the levels of mature and pre-let-7 (**Figure II-9B**, lanes 3-5). Since my uridylation assay using immunoprecipitates indicated that TUT7/4/2 have similar activities (**Figure II-7**), I reasoned that the three TUTs may function redundantly in cells. Consistent with this notion, when all the three TUTs were depleted simultaneously with a mixture of three siRNAs (siTUT mix), pre-let-7 increased significantly while mature let-7 decreased (**Figure II-9B**, lane 2; **Figure II-9C**). Of note, the let-7a probe may cross-hybridize to other let-7 members due to sequence similarities.

Figure II-9. TUT7, TUT4, and TUT2 redundantly promote biogenesis of let-7.[6]

(A) TUT7, TUT4, and TUT2 proteins were depleted in HeLa cells. GAPDH was detected as a loading control. An asterisk indicates a non-specific band. **(B)** Concurrent knock-down of TUT7, TUT4, and TUT2 increased pre-let-7a levels while decreasing mature let-7a levels (left). The same membrane was probed for miR-16 (right). tRNA-lys was detected as a loading control. **(C)** The levels of mature and precursor of let-7a (left) and miR-16 (right) were quantified from two independent northern blot experiments which include the data shown in **[B]** and normalized against tRNA levels. Error bars indicate SDs. Paired one-tailed t test was used to calculate the statistical significance of decrease in the ratio of mature to pre-let-7a level (* $p<0.05$; ** $p<0.01$).

[6] These experiments were performed in collaboration with Inha Heo and Jaechul Lim.


A Western blotting following TUT knock-down
siLuc
siTUT mix1
siTUT2
siTUT4
siTUT7
siDicer
TUT2
TUT4
TUT7
Dicer
GAPDH
*
1 2 3 4 5 6
B Northern blotting following TUT knock-down
let-7a probed
miR-16 probed
pre-let-7a
mature let-7a
tRNA
pre-miR-16
mature miR-16
1 2 3 4 5 6
7 8 9 10 11 12
C Quantification miRNA levels from northern data
pre-miRNA
mature miRNA
let-7a
miR-16
miRNA level (normalized to tRNA)
5.0
4.0
3.0
2.0
1.0
0
**
*
**
*
siLuc
siTUT mix1
siTUT2
siTUT4
siTUT7
siDicer

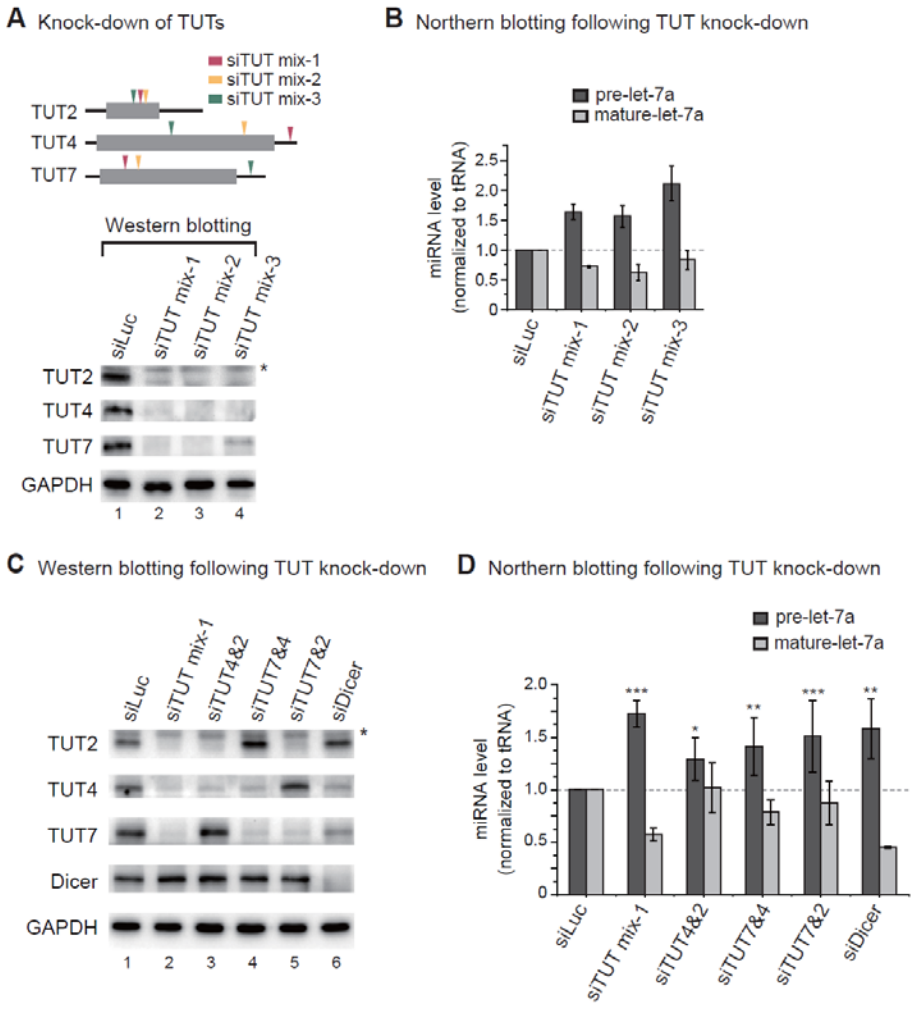
My result indicates that let-7 maturation was blocked at the pre-let-7 level after depletion of TUTs. Unlike let-7, the change in miR-16 level was insignificant in TUT-depleted cells (**Figure II-9B** lane 8**; Figure II-9C**), implicating that TUT7/4/2 may contribute specifically to let-7 maturation. Similar observations were made with two additional sets of siRNA mixtures against TUT7/4/2, ruling out a possibility of off-target effects (**Figures II-10A** and **10B**). Notably, the impact of TUT knock-down on let-7 was as strong as that of Dicer (**Figures II-9** and **II-10**), which indicates that TUT7/4/2 may play an integral role in let-7 biogenesis.

Combinatorial knock-down of two TUTs (siTUT7&4, siTUT7&2, and siTUT4&2) resulted in similar but less prominent effects on let-7 maturation compared to that of TUT7/4/2 (siTUT mix) (**Figures II-10C** and **10D**). Thus, all the three TUTs may act redundantly in let-7 biogenesis. It is noted that TUT7 depletion in single or double knock-down experiments affected let-7 maturation modestly but significantly, while that of TUT4 and TUT2 had less obvious effects (**Figures II-9C** and **10D**). These data implicate that TUT7 may be the major enzyme for pre-let-7 mono-uridylation although I cannot rule out the possibility that TUT4 or TUT2 may function dominantly in other cell types.

Figure II-10. TUT7/4/2 function redundantly in let-7 biogenesis.[7]

(A) (top) Target sites of three different sets of siTUT mixtures within the mRNAs. SiTUT mix-1 was used in other figures. Black line indicates untranslated region (UTR) and grey box marks coding region of mRNAs. (bottom) Western blotting with anti-TUT2, anti-TUT4, anti-TUT7 antibodies. GAPDH was detected as a loading control. An asterisk indicates a non-specific band. **(B)** Depletion of TUTs increased levels of pre-let-7a and decreased mature let-7a levels in the cells treated with siTUT mixtures. The levels of mature and pre-let-7a were quantified and normalized to tRNA-lys from duplicate northern blot experiments. Error bars indicate standard deviations (SD). **(C)** Simultaneous depletion of two TUTs using three different combinations of siRNAs in HeLa cells. GAPDH was detected as a loading control. An asterisk represents a non-specific band. **(D)** Concurrent knock-down of two TUTs suppressed let-7 biogenesis. The levels of mature and pre-let-7a were quantified and normalized to tRNA-lys from triplicate northern blot experiments. Among three different combinations, it was the combination of TUT4 and TUT2 that showed the poorest effect. Error bars indicate SDs. Paired one-tailed t test was used to calculate the statistical significance of increase in the ratio of pre to mature-let-7a level (* $p<0.05$; ** $p<0.01$; *** $p<0.001$).

[7] These experiments were performed in collaboration with Jaechul Lim.


A Knock-down of TUTs
siTUT mix-1
siTUT mix-2
siTUT mix-3
TUT2
TUT4
TUT7
Western blotting
siLuc
siTUT mix-1
siTUT mix-2
siTUT mix-3
TUT2
TUT4
TUT7
GAPDH
1 2 3 4
B Northern blotting following TUT knock-down
pre-let-7a
mature-let-7a
miRNA level
(normalized to tRNA)
2.5
2.0
1.5
1.0
0.5
0
siLuc
siTUT mix-1
siTUT mix-2
siTUT mix-3
C Western blotting following TUT knock-down
siLuc
siTUT mix-1
siTUT4&2
siTUT7&4
siTUT7&2
siDicer
TUT2
TUT4
TUT7
Dicer
GAPDH
1 2 3 4 5 6
D Northern blotting following TUT knock-down
pre-let-7a
mature-let-7a
miRNA level
(normalized to tRNA)
2.0
1.5
1.0
0.5
0
***
*
**
***
**
siLuc
siTUT mix-1
siTUT4&2
siTUT7&4
siTUT7&2
siDicer

TUT7, TUT4, and TUT2 are required for pre-let-7 mono-uridylation in cells

To investigate the uridylation status of pre-let-7 in HeLa cells depleted of TUT7/4/2, I performed sequencing of pre-let-7 (**Figure II-11**). The portion of mono-uridylated pre-let-7 (let-7a-1, d, f-1, f-2, and g) decreased markedly (from 20% to 3%) in siTUT mix-treated cells. This result clearly demonstrates that TUT7/4/2 are indeed required for mono-uridylation of pre-let-7 in cells.

Interestingly, the trimmed forms of pre-let-7 (mostly 1-nt shorter at the 3′ end than unmodified pre-let-7) increased considerably upon TUT knock-down (**Figure II-11** and **Table II-1**), suggesting that mono-uridylation may protect pre-miRNA from 3′-exonuclease-mediated trimming. Because the 3′ trimming enzyme for mammalian miRNA is unknown, it is currently unclear by which mechanism pre-miRNAs are degraded and how mono-uridylation influences trimming.

Mono-uridylation of pre-let-7 enhances Dicer processing

How does mono-uridylation promote let-7 biogenesis? I found that pre-let-7a-1 has an unusual end structure: a 1-nt 5′ overhang and a 2-nt 3′ overhang (**Figure II-12**, unmodified). Because this structure is equivalent to a 1-nt 3′ overhang as far as Dicer processing is concerned (Park et al., 2011), it is expected that pre-let-7a-1 is a suboptimal substrate for Dicer. This unusual end structure is generated by Drosha cleavage (not by trimming), which

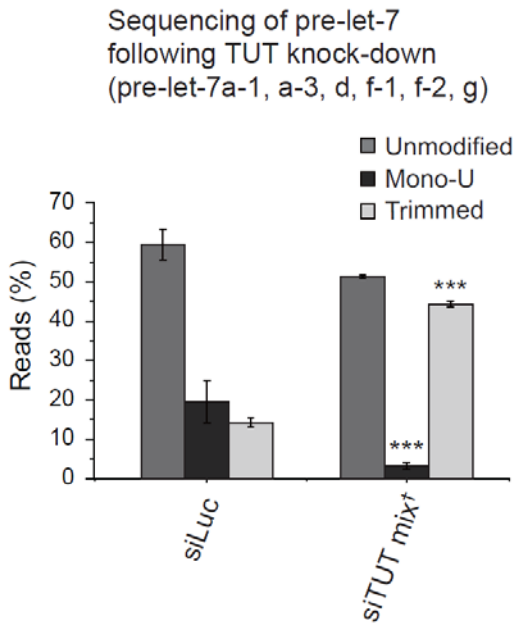


Figure II-11. TUT7, TUT4, and TUT2 are required for mono-uridylation of pre-let-7.[8]

Pre-let-7 was sequenced following the knock-down of TUTs (**Figure II-5**). A proportion of mono-uridylated pre-let-7 significantly decreased in TUT-depleted HeLa cells (*** $p<0.001$, Fisher's exact test). Percentages of each let-7 population were calculated from biological duplicates (**Table II-1**). Error bars indicate SDs. †: siTUT mix represents a mixture of equal amounts of siTUT7, siTUT4, and siTUT2, which applies for all figures.

[8] These experiments were performed in collaboration with Inha Heo and Jaechul Lim.

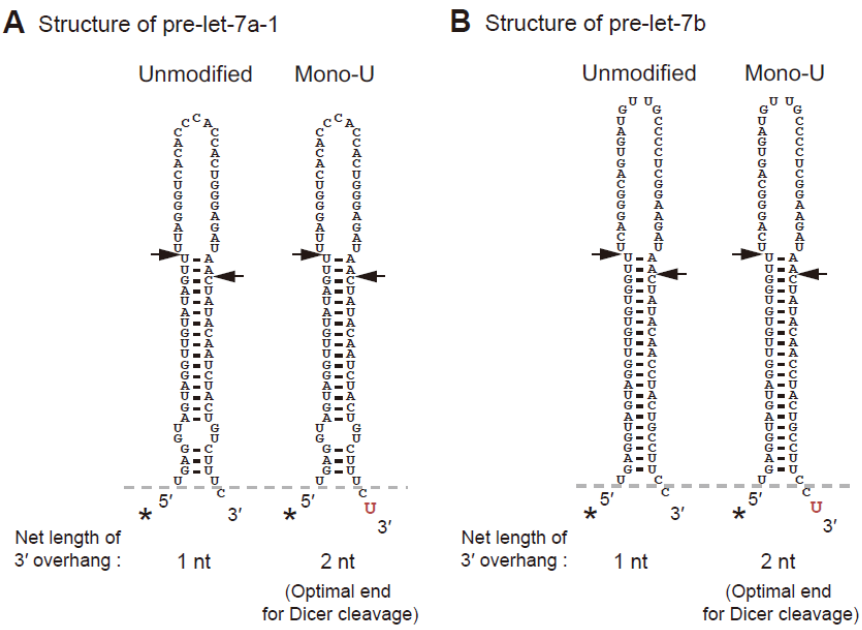


Figure II-12. Structure of pre-let-7.

(A and B) Shown are the structures of human pre-let-7a-1 and pre-let-7b. Mono-uridylation makes pre-let-7 an optimal substrate for Dicer cleavage by elongating the overhang from 1-nt to 2-nt. Arrows indicate Dicer processing sites and untemplated uridine addition is represented in red.

confirmed by in vitro Drosha processing of pri-let-7 and cloning the products from the reaction (data not shown). Given that a 2-nt 3′ overhang of pre-miRNA is favored by Dicer (Park et al., 2011; Zhang et al., 2004), I expected that mono-uridylation of pre-let-7 would create an optimal substrate for Dicer processing (**Figure II-12A**, mono-U). Consistent with this notion, in vitro assay with immunopurified human Dicer demonstrated clearly that mono-uridylated pre-let-7a-1 is processed more efficiently than the unmodified counterpart (**Figure II-13A**). Another family member, pre-let-7b, gave a similar but more dramatic result (**Figures II-12B** and **13A**). Mono-uridylated pre-let-7b was cleaved by Dicer efficiently while unmodified counterpart was barely processed, which clearly shows that mono-uridylation is necessary for efficient Dicer processing (**Figure II-13**). Taken together with the results from the knock-down experiments (**Figure II-9**), these data indicate that mono-uridylation of pre-let-7 by TUT7/4/2 promotes let-7 biogenesis by enhancing Dicer processing.

The let-7 family is subdivided into two groups based on the end structure of the precursor

In humans, nine distinct let-7 members are generated from twelve different precursors. To see whether all let-7 members are regulated by the same mechanism, the end structure of let-7 precursors were examined (**Figure II-14**).

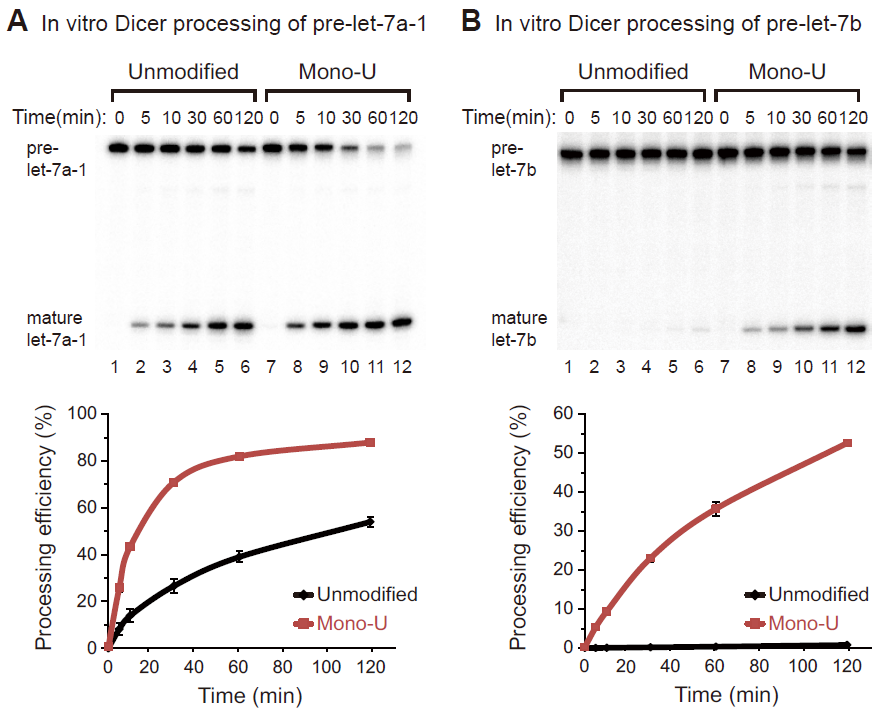


Figure II-13. Mono-uridylation of pre-let-7 enhances Dicer processing.[9]

(A and B) Mono-uridylated pre-let-7a-1 or pre-let-7b was processed more efficiently by purified Dicer than their unmodified counterparts. Processing efficiency was measured from two independent experiments. Error bars indicate SDs.

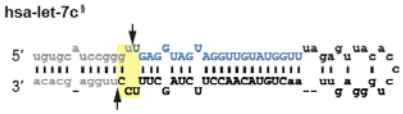
[9] These experiments were carried out by Jaechul Lim.

Figure II-14. The let-7 family is subdivided into two groups.

Structures of group I and group II pri-let-7 in human are shown. (top) The precursors of group I let-7 have a 2-nt 3′ overhang structure. Three let-7 precursors (let-7a-2, c, and e) belong to group I. On the other hand, 9 precursors (let-7a-1, a-3, b, d, f-1, f-2, g, i, and miR-98) belong to group II and carry a 1-nt 3′ overhang (bottom). Mature let-7 sequences are represented in uppercase with blue (group I let-7-5p), red (group II let-7-5p) or black (let-7-3p). Arrows indicate Drosha processing sites. §: The 3′ end of pre-let-7s (mis-annotated in miRBase) were re-determined based on our analysis.

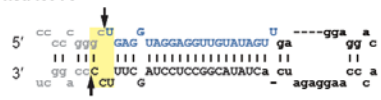
Group I let-7: pre-miRNA with a 2-nt 3′ overhang
(includes let-7a-2, c, e)

hsa-let-7c§



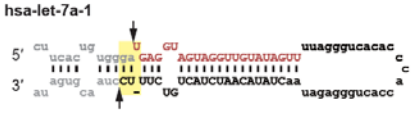
hsa-let-7a-2

hsa-let-7e

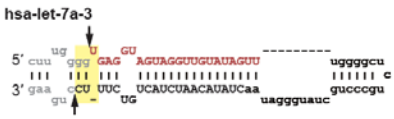


Group II let-7: pre-miRNA with a 1-nt 3′ overhang
(includes let-7a-1, a-3, b, d, f-1, f-2, g, i, miR-98)

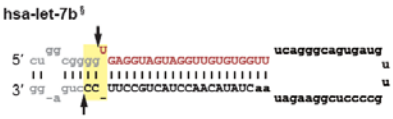
hsa-let-7a-1



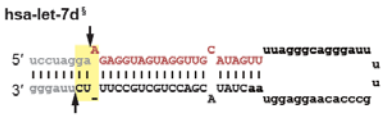
hsa-let-7a-3



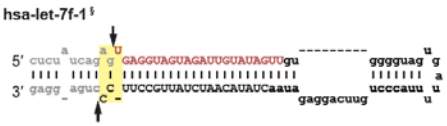
hsa-let-7b§



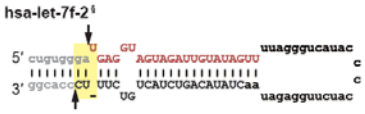
hsa-let-7d§



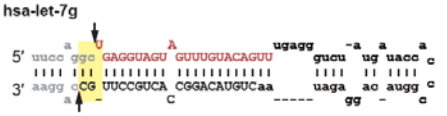
hsa-let-7f-1§



hsa-let-7f-2§



hsa-let-7g



hsa-let-7i§

hsa-miR-98§

The 3′ end of pre-let-7 was inferred based on the 3′ end of mature let-7-3p sequences from multiple small RNA deep sequencing data. In the case of let-7a-1 and let-7d, in vitro Drosha processing was performed and the products were cloned in order to annotate the exact Drosha cleavage sites (data not shown). Based on these analyses, the 3′ end of several let-7 precursors (let-7b, c, d, f-1, f-2, i, and miR-98) were re-determined that appear to be mis-annotated in miRBase database. By analyzing the end structure of precursor, three let-7 sisters (let-7a-2, c, and e) are predicted to carry a typical end structure (2-nt 3′ overhang) as seen in most other pre-miRNAs outside the let-7 family (**Figures II-14**). I refer to this prototypic subset as "group I". On the other hand, the precursors of nine let-7 miRNAs (let-7a-1, a-3, b, d, f-1, f-2, g, i, and miR-98) have a 1-nt 3′ overhang. I name this unusual class as "group II" (**Figures II-14**). The pri-miRNAs of group II let-7 contain a bulged uridine (adenosine in the case of let-7d) next to Drosha processing site. It is likely that Drosha does not recognize this bulged nucleotide which is expected to loop out without disrupting the stem, as often found in structural studies on small bulges in dsRNA (Tian et al., 2004). Thus, Drosha processing of a group II pri-miRNA would result in a 1-nt 5′ overhang (the bulged uridine) and a 2-nt 3′ overhang, which, together, is equivalent to a 1-nt 3′ overhang structure.

When the let-7-3p reads from small RNA deep sequencing libraries from various human tissues (listed in **Table II-2**) were examined, in the case

of group II let-7, mono-uridylated let-7-3p was more abundant than the unmodified let-7-3p (**Figure II-15**). For group I let-7, however, the unmodified let-7-3p was dominant (**Figure II-15**). These data suggest that mono-uridylation may be required specifically for group II let-7 members.

Group II miRNAs, but not group I miRNAs, depend on mono-uridylation

To test whether group I and group II miRNAs are controlled differentially by TUTs, I performed quantitative RT-PCR or northern blotting of miRNAs from HeLa cells depleted of TUT7/4/2. Group II let-7 (let-7a, b, and f) decreased significantly upon TUT knock-down (**Figure II-16**, red bars) while the group I let-7 (let-7c) was unaltered (**Figure II-16**, blue bar). The levels of other group I miRNAs (miR-21, miR-20a, miR-151a-3p, miR-148b, and miR-93) did not change in the TUT knock-down sample (**Figure II-16**, gray bars). This data indicate that TUT7/4/2 play a critical role in the biogenesis of group II let-7.

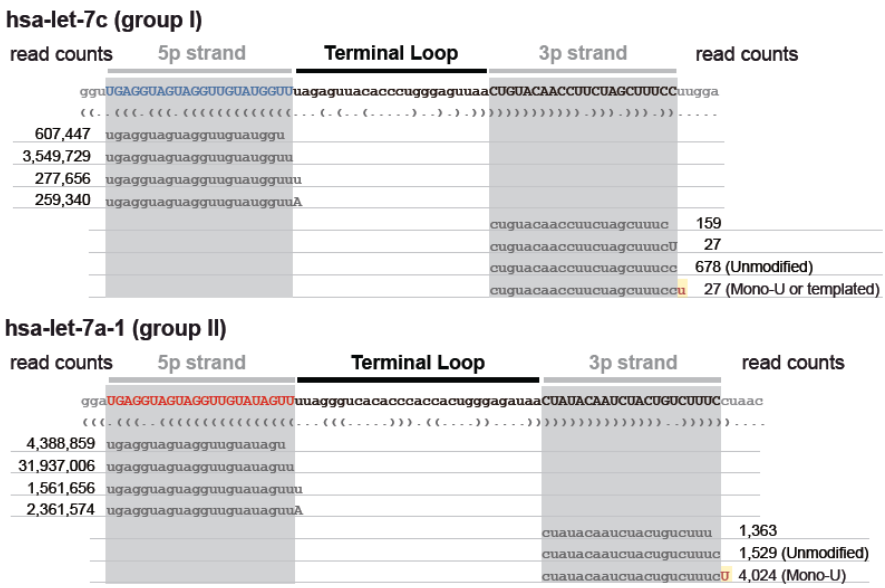
I also performed small RNA deep sequencing to investigate the global changes of miRNA upon knock-down of TUTs (**Figure II-17** and **Table II-3**). Consistent with the quantitative RT-PCR result, all of the group II let-7 decreased after depletion of TUT7/4/2 whereas group I let-7 remained unchanged or modestly increased. It is noted that some of the group I miRNAs appeared to have been affected by TUT depletion in deep sequencing data (e.g. miR-21 and miR-93 in **Figure II-17**) while I could not

Figure II-15. Mono-uridylated sequences in let-7 3p reads.[10]

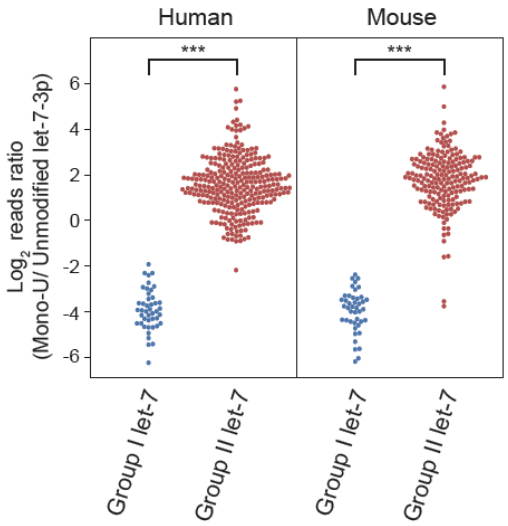
(A) Reads of let-7c (group I) and let-7a-1 (group II) from 100 different deep sequencing libraries from human cells (**Table II-2**) were aligned to their reference sequences. Reads of mono-uridylated let-7a-1* (let-7a-1-3p) are even more abundant than those of unmodified counterparts, while mono-uridylated let-7c* (let-7c-3p) is rarely detected. Reference sequences of 5p and 3p (grey box) were annotated according to the most frequent read among the reads perfectly mapped to genomic sequences in total libraries. Total read numbers of each sequence combined from all libraries are presented on the left and right side. Mono-uridylation on 3p sequence is shown in red color inside a yellow box. Uppercase in sequences indicates untemplated modification. **(B)** Bee-swarm plots show the ratio of mono-uridylated to unmodified let-7-3p read numbers. Read numbers of each let-7-3p from human and mouse libraries (listed in **Table II-2**) were used to calculate the ratio. The ratio of group II let-7-3p is significantly higher than that of group I let-7-3p over 16 folds (***$p<0.001$, Mann-Whitney U test).

[10] The analysis was carried out by Jaechul Lim.

A Reads of let-7c and let-7a-1 in deep sequencing libraries



B Relative abundance of mono-uridylation at the 3′ end of let-7-3p



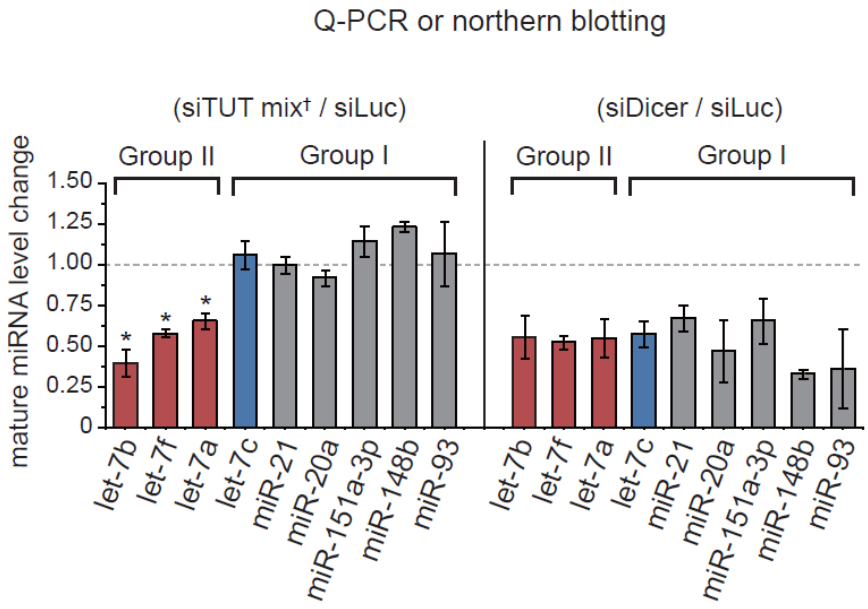


Figure II-16. TUT7, TUT4, and TUT2 are required for the biogenesis of group II let-7.[11]

The levels of mature miRNAs were determined by quantitative RT-PCR (Q-PCR) for let-7, miR-151a-3p, and miR-148b and northern blotting for miR-21, miR-20a, and miR-93. The levels of group II let-7 (let-7b, f, and a) were reduced significantly upon knock-down of TUTs while levels of other miRNAs (group I) remained largely unaltered (* $p<0.05$, Paired one-tailed t test). Note that all detected miRNAs were reduced upon knock-down of Dicer (right side). SDs are from biological duplicates.

[11] These experiments were performed in collaboration with Inha Heo and Jaechul Lim.

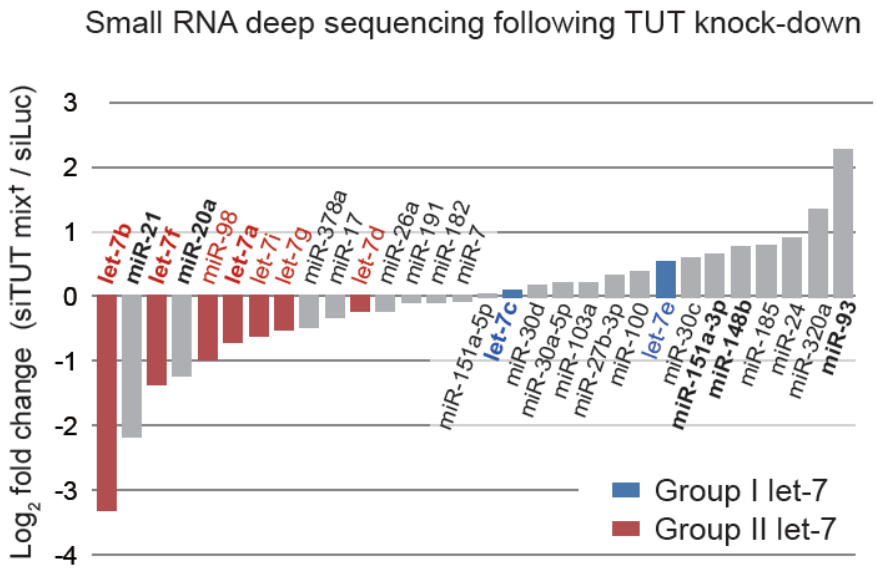


Figure II-17. Group II let-7 is specifically regulated through mono-uridylation by TUT7/4/2.[12]

Small RNA deep sequencing was performed upon TUT knock-down. Shown is the fold change of read numbers (reads per million, RPM) between siTUT mix- and siLuc-treated samples for 30 most abundant miRNAs in HeLa. Average fold changes were calculated from biological duplicates and median normalized (**Table II-3**).

[12] These experiments were performed in collaboration with Inha Heo and Hyeshik Chang.

confirm such changes by quantitative RT-PCR and northern blotting (**Figure II-16**), indicating that my sequencing data may not have been highly quantitative. Nevertheless, my results collectively demonstrate that TUT7/4/2 are required specifically for the biogenesis of group II let-7.

Note that let-7a (which are produced from three loci, let-7a-1, a-2, and a-3) was classified as group II miRNA in **Figures II-16** and **II-17** because the contribution of let-7a-2 (group I) to total let-7a population is expected to be trivial (~1%), judging from the relative abundance of let-7a-2* compared to let-7a-1*/3* from small RNA sequencing data in HeLa cells (data not shown).

TUTs specifically mono-uridylate group II miRNA

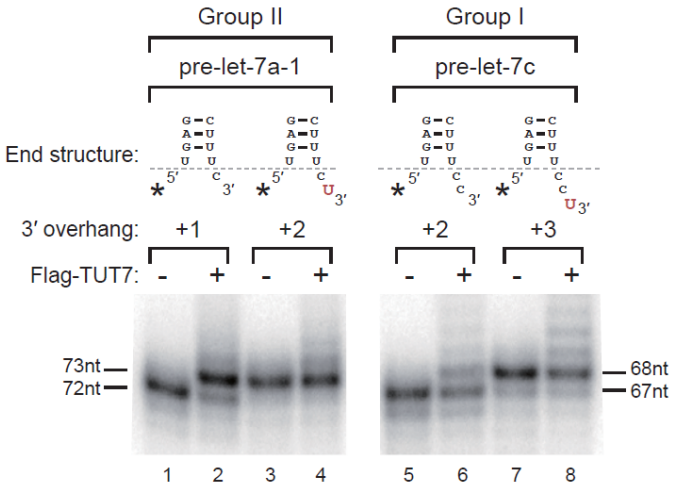
In order to understand the mechanism underlying the specificity of uridylation, I compared a group II precursor (let-7a-1) and a group I precursor (let-7c) in uridylation assay. Interestingly, pre-let-7a-1 was mono-uridylated more efficiently than pre-let-7c (**Figure II-18A**, lanes 1-2, 5-6). Another group II miRNA, pre-let-7b, was also mono-uridylated with higher efficiency compared to pre-let-7c (**Figure II-18B**, lane 1-2, 7-8). Thus, TUT7 acts selectively on group II pre-miRNAs.

I hypothesized that TUT7 may discriminate the substrates based on the 3′ overhang structure. Consistent with this notion, mono-uridylated pre-let-7a-1 was not uridylated efficiently by TUT7 (**Figure II-18A**, lanes 3-4),

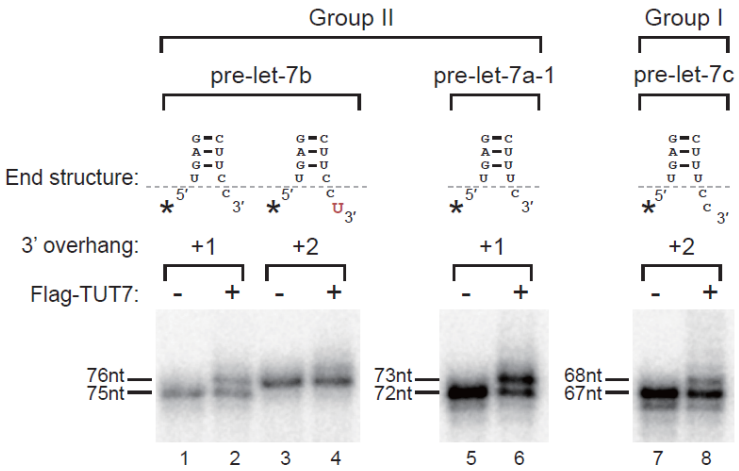
Figure II-18. TUT7 acts specifically on group II pre-miRNAs.

(A) Unmodified pre-let-7a-1 bearing a 1-nt 3′ overhang (group II) underwent mono-uridylation more efficiently than pre-let-7c (group I) and their mono-uridylated variants. In vitro uridylation assay was performed by incubating immunopurified TUT7 with 5′-end labeled RNA. **(B)** Group II pre-let-7 (unmodified pre-let-7b and a-1) with a 1-nt 3′ overhang was mono-uridylated by immunopurified TUT7 more efficiently than other substrates including group I pre-let-7c with a 2-nt 3′ overhang and their mono-uridylated variants.

A In vitro uridylation of pre-let-7a-1, and c



B In vitro uridylation of pre-let-7a-1, b, and c



indicating that TUT7 recognizes the 3′ overhang and transfers one uridylyl residue only when the pre-existing overhang is 1-nt in net length. Mono-uridylated pre-let-7a-1 may not be extended further because a 2-nt 3′ overhang is disfavored by TUT7. TUT4 and TUT2 also showed similar preference for unmodified pre-let-7a-1 (with 1-nt 3′ overhang) compared to its mono-uridylated counterpart (with 2-nt 3′ overhang) (**Figure II-19**). To further validate our hypothesis, I incubated immunopurified TUT7 with dsRNAs containing 3′ overhang of variable length (0 to 4 nt). Interestingly, only the substrate with 1-nt 3′ overhang was significantly mono-uridylated (**Figure II-20**, lane 4). My results demonstrate that TUT7 specifically recognizes dsRNAs with 1-nt 3′ overhang, to generate an optimal substrate for Dicer.

Consistently, when I cloned precursors of group I miRNAs (pre-let-7c and a-2), I found that only 2% of the precursors have an extra U at the 3′ end and that the U residue is genome-templated (**Figure II-21** and **Table II-4**). Furthermore, the proportion of mono-uridylated group I let-7 was not reduced in TUT-depleted cells (**Figure II-21**), indicating that mono-uridylation of group I is a rare event, if any.

Functional significance and evolutionary conservation of group II let-7

Group II let-7 is expressed more abundantly than group I let-7 in HeLa cells (accounting for 98% of total let-7 reads) (**Table II-3**). According to the deep

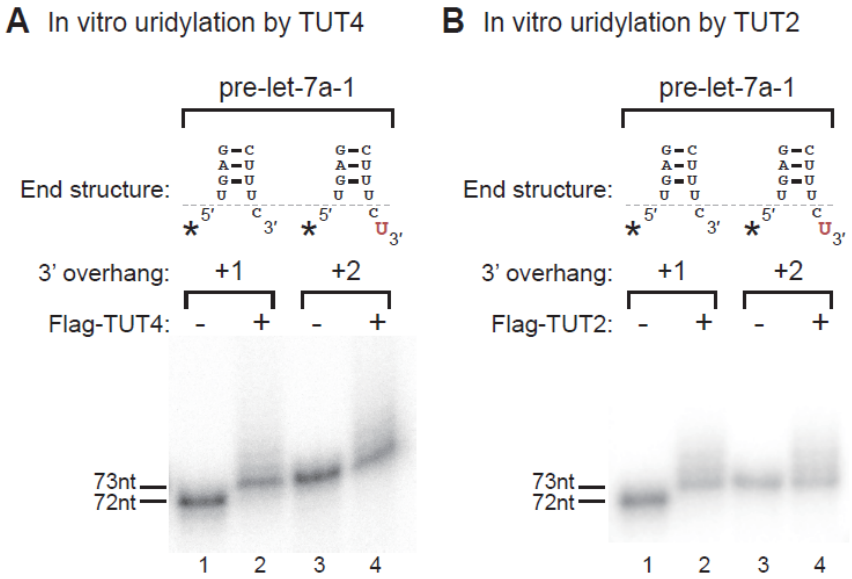


Figure II-19. TUT4 and TUT2 act specifically on pre-let-7a-1 with a 1-nt 3′ overhang.

(A) Immunopurified TUT4 specifically mono-uridylated unmodified pre-let-7a-1 with a 1-nt 3′ overhang while mono-uridylated counterpart with a 2-nt 3′ overhang was rarely mono-uridylated. **(B)** Immunopurified TUT2 mono-uridylated unmodified pre-let-7a-1 more efficiently than the mono-uridylated counterpart.

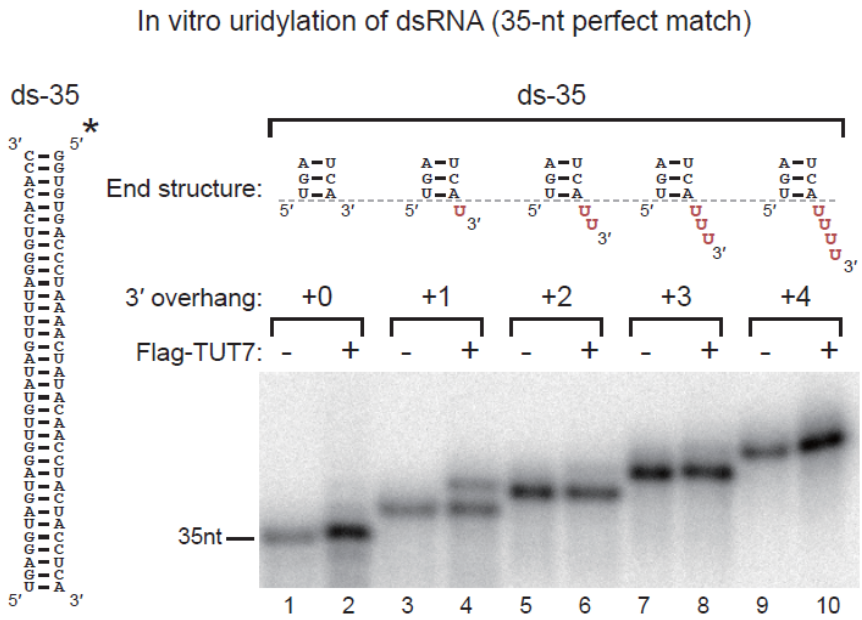


Figure II-20. TUTs act specifically on dsRNAs with a 1-nt 3′ overhang.

The optimal substrate for mono-uridylation of TUT are dsRNAs with a 1-nt 3′ overhang. Immunopurified TUT7 was incubated with dsRNAs (ds-35) carrying an extra 3′ tail of different lengths. An asterisk indicates a radio-labeled phosphate.

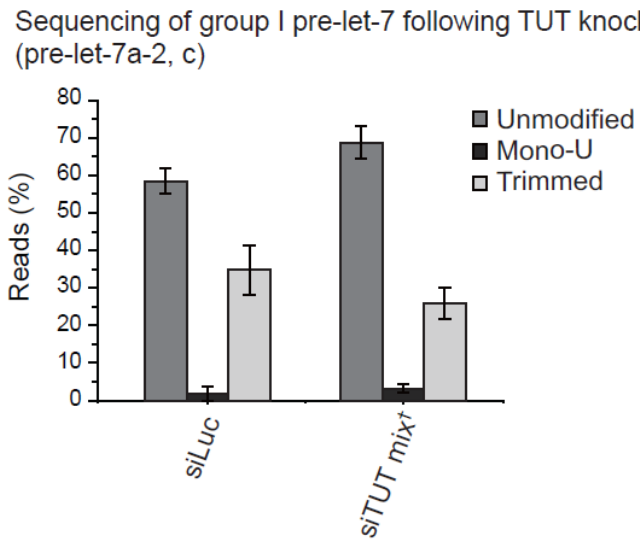


Figure II-21. Sequencing of group I pre-let-7 following TUT knock-down. Cloning of group I pre-let-7 shows that only a small portion of group I pre-let-7 carries an extra uridine in HeLa cells (2%, 2 out of 99) (See also **Figure II-6**). Percentages of each let-7 population were calculated from biological duplicates (**Table II-4**). Error bars indicate SDs.

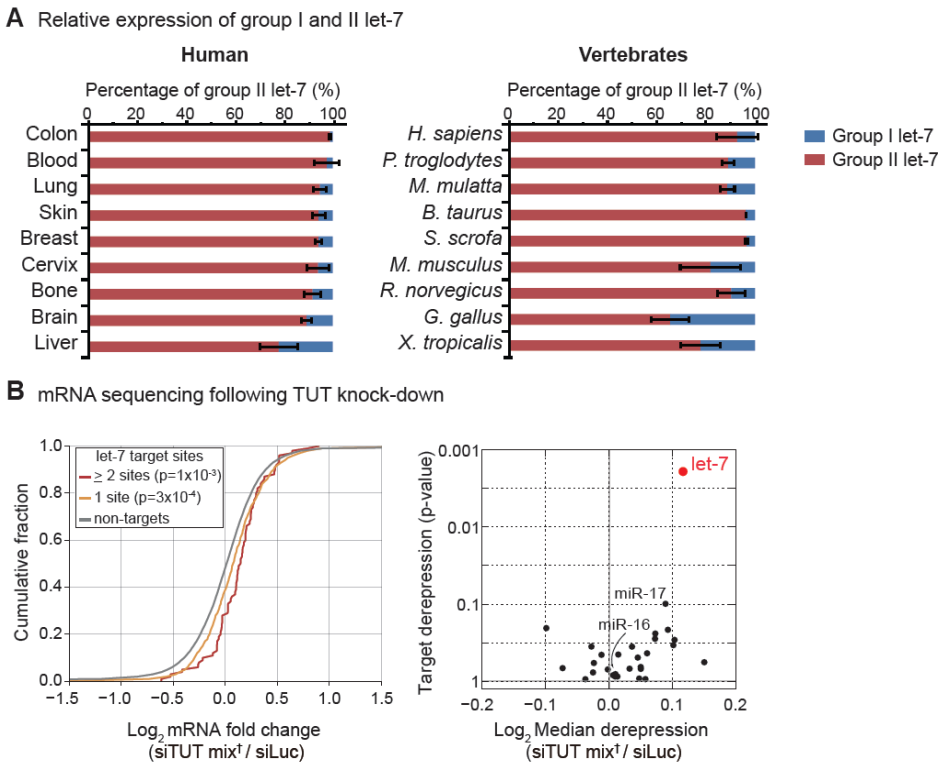
sequencing data from 104 different human cell lines that had originated from 9 different organs (**Table II-2**), group II let-7 accounts for more than 90% of total let-7 population in most human cell types (**Figure II-22A**, left panel). When other vertebrate species are examined, group II let-7 is also expressed dominantly over group I (**Figure II-22A**, right panel). Therefore, mono-uridylation-dependent control of group II let-7 is expected to have a substantial impact on the overall activity of the let-7 family. Consistently, in the mRNA sequencing from TUT7/4/2-depleted HeLa cells, the let-7 target genes (with either one or multiple target sites) were up-regulated whereas non-target genes remained unaffected (**Figure II-22B**, left panel). In addition, only the let-7 target genes were de-repressed while targets of other miRNAs did not change significantly (**Figure II-22B**, right panel). I also validated the increase of known let-7 targets (HMGA2 and NRAS) by quantitative RT-PCR upon TUT knock-down (**Figure II-22C**). Thus, TUT7/4/2 are necessary to maintain the functionality of let-7.

The structure and sequences of group II let-7 precursors are conserved in chordates (including vertebrates, tunicates, and cephalochordates), suggesting that group II may have evolved in the common ancestor of chordates (**Figure II-23**). Group I let-7 loci are found in all bilaterian animals and seem more ancient than group II. However, there are more group II let-7 loci than group I loci in vertebrates, implying that group II

Figure II-22. Mono-uridylation is critical for functionality of group II let-7.[13]

(A) Group II let-7 miRNAs are expressed predominantly in most human cells as well as in other vertebrates. Let-7 reads were analyzed in sequencing libraries from 104 different human cells (from 9 distinct organs) and from various tissues of vertebrate species (listed in **Table II-2**). Average percentages are presented with error bars corresponding to SDs. (Number of libraries used - Blood: 37, Colon: 2, Lung: 2, Cervix: 14, Skin: 25, Breast: 3, Bone: 2, Brain: 4, Liver: 15, *H. sapiens* : 104, *P. troglodytes* : 3, *M. mulatta* : 3, *B. taurus* : 1, *S. scrofa* : 7, *M. musculus* : 75, *R. norvegicus* : 16, *G. gallus* : 2, *X. tropicalis* : 2) **(B)** Let-7 target genes were specifically de-repressed in mRNA deep sequencing from TUT-depleted HeLa cells. (left) Shown is the cumulative distribution of mRNA changes. Two-sided Kolmogorov-Smirnov test was used to calculate the statistical significance of mRNA de-repression. (right) The x-axis represents a median de-repression of mRNAs which contain target sites of each miRNA. The y-axis represents a significance of mRNA de-repression calculated using two-sided Kolmogorov-Smirnov test. Target genes with 7-mer and 8-mer sites were predicted from TargetScan human release 6.1. **(C)** Q-PCR showed that mRNA levels of HMGA2 and NRAS, targets of let-7, accumulated significantly upon knock-down of TUT7/4/2 (** $p<0.01$, Paired one-tailed t test). SDs are from four independent experiments.

[13] These experiments were performed in collaboration with Jaechul Lim and Hyeshik Chang.


A Relative expression of group I and II let-7
Human
Vertebrates
Percentage of group II let-7 (%)
0 20 40 60 80 100
Colon
Blood
Lung
Skin
Breast
Cervix
Bone
Brain
Liver
H. sapiens
P. troglodytes
M. mulatta
B. taurus
S. scrofa
M. musculus
R. norvegicus
G. gallus
X. tropicalis
Group I let-7
Group II let-7
B mRNA sequencing following TUT knock-down
let-7 target sites
≥ 2 sites (p=1x10-3)
1 site (p=3x10-4)
non-targets
Cumulative fraction
Log2 mRNA fold change
(siTUT mix† / siLuc)
Target derepression (p-value)
let-7
miR-17
miR-16
Log2 Median derepression
(siTUT mix† / siLuc)
C Q-PCR of let-7 targets upon TUT knock-down
mRNA level
(normalized to GAPDH)
HMGA2
NRAS
**
siLuc
siTUT mix†

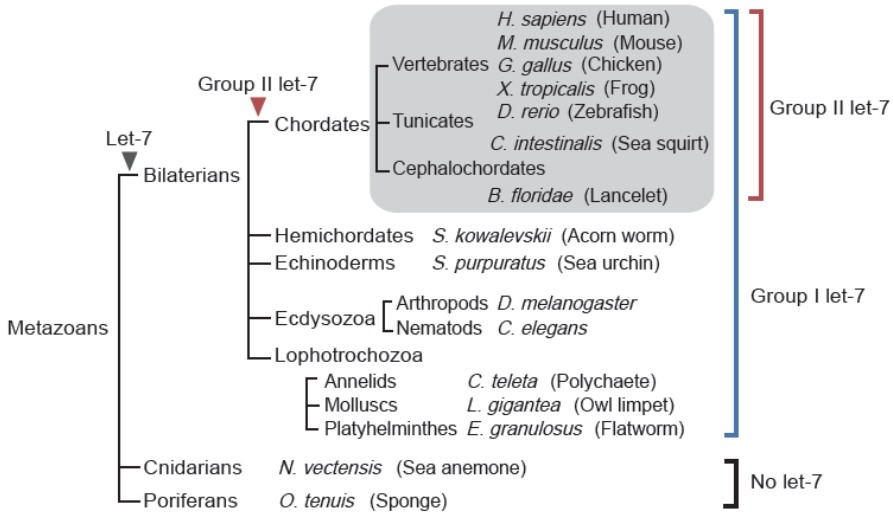


Figure II-23. Phylogenetic distribution of the let-7 miRNAs in metazoans.

Group II let-7 is conserved in chordates (grey box) while group I let-7 is present in all bilaterians, suggesting that group II let-7 may have duplicated and diverged from ancient group I let-7. Precursor structures of all let-7 miRNAs expressed in metazoans were analyzed.

may have expanded more rapidly than group I in vertebrates (Bompfunewerer et al., 2005; Roush and Slack, 2008). Thus, group II is dominant over group I in vertebrates, not only in terms of expression levels but also in gene numbers, suggesting that group II let-7 may have played an important role during vertebrate evolution. Of note, TUT7/4/2 homologs are found in all examined vertebrate species (data not shown).

II-3. Discussion

Here, I identified TUT7, TUT4, and TUT2 as novel components of the miRNA biogenesis pathway. TUT7/4/2 mono-uridylate group II pre-miRNAs and thereby play an integral part in group II miRNA maturation. It is interesting that TUT7/4/2 specifically recognize a 1-nt 3′ overhang structure and transfer only one U residue to the 3′ end. This property enables TUTs to re-shape the dsRNA end structure to provide optimal substrates for Dicer processing. Because TUT4 interacts only transiently with pre-let-7a-1 (Yeom et al., 2011), the duration of the interaction may be just long enough to transfer one nucleotide. After the first uridylation event, TUTs may dissociate quickly and cannot re-bind to the RNA due to the low affinity towards the 2-nt overhang.

Our study highlights the importance of the 2-nt 3′ overhang structure in miRNA biogenesis. Drosha talks to the downstream factors (particularly Dicer), by creating the 2-nt 3′ overhang structure which marks the RNAs involved in the miRNA pathway. This is an intriguing example of two physically separated enzymes communicating through an RNA motif. The dependence on 2-nt 3′ overhang streamlines the miRNA maturation pathway and also offers a ground for regulation through pre-miRNA uridylation. It will be interesting to understand how Dicer distinguishes exquisitely the subtle difference in the overhang structure. Structural analysis of vertebrate Dicer

protein will be necessary as the only published structure of Dicer is from Giardia intestinalis which does not have canonical miRNA pathway (Macrae et al., 2006). Giardia Dicer does not possess a preference for the 2-nt 3′ overhang structure (MacRae et al., 2007; Park et al., 2011).

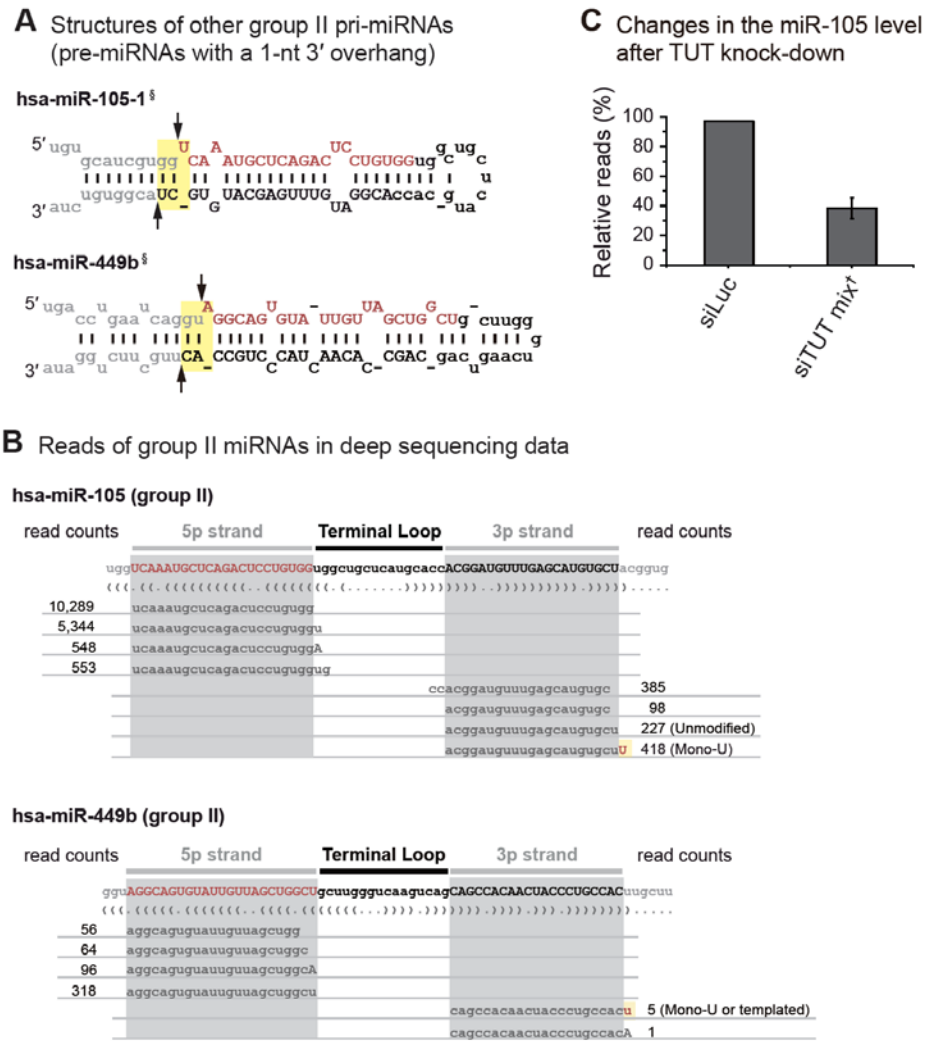
MiRNAs can be classified into group I and group II, based upon the end structure of pre-miRNA and the dependence on mono-uridylation. It is currently unclear which miRNAs outside the let-7 family belong to group II. But based on secondary structure prediction, the pri-miRNAs of miR-105 and miR-449b have a bulged nucleotide next to Drosha cleavage site, suggesting that the precursors may acquire a 1-nt 3′ overhang structure from Drosha cleavage (**Figure II-24A**). Deep sequencing data from 104 different human cells (**Table II-2**) indicated that miR-105-3p and miR-449b-3p are frequently mono-uridylated (**Figure II-24B**). Furthermore, the miR-105 level, albeit very low, decreased considerably in HeLa cells following TUT knock-down, suggesting that miR-105 is a bona fide member of group II (**Figure II-24C**). The level of miR-449b was below detection limit in HeLa cells. Further investigation will be needed to identify additional group II members.

My data suggest that the group I pre-let-7 is rarely uridylated both in HeLa cells and in vitro assays. In addition, although mono-uridylation has been observed on other group I pre-miRNAs (Newman et al., 2011), the frequency of such uridylation was low and the functional significance of the modification is unclear. Group I miRNAs remained largely unaffected upon

Figure II-24. Other group II miRNAs.[14]

(A) Pri-miRNA structure of human miR-105 and miR-449b which are group II miRNA candidates. Mature miRNA sequences are in uppercase with red (miRNA-5p) or black (miRNA-3p). Arrows indicate Drosha processing sites. § indicates re-annotation of Drosha cleavage site according to our analysis. **(B)** Reads of miR-105 from 104 different human libraries (**Table II-2**) were aligned to its reference sequences (top). Reads of mono-uridylated miR-105* (miR-105-3p) were more abundant than those of unmodified form similarly to group II let-7 (**Figure II-15**). Although reads of miR-449b were obtained from 104 different human libraries (bottom), read numbers were too low to analyze. Reference sequences of 5p and 3p (grey box) were annotated based on the most frequent read which is perfectly mapped to genomic sequences. Sequences of the reads are shown in lowercase and total read numbers of each sequence from all libraries are presented on left and right side of 5p and 3p sequences, respectively. Mono-uridylation on 3p sequence is shown in red. Uppercases indicates untemplated modification. **(C)** miR-105 decreased in deep sequencing libraries from TUT-depleted HeLa, implying that miR-105 may also require the TUTs for its efficient maturation. Reads per million (RPM) of miR-105 in siTUT mix-treated cells were compared to that in control cells. Error bars indicate SDs from biological duplicates.

[14] The experiments and analysis were performed in collaboration with and Inha Heo and Jaechul Lim.


A Structures of other group II pri-miRNAs (pre-miRNAs with a 1-nt 3′ overhang)
hsa-miR-105-1 §
hsa-miR-449b §
C Changes in the miR-105 level after TUT knock-down
Relative reads (%)
100
80
60
40
20
0
siLuc
siTUT mix†
B Reads of group II miRNAs in deep sequencing data
hsa-miR-105 (group II)
read counts
5p strand
Terminal Loop
3p strand
read counts
10,289
5,344
548
553
385
98
227 (Unmodified)
418 (Mono-U)
hsa-miR-449b (group II)
read counts
5p strand
Terminal Loop
3p strand
read counts
56
64
96
318
5 (Mono-U or templated)
1

TUT7/4/2 depletion in our experiments (**Figures II-16** and **II-17**). Thus, although it is possible that other TUTases may control group I miRNAs in certain conditions or cell types, mono-uridylation of group I pre-miRNAs may not be frequent enough to exert a meaningful effect in HeLa cells.

TUT7/4/2 mediate mono-uridylation of pre-let-7 in cells lacking Lin28, and thereby facilitate Dicer processing and compete with exonucleases. This mechanism would effectively shift the balance in favor of biogenesis in the absence of Lin28. Because uridylation is generally associated with RNA degradation (Ji and Chen, 2012; Kim et al., 2010; Ren et al., 2012; Wickens and Kwak, 2008; Wilusz and Wilusz, 2008; Zhao et al., 2012), it was unexpected that TUT knock-down resulted in an increase of trimmed pre-let-7 (**Figure II-11**), providing an intriguing case where uridylation may have a protective effect.

II-4. Materials and Methods

Pre-let-7 Sequencing

20-50 μg of HeLa total RNA were resolved on 15 % urea-polyacrylamide gel and RNAs of 50-100 nt were gel purified. Size fractionated RNAs were ligated to 3′ adapter using T4 RNA ligase 2, truncated K227Q (NEB). The ligated RNAs were reverse transcribed with a RT primer that is complementary to the 3′ adapter using superscript II (Life Technologies), followed by PCR amplification with the RT primer and a let-7a-specific (or let-7c-specific) forward primer which contains BanI restriction enzyme sites. Because of sequence similarities, the forward primers can hybridize to other let-7 members. The PCR products were cleaved using BanI (NEB) and then concatamerized using T4 DNA ligase (NEB). The concatamerized DNAs were cloned for Sanger-sequencing. The sequences of 3′ adapter and primers are listed in **Table II-5**.

Western Blotting

Cells were lysed in 0.1 % Triton-X100-Buffer D (200 mM KCl, 10 mM Tris-HCl [pH8.0], 0.2 mM EDTA, 0.1 % Triton-X100). 35-50 μg of each protein sample were separated on 6 % or 10% SDS-PAGE gel and transferred to a methanol-activated PVDF membrane (GE Healthcare). The membrane was blocked for 1 hr in PBS-T containing 5% milk and subsequently probed with

primary antibody for 2 hrs at room temperature. After washing three times with PBS-T, the membrane was probed with anti-mouse or anti-rabbit HRP-conjugated secondary antibodies (Jackson ImmunoResearch Laboratories). The protein levels were detected with West Pico Luminol reagents (Thermo Scientific) using ChemiDoc™ XRS+ System (Biorad). Anti-TUT7 (HPA020620, Sigma), anti-TUT2 (ab103884, Abcam), anti-GAPDH (sc-32233, Santa Cruz), anti-Lin28A (ab46020, Abcam), and anti-Tubulin (ab52866, Abcam) antibodies were used for primary antibodies. Anti-TUT4, anti-Dicer, and anti-Lin28B polyclonal antibody were generated from rabbit.

Immunoprecipitation and In vitro Uridylation

For immunoprecipitation of Flag-TUTases, HEK293T cells were collected 48 hrs after transfection with Flag-TUTase expression plasmids. The cells were incubated in Buffer D (200 mM KCl, 10 mM Tris-HCl [pH8.0], 0.2 mM EDTA) for 20 min followed by sonication on ice and centrifugation twice for 15 min at 4 ºC. The supernatant was incubated with 5 μL of anti-Flag antibody-conjugated agarose beads (anti-Flag M2 affinity gel, Sigma) with constant rotation for 1 hr at 4ºC. The beads were washed six times with Buffer D. In vitro uridylation reaction was performed in a total volume of 30 μL in 3.2 mM $MgCl_2$, 1 mM DTT, 0.25 mM UTP, 5′ end labeled pre-miRNA or dsRNA of $1x10^4$ - $1x10^5$ cpm, and 15 μL of immunopurified proteins on beads (**Figures II-7** and **II-8**) or 3x Flag-peptide (Sigma) eluted proteins

(**Figures II-18 - II-20**) in Buffer D. The reaction mixture was incubated at 37 ºC for 10 (**Figures II-18 - II-20**) or 20 min (**Figures II-7** and **II-8**). The RNA was purified from the reaction mixture by phenol extraction and run on 6 % urea polyacrylamide sequencing gel (20 x 40 cm, 0.4mm thick) at constant 1500 V for 2 hrs. Unmodified and mono-uridylated form of pre-let-7 (let-7a-1, b, and c) and dsRNAs were synthesized by ST Pharm. The pre-miRNAs or dsRNAs were radio-labeled at the 5′ end with T4 polynucleotide kinase (Takara) and [γ-32P] ATP. The sequences of pre-miRNAs and dsRNAs are listed in **Table II-5.**

Mutagenesis

Point mutations were introduced into the conserved catalytic sites of TUT7 and TUT2 using QuickChange Site-Directed Mutagenesis Kit (Stratagene). Aspartic acid at either 1060 in TUT7 or 215 in TUT2 was changed to alanine. The catalytic mutant of TUT4 was described in the previous study (Heo et al., 2009). The primer sequences for the mutagenesis are listed in **Table II-5**.

Cell Culture and Transfection

HeLa and HEK293T cells were maintained in DMEM (Welgene) supplemented with 10% fetal bovine serum (Welgene). For RNAi, HeLa cells were transfected with 42 nM of siRNA using Lipofectamine 2000 (Life Technologies). For simultaneous knock-down of three TUTs, equal amounts

(14 nM) of siTUT7, siTUT4, and siTUT2 were combined. For single knock-down, 14 nM of each siRNA against TUTs was mixed with 28 nM of siLuciferase (siLuc) to adjust final concentration of siRNA (42 nM). For concurrent knock-down of two TUTs, 14 nM of each siRNAs against TUT were combined with 14 nM of siLuc. Transfection was performed two times over 4-5 days. For ectopic expression of proteins, HEK293T cells were transfected with plasmids using calcium-phosphate method. The sequences of siRNAs are listed in **Table II-5**.

Northern Blotting

10-50 μg of total RNA were resolved on 15 % urea-PAGE gel and transferred to Zeta-probe GT membrane (Biorad). Ultra-violet was used for crosslinking of RNA to the membrane. 5′ end labeled oligonucleotides with [γ-32P] ATP were used as probes for northern blotting. The membrane was exposed to phosphor imaging plate (Fujifilm) and read by BAS-2500 system (Fujifilm). The sequences of northern probes are listed in **Table II-5**.

In vitro Dicer Processing

To obtain human Dicer protein with high purity, we carried out tandem purification methods following ectopic expression of Flag-tagged Dicer in HEK293T cells. Dicer cleavage reactions were performed in a total volume of 60 μl in 10 mM Tris [pH 8.0], 0.1 mM EDTA, 100 mM KCl, 2 mM $MgCl_2$, 1

mM DTT, 1 unit/ul Ribonuclease inhibitor (Takara), 5′ end labeled pre-miRNA of 2.5nM (or 5.0nM), and 2 μl of the purified human Dicer. The reaction mixture was incubated at 37 ºC, and 10 μl aliquots were taken after 5, 10, 30, 60, and 120 min. The RNA was purified from the reaction mixture by phenol extraction and separated on 15 % urea polyacrylamide gel. Unmodified and mono-uridylated form of pre-let-7a-1 and pre-let-7b were synthesized by ST Pharm. The pre-miRNAs were labeled at the 5′ end with T4 polynucleotide kinase (Takara) and [γ-32P] ATP. The sequences of pre-miRNAs are listed in **Table II-5**.

Analysis of the End Structure of Pre-miRNAs

To infer the end sequence of pre-miRNAs, we used mature miRNA sequences from multiple small RNA deep sequencing libraries (listed in **Table II-2**). We predicted the 5′ end of pre-miRNAs based on the 5′ end of the most frequent read of miRNA-5p and inferred the 3′ end of pre-miRNAs based on the 3′ end of the most frequent read of miRNA-3p. When the most frequent read of miRNA-3p carries untemplated nucleotide addition at its 3′ end (in the case of most group II let-7), we considered the 3′ end of templated part as the 3′ end of pre-miRNA. In the case of let-7d, the most frequent read of let-7d-3p carries a templated uridyl residue as the last nucleotide. To verify the exact 3′ end of pre-let-7d, we performed in vitro Drosha processing and sequenced the produced pre-miRNA. The most frequent read of let-7i also contains a

templated uridine at its 3′ end. Although we did not carry out sequencing of pre-let-7i processed by in vitro Drosha processing, the secondary structure (a bulged uridine at the 5′ strand) next to Drosha cleavage site is highly similar to that of the other group II let-7. Thus, the last uridyl residue is likely to be generated by mono-uridylation. The secondary structures of pri- and pre-miRNAs were predicted using M-fold program and miRBase 18 (http://www.mirbase.org/).

Meta-analysis of High-throughput Small RNA Sequencing

Small RNA sequencing data were downloaded from NCBI GEO datasets, which include libraries from normal tissue and cell lines (listed in **Table II-2**). All small RNA sequences except low-quality reads were used in analysis. Adaptor sequences of each library were clipped and the same sequences were collapsed. Reads which perfectly matched to nucleotide position 1 to 20 of each miRNA were employed in analysis. Group I let-7 and group II let-7 were estimated based on end structure of pre-let-7 family in each species using miRBase release 18 and UCSC genome browser. If mature miRNAs are generated from different precursors which belong to different groups, we divided the mature miRNA reads to either group based on the expression ratio of 3p miRNA (let-7*). For instance, reads of mature let-7a which are originated from pre-let-7a-1 (group II), pre-let-7a-2 (group I), and pre-let-7a-3 (group II) were distributed to each group according the relative expression of

let-7a-3p reads. The list below is the classification of vertebrate let-7 family into group I and II miRNA.

P. troglodytes: Group I (let-7a-2, let-7c, and let-7e), Group II (let-7a-1, let-7a-3, let-7b, let-7d, let-7f-1, let-7f-2, let-7g, let-7i, and miR-98)

M. mulatta: Group I (let-7a-2, let-7c, and let-7e), Group II (let-7a-1, let-7a-3, let-7b, let-7d, let-7f-1, let-7f-2, let-7g, let-7i, and miR-98)

B. taurus: Group I (let-7a-2, let-7c, and let-7e), Group II (let-7a-1, let-7a-3, let-7b, let-7d, let-7f-1, let-7f-2, let-7g, let-7i, and miR-98)

S. scrofa: Group I (let-7a-1, let-7c, and let-7e), Group II (let-7a-2, let-7f, let-7g, let-7i, and miR-98)

M. musculus: Group I (let-7a-2, let-7c-1, let-7c-2, and let-7e), Group II (let-7a-1, let-7b, let-7d, let-7f-1, let-7f-2, let-7g, let-7i, and miR-98)

R. norvegicus: Group I (let-7a-2, let-7c-1, let-7c-2, and let-7e), Group II (let-7a-1, let-7b, let-7d, let-7f-1, let-7f-2, let-7i, and miR-98)

G. gallus: Group I (let-7a-2 and let-7c), Group II (let-7a-1, let-7a-3, let-7b, let-7d, let-7f, let-7g, let-7i, let-7j, and let-7k)

X. tropicalis: Group I (let-7a and let-7c), Group II (let-7b, let-7e-1, let-7e-2, let-7f, let-7g, let-7i, and miR-98)

The full genus and species names are *Homo sapiens, Pan troglodytes, Macaca mulatta, Bos Taurus, Sus scrofa, Mus musculus, Rattus norvegicus, Gallus Gallus, Xenopus tropicalis.*

Quantitative RT-PCR

Total RNA was extracted from cells using TRIzol® reagent (Life Technologies) and treated with DNase I (Takara). For reverse transcription, 0.8-3 μg of total RNA was used with M-MuLV Reverse Transcriptase (Fermentas). Mature miRNA levels were determined by TaqMan MicroRNA Assays (Applied Biosystems) and normalized to the level of U6 snRNA. The levels of mRNA were analyzed by SYBR® Green assays (Applied Biosystems) using ABI Step One real-time PCR machine and normalized to the level of GAPDH. The comparative cycle threshold (Ct) method was applied to determine miRNA and mRNA levels. The primers used for PCR are presented in **Table II-5**.

Small RNA Library Preparation

To prepare small RNA cDNA libraries, total RNAs were separated on 15 % urea-PAGE and RNAs of 17-30 nt were gel-purified. Size fractionated RNAs were ligated to 3′ adapter using T4 RNA ligase 2, truncated K227Q (NEB). The 3′ adaptor-ligated RNA was hybridized with an RT primer which is complementary to the 3′ adapter and then the RNA was ligated to 5′ adapter using T4 RNA ligase 1 (NEB). The ligation products were reverse transcribed using superscript II (Life Technologies) and the cDNA was PCR-amplified with Phusion DNA polymerase (NEB). The cDNA libraries were sequenced using Illumina Genome Analyzer II.

Analysis of Small RNA Sequencing Data

To generate quality filtered short reads, the 3′ adaptor sequences were clipped and reads which have short (≤17 nucleotides) or low quality bases (<25 in phred quality) were removed using FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/). The filtered sequences were mapped to the human genome hg18 assembly, which is downloaded from the University of California at Santa Cruz (UCSC; http:// genome.cse.ucsc.edu/) using the GSNAP (Wu and Nacu, 2010) with 6% of maximum mismatches. Annotations for aligned regions were retrieved using in-house software from RefSeq, RepeatMasker and miRBase (downloaded from UCSC or miRBase on Aug 25, 2011). Data obtained in this study have been deposited under GEO accession number GSE 35835 in the National Center for Biotechnology Information Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/).

mRNA Library Preparation

mRNAs were purified from total RNA using Dynabeads mRNA Purification Kit (Life Technologies, 61011). Purified mRNAs were fragmented by RNA Fragmentation Reagents (Life Technologies, AM8740). After fragmentation, phosphate group at 3′ end was removed by Antarctic phosphatase (NEB, M0289L), and RNAs were 5′ phosphorylated by T4 PNK (NEB). Directional and multiplexed mRNA libraries were generated using TruSeq Small RNA

Sample Preparation Kit (Illumina, RS-200–0012) and sequenced using Illumina HiSeq 2000.

Analysis of mRNA Sequencing Data

Sequence reads were aligned to UCSC hg19 genome assembly using bowtie version 2.0.0-beta5 (Langmead and Salzberg, 2012) and tophat version 2.0.0 (Trapnell et al., 2009) with RefSeq annotation downloaded on Sep 28, 2011. The differentially regulated genes were identified using cuffdiff tool in cufflinks 2.0.0 (Roberts et al., 2011) with default options for non-discovery tasks. The predicted targets of microRNAs were retrieved from TargetScan human release 6.1 (http://www.targetscan.org/). All genes expressed more than 1 FPKM (fragments per kilobase of exon per million fragments mapped) in siLuc-treated sample were divided into non-targets, 1 site and ≥ 2 sites groups based on numbers of 7-mer and 8-mer target sites on 3′ UTR. Indirect miRNA target derepression effect of knockdown of TUTases were examined by two-sided Kolmogorov-Smirnov test between distributions of log2 fold changes of targets (1 site or ≥ 2 sites group) and non-targets using R stats package.

Phylogenetic Analysis of the Structures of let-7 Precursor

All sequences and structures of the let-7 stem-loop in bilaterians were downloaded from miRBase release 18 and UCSC genome browser. The end

structures of each pre-let-7 were manually classified into group I and group II. The phylogenetic tree illustrated in **Figure II-23** was modified from the earlier studies (Niwa and Slack, 2007; Pasquinelli et al., 2003; Pasquinelli et al., 2000) and miRBase 18. The full genus and species names are as follows: *Homo sapiens, Mus musculus, Gallus gallus, Xenopus tropicalis, Danio rerio, Ciona intestinalis, Branchiostoma floridae, Saccoglossus kowalevskii, Strongylocentrotus purpuratus, Drosophila melanogaster, Caenorhabditis elegans, Capitella teleta, Lottia gigantean, Echinococcus granulosus, Nematostella vectensis, Ophlitaspongia tenuis.*

Accession Numbers

The GEO accession number for the small RNA sequencing and mRNA sequencing data is GSE40236.

Table II-1. Results of pre-let-7 sequencing following knock-down of TUTs in HeLa.

(top) Shown are the sequences of each read and its read numbers in the results of pre-let-7 sequencing. Although let-7a specific primer (purple box) was used in the amplification of sequences by PCR, not only pre-let-7a but also other let-7 family precursors were detected due to the high sequence homology. Trimmed sequences indicate reads which are shorter than unmodified pre-let-7. The reads that do not belong to unmodified, mono-U, mono-A, or trimmed are described in 'others'. **(bottom)** Shown are the detailed sequences of trimmed and the other pre-let-7 reads. Red indicates mono-uridylation, and blue represents untemplated modification.

			read numbers			
let-7a specific PCR primer	GTAGTAGGTTGTATAGTTTTAG		siLuc		siTUT mix†	
pre-let-7a-1	UGAGGUAGUAGGUUGUAUAGUUUUAGGGUCACACCCACCACUGGGAGAUAACUAUACAAUCUACUGUCUUUC	Description	Set-1	Set-2	Set-1	Set-2
reads	GUAGUAGGUUGUAUAGUUUUAGGGUCACACCCACCACUGGGAGAUAACUAUACAAUCUACUGUCUUUC	Unmodified	23	20	21	31
	GUAGUAGGUUGUAUAGUUUUAGGGUCACACCCACCACUGGGAGAUAACUAUACAAUCUACUGUCUUUCU	Mono-U	6	4	0	0
	GUAGUAGGUUGUAUAGUUUUAGGGUCACACCCACCGCUGGGAGAUAACUAU …	Trimmed	4	6	15	11
	GUAGUAGGUUGUAUAGUUUUAGGGUCACACCCACCACUGGGAGAUAACUAUACAAUCUACUGUCUUUCA	Mono-A	0	1	0	0
pre-let-7a-3	UGAGGUAGUAGGUUGUAUAGUUUGGGGCUCUGCCCUGCUAUGGGAUAACUAUACAAUCUACUGUCUUUC	Description	Set-1	Set-2	Set-1	Set-2
reads	GUAGUAGGUUGUAUAGUUUUAGGCUCUGCCCUGCUAUGGGAUAACUAUACAAUCUACUGUCUUUC	Unmodified	0	0	0	0
	GUAGUAGGUUGUAUAGUUUUAGGCUCUGCCCUGCUAUGGGAUAACUAUACAAUCUACUGUCUUUCU	Mono-U	0	0	0	0
	GUAGUAGGUUGUAUAGUUUUAGGCUCUGCCCUGCUAUGGGAUAACUAUACAAU	Trimmed	0	0	0	1
pre-let-7d	AGAGGUAGUAGGUUGCAUAGUUUUAGGGCAGGGAUUUUGCCCACAAGGAGGUAACUAUACGACCUGCUGCCUUUC	Description	Set-1	Set-2	Set-1	Set-2
reads	GUAGUAGGUUGUAUAGUUUUAGGGCAGGGAUUUUGCCCACAAGGAGGUAACUAUACGACCUGCUGCCUUUC	Unmodified	7	3	4	2
	GUAGUAGGUUGUAUAGUUUUAGGGCAGGGAUUUUGCCCACAAGGAGGUAACUAUACGACCUGCUGCCUUUCU	Mono-U	8	2	3	2
	GUAGUAGGUUGUAUAGUUUUAGGGCAGGGAUUUUGCCCACAAGGAGGUAACUAUACGAC …	Trimmed	1	0	3	3
	GUAGUAGGUUGUAUAGUUUUAGGGCAGGGAUUUUGCCCACAAGGAGGUAACUAUACGACCUGCUGCCUUUC…	Others	4	3	0	1
pre-let-7f-1	UGAGGUAGUAGAUUGUAUAGUUGUGGGGUAGUGAUUUUACCCUGUUCAGGAGAUAACUAUACAAUCUAUUGCCUUCC	Description	Set-1	Set-2	Set-1	Set-2
reads	GUAGUAGGUUGUAUAGUUUUAGGGUAGUGAUUUUACCCUGUUCAGGAGAUAACUAUACAAUCUAUUGCCUUCC	Unmodified	2	2	1	0
	GUAGUAGGUUGUAUAGUUUUAGGGUAGUGAUUUUACCCUGUUCAGGAGAUAACUAUACAAUCUAUUGCCUUCCU	Mono-U	0	3	0	0
	GUAGUAGGUUGUAUAGUUUUAGGGUAGUGAUUUUACCCUGUUCAGGAGAU …	Trimmed	0	5	7	6
	GUAGUAGGUUGUAUAGUUUUAGGGUAGUGAUUUUACCCUGUUCAGGAGAUAACUAUACAAUCUAUUGCCUUCCA	Mono-A	0	1	0	0
pre-let-7f-2	UGAGGUAGUAGAUUGUAUAGUUUUAGGGUCAUACCCCAUCUUGGAGAUAACUAUACAGUCUACUGUCUUUC	Description	Set-1	Set-2	Set-1	Set-2
reads	GUAGUAGGUUGUAUAGUUUUAGGGUCAUACCCCAUCUUGGAGAUAACUAUACAGUCUACUGUCUUUC	Unmodified	4	20	7	9
	GUAGUAGGUUGUAUAGUUUUAGGGUCAUACCCCAUCUUGGAGAUAACUAUACAGUCUACUGUCUUUCU	Mono-U	1	0	0	0
	GUAGUAGGUUGUAUAGUUUUAGGGUCAUACCCCAUCUUGGAGAUAACUAUACAGUCUACUGUCUU…	Trimmed	3	1	7	17

pre-let-7g	UGAGGUAGUAGUUUGUACAGUUUGAGGGUCUAUGAUACCACCCGGUACAGGAGAUAACUGUACAGGCCACUGCCUUGC	Description	Set-1	Set-2	Set-1	Set-2
reads	GUAGUAGGUUGUAUAGUUUUAGGGUCUAUGAUACCACCCGGUACAGGAGAUAACUGUACAGGCCACUGCCUUGC	Unmodified	2	4	5	1
	GUAGUAGGUUGUAUAGUUUUAGGGUCUAUGAUACCACCCGGUACAGGAGAUAACUGUACAGGCCACUGCCUUGCU	Mono-U	2	2	0	0
	GUAGUAGGUUGUAUAGUUUUAGGGUCUAUGAUACCACCCGGUACAGGAGAUA	Trimmed	1	0	0	0

Total	Set-1	Set-2	Set-1	Set-2
Unmodified	38	49	38	43
Mono-U	17	11	3	2
Mono-A	0	2	0	0
Trimmed	9	12	32	38
Others	4	3	0	1
Total	68	77	73	84

Sequences of trimmed or other reads			siLuc		siTUT mix†	
pre-let-7a-1	UGAGGUAGUAGGUUGUAUAGUUUUAGGGUCACACCCACCACUGGGAGAUAACUAUACAAUCUACUGUCUUUC	Description	Set-1	Set-2	Set-1	Set-2
reads	GUAGUAGGUUGUAUAGUUUUAGGGUCACACCCACCGCUGGGAGAUAACUAUAUUU	Trimmed	1	0	0	0
	GUAGUAGGUUGUAUAGUUUUAGGGUCACACCCACCACUGGGAGAUAACUAUAUUUUUUUU	Trimmed	1	0	0	0
	GUAGUAGGUUGUAUAGUUUUAGGGUCACACCCACCACUGGGAGAUAACUAUACAAU	Trimmed	1	0	0	0
	GUAGUAGGUUGUAUAGUUUUAGGGUCACACCCACCACUGGGAGAUAACUAUACAAUUUUUUAU	Trimmed	0	1	0	0
	GUAGUAGGUUGUAUAGUUUUAGGGUCACACCCACCACUGGGAGAUAACUAUACAAUC	Trimmed	0	0	0	1
	GUAGUAGGUUGUAUAGUUUUAGGGUCACACCCACCACUGGGAGAUAACUAUACAAUCUUU	Trimmed	1	0	0	0
	GUAGUAGGUUGUAUAGUUUUAAGGUCACACCCACCACUGGGAGAUAACUAUACAAUCUUUUUUU	Trimmed	0	1	0	0
	GUAGUAGGUUGUAUAGUUUUAGGGUCACACCCACCACUGGGAGAUAACUAUACAGUCUACUGUC	Trimmed	0	0	2	0
	GUAGUAGGUUGUAUAGUUUUAGGGUCACACCCACCACUGGGAGAUAACUAUACAAUCUUUUUUUAU	Trimmed	0	1	0	0
	GUAGUAGGUUGUAUAGUUUUAGGGUCACACCCACCACUGGGAGAUAACUAUACAAUCUACUGUCUUU	Trimmed	0	2	13	10
	GUAGUAGGUUGUAUAGUUUUAGGGUCACACCCACCACUGGGAGAUAACUAUACAAUCUACUGUCUUUU	Trimmed	0	1	0	0
pre-let-7d	AGAGGUAGUAGGUUGCAUAGUUUUAGGGCAGGGAUUUUGCCCACAAGGAGGUAACUAUACGACCUGCUGCCUUUC	Description	Set-1	Set-2	Set-1	Set-2
reads	GUAGUAGGUUGUAUAGUUUUAGGGCAGGGAUUUUGCCCACAAGGAGGUAACUAUACGACUAUUU	Trimmed	0	0	1	0
	GUAGUAGGUUGUAUAGUUUUAGGGCAGGGAUUUUGCCCACAAGGAGGUAACUAUACGACCUGCUGUCUUU	Trimmed	1	0	0	3

	GUAGUAGGUUGUAUAGUUUUAGGGCAGGGAUUUUGCCCACAAGGAGGUAACUAUACGACCUGCUGCCUUUU	Trimmed	0	0	1	0
	GUAGUAGGUUGUAUAGUUUUAGGGCAGGGAUUUUGCCCACAAGGAGGUAACUAUACGACCUGCUGCCUUUUU	Trimmed	0	0	1	0
	GUAGUAGGUUGUAUAGUUUUAGGGCAGGGAUUUUGCCCACAAGGAGGUAACUAUACGACCUGCUGCCUUUCUU	Others	1	0	0	0
	GUAGUAGGUUGUAUAGUUUUAGGGCAGGGAUUUUGCCCACAAGGAGGUAACUAUACGACCUGCUGCCUUUCUCU	Others	1	0	0	0
	GUAGUAGGUUGUAUAGUUUUAGGGCAGGGAUUUUGCCCACAAGGAGGUAACUAUACGACCUGCUGCCUUUCUUU	Others	1	2	0	0
	GUAGUAGGUUGUAUAGUUUUAGGGCAGGGAUUUUGCCCACAAGGAGGUAACUAUACGACCUGCUGCCUUUCUUUU	Others	0	1	0	1
	GUAGUAGGUUGUAUAGUUUUAGGGCAGGGAUUUUGCCCACAAGGAGGUAACUAUACGACCUGCUGCCUUUCUUUU	Others	1	0	0	0
pre-let-7f-1	UGAGGUAGUAGAUUGUAUAGUUGUGGGGUAGUGAUUUUACCCUGUUCAGGAGAUAACUAUACAAUCUAUUGCCUUCC	Description	Set-1	Set-2	Set-1	Set-2
reads	GUAGUAGGUUGUAUAGUUUUAGGGUAGUGAUUUUACCCUGUUCAGGAGAU	Trimmed	0	0	1	0
	GUAGUAGGUUGUAUAGUUUUAGGGUAGUGAUUUUACCCUGUUCAGGAGAUAACUAUACAAU	Trimmed	0	0	1	1
	GUAGUAGGUUGUAUAGUUUUAGGGUAGUGAUUUUACCCUGUUCAGGAGAUAACUAUACAAUC	Trimmed	0	1	0	2
	GUAGUAGGUUGUAUAGUUUUAGGGUAGUGAUUUUACCCUGUUCAGGAGAUAACUAUACAAUAUUUUUUUU	Trimmed	0	1	0	0
	GUAGUAGGUUGUAUAGUUUUAGGGUAGUGAUUUUACCCUGUUCAGGAGAUAACUAUACAAUCUAUUGCCUUC	Trimmed	0	1	3	2
	GUAGUAGGUUGUAUAGUUUUAGGGUAGUGAUUUUACCCUGUUCAGGAGAUAACUAUACAAUCUAUUGCCUUCU	Trimmed	0	1	2	1
	GUAGUAGGUUGUAUAGUUUUAGGGUAGUGAUUUUACCCUGUUCAGGAGAUAACUAUACAAUCUAUUGCCUUUU	Trimmed	0	1	0	0
pre-let-7f-2	UGAGGUAGUAGAUUGUAUAGUUUUAGGGUCAUACCCCAUCUUGGAGAUAACUAUACAGUCUACUGUCUUUC	Description	Set-1	Set-2	Set-1	Set-2
reads	GUAGUAGGUUGUAUAGUUUUAGGGUCAUACCCCAUCUUGGAGAUAACUAUACAGUCUACUGUCUUA	Trimmed	0	0	0	1
	GUAGUAGGUUGUAUAGUUUUAGGGUCAUACCCCAUCUUGGAGAUAACUAUACAGUCUACUGUCUUU	Trimmed	3	0	7	14
	GUAGUAGGUUGUAUAGUUUUAGGGUCAUACCCCAUCUUGGAGAUAACUAUACAGUCUACUGUCUUUU	Trimmed	0	1	0	2

† siTUT mix = mixture of equal amounts of siTUT7, siTUT4, and siTUT2

Table II-2. List of small RNA sequencing data used in miRNA analysis.

Data were collected from the NCBI's Gene Expression Omnibus (GEO) database, all of which are publicly available. Libraries are shown in either SRA number or GSM number. For the expression profiles of human tissues, each human library was sorted into tissue of origin.

Species	Library	Description	Tissue
H. sapiens	SRR015358	Naïve B Cell	
	SRR015359	Germinal Center B cell	
	SRR015360	Plasma B cell	
	SRR015361	Memory B cells	
	SRR015362	Naïve B Cell	
	SRR015363	Germinal Center B cell	
	SRR015364	Plasma B cell	
	SRR015365	Memory B cells	
	SRR033710	b cell type: Transformed follicular lymphoma	
	SRR033711	b cell type: Diffuse Large B cell lymphoma	
	SRR033712	Burkitt Lymphoma	Blood
	SRR033713	Burkitt Lymphoma	
	SRR033714	Burkitt Lymphoma	
	SRR033715	Mantle cell Lymphoma	
	SRR033716	Mantle cell Lymphoma	
	SRR033717	Mantle cell Lymphoma	
	SRR033718	Multiple Myeloma	
	SRR033719	6hr Activated B cell line	
	SRR033721	b cell type: Diffuse Large B cell lymphoma	
	SRR033722	Lymphoid organ(Multiple Myeloma)	
	SRR033723	b cell type: Hodkgin lymphoma	

	SRR033724	B cell type: Hodkgin lymphoma	
	SRR033725	Unmutated CLL	
	SRR033726	Mututated CLL	
	SRR033727	HIV lymphoma	
	SRR033728	Marginal Zone lymphoma	
	SRR033729	Marginal Zone lymphoma	
	SRR033730	Lymphoblastic lymphoma	
	SRR033731	Multiple Myeloma	
	SRR033732	b cell type: Diffuse Large B cell lymphoma	
	SRR039190	Peripheral blood mononuclear cells	
	SRR039191	peripheral blood mononuclear cell	
	SRR039192	human erythroleukemic cell	
	SRR039193	Acute Promyelogenous Leukemia cell line	
	SRR039637	Human acute monocytic leukemia	
	SRR060981	Human Centroblasts (Germinal Center B cells)	
	SRR096521	EBV activated B cell line	
	SRR029129	Human colon adenocarcinoma	Colon
	SRR029130	Colon Carcinoma	
	SRR029127	adenocarcinomic human alveolar basal epithelial cells	Lung
	SRR029128	Non-small-cell lung carcinoma (NSCLC)	
	SRR020286	human diploid fibroblast	Skin
	SRR037876	fibroblast	
	SRR038852	primary melanoblast	
	SRR038853	primary melanocyte	
	SRR038854	primary giant congental nevus	
	SRR038855	secondary acral melanoma	
	SRR038856	secondary mucosal melanoma	
	SRR038857	secondary cutaneous melanoma	
	SRR038858	primary uveal melanoma	

	SRR038859	secondary cutaneous melanoma	
	SRR038860	secondary cutaneous melanoma	
	SRR038861	secondary cutaneous melanoma	
	SRR038862	secondary cutaneous melanoma from a chronically sun-damaged body site	
	SRR038863	secondary cutaneous melanoma	
	SRR107296	human diploid fibroblast	
	SRR107297	human diploid fibroblast	
	SRR330866	Skin psoriasis biopsy	
	SRR330867	Skin psoriasis biopsy	
	SRR330868	Skin psoriasis biopsy	
	SRR330901	Skin uninvolved psoriasis biopsy	
	SRR330902	Skin uninvolved psoriasis biopsy	
	SRR330903	Skin uninvolved psoriasis biopsy	
	SRR330905	Skin biopsy	
	SRR330906	Skin biopsy	
	SRR330907	Skin biopsy	
	SRR029054	human breast adenocarcinoma	Breast
	SRR029131	human breast adenocarcinoma	
	SRR029132	human breast carcinoma	
	GSM416753	human cervical adenocarcinoma (HeLa)	Cervix
	GSM755025	human cervical adenocarcinoma (HeLa)	
	GSM876017	HeLa_siGFP-1	
	GSM876013	HeLa_siLuc-1	
	GSM876014	HeLa_siLuc-2	
	GSM870233	HeLa_Thymidine-8hr	
	GSM870234	HeLa_Thymidine-10hr	
	GSM870235	HeLa_Thymidine-14hr	
	GSM870236	HeLa_Thymidine-16hr	

	GSM870237	HeLa_Thymidine-18hr	
	GSM870238	HeLa_Thymidine-20hr	
	GSM721078	HeLa S2	
	GSM721075	HeLa S2_AGO1/2	
	GSM721076	HeLa S2_AGO1/2_RRP40	
	SRR029125	human osteosarcoma	Bone
	SRR029126	human osteosarcoma	
	GSM652847	Human cerebellar cortex rep1 (microRNA-Seq)	Brain
	GSM652848	Human cerebellar cortex rep2 (microRNA-Seq)	
	GSM652851	Human superior frontal gyrus rep1 (microRNA-Seq)	
	GSM652852	Human superior frontal gyrus rep2 (microRNA-Seq)	
	SRR039611	Human Normal Liver Tissue	Liver
	SRR039612	Human Normal Liver Tissue	
	SRR039613	Human Normal Liver Tissue	
	SRR039614	HBV-infected Liver Tissue	
	SRR039615	Severe Chronic Hepatitis B Liver Tissue	
	SRR039616	HBV(+) Distal Tissue	
	SRR039617	HBV(+) Adjacent Tissue	
	SRR039618	HBV(+) Side Tissue	
	SRR039619	HBV(+) HCC Tissue	
	SRR039620	HBV(+) Adjacent Tissue	
	SRR039621	HBV(+) HCC Tissue	
	SRR039622	HCV(+) Adjacent Tissue	
	SRR039623	HCV(+) HCC Tissue	
	SRR039624	HBV(-) HCV(-) Adjacent Tissue	
	SRR039625	HBV(-) HCV(-) HCC Tissue	
P. troglodytes	GSM652849	Chimpanzee cerebellar cortex rep1 (microRNA-Seq)	
	GSM652853	Chimpanzee superior frontal gyrus rep1 (microRNA-Seq)	

	GSM652854	Chimpanzee superior frontal gyrus rep2 (microRNA-Seq)
M. mulatta	GSM652850	Rhesus cerebellar cortex rep1 (microRNA-Seq)
	GSM652855	Rhesus superior frontal gyrus rep1 (microRNA-Seq)
	GSM652856	Rhesus superior frontal gyrus rep2 (microRNA-Seq)
B. taurus	GSM387590	MDBK_mock_infection_control
S. scrofa	GSM364843	Muscle_animal_4363
	GSM364844	Muscle_animal_4365
	GSM364845	Muscle_animal_5342
	GSM364846	Muscle_animal_147
	GSM364847	Muscle_animal_3713
	GSM364848	Muscle_animal_4905
	GSM364849	Muscle_animal_4908
M. musculus	SRR029123	NIH3T3
	SRR039152	HL1_siNon_smallRNAseq(mouse cardiomyocytes)
	SRR039634	mouse_granulocyte_nuclei_smallRNAs
	SRR042443	mouse pro B cells [09-002]
	SRR042444	mouse pre B cells [09-002]
	SRR042445	mouse immature B cells [09-002]
	SRR042446	mouse mature B cells (spleen) replicate 1 [09-002]
	SRR042447	mouse mature B cells (spleen) replicate 2 [09-002]
	SRR042448	mouse B cells activated in vitro with LPS/IL4 replicate 1 [09-002]
	SRR042449	mouse B cells activated in vitro with LPS/IL4 replicate 2 [09-002]
	SRR042450	mouse B cells activated in vitro with LPS/IL4 replicate 3 [09-002]
	SRR042451	mouse B cells activated in vitro with LPS/aIgD-Dextran [09-002]
	SRR042452	mouse B1 B cells [09-002]

	SRR042453	mouse marginal zone B cells (spleen) [09-002]
	SRR042454	mouse germinal center B cells (lymph node) [09-002]
	SRR042455	mouse plasma cells [09-002]
	SRR042456	mouse hematopoetic progenitor cells [09-002]
	SRR042457	mouse mast cells [09-002]
	SRR042458	mouse basophil cells [09-002]
	SRR042459	mouse neutrophil cells replicate 1 [09-002]
	SRR042460	mouse neutrophil cells replicate 2 [09-002]
	SRR042461	mouse dendritic cells [09-002]
	SRR042462	mouse macrophages [09-002]
	SRR042463	mouse natural killer cells [09-002]
	SRR042464	mouse CD4+CD8+ T cells (thymus) [09-002]
	SRR042465	mouse CD4+ T cells (spleen) [09-002]
	SRR042466	mouse CD8+ T cells (spleen) [09-002]
	SRR042467	mouse T cells activated in vitro with Concanavalin A [09-002]
	SRR042468	mouse Th1 cells [09-002]
	SRR042469	mouse Th2 cells [09-002]
	SRR042470	mouse Th17 cells [09-002]
	SRR042471	mouse TFHS cells [09-002]
	SRR042472	mouse iTreg cells [09-002]
	SRR042473	mouse nTreg cells [09-002]
	SRR042474	mouse embryonic stem cells [09-002]
	SRR042475	mouse embryonic fibroblast cells [09-002]
	SRR042476	mouse heart tissue [09-002]
	SRR042477	mouse brain tissue [09-002]
	SRR042478	mouse lung tissue [09-002]
	SRR042479	mouse liver tissue [09-002]
	SRR042480	mouse kidney tissue [09-002]
	SRR042481	mouse pancreatic tissue [09-002]

	SRR042482	mouse skin tissue [09-002]
	SRR042483	mouse muscle tissue [09-002]
	SRR042484	mouse salivary gland tissue [09-002]
	SRR042485	mouse testicular tissue [09-002]
	SRR042486	mouse ovaries [09-002]
	SRR042487	mouse BclXL B activated with LPS/IL4 [BalbC] [09-002]
	SRR042488	mouse BclXL germinal center B cells [BalbC] [09-002]
	SRR059765	DN3_control
	SRR059768	Treg_control
	SRR059774	MEF_control
	SRR073954	microRNA expression in CFU-E late erythroid progenitors
	SRR073955	microRNA expression in Ter119+ mature erythroblasts
	SRR306526	Cerebellum_1
	SRR306527	Cerebellum_2
	SRR306528	Cerebellum_3
	SRR306529	Purkinje_1
	SRR306530	Purkinje_2
	SRR306531	Purkinje_3
	SRR306532	Camk2a_1
	SRR306533	Camk2a_2
	SRR306534	Camk2a_3
	SRR306535	Gad2_1
	SRR306536	Gad2_2
	SRR306537	Gad2_3
	SRR306538	Neocortex_1
	SRR306539	Neocortex_2
	SRR306540	Neocortex_3
	SRR306541	PV_1
	SRR306542	PV_2

	SRR306543	SST_1
	SRR306544	SST_2
	SRR345196	Untreated mouse hippocampus replicate 1
	SRR351925	Mouse_TotalHeart_miRNA_FEMALE
	SRR351926	Mouse_TotalHeart_miRNA_MALE
R. norvegicus	GSM471497	BN-Lx_adult_male_whole_brain
	GSM471498	BN-Lx_adult_male_whole_liver
	GSM471499	BN-Lx_adult_male_whole_spleen
	GSM471500	BN-Lx_adult_male_whole_heart
	GSM471501	BN-Lx_adult_male_whole_testis
	GSM471502	BN-Lx_adult_male_whole_kidney
	GSM471503	SHR_adult_male_whole_brain
	GSM471504	SHR_adult_male_whole_liver
	GSM471505	SHR_adult_male_whole_spleen
	GSM471506	SHR_adult_male_whole_heart
	GSM471507	SHR_adult_male_whole_testis
	GSM471508	SHR_adult_male_whole_kidney
	GSM471509	BN-Lx_adult_male_whole_liver, replicate 2
	GSM471510	BN-Lx_adult_male_whole_liver, replicate 3
	GSM471511	SHR_adult_male_whole_liver, replicate 2
	GSM471512	SHR_adult_male_whole_liver, replicate 3
G. gallus	GSM523740	layer_reads
	GSM523741	broiler_reads
X. tropicalis	GSM372598	adult liver
	GSM372601	adult skin

Table II-3. Results of small RNA deep sequencing following knock-down of TUTs in HeLa.

Shown is the list of deep sequencing results for 30 most abundant miRNAs in HeLa. The sequence and read number of most frequent read for each miRNA is shown. Reads per million (RPM) of each miRNA was used to calculate fold change. The median log2 fold change of 212 miRNAs was used to normalize entire distribution. Blue indicates group I let-7, and red indicates group II let-7.

miRNA name	Sequences	Len	siLuc raw reads		siTUT mix† raw reads		siLuc RPM		siTUT mix† RPM		Log_2 fold change (siTUT mix† / siLuc)		Log_2 fold change_median normalized (siTUT mix† / siLuc)		Average log_2 fold change_median normalized (siTUT mix† / siLuc)
			Set-1	Set-2	Set-1	Set-2	Set-1	Set-2	Set-1	Set-2	Set-1	Set-2	Set-1	Set-2	
let-7b-5p	UGAGGUAGUAGGUUGUGUGGUU	22	53453	103369	9683	9571	2685	5478	495	603	-2.439	-3.184	-3.067	-3.532	-3.299
miR-21-5p	UAGCUUAUCAGACUGAUGUUGAC	23	913075	882820	279631	235472	45873	46783	14304	14827	-1.681	-1.658	-2.310	-2.005	-2.157
let-7f-5p	UGAGGUAGUAGAUUGUAUAGUU	22	2327620	2301748	1395869	944883	116940	121976	71403	59498	-0.712	-1.036	-1.340	-1.383	-1.362
miR-20a-5p	UAAAGUGCUUAUAGUGCAGGUAG	23	141296	144829	92698	66254	7099	7675	4742	4172	-0.582	-0.879	-1.211	-1.227	-1.219
miR-98	UGAGGUAGUAAGUUGUAUUGUU	22	74690	69188	61004	36574	3752	3666	3121	2303	-0.266	-0.671	-0.894	-1.019	-0.956
let-7a-5p	UGAGGUAGUAGGUUGUAUAGUU	22	1260808	1741254	1232757	1108053	63343	92274	63059	69772	-0.006	-0.403	-0.635	-0.751	-0.693
let-7i-5p	UGAGGUAGUAGUUUGUGCUGUU	22	1940717	1857099	1881165	1336288	97502	98413	96227	84144	-0.019	-0.226	-0.647	-0.574	-0.610
let-7g-5p	UGAGGUAGUAGUUUGUACAGUU	22	1550792	1381794	1705791	1046818	77912	73225	87256	65916	0.163	-0.152	-0.465	-0.499	-0.482
miR-378a-3p	ACUGGACUUGGAGUCAGAAGGC	22	235516	261876	268773	196990	11832	13878	13749	12404	0.217	-0.162	-0.412	-0.510	-0.461
miR-17-5p	CAAAGUGCUUACAGUGCAGGUAG	23	46133	45725	62074	36626	2318	2423	3175	2306	0.454	-0.071	-0.174	-0.419	-0.297
let-7d-5p	AGAGGUAGUAGGUUGCAUAGUU	22	215996	219988	271773	212182	10852	11658	13902	13361	0.357	0.197	-0.271	-0.151	-0.211
miR-26a-5p	UUCAAGUAAUCCAGGAUAGGCU	22	388004	484915	552678	417998	19493	25697	28271	26321	0.536	0.035	-0.092	-0.313	-0.203
miR-191-5p	CAACGGAAUCCCAAAAGCAGCUG	23	133524	152697	210055	144856	6708	8092	10745	9121	0.680	0.173	0.051	-0.175	-0.062
miR-182-5p	UUUGGCAAUGGUAGAACUCACACU	24	44469	41622	69653	39817	2234	2206	3563	2507	0.673	0.185	0.045	-0.163	-0.059
miR-7-5p	UGGAAGACUAGUGAUUUUGUUGUU	24	283709	279098	458941	262856	14254	14790	23476	16552	0.720	0.162	0.091	-0.185	-0.047
miR-151a-5p	UCGAGGAGCUCACAGUCUAGU	21	50002	45614	64253	58593	2512	2417	3287	3689	0.388	0.610	-0.241	0.262	0.011
let-7c	UGAGGUAGUAGGUUGUAUGGUU	22	49676	84720	86120	91876	2496	4490	4405	5785	0.820	0.366	0.191	0.018	0.105
miR-30d-5p	UGUAAACAUCCCCGACUGGAAGC	23	68111	114497	115754	138967	3422	6068	5921	8751	0.791	0.528	0.163	0.181	0.172
miR-30a-5p	UGUAAACAUCCUCGACUGGAAGC	23	233626	224802	432322	264565	11737	11913	22115	16659	0.914	0.484	0.285	0.136	0.211
miR-103a-3p	AGCAGCAUUGUACAGGGCUAUGA	23	659329	675295	1212048	805751	33125	35786	62000	50737	0.904	0.504	0.276	0.156	0.216
miR-27b-3p	UUCACAGUGGCUAAGUUCUGC	21	244944	223337	426433	325895	12306	11835	21813	20521	0.826	0.794	0.197	0.446	0.322
miR-100-5p	AACCCGUAGAUCCGAACUUGUG	22	128405	167701	282809	209902	6451	8887	14466	13217	1.165	0.573	0.537	0.225	0.381
let-7e-5p	UGAGGUAGGAGGUUGUAUAGUU	22	31262	52105	61983	88470	1571	2761	3171	5571	1.013	1.013	0.385	0.665	0.525
miR-30c-5p	UGUAAACAUCCUACACUCUCAGC	23	40963	63163	108739	89214	2058	3347	5562	5618	1.434	0.747	0.806	0.399	0.603
miR-151a-3p	CUAGACUGAAGCUCCUUGAGG	21	25482	29537	54779	54644	1280	1565	2802	3441	1.130	1.136	0.502	0.789	0.645
miR-148b-3p	UCAGUGCAUCACAGAACUUUGU	22	103277	105795	260091	194807	5189	5606	13304	12267	1.358	1.130	0.730	0.782	0.756
miR-185-5p	UGGAGAGAAAGGCAGUUCCUGA	22	28238	30487	69140	59209	1419	1616	3537	3728	1.318	1.206	0.689	0.859	0.774
miR-24-3p	UGGCUCAGUUCAGCAGGAAC	20	107320	50620	218745	139896	5392	2683	11189	8809	1.053	1.715	0.425	1.368	0.896
miR-320a	AAAAGCUGGGUUGAGAGGGCGA	22	11064	35994	56789	73178	556	1907	2905	4608	2.386	1.272	1.757	0.925	1.341
miR-93-5p	CAAAGUGCUGUUCGUGCAGGUAG	23	19952	20183	126013	116692	1002	1070	6446	7348	2.685	2.780	2.057	2.433	2.245

† siTUT mix = mixture of equal amounts of siTUT7, siTUT4, and siTUT2

Median log_2 fold change (212 miRNAs)	
Set-1	Set-2
0.628	0.348

Table II-4. Results of group I pre-let-7 sequencing following knock-down of TUTs in HeLa.

(top) Shown are the sequences of each read and its read numbers in the results of group I pre-let-7 sequencing. Although let-7c specific primer (yellow box) was used in the amplification of sequences by PCR, not only pre-let-7c but also other group I let-7 family precursors (pre-let-7a-2) were detected due to the high sequence homolgy. Trimmed sequences indicate reads which are shorter than unmodified pre-let-7. The reads that do not belong to unmodified, mono-U, or trimmed are described in 'others'. **(bottom)** Shown are the detailed sequences of trimmed and the other pre-let-7 reads. Red indicates mono-uridylation, and blue represents untemplated modification.

let-7c specific PCR primer	GTAGTAGGTTGTATGGTTTAGA		read numbers			
			siLuc		siTUT mix†	
pre-let-7c	UGAGGUAGUAGGUUGUAUGGUUUAGAGUUACACCCUGGGAGUUAACUGUACAACCUUCUAGCUUUCC	Description	Set-1	Set-2	Set-1	Set-2
reads	GUAGUAGGUUGUAUAGUUUAGAGUUACACCCUGGGAGUUAACUGUACAACCUUCUAGCUUUCC	Unmodified	10	21	8	14
	GUAGUAGGUUGUAUAGUUUAGAGUUACACCCUGGGAGUUAACUGUACAACCUUCUAGCUUUCCU	Mono-U	0	2	1	1
	GUAGUAGGUUGUAUAGUUUAGAGUUACACCCUGGGAGUUAACUGUACAACCUUCUAGCUUU …	Trimmed	9	7	6	4
	GUAGUAGGUUGUAUAGUUUAGAGUUACACCCUGGGAGUUAACUGUACAACCUUCUAGCUUUCC …	Others	1	2	0	1
pre-let-7a-2	UGAGGUAGUAGGUUGUAUAGUUUAGAAUUACAUCAAGGGAGAUAACUGUACAGCCUCCUAGCUUUCC	Description	Set-1	Set-2	Set-1	Set-2
reads	GUAGUAGGUUGUAUAGUUUAGAAUUACAUCAAGGGAGAUAACUGUACAGCCUCCUAGCUUUCC	Unmodified	14	14	7	23
	GUAGUAGGUUGUAUAGUUUAGAAUUACAUCAAGGGAGAUAACUGUACAGCCUCCUAGCUUUCCU	Mono-U	0	0	0	0
	GUAGUAGGUUGUAUAGUUUAGAAUUACAUCAAGGGAGAUAACUGUACAGCCUCCUAGCUUU …	Trimmed	9	9	1	7
	GUAGUAGGUUGUAUAGUUUAGAAUUACAUCAAGGGAGAUAACUGUACAGCCUCCUAGCUUUCC …	Others	0	1	0	0

Total	Set-1	Set-2	Set-1	Set-2
Unmodified	24	35	15	37
Mono-U	0	2	1	1
Trimmed	18	16	7	11
Others	1	3	0	1
Total	43	56	23	50

† siTUT mix = mixture of equal amounts of siTUT7, siTUT4, and siTUT2

Sequences of trimmed or other reads			siLuc		siTUT mix†	
pre-let-7c	UGAGGUAGUAGGUUGUAUGGUUUAGAGUUACACCCUGGGAGUUAACUGUACAACCUUCUAGCUUUCC	Description	Set-1	Set-2	Set-1	Set-2
reads	GUAGUAGGUUGUAUGGUUUAGAGUUACACCCUGGGAGUUAACUGUACAACC	Trimmed	0	0	1	0
	GUAGUAGGUUGUAUGGUUUAGAGUUACACCCUGGGAGUUAACUGUACAACCUUUUU	Trimmed	1	0	0	0
	GUAGUAGGUUGUAUGGUUUAGAGUUACACCCUGGGAGUUAACUGUACAACCUUUAUUUU	Trimmed	0	1	0	0
	GUAGUAGGUUGUAUGGUUUAUAAUAACACCCUGCGAGUUAACUGUAGAACCUUCUACCUU	Trimmed	0	0	1	0
	GUAGUAGGUUGUAUGGUUUAGAGUUACACCCUGGGAGUUAACUGUACAACCUUCUAGCU	Trimmed	0	0	0	1
	GUAGUAGGUUGUAUGGUUUAGAGUUACACCCUGGGAGUUAACUGUACAACCUUCUAGCUUU	Trimmed	0	3	2	1
	GUAGUAGGUUGUAUGGUUUAGAGUUACACCCUGGGAGUUAACUGUACAACCUUCUAGCUUUA	Trimmed	0	0	1	0
	GUAGUAGGUUGUAUAGUUUAGAGUUACACCCUGGGAGUUAACUGUACAACCUUCUAGCUUUC	Trimmed	2	2	0	2
	GUAGUAGGUUGUAUAGUUUAGAGUUACACCCUGGGAGUUAACUGUACAACCUUCUAGCUUUCU	Trimmed	1	0	0	0
	GUAGUAGGUUGUAUAGUUUAGAGUUACACCCUGGGAGUUAACUGUACAACCUUCUAGCUUUCA	Trimmed	5	1	1	0
	GUAGUAGGUUGUAUGGUUUAGAGUUACACCCUGGGAGUUAACUGUACAACCUUCUAGCUUUCCAAU	Others	1	0	0	0
	GUAGUAGGUUGUAUGGUUUAGAGUUACACCCUGGGAGUUAACUGUACAACCUUCUAGCUUUCCUU	Others	0	1	0	1
	GUAGUAGGUUGUAUGGUUUAGAGUUACACCCUGGGAGUUAACUGUACAACCUUCUAGCUUUCCUUUU	Others	0	1	0	0
pre-let-7a-2	UGAGGUAGUAGGUUGUAUAGUUUAGAAUUACAUCAAGGGAGAUAACUGUACAGCCUCCUAGCUUUCC	Description	Set-1	Set-2	Set-1	Set-2
reads	GUAGUAGGUUGUAUGGUUUAGAAUUACAUCAAGGGAGAUAACUGUACAGCCUCCUAU	Trimmed	0	0	0	1
	GUAGUAGGUUGUAUGGUUUAGAAUUACAUCAAGGGAGAUAACUGUACAGCCUCCUAUUUUUU	Trimmed	0	1	0	0
	GUAGUAGGUUGUAUGGUUUAGAAUUACAUCAAGGGAGAUAACUGUACAGCCUCCUAGUA	Trimmed	0	0	0	1
	GUAGUAGGUUGUAUGGUUUAGAAUUACAUCAAGGGAGAUAACUGUACAGCCUCCUAGCU	Trimmed	1	0	0	0
	GUAGUAGGUUGUAUGGUUUAGAAUUACAUCAAGGGAGAUAACUGUACAGCCUCCUAGCUU	Trimmed	0	0	1	0
	GUAGUAGGUUGUAUGGUUUAGAAUUACAUCAAGGGAGAUAACUGUACAGCCUCCUAGCUUU	Trimmed	3	2	0	3
	GUAGUAGGUUGUAUGGUUUAGAAUUACAUCAAGGGAGAUAACUGUACAGCCUCCUAGCUUUU	Trimmed	1	1	0	0
	GUAGUAGGUUGUAUGGUUUAGAAUUACAUCAAGGGAGAUAACUGUACAGCCUCCUAGCUUUUUA	Trimmed	0	1	0	0
	GUAGUAGGUUGUAUGGUUUAGAAUUACAUCAAGGGAGAUAACUGUACAGCCUCCUAGCUUUA	Trimmed	4	0	0	0
	GUAGUAGGUUGUAUGGUUUAGAAUUACAUCAAGGGAGAUAACUGUACAGCCUCCUAGCUUUC	Trimmed	0	2	0	1
	GUAGUAGGUUGUAUGGUUUAGAAUUACAUCAAGGGAGAUAACUGUACAGCCUCCUAGCUUUCU	Trimmed	0	1	0	0
	GUAGUAGGUUGUAUGGUUUAGAAUUACAUCAAGGGAGAUAACUGUACAGCCUCCUAGCUUUCA	Trimmed	0	1	0	1
	GUAGUAGGUUGUAUGGUUUAGAAUUACAUCAAGGGAGAUAACUGUACAGCCUCCUAGCUUUCCCU	Others	0	1	0	0

Table II-5. Oligonucleotides Used in This Study.

Adapter and Primers for Pre-let-7 Sequencing

3′ adapter	5′-rApp CTG TAG GCA CCA TCA AT/ddC/-3′ (rApp, adenylated; ddC, dideoxy-C)
RT primer	5′-TGC ATT GAT GGT GCC TAC AG-3′
Let-7a specific forward primer (with BanI site)	5′-ATC GTA GGC ACC GTA GTA GGT TGT ATA GTT TTA G-3′
Let-7c specific forward primer (with BanI site)	5′-ATC GTA GGC ACC GTA GTA GGT TGT ATG GTT TAG A-3′

Q-PCR Primers

hsa-HMGA2 (forward)	5′-GCG CCT CAG AAG AGA GGA C-3′
hsa-HMGA2 (reverse)	5′-GGT CTC TTA GGA GAG GGC TCA-3′
hsa-NRAS (forward)	5′-TGA GAG ACC AAT ACA TGA GGA CA-3′
hsa-NRAS (reverse)	5′-CCC TGT AGA GGT TAA TAT CCG CA-3′

The Sequences of DNA Oligonucleotide for Northern Blotting

hsa-let-7a	5′-ACT ATA CAA CCT ACT ACC TCA-3′
hsa-miR-16	5′-GCC AAT ATT TAC GTG CTG CTA-3′
hsa-miR-20a	5′-TAC CTG CAC TAT AAG CAC TTT A-3′
hsa-miR-21	5′-TCA ACA TCA GTC TGA TAA GCT A-3′
hsa-miR-93	5′-TAC CTG CAC GAA CAG CAC TTT G-3′
tRNA-lys-AAG	5′-GAG ATT AAG AGT CTC ATG CTC-3′

Mutagenesis Primers

TUT7 D1060A (forward)	5′-CAA ACA GAG TGA CCT TGC CGT CTG TAT GAC AAT TAA TG-3′
TUT7 D1060A (reverse)	5′-CAT TAA TTG TCA TAC AGA CGG CAA GGT CAC TCT GTT TG-3′
TUT2 D215A (forward)	5′-CCG GAG CAG TGA TGG TGC TTT ATG CCT AGT TGT TAA G-3′
TUT2 D215A (reverse)	5′-CTT AAC AAC TAG GCA TAA AGC ACC ATC ACT GCT CCG G-3′

Target site Sequences of siRNA

siLuciferase	5′-CUU ACG CUG AGU ACU UCG A-3′
siTUT7-1	5′-GAA AAG AGG CAC AAG AAA A-3′
siTUT7-2	5′-GAU AAG UAU UCG UGU CAA A-3′
siTUT7-3	5′-GAA CAU GAG UAC CUA UUU A-3′
siTUT4-1	5′-GGU UGC UUC AGA CUU UAU A-3′
siTUT4-2	5′-GGA AUG AAG AAG AGA AAG A-3′
siTUT4-3	5′-GGU UAG AGC UGC UUA AAU U-3′
siTUT2-1	5′-CGU UAG UGC UGG UGA UUA A-3′
siTUT2-2	5′-ACA UCA AGC UCC AUG UAA U-3′
siTUT2-3	5′-GAG CAG UGA UGG UGA UUU A-3
siDicer	5′-UGC UUG AAG CAG CUC UGG A-3′

Pre-miRNA and ds-35 Sequences for in vitro Dicer Processing or in vitro Uridylation

hsa-pre-let-7a-1 (Unmodified)	5′-UGA GGU AGU AGG UUG UAU AGU UUU AGG GUC ACA CCC ACC ACU GGG AGA UAA CUA UAC AAU CUA CUG UCU UUC-3′
hsa-pre-let-7a-1 (Mono-U)	5′-UGA GGU AGU AGG UUG UAU AGU UUU AGG GUC ACA CCC ACC ACU GGG AGA UAA CUA UAC AAU CUA CUG UCU UUC U-3′
hsa-pre-let-7b (Unmodified)	5′-UGA GGU AGU AGG UUG UGU GGU UUC AGG GCA GUG AUG UUG CCC CUC GGA AGA UAA CUA UAC AAC CUA CUG CCU UCC -3′
hsa-pre-let-7b (Mono-U)	5′-UGA GGU AGU AGG UUG UGU GGU UUC AGG GCA GUG AUG UUG CCC CUC GGA AGA UAA CUA UAC AAC CUA CUG CCU UCC U -3′
hsa-pre-let-7c (Unmodified)	5′-UGA GGU AGU AGG UUG UAU GGU UUA GAG UUA CAC CCU GGG AGU UAA CUG UAC AAC CUU CUA GCU UUC C -3′
hsa-pre-let-7c (Mono-U)	5′-UGA GGU AGU AGG UUG UAU GGU UUA GAG UUA CAC CCU GGG AGU UAA CUG UAC AAC CUU CUA GCU UUC CU -3′
ds-35 (antisense)	5′-UGA GGU AGU AGG UUG UAU AGU UUU AGG GUC ACA CC -3′
ds-35 (sense, +0)	5′-GGU GUG ACC CUA AAA CUA UAC AAC CUA CUA CCU CA -3′
ds-35 (sense, +1)	5′-GGU GUG ACC CUA AAA CUA UAC AAC CUA CUA CCU CAU -3′

ds-35 (sense, +2)	5′-GGU GUG ACC CUA AAA CUA UAC AAC CUA CUA CCU CAU U -3′
ds-35 (sense, +3)	5′-GGU GUG ACC CUA AAA CUA UAC AAC CUA CUA CCU CAU UU -3′
ds-35 (sense, +4)	5′-GGU GUG ACC CUA AAA CUA UAC AAC CUA CUA CCU CAU UUU -3′

CHAPTER III

Conclusion

MicroRNAs are single-stranded RNAs of ~22nt in length that are associated with Argonaute family proteins to repress their complementary target mRNAs. As key regulators in various cellular functions, miRNAs themselves need to be tightly controlled. MiRNAs are generated by stepwise processing of RNase III enzymes, Drosha and Dicer, which provide regulatory points of miRNA maturation.

In this study, I identified TUT7, TUT4, and TUT2 as novel regulators of the miRNA biogenesis pathway that act between Drosha and Dicer steps (**Figure III-1** for a model). TUT7/4/2 mono-uridylate majority of pre-let-7 members, which have a 1-nt 3' overhang as a result of Drosha processing. Mono-uridylation creates a 2-nt 3' overhangs of pre-miRNAs, which is optimal for Dicer processing. Based on the end structure of the precursor and dependency for mono-uridylation, miRNAs is subdivided into two groups, group I (which has a 2-nt 3' overhang) and group II (which has a 1-nt 3' overhang). The vast majority of let-7 miRNAs and miR-105 belong to group II.

Using both biochemical assay and analysis of miRNA levels in cells, I showed that TUT7/4/2 specifically recognize and mono-uridylate a 1-nt 3′ overhang structure of group II miRNAs. It will be of interest to investigate how TUTs specifically recognize the 1-nt 3′ overhang structure and transfer only one U residue. Recently, crystal structure of Cid1, a homolog of TUT7/4/2 in *Schizosaccharomyces pombe* was reported revealing the

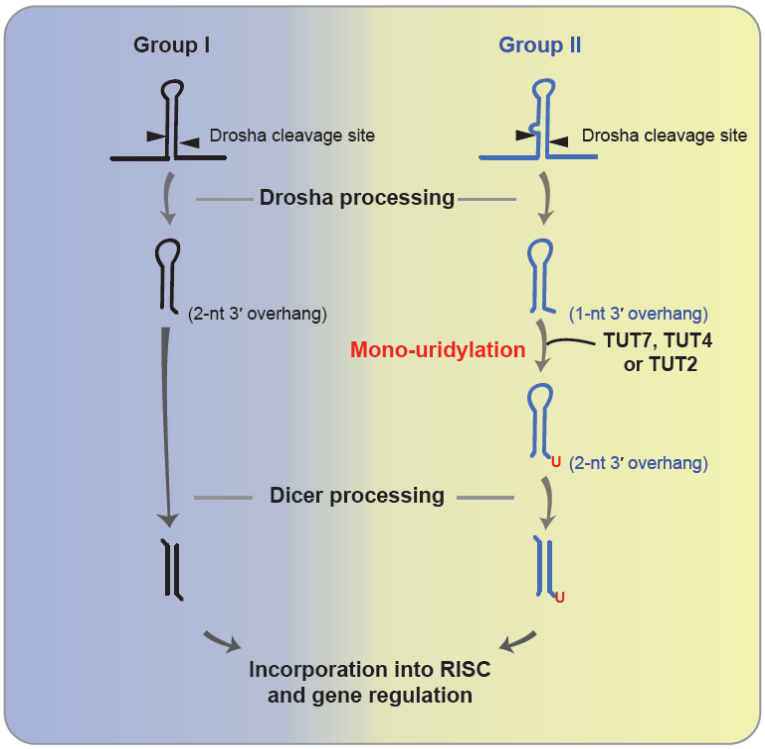


Figure III-1. Mono-uridylation is an integral step for group II miRNA biogenesis.

Unlike typical pre-miRNAs which have a 2-nt overhang at the 3′ end (group I), group II pre-miRNAs acquire a 1-nt 3′ overhang end structure due to a bulged nucleotide next to Drosha cleavage site. TUT7, TUT4, and TUT2 redundantly carry out mono-uridylation of group II pre-miRNAs, transforming them to optimal substrates for Dicer cleavage.

molecular mechanism of UTP selectivity of the enzymes (Lunde et al., 2012; Munoz-Tello et al., 2012; Yates et al., 2012). However, human TUT7 or TUT4 is much larger than Cid1 as it contains a duplication of both motifs (a nucleotidyl transferase motif and a poly (A) polymerase associated motif) as well as four zinc finger motifs. Structural study of human TUT7/4/2 in complex with its RNA substrate is required to uncover the molecular determinants of substrate specificity. In addition, biochemical property of TUTs awaits further examination to provide an insight into the molecular mechanism of miRNA uridylation.

It was previously reported that Lin28 increases the dwelling time of TUT4 on pre-let-7 (Yeom et al., 2011). Thus, Lin28 enhances the processivity of TUT4 to induce oligo-uridylation, which in turn blocks Dicer processing and facilitates pre-let-7 decay (Hagan et al., 2009; Heo et al., 2008; Heo et al., 2009). Therefore,

Lin28 serves as a molecular switch that converts TUT4 (and TUT7) from key biogenesis factors into negative regulators (**Figure III-2**). In parallel with my group's previous study, my work unveils two opposing functions of TUTs in miRNA biogenesis. Let-7 miRNAs is crucial for cell differentiation and proliferation. The functional duality of uridylation may contribute to the tight control of let-7 expression during developmental transition and tumorigenesis.

To conclude, miRNA precursors are regulated post-transcriptionally through uridylation (**Figure III-3**). When Lin28 expresses, TUT4 (and TUT7)

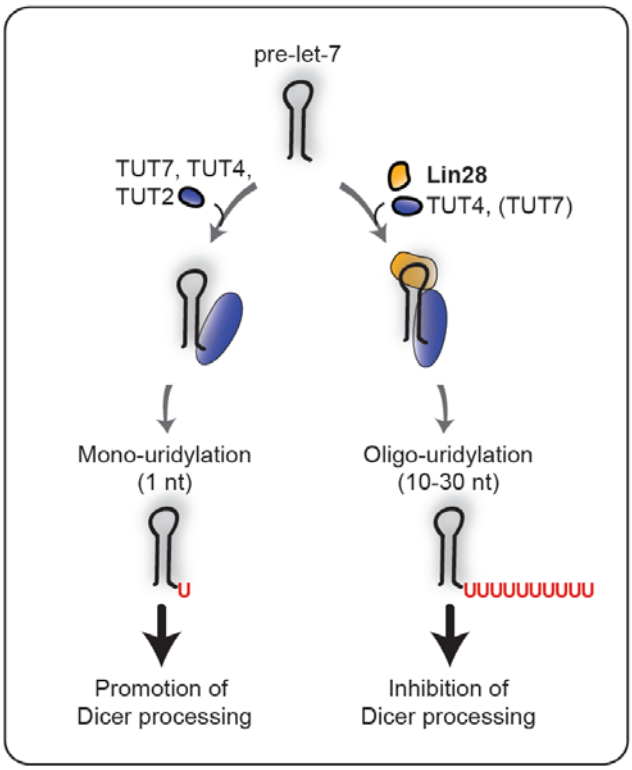


Figure II-2. A model for dual effects of uridylation

Terminal uridylation of pre-let-7 has two opposite effects on let-7 biogenesis depending on the length of uridyl tail. In somatic cells lacking Lin28, TUT7/4/2 redundantly mono-uridylate group II pre-let-7, which promotes Dicer processing. When Lin28 is expressed in embryonic cells, Lin28 increases processivity of TUT4 (and TUT7) and thereby TUT4/7 induce oligo-uridylation (10-30 nt) of pre-let-7, which inhibits Dicer processing. TUT4 play a dominant role over TUT7 in oligo-uridylation. Lin28 acts as a molecular switch by converting TUT4 (and TUT7) from biogenesis factors to negative regulators.

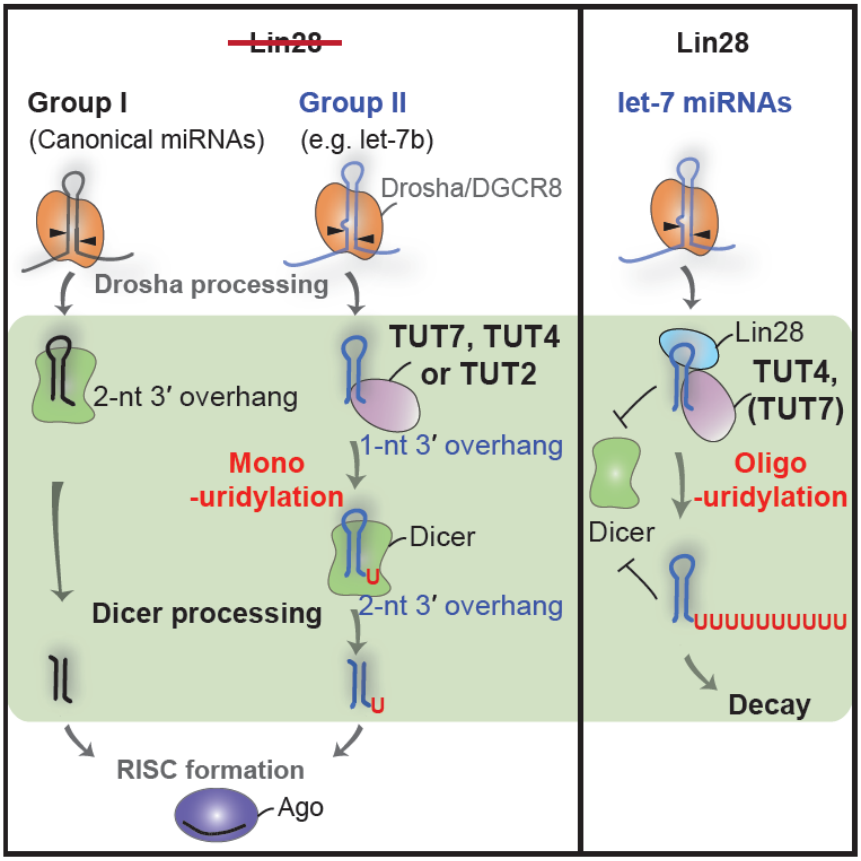


Figure III-3. Post-transcriptional control of miRNA precursors

In the absence of Lin28, TUT7/4/2 specifically mono-uridylate group II pre-miRNAs that have a 3' 1-nt overhang. In Lin28 expressing cells, TUT4 (and TUT) induce oligouridylation of pre-let-7 regardless of their 3' overhang length.

induce oligouridylation of all members of pre-let-7 regardless of their 3' overhang length. Whereas in the absence of Lin28, TUT7/4/2 specifically mono-uridylate group II pre-miRNAs that have a 3' 1-nt overhang. Although group I pre-miRNAs are not regulated by TUT7/4/2 in HeLa cells, it is possible that other TUTs may regulate group I miRNAs in other cell types.

It is clear that depletion of TUT7/4/2 disrupts let-7 functions as demonstrated by target mRNA sequencing analysis, implicating that dependence upon TUTs for expression of let7 miRNAs will be related to biology in the future. Extensive works have been performed on the crucial role of the Lin28 - let-7 control in development and disease (Thornton and Gregory, 2012). However, few studies have reported about the role of TUT7/4/2 in development or cancer. Expression profiling of TUTs and let-7 miRNAs in developmental stages or various tissue will give clues to physiological context where TUTs promote let-7 activity. Furthermore, knockout mouse study of TUTs can be utilized to uncover the role of TUTs in development which may interplay with miRNA pathway.

REFERENCES

Backes, S., Shapiro, J.S., Sabin, L.R., Pham, A.M., Reyes, I., Moss, B., Cherry, S., and tenOever, B.R. (2012). Degradation of host microRNAs by poxvirus poly(A) polymerase reveals terminal RNA methylation as a protective antiviral mechanism. Cell host & microbe *12*, 200-210.

Balzer, E., and Moss, E.G. (2007). Localization of the developmental timing regulator Lin28 to mRNP complexes, P-bodies and stress granules. RNA biology 4, 16-25.

Bartel, D.P. (2009). MicroRNAs: target recognition and regulatory functions. Cell 136, 215-233.

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.

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.

Bompfunewerer, A.F., Flamm, C., Fried, C., Fritzsch, G., Hofacker, I.L., Lehmann, J., Missal, K., Mosig, A., Muller, B., Prohaska, S.J., et al. (2005).

Evolutionary patterns of non-coding RNAs. Theory in biosciences = Theorie in den Biowissenschaften 123, 301-369.

Burns, D.M., D'Ambrogio, A., Nottrott, S., and Richter, J.D. (2011). CPEB and two poly(A) polymerases control miR-122 stability and p53 mRNA translation. Nature 473, 105-108.

Burroughs, A.M., Ando, Y., de Hoon, M.J., Tomaru, Y., Nishibu, T., Ukekawa, R., Funakoshi, T., Kurokawa, T., Suzuki, H., Hayashizaki, Y., et al. (2010). A comprehensive survey of 3' animal miRNA modification events and a possible role for 3' adenylation in modulating miRNA targeting effectiveness. Genome research 20, 1398-1410.

Bussing, I., Slack, F.J., and Grosshans, H. (2008). let-7 microRNAs in development, stem cells and cancer. Trends in molecular medicine 14, 400-409.

Chatterjee, S., Fasler, M., Bussing, I., and Grosshans, H. (2011). Target-mediated protection of endogenous microRNAs in C. elegans. Developmental cell 20, 388-396.

Chatterjee, S., and Grosshans, H. (2009). Active turnover modulates mature microRNA activity in Caenorhabditis elegans. Nature 461, 546-549.

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 & development 24, 992-1009.

Das, S.K., Sokhi, U.K., Bhutia, S.K., Azab, B., Su, Z.Z., Sarkar, D., and Fisher, P.B. (2010). Human polynucleotide phosphorylase selectively and preferentially degrades microRNA-221 in human melanoma cells. Proceedings of the National Academy of Sciences of the United States of America 107, 11948-11953.

Davis-Dusenbery, B.N., and Hata, A. (2010). Mechanisms of control of microRNA biogenesis. Journal of biochemistry 148, 381-392.

Davis, B.N., Hilyard, A.C., Lagna, G., and Hata, A. (2008). SMAD proteins control DROSHA-mediated microRNA maturation. Nature 454, 56-61.

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.

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.

Griffiths-Jones, S., Saini, H.K., van Dongen, S., and Enright, A.J. (2008). miRBase: tools for microRNA genomics. Nucleic acids research 36, D154-158.

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.

Grosshans, H., Johnson, T., Reinert, K.L., Gerstein, M., and Slack, F.J. (2005). The temporal patterning microRNA let-7 regulates several transcription factors at the larval to adult transition in C. elegans. Developmental cell 8, 321-330.

Guil, S., and Caceres, J.F. (2007). The multifunctional RNA-binding protein hnRNP A1 is required for processing of miR-18a. Nature structural & molecular biology 14, 591-596.

Guo, Y., Chen, Y., Ito, H., Watanabe, A., Ge, X., Kodama, T., and Aburatani, H. (2006). Identification and characterization of lin-28 homolog B (LIN28B) in human hepatocellular carcinoma. Gene 384, 51-61.

Hagan, J.P., Piskounova, E., and Gregory, R.I. (2009). Lin28 recruits the TUTase Zcchc11 to inhibit let-7 maturation in mouse embryonic stem cells. Nature structural & molecular biology 16, 1021-1025.

Hammond, S.M., Boettcher, S., Caudy, A.A., Kobayashi, R., and Hannon, G.J. (2001). Argonaute2, a link between genetic and biochemical analyses of RNAi. Science 293, 1146-1150.

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 & development 18, 3016-3027.

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.

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. Molecular cell 32, 276-284.

Heo, I., Joo, C., Kim, Y.K., Ha, M., Yoon, M.J., Cho, J., Yeom, K.H., Han, J., and Kim, V.N. (2009). TUT4 in concert with Lin28 suppresses microRNA biogenesis through pre-microRNA uridylation. Cell 138, 696-708.

Huntzinger, E., and Izaurralde, E. (2011). Gene silencing by microRNAs: contributions of translational repression and mRNA decay. Nature reviews Genetics 12, 99-110.

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.

Ji, L., and Chen, X. (2012). Regulation of small RNA stability: methylation and beyond. Cell research.

Jones, M.R., Quinton, L.J., Blahna, M.T., Neilson, J.R., Fu, S., Ivanov, A.R., Wolf, D.A., and Mizgerd, J.P. (2009). Zcchc11-dependent uridylation of microRNA directs cytokine expression. Nature cell biology 11, 1157-1163.

Katoh, T., Sakaguchi, Y., Miyauchi, K., Suzuki, T., Kashiwabara, S., and Baba, T. (2009). Selective stabilization of mammalian microRNAs by 3' adenylation mediated by the cytoplasmic poly(A) polymerase GLD-2. Genes & development 23, 433-438.

Kawahara, Y., Zinshteyn, B., Chendrimada, T.P., Shiekhattar, R., and Nishikura, K. (2007). RNA editing of the microRNA-151 precursor blocks cleavage by the Dicer-TRBP complex. EMBO reports 8, 763-769.

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 & development 15, 2654-2659.

Kim, V.N., Han, J., and Siomi, M.C. (2009). Biogenesis of small RNAs in animals. Nature reviews Molecular cell biology 10, 126-139.

Kim, Y.K., Heo, I., and Kim, V.N. (2010). Modifications of small RNAs and their associated proteins. Cell 143, 703-709.

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.

Krol, J., Busskamp, V., Markiewicz, I., Stadler, M.B., Ribi, S., Richter, J., Duebel, J., Bicker, S., Fehling, H.J., Schubeler, D., et al. (2010a). Characterizing light-regulated retinal microRNAs reveals rapid turnover as a common property of neuronal microRNAs. Cell 141, 618-631.

Krol, J., Loedige, I., and Filipowicz, W. (2010b). The widespread regulation of microRNA biogenesis, function and decay. Nature reviews Genetics 11, 597-610.

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. Current biology : CB 14, 2162-2167.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nature methods 9, 357-359.

Lee, R.C., Feinbaum, R.L., and Ambros, V. (1993). The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell 75, 843-854.

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.

Lewis, B.P., Burge, C.B., and Bartel, D.P. (2005). Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell 120, 15-20.

Loughlin, F.E., Gebert, L.F., Towbin, H., Brunschweiger, A., Hall, J., and Allain, F.H. (2012). Structural basis of pre-let-7 miRNA recognition by the zinc knuckles of pluripotency factor Lin28. Nature structural & molecular biology 19, 84-89.

Lund, E., Guttinger, S., Calado, A., Dahlberg, J.E., and Kutay, U. (2004). Nuclear export of microRNA precursors. Science 303, 95-98.

Lunde, B.M., Magler, I., and Meinhart, A. (2012). Crystal structures of the Cid1 poly (U) polymerase reveal the mechanism for UTP selectivity. Nucleic acids research 40, 9815-9824.

MacRae, I.J., Zhou, K., and Doudna, J.A. (2007). Structural determinants of RNA recognition and cleavage by Dicer. Nature structural & molecular biology 14, 934-940.

Macrae, I.J., Zhou, K., Li, F., Repic, A., Brooks, A.N., Cande, W.Z., Adams, P.D., and Doudna, J.A. (2006). Structural basis for double-stranded RNA processing by Dicer. Science 311, 195-198.

Martin, G., and Keller, W. (2007). RNA-specific ribonucleotidyl transferases.

RNA 13, 1834-1849.

Meneely, P.M., and Herman, R.K. (1979). Lethals, steriles and deficiencies in a region of the X chromosome of Caenorhabditis elegans. Genetics 92, 99-115.

Michlewski, G., Guil, S., Semple, C.A., and Caceres, J.F. (2008). Posttranscriptional regulation of miRNAs harboring conserved terminal loops. Molecular cell 32, 383-393.

Mourelatos, Z., Dostie, J., Paushkin, S., Sharma, A., Charroux, B., Abel, L., Rappsilber, J., Mann, M., and Dreyfuss, G. (2002). miRNPs: a novel class of ribonucleoproteins containing numerous microRNAs. Genes & development 16, 720-728.

Munoz-Tello, P., Gabus, C., and Thore, S. (2012). Functional implications from the Cid1 poly(U) polymerase crystal structure. Structure 20, 977-986.

Nakanishi, T., Kubota, H., Ishibashi, N., Kumagai, S., Watanabe, H., Yamashita, M., Kashiwabara, S., Miyado, K., and Baba, T. (2006). Possible role of mouse poly(A) polymerase mGLD-2 during oocyte maturation. Developmental biology 289, 115-126.

Nam, Y., Chen, C., Gregory, R.I., Chou, J.J., and Sliz, P. (2011). Molecular basis for interaction of let-7 microRNAs with Lin28. Cell 147, 1080-1091.

Newman, M.A., Mani, V., and Hammond, S.M. (2011). Deep sequencing of

microRNA precursors reveals extensive 3' end modification. RNA 17, 1795-1803.

Newman, M.A., Thomson, J.M., and Hammond, S.M. (2008). Lin-28 interaction with the Let-7 precursor loop mediates regulated microRNA processing. RNA 14, 1539-1549.

Niwa, R., and Slack, F.J. (2007). The evolution of animal microRNA function. Current opinion in genetics & development 17, 145-150.

Park, J.E., Heo, I., Tian, Y., Simanshu, D.K., Chang, H., Jee, D., Patel, D.J., and Kim, V.N. (2011). Dicer recognizes the 5' end of RNA for efficient and accurate processing. Nature 475, 201-205.

Pasquinelli, A.E., McCoy, A., Jimenez, E., Salo, E., Ruvkun, G., Martindale, M.Q., and Baguna, J. (2003). Expression of the 22 nucleotide let-7 heterochronic RNA throughout the Metazoa: a role in life history evolution? Evolution & development 5, 372-378.

Pasquinelli, A.E., Reinhart, B.J., Slack, F., Martindale, M.Q., Kuroda, M.I., Maller, B., Hayward, D.C., Ball, E.E., Degnan, B., Muller, P., et al. (2000). Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. Nature 408, 86-89.

Piskounova, E., Polytarchou, C., Thornton, J.E., LaPierre, R.J., Pothoulakis, C., Hagan, J.P., Iliopoulos, D., and Gregory, R.I. (2011). Lin28A and

Lin28B inhibit let-7 microRNA biogenesis by distinct mechanisms. Cell 147, 1066-1079.

Polesskaya, A., Cuvellier, S., Naguibneva, I., Duquet, A., Moss, E.G., and Harel-Bellan, A. (2007). Lin-28 binds IGF-2 mRNA and participates in skeletal myogenesis by increasing translation efficiency. Genes & development 21, 1125-1138.

Ramachandran, V., and Chen, X. (2008). Degradation of microRNAs by a family of exoribonucleases in Arabidopsis. Science 321, 1490-1492.

Reinhart, B.J., Slack, F.J., Basson, M., Pasquinelli, A.E., Bettinger, J.C., Rougvie, A.E., Horvitz, H.R., and Ruvkun, G. (2000). The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. Nature 403, 901-906.

Ren, G., Chen, X., and Yu, B. (2012). Uridylation of miRNAs by hen1 suppressor1 in Arabidopsis. Current biology : CB 22, 695-700.

Richards, M., Tan, S.P., Tan, J.H., Chan, W.K., and Bongso, A. (2004). The transcriptome profile of human embryonic stem cells as defined by SAGE. Stem cells 22, 51-64.

Roberts, A., Pimentel, H., Trapnell, C., and Pachter, L. (2011). Identification of novel transcripts in annotated genomes using RNA-Seq. Bioinformatics 27, 2325-2329.

Rouhana, L., Wang, L., Buter, N., Kwak, J.E., Schiltz, C.A., Gonzalez, T., Kelley, A.E., Landry, C.F., and Wickens, M. (2005). Vertebrate GLD2 poly(A) polymerases in the germline and the brain. RNA 11, 1117-1130.

Roush, S., and Slack, F.J. (2008). The let-7 family of microRNAs. Trends in cell biology 18, 505-516.

Ruegger, S., and Grosshans, H. (2012). MicroRNA turnover: when, how, and why. Trends in biochemical sciences 37, 436-446.

Rybak, A., Fuchs, H., Smirnova, L., Brandt, C., Pohl, E.E., Nitsch, R., and Wulczyn, F.G. (2008). A feedback loop comprising lin-28 and let-7 controls pre-let-7 maturation during neural stem-cell commitment. Nature cell biology 10, 987-993.

Saj, A., and Lai, E.C. (2011). Control of microRNA biogenesis and transcription by cell signaling pathways. Current opinion in genetics & development 21, 504-510.

Siomi, H., and Siomi, M.C. (2010). Posttranscriptional regulation of microRNA biogenesis in animals. Molecular cell 38, 323-332.

Stevenson, A.L., and Norbury, C.J. (2006). The Cid1 family of non-canonical poly(A) polymerases. Yeast 23, 991-1000.

Suzuki, H.I., Arase, M., Matsuyama, H., Choi, Y.L., Ueno, T., Mano, H., Sugimoto, K., and Miyazono, K. (2011). MCPIP1 ribonuclease

antagonizes dicer and terminates microRNA biogenesis through precursor microRNA degradation. Molecular cell 44, 424-436.

Tabara, H., Sarkissian, M., Kelly, W.G., Fleenor, J., Grishok, A., Timmons, L., Fire, A., and Mello, C.C. (1999). The rde-1 gene, RNA interference, and transposon silencing in C. elegans. Cell 99, 123-132.

Thornton, J.E., and Gregory, R.I. (2012). How does Lin28 let-7 control development and disease? Trends in cell biology 22, 474-482.

Tian, B., Bevilacqua, P.C., Diegelman-Parente, A., and Mathews, M.B. (2004). The double-stranded-RNA-binding motif: interference and much more. Nature reviews Molecular cell biology 5, 1013-1023.

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.

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinformatics 25, 1105-1111.

Upton, J.P., Wang, L., Han, D., Wang, E.S., Huskey, N.E., Lim, L., Truitt, M., McManus, M.T., Ruggero, D., Goga, A., et al. (2012). IRE1alpha cleaves select microRNAs during ER stress to derepress translation of proapoptotic Caspase-2. Science 338, 818-822.

Vermeulen, A., Behlen, L., Reynolds, A., Wolfson, A., Marshall, W.S., Karpilow, J., and Khvorova, A. (2005). The contributions of dsRNA structure to Dicer specificity and efficiency. RNA 11, 674-682.

Viswanathan, S.R., Daley, G.Q., and Gregory, R.I. (2008). Selective blockade of microRNA processing by Lin28. Science 320, 97-100.

Wickens, M., and Kwak, J.E. (2008). Molecular biology. A tail tale for U. Science 319, 1344-1345.

Wilusz, C.J., and Wilusz, J. (2008). New ways to meet your (3') end oligouridylation as a step on the path to destruction. Genes & development 22, 1-7.

Wu, H., Ye, C., Ramirez, D., and Manjunath, N. (2009). Alternative processing of primary microRNA transcripts by Drosha generates 5' end variation of mature microRNA. PloS one 4, e7566.

Wu, T.D., and Nacu, S. (2010). Fast and SNP-tolerant detection of complex variants and splicing in short reads. Bioinformatics 26, 873-881.

Wyman, S.K., Knouf, E.C., Parkin, R.K., Fritz, B.R., Lin, D.W., Dennis, L.M., Krouse, M.A., Webster, P.J., and Tewari, M. (2011). Post-transcriptional generation of miRNA variants by multiple nucleotidyl transferases contributes to miRNA transcriptome complexity. Genome research 21, 1450-1461.

Xhemalce, B., Robson, S.C., and Kouzarides, T. (2012). Human RNA Methyltransferase BCDIN3D Regulates MicroRNA Processing. Cell 151, 278-288.

Yang, D.H., and Moss, E.G. (2003). Temporally regulated expression of Lin-28 in diverse tissues of the developing mouse. Gene expression patterns : GEP 3, 719-726.

Yang, W., Chendrimada, T.P., Wang, Q., Higuchi, M., Seeburg, P.H., Shiekhattar, R., and Nishikura, K. (2006). Modulation of microRNA processing and expression through RNA editing by ADAR deaminases. Nature structural & molecular biology 13, 13-21.

Yates, L.A., Fleurdepine, S., Rissland, O.S., De Colibus, L., Harlos, K., Norbury, C.J., and Gilbert, R.J. (2012). Structural basis for the activity of a cytoplasmic RNA terminal uridylyl transferase. Nature structural & molecular biology 19, 782-787.

Yeom, K.H., Heo, I., Lee, J., Hohng, S., Kim, V.N., and Joo, C. (2011). Single-molecule approach to immunoprecipitated protein complexes: insights into miRNA uridylation. EMBO reports 12, 690-696.

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 & development 17, 3011-3016.

Zhang, H., Kolb, F.A., Brondani, V., Billy, E., and Filipowicz, W. (2002). Human Dicer preferentially cleaves dsRNAs at their termini without a requirement for ATP. The EMBO journal 21, 5875-5885.

Zhang, H., Kolb, F.A., Jaskiewicz, L., Westhof, E., and Filipowicz, W. (2004). Single processing center models for human Dicer and bacterial RNase III. Cell 118, 57-68.

Zhao, Y., Yu, Y., Zhai, J., Ramachandran, V., Dinh, T.T., Meyers, B.C., Mo, B., and Chen, X. (2012). The Arabidopsis nucleotidyl transferase HESO1 uridylates unmethylated small RNAs to trigger their degradation. Current biology : CB 22, 689-694.

Zhu, H., Shah, S., Shyh-Chang, N., Shinoda, G., Einhorn, W.S., Viswanathan, S.R., Takeuchi, A., Grasemann, C., Rinn, J.L., Lopez, M.F., et al. (2010). Lin28a transgenic mice manifest size and puberty phenotypes identified in human genetic association studies. Nature genetics 42, 626-630.

Zhu, H., Shyh-Chang, N., Segre, A.V., Shinoda, G., Shah, S.P., Einhorn, W.S., Takeuchi, A., Engreitz, J.M., Hagan, J.P., Kharas, M.G., et al. (2011). The Lin28/let-7 axis regulates glucose metabolism. Cell 147, 81-94.

PUBLICATIONS

Mono-Uridylation of Pre-MicroRNA as a Key Step in the Biogenesis of Group II let-7 MicroRNAs, I. Heo*, M. Ha*, J. Lim, M. J. Yoon, J. E. Park, S. C. Kwon, H. Chang and V. N. Kim (2012) (*co-first authors) *Cell* 151:521 – 532

LIN28A is a suppressor of ER-associated translation in embryonic stem cells, J. Cho, H. Chang, S. C. Kwon, B. Kim, Y. Kim, J. Choe, M. Ha, Y. K. Kim and V. N. Kim (2012) *Cell* 151: 765-777

Short Structured RNAs with Low GC Content Are Selectively Lost during Extraction from a Small Number of Cells, Y. K. Kim, J. Yeo, B. Kim, M. Ha and V. N. Kim (2012) *Molecular Cell* 46:893-895

Human UPF1 modulates small RNA-induced mRNA down-regulation, H. Jin, M. R. Suh, J. Han, K. H. Yeom, Y. Lee, I. Heo, M. Ha, S. Hyun and V. N. Kim (2009) *Mol. and Cell. Biol.* 20(21):5789-99

TUT4 in Concert with Lin28 Suppresses MicroRNA Biogenesis through Pre-MicroRNA Uridylation, I. Heo, C. Joo, Y. K. Kim, M. Ha, M. J. Yoon, J. Cho, K. H. Yeom, J. Han, and V. N. Kim (2009) *Cell* 138:696-708

miR-29 miRNAs activate p53 by targeting p85a and CDC42, S. Y. Park, J. H. Lee, M. Ha, J. W. Nam and V. N. Kim (2009) *Nature Structural and Molecular Biology* 16(1):23-9

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

국문 초록

마이크로RNA는 약 22개 뉴클레오티드 길이의 단일 가닥 RNA로서, 타겟 유전자 발현을 억제하는 기능을 한다. 마이크로RNA는 세포 안에서 여러 단계의 효소 작용을 거쳐서 만들어진다. 마이크로RNA의 첫번째 전구체인 pri-마이크로RNA는 RNase III type에 속하는 Drosha 효소에 의해서 잘려서, 3'쪽 말단에 2개 뉴클레오티드가 돌출된 모양 (overhang) 을 가지는 두번째 전구체 pre-마이크로RNA를 만들어낸다. Dicer 효소는 이러한 3'쪽 2개 뉴클레오티드 overhang을 인지하고 pre-마이크로RNA에 선택적으로 작용하여 성숙한 마이크로RNA를 만들어낸다.

Let-7 마이크로RNA는 가장 오래되고 잘 보존된 마이크로RNA 중 하나로서, 발생단계 중 배아 시기와 일부 암에서 그 발현이 억제되어 있다. 기존 연구 결과로 RNA 말단에 유리딘화 변형을 담당하는 효소인 Terminal uridylyl transferase 4 (TUT4) 가

let-7 pre-마이크로RNA (pre-let-7)의 말단을 올리고-유리딘화하여 let-7 마이크로RNA의 생성을 억제함이 알려져 있었다. 이러한 TUT4 단백질은 줄기세포 유지에 중요한 단백질인 Lin28이 발현하는 경우에만 pre-let-7에 올리고-유리딘화 변형을 가한다.

본 연구에서 나는 TUT 단백질들이 pre-let-7을 모노-유리딘화하여 마이크로RNA의 생성을 촉진한다는 사실을 보고하였다. Lin28 단백질이 발현하지 않는 경우, TUT7/ZCCHC6, TUT4/ZCCHC11, TUT2/PAPD4/GLD2 효소들이 이러한 모노유리딘화를 담당한다. Pre-let-7이 모노-유리딘화 되면 Dicer 효소에 적합한 구조인 3'쪽 2개 뉴클레오티드 overhang 구조를 갖게 되기 때문에 결과적으로 let-7 마이크로RNA가 더 잘 만들어지게 된다.

더 나아가, 마이크로RNA는 pre-마이크로RNA의 말단 구조와 모노-유리딘화에 의존성 여부에 따라서 두가지 그룹, 그룹 I (pre-마이크로RNA가 3'쪽 말단에 2개 뉴클레오티드 overhang

을 가짐) 과 그룹 II (pre-마이크로RNA가 3'쪽 말단에 1개 뉴클레오티드 overhang을 가짐) 로 나뉠 수 있었다. 일반적인 pre-마이크로RNA (그룹 I) 와 달리, 그룹 II 에 속하는 pre-마이크로RNA는 Drosha 효소가 3'쪽 말단에 짧은 (1개 뉴클레오티드) overhang 구조를 만들어내기 때문에, 모노-유리딘화 변형이 추가적으로 일어나야만 Dicer 효소가 잘 작용할 수 있는 구조로 만들어진다. 대부분의 let-7 마이크로RNA와 miR-105가 그룹 II 에 속한다. TUT7/4/2 단백질은 이중가닥 RNA의 3'쪽 말단의 1개 뉴클레오티드 overhang 을 선택적으로 인지하고 모노-유리딘화하여 3'쪽 2개 뉴클레오티드 overhang 구조로 변형시켜 준다. 세포내에서 TUT 단백질들의 양을 감소시키면 let-7 마이크로RNA의 양이 감소하고, let-7 마이크로RNA의 기능이 저해된다.

Let-7 억제인자인 Lin28 단백질이 줄기세포에서 마이크로RNA의 생성을 억제하는 올리고-유리딘화를 유도함에 반해서, 모노-유리딘화는 Lin28 단백질이 발현하지 않는 체세포에서 let-

7 마이크로RNA의 생성을 촉진한다. 나는 본 연구에서 유리딘화 변형이 길이에 따라 기능적 이중성을 가진다는 것과 마이크로RNA 생성 과정에 관여하는 새로운 단백질들인 TUT7/4/2 의 역할을 보고하였다.

주요어

마이크로RNA/ 마이크로RNA 생성 과정/ 전사후 조절기작/ pre-마이크로RNA/ 마이크로RNA 전구체/ let-7/ 유리딘화/ RNA 말단 유리딘화 효소 (TUT)/ TUT7/ TUT4/ TUT2/ Dicer