Thesis for the *Master of Science*

The synergistic effect of UPF1 binding to 3'UTR on microRNA targeting

Jwa-Won SEO

Graduate School of Hanyang University

August 2016

Thesis for the *Master of Science*

The synergistic effect of UPF1 binding to 3'UTR on microRNA targeting

Thesis Supervisor: Jin-Wu NAM

A Thesis submitted to the graduate
school of Hanyang University in partial
fulfillment of the requirements for the degree of
Master of Science

Jwa-Won SEO

August 2016

Department of Life Science
Graduate School of Hanyang University

This thesis, written by *Jwa-Won SEO*,
has been approved as a thesis for the degree of
Master of Science.

August 2016

Committee Chairman: Chul Geun Kim (Signature)

Committee member: Jungwook Hwang (Signature)

Committee member: Jin-Wu Nam (Signature)

Graduate School of Hanyang University

Table of Contents

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

The synergistic effect of UPF1 binding to 3'UTR on microRNA targeting

Jwa-Won Seo

Department of Life Science

Graduate School of

Hanyang University

UPF1 is a well-known RNA helicase with essential roles in nonsense-mediated mRNA decay (NMD), a surveillance pathway. However, UPF1 binds predominantly not to coding region but to 3’UTR of mature mRNA and regulates the mRNA abundance depending on the length of 3’UTR. UPF1 also interacts with Argonaute (AGO) proteins, which are major components of microRNA-mediated mRNA decay. The global relationship between UPF1 binding and miRNA-mediated mRNA decay in 3’UTR is barely known. In our study, we reanalyzed available RNA-seq which data followed UPF1 down-regulation to confirm a NMD independent repression pathway. In selected non-NMD target genes, we found that UPF1 repressed mRNA abundance along the 3’UTR

length. We identified binding sites of UPF1 and AGO2 using available CLIP-seq data. A majority of AGO2 binding sites co-localized with UPF1 proteins in which endogenous miRNA (endo-miRNA) target sites in mouse embryonic stem (mES) cells are embedded. In the UPF1 knock-downed mES cells, co-localized UPF1 and AGO2 with endo-miRNA targets were significantly more derepressed than only AGO2 localized targets and nontargets, suggesting a synergistic effect of UPF1 on miRNA targeting. In mRNAs with no endo-miRNA target sites, the 3' UTR dependency of UPF1-mediated mRNA down-regulation was greatly reduced. Additionally, we suggested that SMG7 protein was associated with the downstream mechanism of the UPF1-mediated mRNA decay. In SMG7 knock-downed HeLa cells, we confirmed that non-NMD target genes showed positive correlation between derepression and the length of 3' UTR. Moreover, SMG7 derepressed endo-miRNA targets like UPF1-mediated mRNA decay. Taken together, it suggests that the synergistic role of UPF1 interacting with AGO2 in miRNA-mediated mRNA decay is evolutionally conserved, and UPF1-mediated mRNA decay is associated with SMG7.

Chapter 1. Introduction

Nonsense-mediated mRNA decay (NMD) is a general cellular quality control system to prevent the cell from producing aberrant proteins. Mammalian NMD pathway degrades mRNAs including premature termination codons (PTCs) which is positioned more than 50-55 nt upstream of the last exon-exon junction (Kervestin and Jacobson 2012; Schweingruber et al. 2013; Popp and Maquat 2013; Smith and Baker 2015).

UPF1 is a well-known RNA helicase as an essential factor in NMD, because it plays a role as a core bridge in NMD machinery. Particularly, UPF1 can identify PTC-containing mRNAs by interaction with a ribosome which stalls at the PTC. In this selective recognition, ATPase and helicase activities of UPF1 is necessary. (Weng and Peltz 1996; Kashima et al. 2006; Franks and Lykke 2010; Kurosaki et al. 2014)

To mediated mRNA decay, several different downstream mechanisms are triggered by the UPF1 in Mammalian NMD. For example, phosphorylated UPF1 recruits endonuclease SMG6 which splits PTC-proximal sites on mRNA. Then, 5' and 3' cleavage fragments are degraded by 5'-3' exonuclease and 3'-5' exonuclease (Gatfield and Izaurralde 2004; Glavan et al. 2006; Huntzinger et al. 2009; Schmidt et al. 2014; Lykke-Andersen et al. 2014). On the other hand, when phosphorylated UPF1 binds to SMG5-SMG7

heterodimer, which cooperates with the carbon catabolite repressor protein 4 (CCR4)-NOT deadenylase complex to promote deadenylation and subsequent decapping of NMD targets (Unterholzner and Izaurralde 2004; Yamashita et al. 2005; Loh and Izaurralde 2013).

Several recent studies suggest that UPF1 is predominantly positioned in the 3' UTR of actively translated mRNAs without distinction of NMD-targeting and normal mRNAs. This non-canonical features have received much attention in recent years due to long 3' UTR shown to be sufficient for decay (Kurosaki et al. 2014; Hogg and Coff 2010; Kurosaki and Maquat 2013; Hurt et al. 2013; Zund et al. 2013; Gregersen et al. 2014).

However, non-NMD target mRNAs which do not have exon-exon junction more than 50nt downstream from an annotated stop codon are derepressed by depletion of UPF1. Also, increasing non-NMD target mRNAs 3' UTR length was correlated with increasing derepression (Hurt et al. 2013).

The indiscriminate binding of UPF1 on 3' UTR of mRNAs suggests that long 3' UTR feature is not enough to explain as NMD factor (Kashima et al. 2006; Czaplinski et al. 1998; Wang et al. 2001; Amrani et al. 2004; Hug and Caceres et al. 2014). Especially, regulation of non-NMD target mRNAs by UPF1 on 3' UTR is still unknown mechanism.

In RNA silencing, the typical decay pathway on 3' UTR is microRNA(miRNA)-mediated mRNA decay. miRNAs are small non-coding RNAs

that have ~22 nucleotides length. On recognition of targeted mRNAs, seed region defined as 5' end of miRNAs from nucleotide position 2 to 7 is essential (Lewis et al 2003; Lewis et al 2005). By Watson-Crick base pairing of seed region, miRNA plays a role as a regulator on 3' UTR (Ghildiyal and Zamore 2009; Huntzinger and Izaurralde 2011; Bartel 2009). In miRNA targeting, mature miRNAs are loaded into Argonaute (AGO) proteins of the silencing complex (Hutvagner and Zamore 2002; Mourelatos et al. 2002).

The AGO-miRNA complexes recruit factors that induce translational repression, mRNA deadenylation and mRNA decay, and constitute miRNA-induced silencing complexes (miRISCs). Concretely, AGO proteins interact with GW182 proteins which function as flexible scaffolds to bridge the interaction between AGO proteins and the cytoplasmic deadenylase complexes PAN2-PAN3 and CCR4-NOT (Huntzinger and Izaurralde 2011; Fabian and Sonenberg 2012).

Unlike the traditional miRNA targeting, UPF1 is suggested as a novel factor that is involved in miRNA-induced mRNA decay and interacts with the AGO proteins and colocalizes with the AGO proteins in P bodies (Jin et al. 2009).

In this study, we found that UPF1 destabilizes non-NMD target mRNAs depending on length of 3' UTR. Besides, UPF1 affects non-NMD target mRNAs

which could be targeted by highly expressed endogenous miRNA. Additionally, UPF1 is co-localized with AGO2 proteins, and AGO2-UPF1 complexes are associated with miRNA targeting. Interestingly, we show that SMG7 is also associated with the UPF1-mediated mRNA decay. Our results enable us to describe alternative miRNA targeting which is associated with UPF1 and SMG7, and to better understand the role of UPF1 bound on 3' UTR.

Chapter 2. UPF1 affects endogenous miRNA targeting with AGO2 proteins

Section 1. UPF1 destabilizes Non-NMD target mRNAs depending on length of 3' UTR

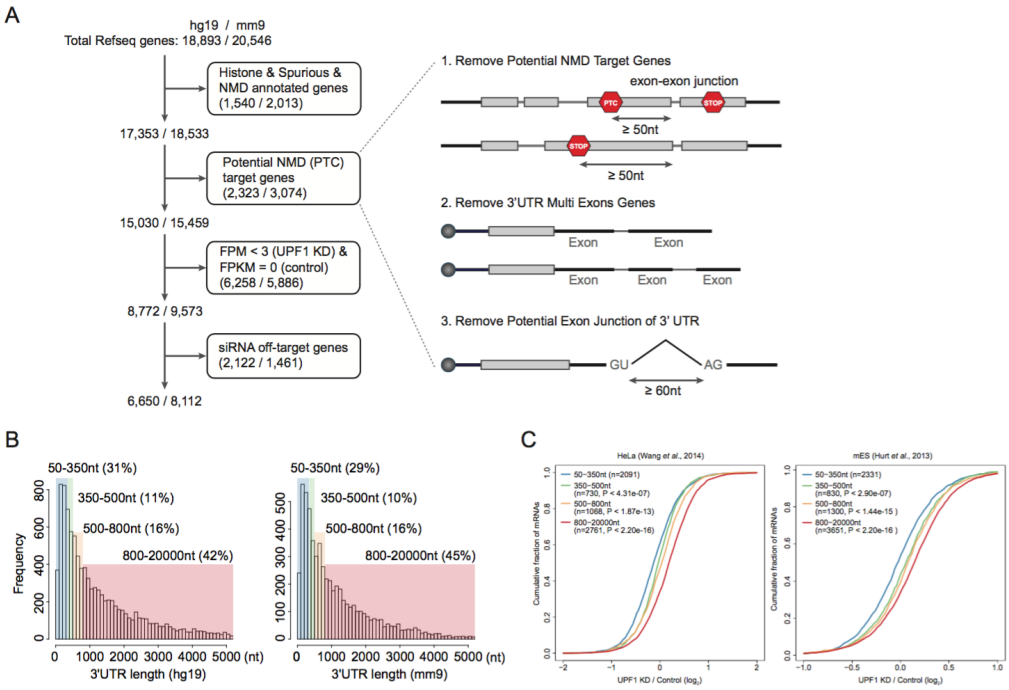

Figure 1. UPF1 derepresses non-NMD target mRNAs depending on length of 3' UTR. (A) A schematic flow for selection of non-NMD target mRNAs. (B) The proportion of 3' UTR length bins. (C) Cumulative distribution fuctions (CDFs) of changes in mRNA expression following UPF1 depletion for non-NMD target mRNAs. All fold change values and ratios and plotted on a $\log_2$ scale. P-value was calculated with the Kolmogorov-Smirnov test.

We first developed a pipeline for selecting non-NMD target mRNAs to confirm whether increasing derepression was correlated with increasing 3' UTR, because derepression of mRNA which did not have downstream exon-exon junction were correlated with 3' UTR length (Hurt et al. 2013). For this study, we used two UPF1 knockdown RNA-seq data (HeLa: Wang et al. 2014; mES: Hurt et al. 2013). First, both histone genes and genes annotated as NMD targets in GENCODE (v19) were filtered because UPF1 was associated with mRNA decay pathways such as NMD and replication dependent histone mRNA decay. Second, to select genes which are not potential NMD targets, we got rid of genes that have premature termination codons located > 50nt from the annotated exon junctions. Next, genes whose 3' UTR consisted of multi exons were eliminated. In order to eliminate potential alternative splicing effects on 3' UTR, we also checked the distribution of the distance between splicing signal (GU-AG). The top 1% of distance between splicing signal was under 60nt. Therefore, we removed genes with distance longer than 60nt, and genes which had RNA-seq based junction reads on 3' UTR. For expression quantification, each RNA-seq reads were mapped to the mm9 and hg19 using TopHat (v2.0.6), and gene and isoform expression levels were quantitated using Cufflinks (v2.1.1). Our previous study shows that more than 3 FPM could be cutoff value for significantly differential expression in knock-down experiments. siRNAs have off target effect like

miRNA which suppress target mRNA (Jackson et al. 2010). Therefore, mRNAs which have siRNA seed sites (>=7A1) on 3' UTR is filtered. Finally, we selected 6,650 and 8,112 of genes in hg19 and mm9 RefSeq, respectively (Figure 1A). Additionally, to minimize side effect of minor observation error, selected genes were divided by four bins such as 50nt to 350nt and 800nt to 20000nt (Figure 1B).

We checked that increasing 3' UTR length of non-NMD targets was correlated with increasing derepression (Figure 1C). Furthermore, changes of derepression between in mRNAs with long 3' UTR length (800-20000nt) and with short 3' UTR length (50-350nt) were significantly different in both HeLa and mES RNA-seq.

Together, these data suggest that UPF1 repressed mRNA abundance along 3' UTR length regardless of NMD.

Section 2. UPF1 affects endogenous miRNA targets

Little is known about the role of UPF1 on the 3'UTR for mRNA decay. Although UPF1 can participate miRNA-mediated mRNA decay on 3' UTR (Jin et al. 2009), it had not been investigated genome-wide in mammalian. We predicted that derepression on 3' UTR by knock-downed UPF1 was probably associated with miRNA-mediated mRNA decay. To address this question, we checked expression level of mRNAs which were targeted by highly expressed

endogenous miRNA. To obtain highly expressed miRNAs, published sRNA-seq data from HeLa (Shin et al. 2010) and mES (Chiang et al. 2010) was used. Because small number of miRNAs have been accounted for most of miRNA expression, we classified into top 10, 30, and 50 miRNA families along with expression values of small RNA-seq. The expression level of top 10 endogenouse miRNAs were validated by semiquantitative real-time PCR (RT-sqPCR) analyses in HeLa cells (Figure 2A). As a control, mean fold changes were calculated for 100 cohorts of random generated miRNA (Figure 2B).

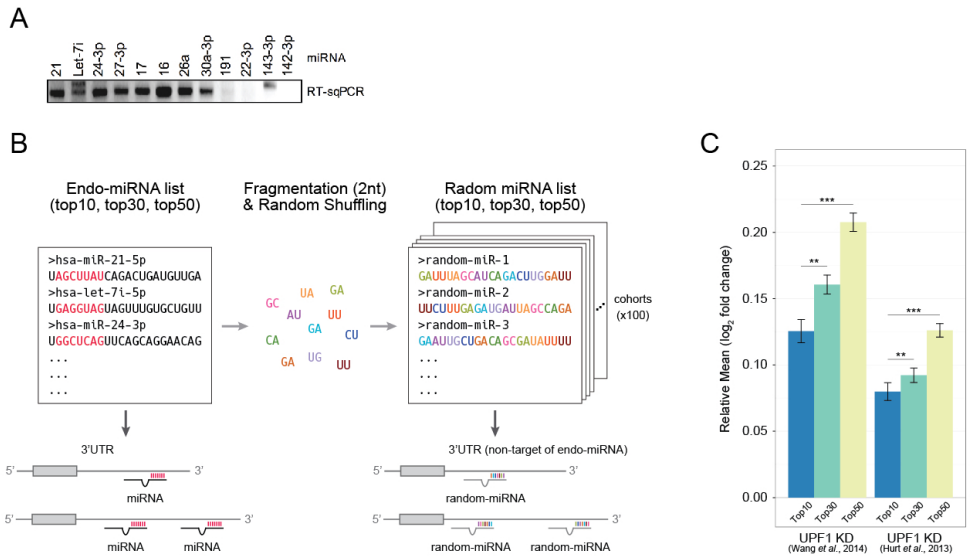

Figure 2. Analysis of highly expressed endo-miRNA targeting. (A) Semiquantative RT-PCR were performed for quantitate top 10 miRNAs. 143-3p and 142-3p were used as negative controls. (B) A schematice flow for generating the randomized miRNA list. A miRNA was selected among the endo-miRNAs for fragmentation. Fragmentated dinucleotides were randomly shuffled. For statistical comparison, 100 cohorts were generated considering endo-miRNA seed sites. (C) Derepression of endo-miRNA targets following UPF1 depletion for non-NMD

target mRNAs. In this relative mean, fold changes (log_2) of random miRNA targets were used as negative control compared with fold changes (log_2) of endo-miRNA targets. Significance was calculated with the Mann-Whitney U test (*p<0.05, **p<0.01, ***p<0.001).

In each cell lines, mRNAs targeted by miRNA top groups were up-regulated (Figure 2C). From top10 to top50, relative mean fold change was gradually increased. Interestingly, when mean fold changes were calculated without normalization by control, top10 group had the highest up-regulated effect on targeted mRNA (not shown). This inverse data is consistent with previous results that increasing number of miRNA reduced down-regulation of mRNAs due to competing binding to AGO proteins.

To confirm association between UPF1 and miRNA targeting, we performed two kinds of RNA-seq. First RNA-seq was UPF1 depletion experiment for validation of previous results (Figure 1C, 2C) in our HeLa cells. Next RNA-seq was co-depleted of UPF1 and DICER, and depleted of DICER as control. DICER is a major component of miRNA biogenesis to process pre-miRNAs into mature miRNA in cytoplasm (Koscianska et al. 2011). In DICER-deficient cell, most canonical miRNAs were prominently repressed (Kim et al. 2016). If up-regulation of non-NMD target mRNAs by UPF1 knockdown were related with miRNA targeting, this effect would be disappeared in DICER knockdown cell lines. Our knockdown experiments of UPF1 and DICER were confirmed by western blot. We could not detect DICER protein in siRNA transfection of DICER and of DICER&UPF1. UPF1 protein was slightly

detected, but UPF1 mRNAs levels were reduced to 11%. On the other hand, UPF1 protein were not affected in DICER knock-downed cells (Figure 3A).

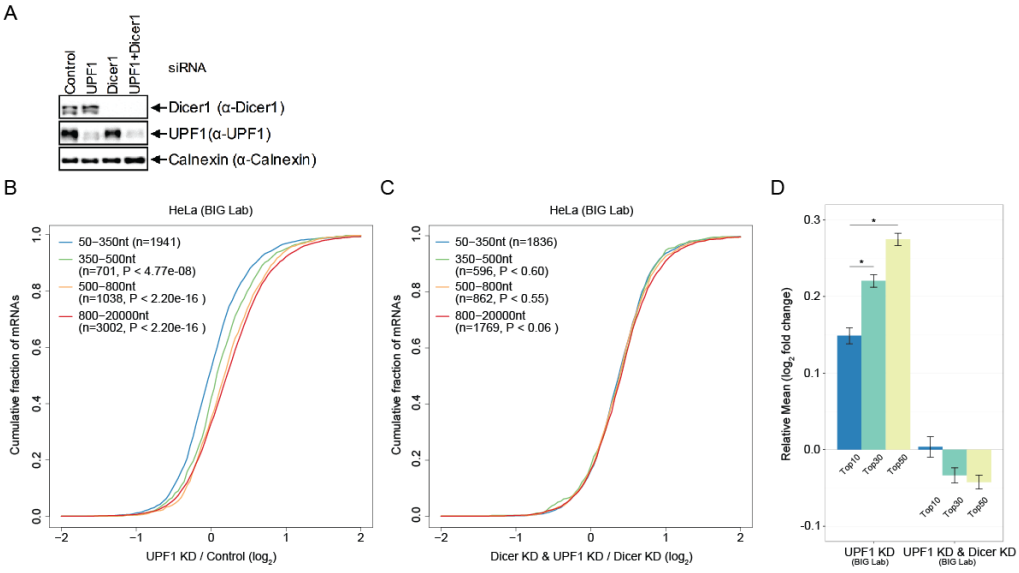

Figure 3. Derepression by UPF1 in the DICER knock-downed HeLa cells. (A) Western blot to demonstrate specific downregulation by siRNAs. Western blotting was performed using the specified antibody (left panel). (B-C) CDFs of changes in mRNA expression following UPF1 depletion and UPF1 & DICER co-depletion for non-NMD target mRNAs. (D) Derepression of endo-miRNA targets following UPF1 depletion and UPF1 & DICER co-depletion. Significance was calculated with the Mann-Whitney U test (*p<0.05, **p<0.01, ***p<0.001).

In order to validate depletion of UPF1 in experiments, we checked FPKM values of UPF1 in each RNA-seq. As expected, UPF1 in all RNA-seq were downregulated to 2.05-, 2.84-, and 2.47-fold (log_2), respectively (Table 1). In RNA-seq for co-depletion of DICER and UPF1, UPF1 was only downregulated to 2.34-fold (log_2) compared to fold change of DICER (Table2).

Table 1. UPF1 expression level in UPF1 KD RNA-seq

	UPF1 (Wang et al. 2014)	UPF1 (Hurt et al. 2013)	UPF1 (BIG Lab.)
UPF1 KD	4.848	4.746	1.194
control	20.129	34.002	6.614
fold change (log_2)	-2.054	-2.841	-2.470

Table 2. UPF1 and DICER expression level in UPF1 KD and DICER KD RNA-seq

	UPF1 (BIG Lab.)	DICER (BIG Lab.)
UPF1 KD & DICER KD	1.020	0.199
DICER KD	5.191	0.248
fold change (log_2)	-2.347	-0.317

To test this hypothesis, we analysed abundance of non-NMD target mRNAs with same process. Consistently, we confirmed that UPF1 affected abundance of non-NMD target mRNAs following 3' UTR length (Figure 3B). Unlike UPF1 only knockdown cell, up-regulation by increasing 3' UTR length was disappeared in DICER knockdown cell (Figure 3C). Also, we reanalysed highly expressed miRNA targeting effect on each experiment (Figure 3D). Similar results were obtained when we knock-donwed UPF1 in HeLa. Strikingly, relative mean fold change (log_2) of miRNA targeted genes was significantly decreased on top10, top30 and top50 in co-depleted of UPF1 and DICER cells compared with depleted of UPF1.

To confirm miRNA targeting efficacy, we examined seed effect from 8mer to 6mer (Bartel 2009). In all UPF1 knock-downed HeLa cells, genes with 8mer sites is significantly derepressed compared with genes with random sites as control (Figure 4). Furthermore, we sought to similar efficacy of canonical miRNA seed types with published results (Bartel 2009). In

contrast, we did not observe miRNA seed efficacy in UPF1 and DICER co-depleted HeLa cells.

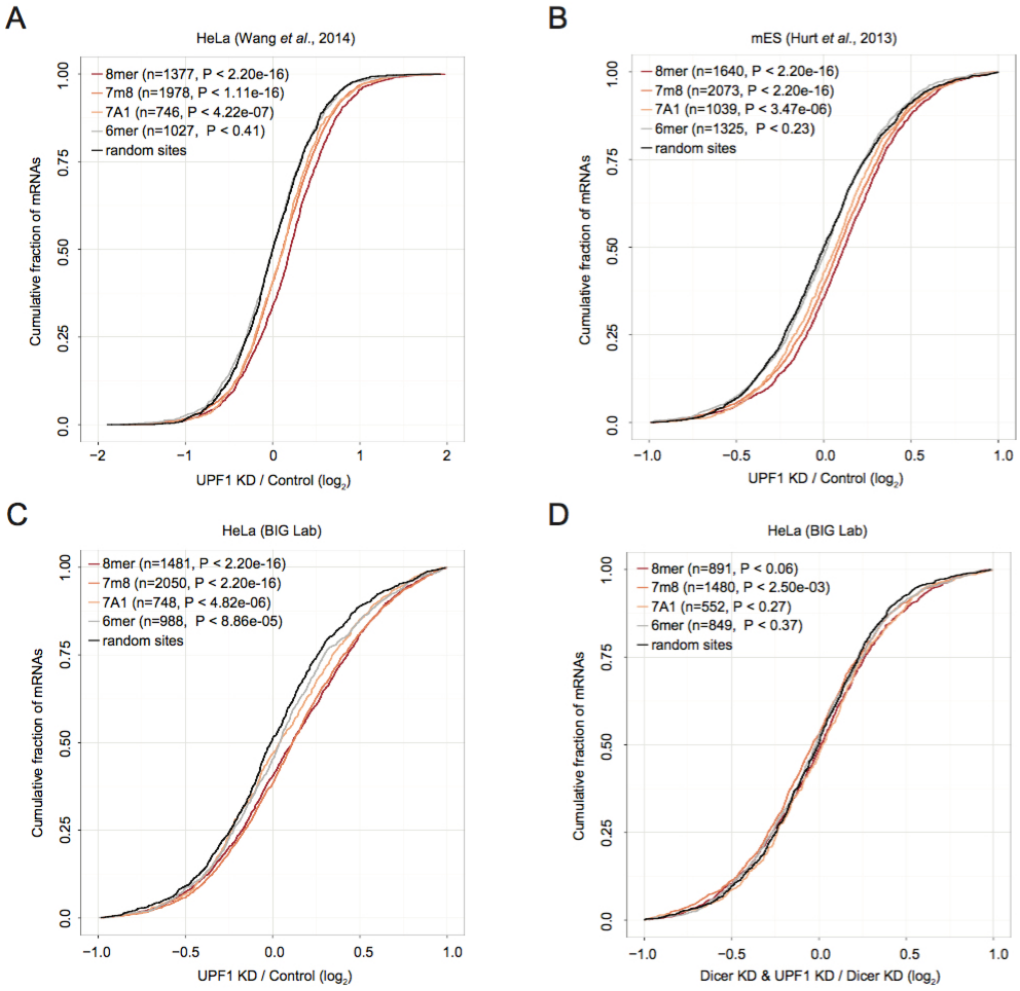

Figure 4. Efficacy of canonical seed sites of non-NMD target mRNAs. (A-C) CDFs of changes in mRNA expression following UPF1 depletion. (D) in miRNA expression following UPF1 and DICER depletion based on DICER depletion. All fold change values and ratios and plotted on a $\log_2$ scale. P-value was determined using the Kolmogorov-Smirnov test.

Additionally, to reveal the effect of 3' UTR length on UPF1 knock-downed cells, we made a hypothesis that correlation between 3' UTR length and

endo-miRNA sites could affect UPF1-mediated miRNA decay. With top 10 highly expressed endo-miRNAs in HeLa and mES cells, we confirmed the correlation between 3' UTR length and the number of miRNA in selected mRNAs (Figure 5A). This result was consistent in previous report (Kim et al. 2014).

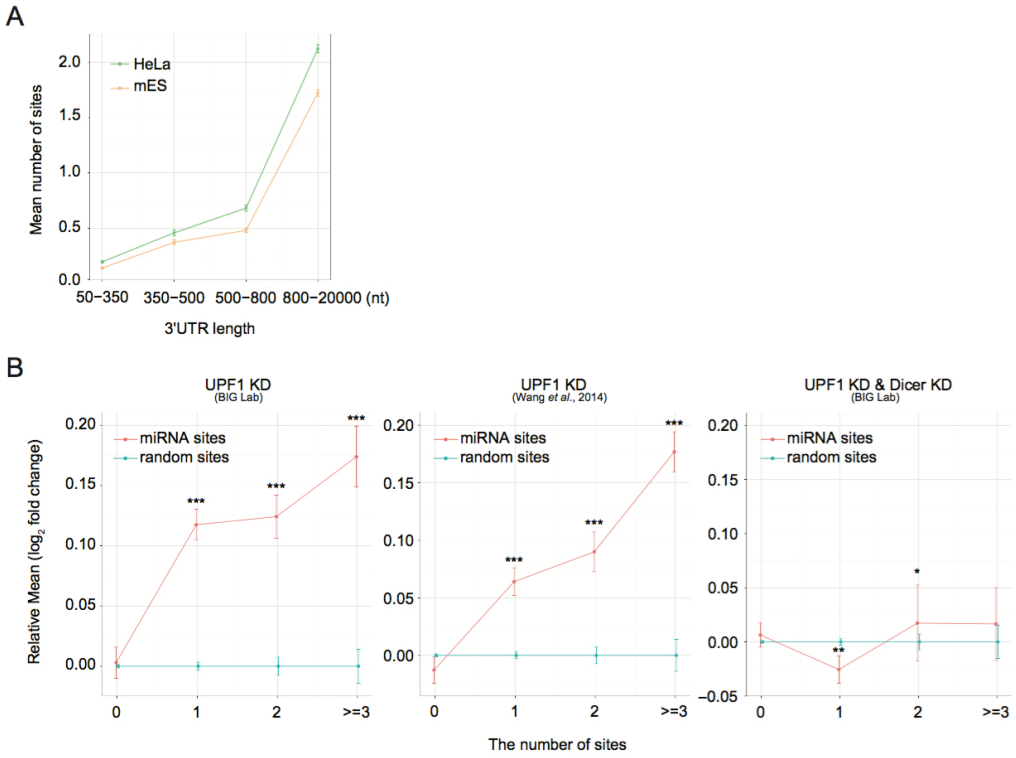

Figure 5. Longer 3' UTR contains more miRNA target sites and has stronger influence on miRNA target effects. (A) The average number of top 10 endo-miRNA sites on 3' UTR length bins in HeLa and mES cells. (B) The correlation between number of miRNA sites and derepression. Mean fold change (log_2) of endo-miRNA were normalized by mean values of random sites. P-value was determined using the Kolmogorov-Smirnov test (*p<0.05, **p<0.01, ***p<0.001).

To extend our result to the number of endo-miRNA sites, we counted endo-miRNA sites on 3' UTR of selected mRNAs. We found that the more increasing endo-miRNA sites, the more derepression effect on targeted mRNAs, also the differences between endo-miRNA sites and random sites as negative control were significant (Figure 5B). Meanwhile, gradually increasing of derepression was disappeared in DICER depleted RNA-seq.

Overall, these results indicate that UPF1 can affects on miRNA targeting in selected non-NMD target genes, and UPF1 repressed mRNA abundance along the abundance of miRNA.

Section 3. UPF1 is co-localized with AGO2 proteins and associated with miRNA targeting

For Watson-Crick pairing of miRNA on 3' UTR of miRNA, miRNAs were bound by AGO proteins (Bartel 2009). Also, previous studies reported that AGOs can bind with UPF1 through protein-protein interaction (Jin et al. 2009; Höck et al. 2007). Therefore, we thought that UPF1 can affect on miRNA targeting with AGO2. To examine this hypothesis, we identified genome-wide binding sites of UPF1 and AGO2 using available CLIP-seq data.

We obtained UPF1 and AGO2 CLIP-seq data in two studies of mES cells (Hurt et al. 2013; Leung et al. 2011) and in two studies of HeLa cells (Zünd et al. 2014; Kishore et al. 2013). To identify binding sites, we

used processed mES CLIP-seq and raw HeLa CLIP-seq, and wild type replicates on each data were merged. HeLa CLIP-seq is processed by HOMER (Heinz et al. 2010).

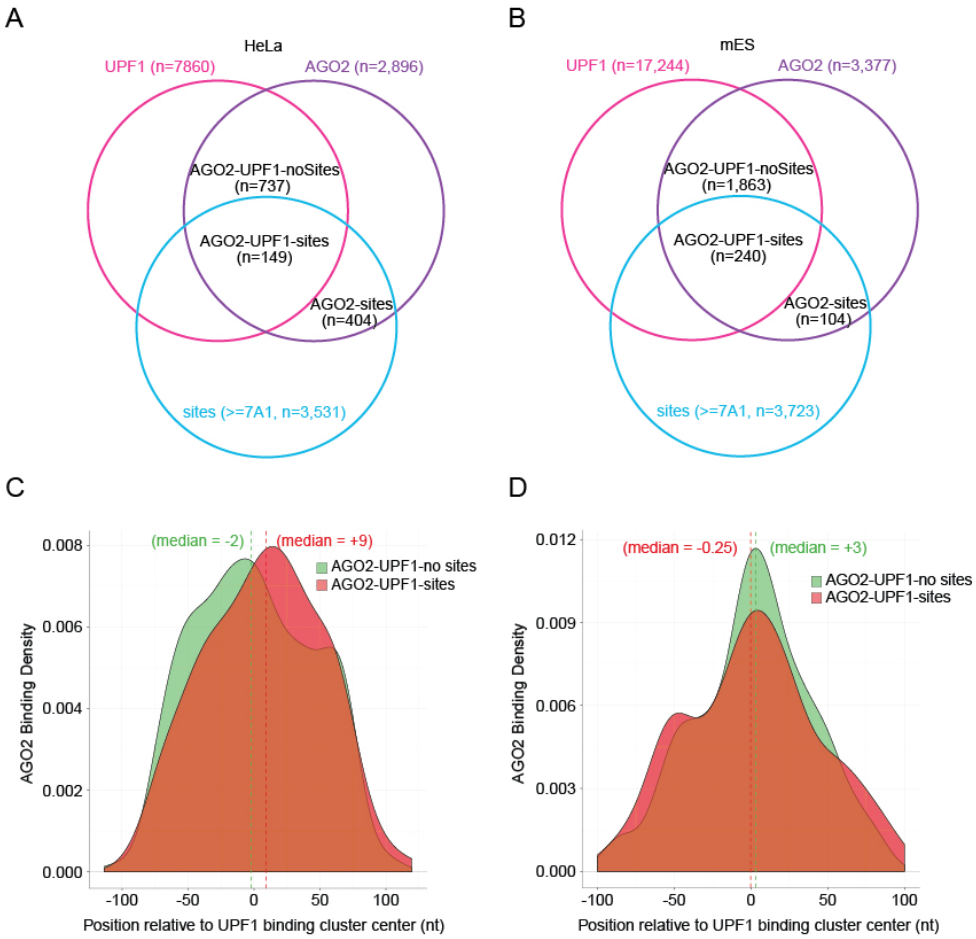

Figure 6. UPF1 proteins are co-localized with AGO2 proteins. (A-B) Venn diagram showing the overlap between UPF1, AGO2, endo-miRNA sites (top10). (C-D) Density plots for comparison of UPF1 and AGO2 binding sites. Density of AGO2 binding peaks at UPF1 binding clusters in 3' UTRs.

We identified 31% and 62% of AGO2 binding sites co-localized with UPF1 on 3' UTR in HeLa cells and mES cells, respectively. Among AGO2-UPF1 binding sites, 13% and 20% sites included miRNA 8mer, 7mer-m8, or 7mer-A1 sites in HeLa cells and mES cells respectively. (Figure 6A, 6B) Besides, overlapped binding sites of AGO2 and UPF1 revealed significantly overlaps (Figure 6C and 6D).

However, identified AGO2 and UPF1 binding sites may include false positive sites because majority of endo-miRNA sites were not included in AGO2 binding sites. Although the overlaps could not indicate direct functional interaction, it could suggest probability of relationship between UPF1 and AGO2.

In order to determine whether overlaps of AGO2 and UPF1 affect on miRNA targeting, we categorized into AGO2-UPF1, AGO2, and no CLIP-seq sites on 3' UTR. Next, each groups are divided by presence of miRNA sites. Groups which included miRNA on AGO2 binding sites are normalized by others which did not include miRNA on 3' UTR but had randomized miRNA sites.

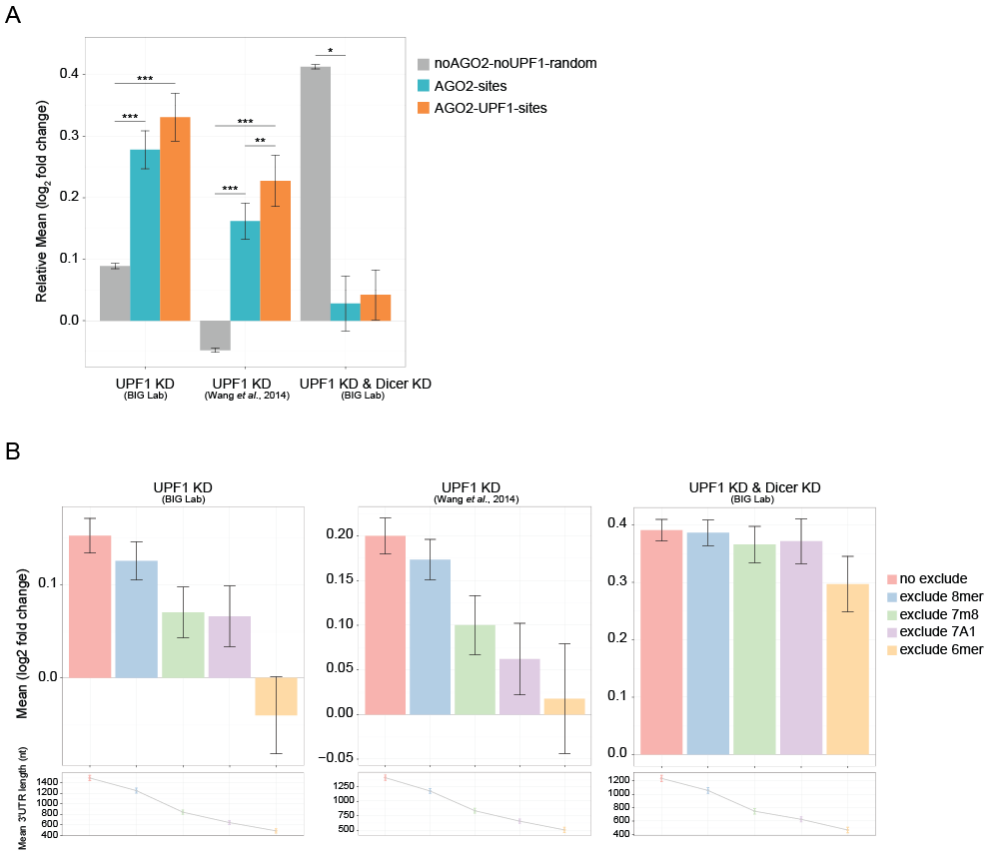

Figure 7. The colocalization of AGO2 and UPF1 affects on miRNA targeting. (A) Derepression analysis of AGO2-UPF1-sites, AGO2-sites, and negative control. Relative mean was calculated by endo-miRNA sites and random sites in AGO2 sites. (B) Analysis of efficacy of miRNA seed sites by the loss of miRNA seed from 8mer to 6mer. Below lines were generated by calculating average of each group 3' UTR length.

Two AGO2-sites and AGO2-UPF1-sites groups were significantly derepressed compared with negative control (Figure 7A). Interestingly, mRNAs with AGO2-UPF1-sites were more derepressed than with AGO2-sites on all UPF1 knock-downed RNA-seq data. On the other hand, DICER and UPF1 co-depleted RNA-seq data showed reversed result. In addition, to validate

miRNA targeting on AGO2-UPF1 bound mRNA, we checked efficacy of seed-matched miRNA sites (Bartel 2009). In mRNAs with AGO2-UPF1 which didn't include miRNA sites, we excluded miRNA sites from 8mer to 6mer. In all UPF1 knock-downed RNA-seq, derepression was reduced following excluding miRNA seed sites (Figure 7B). As expected, co-depleted data did not decrease despite excluding all miRNA sites. However, this decreasing results could be caused by shortening of 3'UTR length (Figure 7B). Nevertheless, our results could suggest that co-localized of AGO2-UPF1 with miRNA seed sites were related with miRNA targeting and had synergistic effect compared with canonical AGO2 with miRNA.

Chapter 3. Alternative miRNA-mediated mRNA decay pathway

Section 1. UPF1 helicase activity and SMG5 does not affect Non-NMD target mRNAs

In order to identify the detailed mechanism of the UPF1-mediate miRNA decay, we focused helicase activity of UPF1 and NMD factors such as SMG5 and SMG7. First, UPF1 is a well known RNA helicase which is able to efficiently translocate through double-stranded structure (Fiorini et al. 2015). Therefore, we assumed that the helicase activity of UPF1 could improve accessibility of AGO2 for miRNA targeting. Second, we found that CCR4-NOT deadenylation complex was common factor for degradation in both miRNA-mediated decay and NMD machinery (Jonas and Izaurralde 2015; Loh et al, 2013). Moreover, SMG5-SMG7 heterodimer which was interacted with UPF1 on NMD could recruit CCR4-NOT (Loh et al, 2013). These previous results implied that AGO2-UPF1-miRNA complex could be interacted with CCR4-NOT through SMG5-SMG7 heterodimer.

In order to assess helicase activity of UPF1, we performed RNA-seq analysis of HeLa cells depleted of UPF1, and transfected helicase mutant

UPF1 or wild type UPF1. We validated overexpression of UPF1 in both RNA-seq 616.68 and 381.51, respectively (Table 3).

Table 3. UPF1 expression level in helicase mutant or wildtype RNA-seq

	UPF1 (BIG Lab.)
Helicase MT overexpression (UPF1 KD)	616.683
Helicase WT overexpression (UPF1 KD)	381.508

To explore second hypothesis, we took advantage of published microarray data sets scrutinizing NMD factors such as SMG5 and SMG7 (Cho et al. 2013). The microarray data was depleted SMG5 or SMG7 in HeLa cells. Using same procedure (Figure 1A), we analyzed derepression effect by increasing 3' UTR length on each data. Interestingly, we showed similar increasing derepression effect by increasing 3' UTR length in SMG7 knock-downed microarray data (Figure 8C). Particularly, long 3' UTR (800-20000nt) was significantly derepressed compared with short 3' UTR (50-350nt) ($P < 3.90e-08$). However, we did not see difference of derepression between shorter 3' UTR length and longer 3' UTR length in helicase mutant RNA-seq and SMG5 knock-downed microarray data (Figure 8A, 8B).

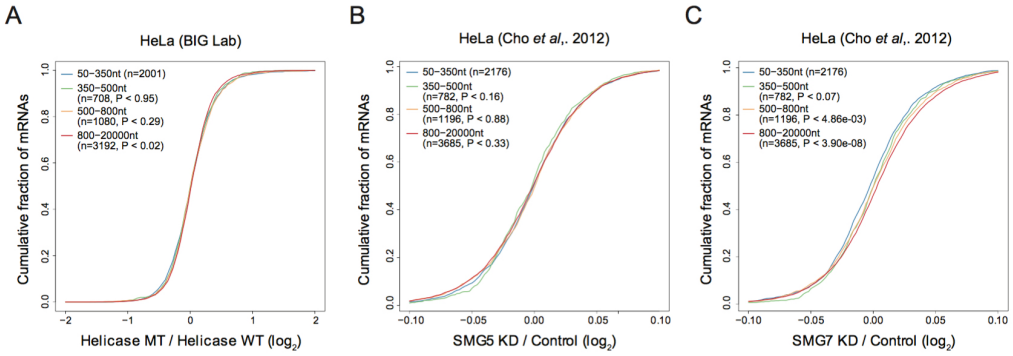

Figure 8. Consistent derepression analysis on three kinds of samples. (A-C) CDFs of changes in mRNA expression following (A) Helicase mutant, (B) SMG5 depletion, or (C) SMG7 depletion for non-NMD target mRNAs. All fold change values and ratios and plotted on a $\log_2$ scale. P-value was calculated with the Kolmogorov-Smirnov test.

Section 2. SMG7 is associated with UPF1-mediated miRNA decay

To validate association of SMG7, we also performed RNA-seq analysis of HeLa cells depleted of SMG7. SMG7 mRNA levels in SMG7 depeleted HeLa cells were reduced to 30% of these in control cells (Figure 9A). In this RNA-seq, SMG7 were downregulated to 1.03-fold ($\log_2$) (Table 4).

Table 4. SMG7 expression level in SMG7 KD RNA-seq

	SMG7 (BIG Lab.)
SMG7 KD	16.991
control	34.724
fold change ($\log_2$)	-1.031

Through repetitive procedure (Figure 1A), we got a consistent result that indicated correlation between derepression and 3' UTR length (Figure 9B).

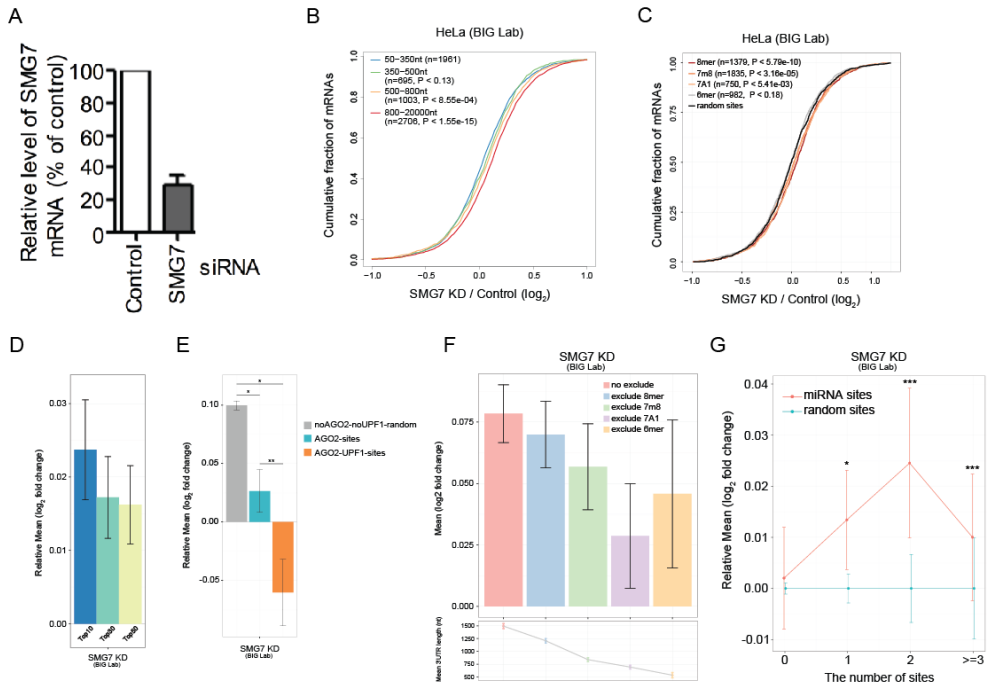

Figure 9. SMG7 is associated with miRNA decay. (A) Semiquantative RT-PCR were performed for quantitate SMG7 to demonstrate specific downregulation by siRNA. (B-G) All results were reproduced in SMG7 depletion. (B) This panel is as in Figure 1C. (C) This panel is as in Figure 4A-4C. (D) This panel is as in Figure 2C. (E) This panel is as in Figure 7A. (F) This panel is as in Figure 7B. (G) This panel is as in Figure 5B.

To ask whether SMG7 influence on miRNA targeting, we performed same analysis for miRNA targeting using highly expressed endo-miRNA set in HeLa. Interestingly, we found that the rule of miRNA seed efficacy worked when SMG7 was knock-downed (Figure 9C). In particular, genes with 8mer sites is significantly derepressed compared with genes with random sites as control ($P < 5.79e-10$). Also, miRNA targeted genes were up-regulated following knock-downed SMG7 (Figure 9D).

Next, we investigated whether co-localized AGO2 and UPF1 could affect on HeLa cells depleted SMG7. Unlike previous results (Figure 3C), genes with AGO2-UPF1-sites were repressed, but negative control sets were rather derepressed (Figure 9E). When we gradually excluded miRNA seed in genes which had AGO2-UPF1, mean fold change (log_2) was decreased before excluding 6mer (Figure 9F). Lastly, we checked derepression effect by number of endo-miRNA sites. In this analysis, gene sets which had more than one miRNA sites were significantly derepressed compared with random sites (Figure 9G). However, genes with more than three miRNA sites were not more derepressed than more than genes with two miRNA sites group. The inconsistent results on SMG7 knock-downed cells might be derived from combination of different CLIP-seq or batch effects.

Section 3. Design of experimental validation

To validate of these UPF1 mediated miRNA decay related to SMG7, we selected up-regulated mRNAs which had AGO2-UPF1-sites complex for miRNA inhibitor experiments. First, we chose mRNAs which had more than 0.2-fold (log_2) in both cells depleted SMG7 and depleted UPF1, and whose AGO2-UPF1-sites had candidate miRNAs (miR-24-3p, miR-26a-5p, miR-17-5p or miR-16-5p) (Table 5).

Table 5. Experiment candidate genes (SMG7KD and UPF1KD)

gene name	Mean_log2_fold_change (SMG7KD)	Mean_log2_fold_change (UPF1KD)	3'UTR_length	miRNA
LASP1	0.367	0.532	2989	hsa-miR-24-3p(7m8)
TEX261	0.307	0.849	2656	hsa-miR-24-3p(7m8)
GREB1L	0.717	0.567	2595	hsa-miR-26a-5p(8mer)
ADD1	0.304	0.659	1887	hsa-miR-17-5p(7m8)
TXLNG	0.214	0.255	1453	hsa-miR-26a-5p(8mer)
ATG16L1	0.310	0.247	1314	hsa-miR-17-5p(8mer)
HECTD1	0.413	0.477	816	hsa-miR-16-5p(7A1)
PRPF19	0.298	0.413	616	hsa-miR-24-3p(8mer), hsa-miR-27a-3p(8mer)
REEP4	0.492	0.772	473	hsa-miR-26a-5p(7m8)
RPS19BP1	0.308	0.634	395	hsa-miR-24-3p(7m8)

Second, we selected mRNAs which had AGO2-UPF1 with miR-26a-5p, and seperated into two mutually exclusive groups as positive control. First group were more than 0.2-fold (log_2) in depleted SMG7 cell, but under 0.2-fold (log_2) in depleted UPF1 cells and second group were more than 0.2-fold (log_2) in depleted UPF1 cell, but under 0.2-fold (log_2) in depleted SMG7 cells (Table 6). Lastly, we selected mRNAs which did not have any miRNA sites in 3' UTR and which had almost 1-fold (Table 6).

Table 6. Experiment candidate genes (Control)

geneName	Mean_log2_fold_change (SMG7KD)	Mean_log2_fold_change (UPF1KD)	3'UTR_length	miRNA
FAM98A	0.343	-0.160	1729	hsa-miR-26a-5p(8mer)
HTATIP2	0.205	-0.132	575	hsa-miR-26a-5p(7A1)
EPAS1	-0.128	1.230	2034	hsa-miR-26a-5p(7A1), hsa-miR-27a-3p(7A1)
HMGA1	0.016	0.478	1344	hsa-miR-26a-5p(7A1), hsa-miR-16-5p(7m8)
SERPINB1	-0.023	-0.022	1275	X
NOC3L	-0.068	0.030	948	X
RPA1	0.048	0.029	892	X

Discussion

Prior works have been focused on the role of UPF1 as core molecule in NMD pathway (Lykke-Andersen and Jensen. 2015). Moreover, most studies have considered that expression change of genes by knock-downed UPF1 was caused by inhibition of NMD (Hurt et al. 2013; Tani et al. 2012). However, these studies have not identified casual relationship between miRNA targeting and UPF1, though UPF1 was predominantly bind on 3' UTR (Kurosaki and Maquat. 2013; Zünd et al. 2013; Hurt et al. 2013).

In this study, we analyzed several RNA-seq and CLIP-seq in HeLa and mES cells to confirm genome-wide correlation between derepression effect and miRNA targeting with AGO2 and UPF1. We found that 3' UTR bound UPF1 could have a role in alternative miRNA targeting with AGO2. In particular, SMG7 might be related in the alternative miRNA pathway as molecule linker. These findings extend UPF1 related regulation system from NMD to miRNA targeting. In addition, UPF1 could have a synergistic effect on miRNA targeting compared with canonical miRNA targeting which is controlled by AGO2.

However, a limitation about CLIP-seq is worth noting. Although we identitied the relationship of AGO2 and UPF1 using CLIP-seq data, it is

hard to assure that two proteins have physical interaction. Furthermore, the detailed mechanism which is related with SMG7 is unclear.

Thus, our further studies should include experiments designed to reveal protein-protein interaction between AGO2, UPF1 and SMG7. Also, functional causality of the complex in miRNA-mediated decay should be established. Through further experiments, we expected that results enable us to describe alternative miRNA targeting which is associated with UPF1 and SMG7. Unlike canonical miRNA targeting, alternative miRNA targeting is interacted by not GW182 but UPF1 which play a role as a bridge between deadenylase complexes. Similar to NMD machinery, UPF1 which is bound to AGO2-miRNA is interacted with SMG7, then this complex recruit CCR4-NOT for deadenylation. Additionally, we expect that UPF1-mediated miRNA targeting has a synergistic effect than canonical miRNA targeting (Figure 10).

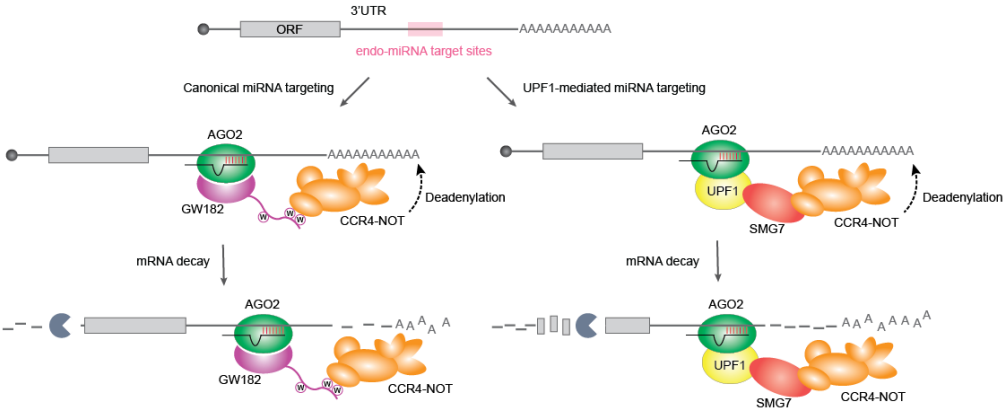

Figure 10. Model illustrating the alternative miRNA targeting. Unlike canonical miRNA targeting, AGO2 interacts with UPF1 on miRNA targeting. AGO2-UPF1 complex recruits SMG7 protein in order to interact with CCR4-NOT. Finally, deadenylation and decapping are triggered by the complex.

Detailed Methods

Data source

For reference genome, NCBI RefSeq gene annotation (hg19: Aug-22-2011, mm9: Feb-15-2014) were used. RNA-seq data for UPF1 knockdown and wild-type in HeLa cells were from GSE63091 (Wang et al. 2014) and in mES cells were from GSE41785 (Hurt et al. 2013). CLIP-seq data for wild-type UPF1 in HeLa cells were from GSE47976 (Zünd et al. 2013) and in mES cells were from GSE41785 (Hurt et al. 2013). CLIP-seq data for wild-type AGO2 in HeLa cells were from GSE43666 (Kishore et al. 2013) and in mES cells were from GSE25310 (Leung et al. 2011). 3P-seq data for HeLa cells and mES cells were from GSE52531 (Nam et al. 2014).

Processing for gene expression analysis

In annotated genomes for human (hg19) and mouse (mm9), 3' UTR were updated using major 3P-seq signal in HeLa and mES cells, repectively (Nam et al. 2014). All RNA-seq reads were mapped to the genome using Tophat version 2.0.6 (Trapnell et al. 2009), allowing at most five genomic matches, 2 mismatches, and novel introns but disallowing splice sites mismatches (--solexa1.3-quals --splice-mismatches 0 --min-intron-length 61 --max-intron-length 265006 -g 5). Expression levels (FPKMs, fragments per

kilobase per million mapped) were quantitated using Cufflinks version 2.1.1 (Trapnell et al. 2010). For accurate expression of mRNAs, we selected transcripts with the longest 3' UTR length among isoforms and applied gene expression levels to seleted mRNAs.

Analysis of AGO2 and UPF1 binding

AGO2 and UPF1 binding sites of mES cells were used processed data. On the other hand, CLIP-seq data in HeLa cells were used raw data. The raw CLIP-seq data were mapped to the human reference genome (hg19) using Bowtie version 4.4.7 (Trapnell et al. 2009) allowing 2 mismatches but disallowing multi loci mapping (-v 2 -m 1 --best --strata). For identifying binding sites of AGO2 and UPF1 bound 3' UTR of mRNA, we used findPeaks program in HOMER (Heinz et al. 2010). All wild-type replicates on each CLIP-seq data were collapsed using Bedtools version 2.17.0 (Quinlan and Hall. 2010).

Cell culture and stable knockdown of UPF1, DICER, and SMG7

HeLa cells were grown in Dulbecco's modified Eagle's medium (DMEM) containing 10% fetal bovine serum and 1% penicillin / streptomycin. Cells were transfected with 75 nM *in vitro*-synthesized small interfering RNA (siRNA) using Lipofectamine 3000 (Invitrogen). Two day later, the cells were harvested.

Western blotting

Proteins were separated by 7-15% sodium dodecyl sulfate-polyarcylamide gel electrophoresis (SDS-PAGE). The following antibodies were used: UPF1 (a gift from Lynne E. Maquat), DICER (Cell Signaling Technology), Calnexin (Santa Cruz Biotechnology)

Semiquantitative RT-PCR

Total RNA was extracted using Trizol (Invitrogen) according to manufacturer's protocol. Then RNA was further treated with RQ DNaseI (Promega) for 30 min at 37 °C to eliminate possible exogenous or endogenous DNA contamination. The RNA was reverse-transcribed using random hexamers (Macrogen) or oligo dT adaptor (System Bioscience) and reverse-transcribed cDNA was amplified by quantitative or semiquantitative RT-PCR (RT-sqPCR).

For RT-PCR, miRNAs, SMG7 and GAPDH mRNAs were amplified as previously described (Hwang et al. 2010). Also, primers of GAPDH, SMG7 and miRNAs were used (Table 7). RT-sqPCR results were analyzed in 1% agarose gel using ethidium bromide staining.

Table 7. Oligonucleotides used in semiquantitative RT-PCR

Genes	Primer
GAPDH	Forward: CTG TGG TCA TGA GTC CTT CC Reverse: CAA GAT CAT CAG CAA TGC C
SMG7	Forward: GGC AGG CAG AAG TCC TGA AG Reverse: AGG CGT GAT TCC AGA GAT CC
hsa-miR-21-5p	Forward: CGC TAG CTT ATC AGA CTG ATG TTG A
hsa-let-7i-5p	Forward: CGC TGA GGT AGT AGT TTG TGC TGT T
hsa-miR-24-3p	Forward: TGGCTCAGTTCAGCAGGAACAG
hsa-miR-27a-3p	Forward: CGC TTC ACA GTG GCT AAG TTC CGC
hsa-miR-17-5p	Forward: CGC CAA AGT GCT TAC AGT GCA GGT AG
hsa-miR-16-5p	Forward: CGC TAG CAG CAC GTA AAT ATT GGC G
hsa-miR-26a-5p	Forward: CGC TTC AAG TAA TCC AGG ATA GGC
hsa-miR-30a-3p	Forward: CGC CTT TCA GTC GGA TGT TTG CAG C
hsa-miR-191-5p	Forward: CGC CAA CGG AAT CCC AAA AGC AGC TG
hsa-miR-22-3p	Forward: CGC AAG CTG CCA GTT GAA GAA CTG
hsa-miR-143-3p	Forward: CGC TGA GAT GAA GCA CTG TAG CTC
hsa-miR-142-3p	Forward: CGC TGT AGT GTT TCC TAC TTT ATG GA

Reference

Ahn, S., Kim, J. & Hwang, J. CK2-mediated TEL2 phosphorylation augments nonsense-mediated mRNA decay (NMD) by increase of SMG1 stability., 1829(10), 1047-1055.

Amrani, N. et al. A faux 3′-UTR promotes aberrant termination and triggers nonsense-mediated mRNA decay. Nature 432, 112-118 (2004).

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

Czaplinski, K. et al. The surveillance complex interacts with the translation release factors to enhance termination and degrade aberrant mRNAs. Genes Dev. 12, 1665-1677 (1998).

Eberle, A. B., Lykke-Andersen, S., Muhlemann, O. & Jensen, T. H. SMG6 promotes endonucleolytic cleavage of nonsense mRNA in human cells. Nat. Struct. Mol. Biol. 16, 49-55 (2009).

Fabian, M. R. & Sonenberg, N. The mechanics of miRNA-mediated gene silencing: a look under the hood of miRISC. Nat. Struct. Mol. Biol. 19, 586-593 (2012).

Franks, T. M., Singh, G. & Lykke-Andersen, J. Upf1 ATPase-dependent mRNP disassembly is required for completion of nonsense-mediated mRNA

decay. Cell 143, 938-950 (2010).

Gatfield, D. & Izaurralde, E. Nonsense-mediated messenger RNA decay is initiated by endonucleolytic cleavage in Drosophila. Nature 429, 575-578 (2004).

Ghildiyal, M. & Zamore, P. D. Small silencing RNAs: an expanding universe. Nature Rev. Genet. 10, 94-108 (2009).

Glavan, F., Behm-Ansmant, I., Izaurralde, E. & Conti, E. Structures of the PIN domains of SMG6 and SMG5 reveal a nuclease within the mRNA surveillance complex. EMBO J. 25, 5117-5125 (2006).

Gregersen, L. H. et al. MOV10 is a 5′ to 3′ RNA helicase contributing to UPF1 mRNA target degradation by translocation along 3′ UTRs. Mol. Cell 54, 573-585 (2014).

Hogg, J. R. & Goff, S. P. Upf1 senses 3′ UTR length to potentiate mRNA decay. Cell 143, 379-389 (2010).

Hug, N. & Caceres, J. F. The RNA helicase DHX34 activates NMD by promoting a transition from the surveillance to the decay-inducing complex. Cell Rep. 8, 1845-1856 (2014).

Huntzinger, E. & Izaurralde, E. Gene silencing by microRNAs: contributions of translational repression and mRNA decay. Nat. Rev. Genet. 12, 99-110 (2011).

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

Huntzinger, E., Kashima, I., Fauser, M., Sauliere, J. & Izaurralde, E. SMG6 is the catalytic endonuclease that cleaves mRNAs containing nonsense codons in metazoan. RNA 14, 2609-2617 (2008).

Hurt, J. A., Robertson, A. D. & Burge, C. B. Global analyses of UPF1 binding and function reveal expanded scope of nonsense-mediated mRNA decay. Genome Res. 23, 1636-1650 (2013).

J. Hwang, H. Sato, Y. Tang, D. Matsuda, L.E. Maquat, UPF1 association with the cap-binding protein, CBP80, promotes nonsense-mediated mRNA decay at two distinct steps, Mol. Cell 39 (2010) 396-409.

Kashima, I. et al. Binding of a novel SMG-1-Upf1- eRF1-eRF3 complex (SURF) to the exon junction complex triggers Upf1 phosphorylation and nonsense- mediated mRNA decay. Genes Dev. 20, 355-367 (2006).

Kashima, I. et al. Binding of a novel SMG-1-Upf1- eRF1-eRF3 complex (SURF) to the exon junction complex triggers Upf1 phosphorylation and nonsense- mediated mRNA decay. Genes Dev. 20, 355-367 (2006).

Kervestin, S. & Jacobson, A. NMD: a multifaceted response to premature translational termination. Nat. Rev. Mol. Cell Biol. 13, 700-712

(2012).

Kim, Y.-K., Kim, B., & Kim, V. N. (2016). Re-evaluation of the roles of DROSHA, Exportin 5, and DICER in microRNA biogenesis. Proceedings of the National Academy of Sciences of the United States of America, 113(13)

Kurosaki, T. & Maquat, L. E. Rules that govern UPF1 binding to mRNA 3′ UTRs. Proc. Natl Acad. Sci. USA 110, 3357-3362 (2013).

Kurosaki, T. et al. A post-translational regulatory switch on UPF1 controls targeted mRNA degradation. Genes Dev. 28, 1900-1916 (2014).

Kurosaki, T. et al. A post-translational regulatory switch on UPF1 controls targeted mRNA degradation. Genes Dev. 28, 1900-1916 (2014).

Loh, B., Jonas, S. & Izaurralde, E. The SMG5-SMG7 heterodimer directly recruits the CCR4-NOT deadenylase complex to mRNAs containing nonsense codons via interaction with POP2. Genes Dev. 27, 2125-2138 (2013).

Lykke-Andersen, S. et al. Human nonsense-mediated RNA decay initiates widely by endonucleolysis and targets snoRNA host genes. Genes Dev. 28, 2498-2517 (2014).

Popp, M. W. & Maquat, L. E. Organizing principles of mammalian nonsense-

mediated mRNA decay. Annu. Rev. Genet. 47, 139-165 (2013).

Schmidt, S. A. et al. Identification of SMG6 cleavage sites and a preferred RNA cleavage motif by global analysis of endogenous NMD targets in human cells. Nucleic Acids Res. 43, 309-323 (2014).

Schweingruber, C., Rufener, S. C., Zund, D., Yamashita, A. & Muhlemann, O. Nonsense-mediated mRNA decay - mechanisms of substrate mRNA recognition and degradation in mammalian cells. Biochim. Biophys. Acta 1829, 612-623 (2013).

Smith, J. E. & Baker, K. E. Nonsense-mediated RNA decay - a switch and dial for regulating gene expression. Bioessays 37, 612-623 (2015).

Unterholzner, L. & Izaurralde, E. SMG7 acts as a molecular link between mRNA surveillance and mRNA decay. Mol. Cell 16, 587-596 (2004).

Wang, W., Czaplinski, K., Rao, Y. & Peltz, S. W. The role of Upf proteins in modulating the translation read-through of nonsense-containing transcripts. EMBO J. 20, 880-890 (2001).

Weng, Y., Czaplinski, K. & Peltz, S. W. Genetic and biochemical characterization of mutations in the ATPase and helicase regions of the Upf1 protein. Mol. Cell. Biol. 16, 5477-5490 (1996).

Yamashita, A. et al. Concerted action of poly(A) nucleases and decapping enzyme in mammalian mRNA turnover. Nat. Struct. Mol. Biol. 12, 1054-

1063 (2005).

Zund, D., Gruber, A. R., Zavolan, M. & Muhlemann, O. Translation-dependent displacement of UPF1 from coding sequences causes its enrichment in 3′ UTRs. Nat. Struct. Mol. Biol. 20, 936-943 (2013).

국문요지

UPF1은 비정상적인 위치에 종결코돈이 존재하는 mRNA를 제거하는 기전(nonsense-mediated mRNA decay, NMD)에서 중요한 역할을 하고 있는 단백질로 알려져 있다. 하지만 UPF1은 코딩영역이 아닌 3’UTR에 대부분이 존재하고 3’UTR 길이가 길어짐에 따라 mRNA의 발현량을 조절한다. 또한 UPF1은 miRNA를 통한 mRNA 조절 기작에서 핵심적인 역할을 하는 AGO 단백질과의 상호작용이 보고되어 있다. 하지만 UPF1이 3’UTR에 붙게 되어 miRNA를 통한 mRNA 조절 기작에 어떠한 역할을 하는지는 알려진 것이 거의 없다. 본 연구에서는 UPF1 유전자의 발현을 억제한 넉다운 실험을 통해 만든 공개된 RNA-seq을 재분석하여 NMD 기전과는 독립적으로 유전자를 조절하는 기전을 확인하였다. UPF1이 NMD 타겟이 되지 않는 유전자를 3’UTR 길이에 따라 발현량을 조절한다는 것을 발견하였다. 또한 CLIP-seq 데이터를 이용하여 UPF1과 AGO2의 3’UTR에서의 위치를 조사하였을때, 대다수의 UPF1이 AGO2와 동일한 위치에 결합자리를 갖고 있는 것을 확인하였다. 그리고 UPF1을 넉다운한 생쥐 배아줄기 (mES) 세포에서, miRNA를 가지고 있는 AGO2와 UPF1의 결합자리가 겹치는 유전자는 miRNA가 타겟이 되지 않는 유전자들에 비해 발현량이 증가되어 있는 것을 확인하였고, 이는 UPF1이 miRNA 조절기작에서 시너지 효과가 있는 것을 의미할 수 있다. 또한 3’UTR의 길이 변화에 따른 UPF1의 mRNA 발현 조절은 miRNA의 타겟이 되지 않는 유전자들에 대해서는 거의 효과가 없었다. 추가적으로, SMG7 단백질이 UPF1에 의한 mRNA 발현 조절

기작에 연관있다는 것을 제시하게 되었다. SMG7을 넉다운한 헬라 (HeLa) 세포에서 NMD의 타겟이 되지 않는 유전자들이 3’UTR 길이가 길어짐에 따라 발현량이 증가되어 있는 것을 밝혔다. 뿐만 아니라, UPF1에 의한 조절기작처럼 miRNA 타겟 유전자들의 발현이 증가되어 있는 것도 확인하였다.

이러한 결과들을 종합하면, UPF1은 AGO2와 상호작용하여 miRNA에 의한 mRNA 조절 기작에서 상승작용 효과를 만들어 내는 것은 진화적으로 보존되어 있고, UPF1에 의한 발현 조절은 SMG7과 연관이 있다는 새로운 기작을 제안할 수 있다.

Acknowledgements

This thesis signifies my first step for career in science. I learned research skills such as logical thinking and data-driven decision through this remarkable experience. Without help of many people whom I wish to thank, it would have been impossible.

First and foremost, I would like to express my deep gratitude to Prof. Dr. Jin-Wu Nam for giving me the opportunity to study under his direct supervision. I am truely grateful for his support, patience and guidance in every aspect of my training.

I appreciate experimental supports of Prof. Dr. Jungwook Hwang. His enormous insight in experiments taught me how to design experiments to prove hypothesis. I also would like to thank the other lab members for their feedback and helpful discussions.

Lastly, I am extremely grateful to my parents who gave me chance to focus on the thesis without anxiety.

연구 윤리 서약서

본인은 한양대학교 대학원생으로서 이 학위논문 작성 과정에서 다음과 같이 연구 윤리의 기본 원칙을 준수하였음을 서약합니다.

첫째, 지도교수의 지도를 받아 정직하고 엄정한 연구를 수행하여 학위논문을 작성한다.

둘째, 논문 작성시 위조, 변조, 표절 등 학문적 진실성을 훼손하는 어떤 연구 부정행위도 하지 않는다.

셋째, 논문 작성시 논문유사도 검증시스템 "카피킬러"등을 거쳐야 한다.

2016년06월04일

학위명 : 석사

학과 : 생명과학과

지도교수 : 남진우

성명 : 서좌원 

한 양 대 학 교 대 학 원 장 귀 하

Declaration of Ethical Conduct in Research

I, as a graduate student of Hanyang University, hereby declare that I have abided by the following Code of Research Ethics while writing this dissertation thesis, during my degree program.

"First, I have strived to be honest in my conduct, to produce valid and reliable research conforming with the guidance of my thesis supervisor, and I affirm that my thesis contains honest, fair and reasonable conclusions based on my own careful research under the guidance of my thesis supervisor.

Second, I have not committed any acts that may discredit or damage the credibility of my research. These include, but are not limited to : falsification, distortion of research findings or plagiarism.

Third, I need to go through with Coppykiller Program(Internet-based Plagiarism-prevention service) before submitting a thesis."

JUNE 04, 2016

Degree : Master

Department : DEPARTMENT OF LIFE SCIENCE

Thesis Supervisor : Nam, Jin-Wu

Name : SEO JWAWON (Signature)