UNIVERSITY OF CALIFORNIA SAN DIEGO

3' end Modifications of RNA Polymerase III Transcripts by Factors TENT2 and Usb1

A thesis submitted in partial satisfaction of the
requirements for the degree Master of Science

in

Biology

by

Anastasia Elizabeth Pimentel

Committee in charge:

Professor Jens Lykke-Andersen, Chair
Professor Lin Chao
Professor Amy Pasquinelli

2021

The thesis of Anastasia Elizabeth Pimentel is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2021

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SUPPLIMENTAL TABLES

ACKNOWLEDGMENTS

There are many people without whom this pandemic thesis would not have been possible: my committee, the Lykke-Andersen lab, and my family. Many thanks to the Lykke-Andersen lab for supporting me in my research. Thanks especially to Dr. Jens Lykke-Andersen for his encouragement and enthusiasm and to Cody Ocheltree for his unwavering mentorship and willingness to answer my many questions about experiments. I'm grateful to have had this opportunity to learn with you all these past two years.

I'd like to thank my family for their support and always being a phone call away when I need them.

ABSTRACT OF THE THESIS

3' end Modifications of RNA Polymerase III Transcripts by Factors TENT2 and Usb1

by

Anastasia Elizabeth Pimentel

Master of Science in Biology

University of California San Diego, 2021

Professor Jens Lykke-Andersen, Chair

Eukaryotic RNA processing and quality control involves factors that add 3' end modifications on RNA transcripts. However, the importance of most modifications of small non-coding RNAs (sncRNAs) is poorly understood. The phosphodiesterase Usb1 and the non-canonical poly(A) polymerase TENT2 are known to modify the 3' ends of RNA polymerase-III (pol-III) transcripts U6 and 7SL, respectively. However, Usb1 and TENT2's targets have not been systematically identified. To better understand the scope

of 3' end modifications of sncRNAs, I analyzed differences in post-transcriptional 3' end modifications between pol-III and pol-II sncRNAs and found that adenylation can be observed at the mature 3' end of most pol-III transcripts at a level significantly higher than for pol-II transcripts. For all pol-III transcripts, only a subset of the population is adenylated at steady state, and adenylation is generally restricted to a single adenosine. By quantifying differences in 3' end adenylation via 3' end sequencing in the presence or absence of TENT2, I identified several pol-III transcripts as targets of TENT2, including 7SL, U6atac, and Y RNAs. As Usb1 is known to add 3' end cyclic phosphates to U6 snRNA, we hypothesized that Usb1 may be responsible for modifying additional pol-III transcripts. However, my experiments did not show evidence for cyclic phosphates on additional tested pol-III transcripts. My results demonstrate that adenylation can be observed for most pol-III sncRNAs and that TENT2 is responsible for the adenylation of a subset of these RNAs, suggesting a general role for 3' monoadenylation in pol-III transcript processing or function.

Introduction

RNAs are essential molecules that perform many varied and necessary tasks. While RNAs are most commonly thought of in terms of the central dogma of DNA to RNA to protein, non-coding RNAs also play many central roles in the cell (Eddy 2001; Costa 2005). Non-coding RNAs are RNA transcripts that, once transcribed, are fully functional in the cell as RNA and are not translated into proteins. An important class of non-coding RNAs are small non-coding RNAs (sncRNAs) (Bunch 2018; Jawdekar and Henry 2008). There are many different categories of sncRNAs, including microRNAs, Y RNAs, and small nuclear and nucelolar RNAs (Mattick and Makunin 2006; Kowalski and Krude 2015). SncRNAs have been discovered to perform important and diverse tasks in the cell, such as serving as scaffolds and catalysts in ribonucleoprotein complexes and regulating gene expression through effects on mRNA and DNA (Mattick and Makunin 2006; Hombach and Kretz 2016). As RNAs are essential to the cell, it is vital that the cell precisely regulates how transcripts are expressed. Improper upregulation and downregulation of sncRNAs have been implicated in various diseases from cancers to Parkinson's (Zhao et al. 2007; Croce 2009; Esteller 2011).

3' end modifications play important roles in regulating RNA processing and function (Munoz-Tello et al. 2015). For example, in eukaryotic mRNAs, 3' end polyadenylation plays an important role in activating translation and maintaining steady-state levels of mRNA by protecting the transcripts from degradation (Meyer et al. 2004). For sncRNAs, many types of 3' end modifications have been observed, including adenylation, uridylation, and formation of 3' end cyclic phosphates (Mroczek and Dziembowski 2013), but what factors make these modifications and how they

contribute to sncRNA processing, function and/or stability is poorly understood for most RNAs (Wlotzka et al. 2011; Song et al. 2015). One aspect that influences how different sncRNA species are post-transcriptionally modified on the 3' end is if they are transcribed by RNA pol-III or pol-II (Core and Adelman 2019; Park et al. 2017; Willis and Moir 2018; Arimbasseri et al. 2013). RNA pol-III and pol-II have distinct methods of transcription initiation and termination which result in different 3' end profiles of the RNA transcript (Dergai and Hernandez 2019). Additionally, there are a multitude of factors in the cell that can work in concordance or opposition with each other to modify the 3' ends of sncRNA, often promoting stabilization or degradation (Menezes et al. 2018). Four factors of interest that have been previously identified to be involved in the 3' end modifications of sncRNAs in humans are the polymerases TENT2 (Kwak et al. 2004), TENT4A and TENT4B (Ogami et al. 2013; Rammelt et al. 2011; Berndt et al. 2012), and the cyclic phosphodiesterase Usb1 (Didychuk et al. 2017).

TENT2 was first identified in *Caenorhabditis elegans* as a non-canonical poly(A)-polymerase and was shown to have mRNA polyadenylation activity in the cytoplasm (Wang et al. 2002). TENT2 was also found to adenylate the RNA pol-III transcribed sncRNA 7SL (Katoh et al. 2009) an essential component of the signal recognition particle (Ullu and Weiner 1984) which has an important role in protein synthesis as it transports newly translated peptides across the endoplasmic reticulum membrane. TENT2 has also been observed to add a single adenosine to the 3' ends of certain miRNA transcripts, which appears to promote their stabilization (D'Ambrogio et al. 2012; Chung et al. 2016; Hojo et al. 2020). In general, previous research on TENT2 has focused on a candidate-based approach instead of considering the whole transcriptome.

Therefore, it is not known to what extent TENT2 adenylates sncRNA transcripts in the cell, how TENT2 identifies its targets, or what the full repercussions of these modifications are.

Two additional human non-canonical polymerases are TENT4A and TENT4B. These two factors share a common homologue in yeast, the RNA polymerase Trf4 (Ogami et al. 2013; Rammelt et al. 2011). Trf4 has been found in yeast to tag sncRNA transcripts with post-transcriptional oligonucleotide tails to facilitate the degradation of those transcripts via the nuclear RNA exosome (LaCava et al. 2005; Berndt et al. 2012). Research in human cells has suggested that both TENT4A and TENT4B similarly associate with the human nuclear exosome (Lubas et al. 2011; Schmidt and Norbury 2010), and show preference for adenylating sncRNAs (Ogami et al. 2013; Rammelt et al. 2011). While TENT4B appears to localize in the nucleolus, TENT4A appears to be excluded from the nucleolus (Ogami et al. 2013), although the purpose of this localization is not fully understood, nor are the particular targets of TENT4A/TENT4B. Thus, creating biological tools that can advance the further investigation of what the unique RNA targets of TENT4A/TENT4B are, what the role TENT4A/TENT4B different localizations play in their target selection, and the part TENT4A/TENT4B play in exosome-mediated degradation pathways will help us better understand the role TENT4A/TENT4B play in the processing and quality control of sncRNAs.

Another factor known to modify sncRNA pol-III transcripts is Usb1, a cyclic phosphodiesterase. Usb1 is responsible for the 3' end modification of U6 snRNA (Didychuk et al. 2017). U6 snRNA is a central component of the spliceosome (Didychuk et al. 2018). Loss of Usb1 function has been linked to poikiloderma with neutropenia, a skin disease (Hilcenko et al. 2013; Mroczek and Dziembowski 2013). U6 snRNA is produced by RNA pol-III with a 3' end tail of uridines that Usb1 trims back to five uridines

and adds a post-transcriptional cyclic phosphate to the final uridine (Nomura et al. 2018, 2019). This cyclic phosphate is thought to protect U6 from degradation (Mroczek et al. 2012). Although Usb1 is not currently known to process other RNA transcripts, it is possible that it has other targets in addition to U6 snRNA not yet discovered. Especially of interest are potential target transcripts with 3' end uridylation similar to U6 when unprocessed, as this uridylation may signal Usb1 to modify the transcript (Nomura et al. 2018).

To gain a better understanding of the processing and quality control of sncRNAs in human cells, this thesis aims to 1) analyze differences in adenylation between sncRNAs species transcribed by RNA pol-III and RNA pol-II, 2) identify species in the sncRNA transcriptome that are adenylated by TENT2, 3) design tools for future identification of TENT4A and TENT4B targets, 4) investigate the possibility of Usb1's addition of 3' end cyclic phosphates to RNA pol-III transcripts in addition to U6 snRNA. These findings help provide insight into the mechanisms and functioning of sncRNA 3' end processing by clarifying the modifications made to particular sncRNA species by certain factors.

To this end, I analyzed an RNA-seq dataset from the NCBI GEO repository (Łabno et al. 2016) to determine sncRNA 3' end post-transcriptional modifications of human HEK293 cell sncRNAs, and detected adenylation of most RNA pol-III transcripts. Knockdown of TENT2 combined with global sequencing of 3' ends of sncRNA transcripts with or without TENT2 rescue revealed that TENT2 adenylates a subset of these pol-III transcripts. So that similar knockdown and rescue experiments with TENT4A and TENT4B can be performed in future experiments, I developed cell lines with exogenous wild-type and catalytic mutant TENT4A/B genes integrated into the genome. To explore

Usb1's possible modification of RNA pol-III transcripts other than U6 snRNA, I performed assays to monitor 3' end cyclic phosphorylation on sncRNA pol-III transcripts U6, 7SL and 7SK, and additionally performed knockdown of Usb1 to analyze changes in 3' end cyclic phosphate modification. These analyses help us understand the role of TENT2 and Usb1 and how both of these factors fit into the landscape of sncRNA processing.

Results

Analysis of sncRNA adenylation

In preliminary analyses, I had observed monoadenylation of a subset of RNA pol-III transcripts (data not shown). To determine if adenylation is a general feature of RNA pol-III transcripts and if there are differences in levels of adenylation between sncRNA species transcribed by RNA pol-III or RNA pol-II, I analyzed a 3' end RNA-seq dataset of wildtype HEK293 cells retrieved from the NCBI GEO repository, which was chosen based on its high read depth and quality (Łabno et al. 2016). The dataset was aligned to the human genome using Spliced Transcripts Alignment to a Reference (STAR) software (Dobin et al. 2013), and in-house python and R scripts were used to analyze thirty-one of the most abundant sncRNA species transcribed by RNA pol-III and pol-II from free-standing genes, excluding tRNAs (i.e. sncRNAs not produced from poyladenylated precursor RNAs, Table 1). The average number of adenosines added per transcript for each RNA species analyzed was plotted using custom R scripts to generate scatterplots (Figure 1). Of these species, only U8 snoRNA was not detected in the sequencing dataset and does not appear in the analyses.

For RNA pol-III transcripts analyzed, the median average number of post-transcriptional adenosines per transcript was 0.47 while the median for RNA pol-II transcripts was only 0.04 (Figure 1A). The observed difference in adenylation is statistically significant, with a p-value from a Kolmogorov–Smirnov test of 0.001. There is also more variance in the adenylation levels of RNA pol-III transcripts with an interquartile range (IQR) of 0.49, while the RNA pol-II transcripts have an IQR of 0.09. These results indicate that RNA pol-III transcribed species, as a whole, have a higher

level of post-transcriptional adenylation than RNA pol-II transcribed species, but that RNA pol-III transcribed species are not universally post-transcriptionally adenylated. The RNA pol-III transcribed species with the highest levels of adenylation was 7SL2 with 0.94 adenosines added per transcript on average. The species 7SL1, 7SL3, 7SK, RNase MRP, Y1, and Y3 were above or at the median level of adenylation for RNA pol-III transcribed species. The RNA pol-II transcript with the highest levels of adenylation was U11, with 0.28 adenosines added per transcript, while every other RNA pol-II transcribed species is clustered around the median of 0.04 adenosines added per transcript.

Since most RNA pol-III transcripts, as well as the highly adenylated RNA pol-II transcript, U11 snRNA, terminate in uridine, I tested whether a correlation exists between the terminal nucleotide of the sncRNAs and their level of post-transcriptional adenylation. Species that have a genomically encoded terminal uridine have the highest levels of post-transcriptional adenylation with a median of 0.19 adenosines added per transcript (Figure 1B), while species that terminate in a genomically encoded adenosine have the lowest median of 0.03. Those species that terminate in a cytidine have a median of 0.09 and those that terminate in a guanosine have a median of 0.04. The species that terminate in a uridine have the largest IQR of 0.48 and the species that terminate in a cytidine have a fairly large IQR as well of 0.26. However, species that terminate in an adenosine or guanine are clustered around low median levels of adenylation, with IQRs of 0.04 and 0.08 respectively.

The species that terminated in a uridine that had levels of adenylation above the median were 7SL2, Y1, 7SL1, RNase MRP, 7SK, 7SL3, U11, and U6atac. Y3 and Y4 RNAs are the only two species that show high levels of adenylation while not terminating with a uridine according to the analysis. Y3 RNA also showed high levels of adenylation of an

average 0.49 adenosines added per transcript and Y4 had 0.35 adenosine added per transcript. As the transcripts that have a terminal adenosine are the least adenylated and have the smallest IQR, the other three groups of transcripts were compared to them via Kolmogorov–Smirnov tests. These resulted in a p-value of 0.16 for species with a terminal uridine, 0.66 for species with a terminal cytidine, and 0.99 for species with a terminal guanine. While none of these p-values indicate statistical significance (i.e., p-value ≤ 0.05), the comparisons do show that the transcripts terminating in a uridine are the least similar when compared to the transcripts terminating with an adenosine. Overall, these results show that RNA pol-III sncRNA transcripts have significantly higher levels of adenylation than RNA pol-II sncRNA transcripts and that this is likely either a result of them being transcribed by RNA pol-III or of terminating in a uridine (Figure 1).

RNA pol-III sncRNA adenylation occurs primarily at the mature 3' end and is mostly restricted to a single adenosine

To test where sncRNAs are adenylated and how many adenosines are added, I investigated the top ten adenylated RNA pol-III sncRNAs in more detail. The three 7SL transcripts have a strong trend of monoadenylation, where a single adenosine is added on up to 80% of transcripts (Figure 2A). The majority of this monoadenylation is added to mature length 7SL1 and 7SL2 RNAs, while much of the monoadenylation on 7SL3 RNA is added onto transcripts extended by one to two nucleotides (Figure 2B). For the Y RNAs, up to 20% of transcripts are post-transcriptionally adenylated, with a lower level (~5%) of a second adenosine added (Figure 2A). Uridine modifications are also observed for Y3 and Y4 RNAs. These modifications are added onto mainly mature length or extended Y RNAs (Figure 2B).

While untemplated guanosines are also observed in these analyses, upon close inspection of the primary reads, these appear to be artifacts coming from an unknown linker, which were also observed at various levels for the other sncRNAs shown. 7SK RNA shows monoadenylation of almost 20% of transcripts as well as some uridylation, which are mainly added to mature or extend length transcripts (Figures 2A and 2B). U6atac shows a low level (~5%) of transcripts that are monoadenylated at the mature length or longer (Figure 2A and 2B). Although RNase P and RNase MRP were among the ten most adenylated pol-III species, these were excluded from this analysis because of a large fraction of reads containing the G-rich linker that was observed at lower levels for the other RNA species.

TENT2 knockdown and analysis of target candidates

TENT2 has previously been implicated in the adenylation of 7SL RNA (Katoh et al. 2009). In order to test whether TENT2 also post-transcriptionally modifies other sncRNAs, I analyzed 3' end sequencing data from an siRNA knockdown of TENT2 performed in the lab. The knockdown experiment had four conditions, 1) a control siRNA condition that targeted an mRNA known to not be produced by the cells, 2) an siTENT2 condition that targets TENT2 mRNA and thus decreases TENT2 levels, 3) a rescue condition, where cells are still treated with siTENT2, but in a cell line that expresses an exogenous wildtype TENT2 mRNA that is resistant to the siTENT2, and 4) a failed rescue condition, where cells are treated with siTENT2, in a cell line that expresses an exogenous TENT2 mutant that is catalytically inactive. These last two conditions of a rescue and a failed rescue are important controls for potential off-target effects of the used siRNA.

The sncRNAs from the cells from each of these conditions were isolated and 3' end sequenced. These sequencing data were analyzed to identify which transcript showed significantly shortened 3' ends (Kolmogorov–Smirnov test p-value ≤ 0.05) in the comparison of both conditions with functional TENT2 present or absent (i.e. control siRNA vs. siTENT2 and TENT2 rescue vs. catalytic inactive TENT2 rescue). Unbiased analysis of all sequenced RNAs identified six RNAs that showed significant differences between each tested condition, 7SL1, 7SL2, 7SL3, Y1, Y3, and U6atac RNAs (Table 3). Notably, these are all RNA pol-III transcripts.

The targets of TENT2 identified by this analysis fall into three groups: 7SL transcripts, Y RNA transcripts, and U6atac (Table 3). Analyses of the 3' end post-transcriptional tails of these RNAs in the absence or presence of TENT2 activity (Figure 3) support the results of the pol-III sncRNA analysis (Figure 1): a large fraction of the populations of the three 7SL RNAs are monoadenylated, while Y1, Y3, and U6atac RNAs show more varied post-transcriptional modifications (Figure 3). The adenylation of 7SL RNAs, Y RNAs, and U6atac snRNA is reduced upon depletion of TENT2 activity (Figures 3A and 3B). Although Y1, Y3, and U6atac RNAs show more varied post-transcriptional modifications, adenylation is seen to decrease when functional TENT2 is absent while other modifications such as uridylation do not show the same pattern of decrease (Figures 3A and 3B). Thus, TENT2 is involved in the post-transcriptional adenylation of at least a subset of RNA pol-III transcripts.

Creation of tools for future experiments on TENT4A and TENT4B

TENT2 may not be the only terminal transferase involved in the adenylation of RNA pol-III transcripts. In order to build tools to analyze two other terminal transferases,

TENT4A and TENT4B, I created 293-tREX cell lines designed for an siRNA knockdown experiment with rescue controls, similar to the cell lines used in the TENT2 knockdown experiments (Figure 3). For both TENT4A and TENT4B, I created a 293-tREX cell lines with exogenous 3xFLAG-tagged open reading frames of TENT4A or TENT4B inserted into the genome under the control of a tetracycline-regulated promoter. These cell lines can be used to test a rescue condition in an siRNA experiment when the siTENT4A/B is designed so it cannot target the exogenous open reading frame versions of the TENT4A/B mRNA. Additionally, I created cell lines for both TENT4A and TENT4B that have a catalytically mutated version, which can be used in an siRNA experiment to test a failed rescue condition.

Are RNA pol-III transcripts additional to U6 snRNA modified with a cyclic phosphate?

Usb1 is known to modify the RNA pol-III transcript U6 snRNA with a 3' end cyclic phosphate, which is important for U6 snRNA stability. To investigate whether Usb1 modifies pol-III transcripts other than U6 snRNA, I developed two siRNAs targeting Usb1 and treated human HeLa cells with these siRNAs as compared to a control siRNA. To confirm knockdown of Usb1 and investigate the effect on a set of sncRNAs, nascent Usb1, U6, U1, 7SK and 7SL RNA transcript levels were measured by qRT-PCR. U1 snRNA is known to not have the addition of a 3' end cyclic phosphate (Nomura et al 2020), while it is unknown if 7SL and 7SK have a 3 'end cyclic phosphate. Ct-values for the qRT-PCR assays were normalized to levels of GAPDH, RPS16, and aldolase mRNAs, as well as to the corresponding RNA levels observed in the control siRNA condition. Nascent levels of Usb1 transcripts in the siUsb1 #1 condition were decreased to about 23% of the control siRNA, and to about 13% of the control levels for the siUsb1 #2 condition (Figure

4). There was no substantial decrease in the level of U6 snRNA in the Usb1 knock-down conditions, suggesting that Usb1 was not sufficiently knocked down to affect U6 snRNA stability, or that Usb1 does not affect U6 snRNA stability in this particular cell line. There was also no observed effect on U1, 7SK, or 7SL transcript levels between the control siRNA and the siUsb1 conditions (Figure 4).

To monitor changes in post-transcriptional cyclic phosphorylation, I performed sequencing and qPCR of T4 Polynucleotide Kinase (PNK)-treated versus non-treated total RNA samples isolated from the control siRNA and siUsb1 #2 conditions. PNK treatment removes any 3' end cyclic phosphorylation. As 3' end cyclic phosphorylation prevents linkers from being ligated to the 3' end of RNA, PNK treatment allows those transcripts that had 3' end cyclic phosphates to now be ligated with linkers and then sequenced or amplified via qPCR. By comparing transcripts treated with PNK, which includes transcripts that had a 3' end cyclic phosphate, to transcripts not treated with PNK, which excludes transcripts with a cyclic phosphate, it is possible to make inferences about the population of transcripts with a 3' end cyclic phosphate.

After PNK or mock PNK treatment, gene-specific 3' end sequencing was performed for U1, U6, 7SL, and 7SK RNAs, and their 3' ends were analyzed. A much higher percentage of U6 snRNAs terminate at mature length upon treatment with PNK as compared to those transcripts not treated with PNK (Figure 5A). This is consistent with a large fraction of mature U6 snRNAs being modified with a terminal cyclic phosphate. However, there is no substantial difference in the transcripts treated or not treated with PNK for U1, 7SL, or 7SK, suggesting that these are not modified by a cyclic phosphate to a level detectable in this assay (Figure 5A). Knockdown of Usb1 does not change the 3'

end profile of any of the transcripts, including U6 snRNA, suggesting that this depletion is insufficient to reduce Usb1 to limiting levels for U6 snRNA modification (Figure 5B).

To analyze the levels of cyclic phosphate more directly, RNA samples were treated in the absence or presence of PNK prior to 3' end oligo ligation and subjected to qRT-PCR targeting U6, 7SL, 7SK RNAs. When normalized to U1 snRNA levels, levels of U6 doubled (2.095) when treated with PNK, with a students t-test p-value of 0.042, consistent with the presence of a cyclic phosphate (Figure 6). There are slight differences between the PNK treated and non-treated conditions for 7SL and 7SK RNAs, but the differences are not statistically significant. Thus, my experiments, which successfully detected a cyclic phosphate on U6 snRNA, found no evidence for cyclic phosphates at the 3' ends of 7SL or 7SK RNAs.

Figures

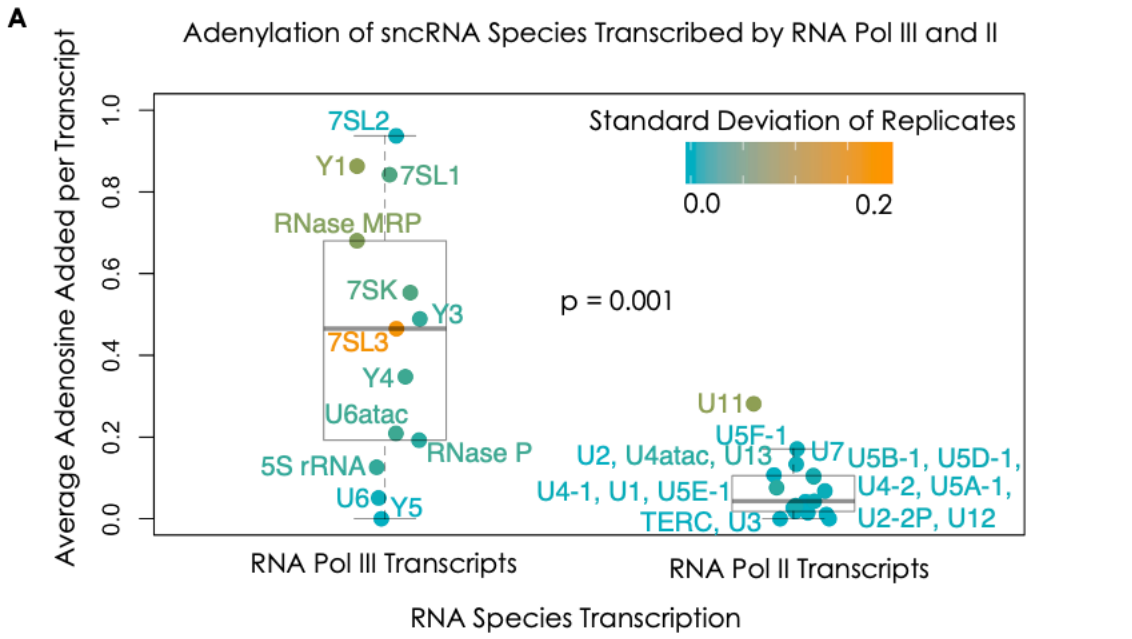
Table 1. Characteristics of analyzed sncRNA species

RNA Species	Pre-modified Genomic 3' end	Transcribed by
U1	CCCUG	RNA pol-II
U2	GCACC	RNA pol-II
U2-2P	GUUUA	RNA pol-II
U3	UUUUU	RNA pol-II
U4-1	UACGG	RNA pol-II
U4-2	UACGG	RNA pol-II
U4atac	AAAUG	RNA pol-II
U5A-1	CCGUG	RNA pol-II
U5B-1	AAGGC	RNA pol-II
U5D-1	AGGGU	RNA pol-II
U5E-1	CAGGG	RNA pol-II
U5F-1	AGGCU	RNA pol-II
U6	UUUUU	RNA pol-III
U6atac	CGUUU	RNA pol-III
U7	AAGCC	RNA pol-II
U8	CGAAU	RNA pol-II
U11	CCCUU	RNA pol-II
U12	GCCCG	RNA pol-II
U13	UGACC	RNA pol-II
7SK	UCUUU	RNA pol-III
7SL1	CUCUU	RNA pol-III
7SL2	CUCUU	RNA pol-III
7SL3	CCUCU	RNA pol-III
Y1	UCUUU	RNA pol-III
Y3	ACAGA	RNA pol-III
Y4	CUAAA	RNA pol-III
Y5	GUUUU	RNA pol-III
5S rRNA	GCUUU	RNA pol-III
RNase MRP	GCUGU	RNA pol-III
RNase P	CCAUU	RNA pol-III
TERC	CCUGA	RNA pol-II

Table 2. Modifications made to 3' end location relative to ENSEMBL annotations

RNA Name	New 3' end relative to ENSEMBL annotated 3' end (0)
7SK	2
7SL1	1
U2 wildtype	-4
U2-2P	-4
U3	-40
U4atac	3
U6atac	-3
Y1	-35
Y3	-5
Y4	-6
Y5	-68
TERC	-90
U4-1	3
U4-2	3
U5A-1	1
U5B-1	1
U5D-1	-14
U5E-1	2
RNase P	4

Figure 1. SncRNA species transcribed by RNA pol-III and terminating in a genomic uridine have higher levels of post-transcriptional 3' end adenylation than species transcribed by RNA pol-II. A) Scatterplots show wildtype 3' end RNA-seq data with each dot representing an RNA species. The Kolmogorov–Smirnov test comparing the two sets of transcript gives a p-value of 0.001. The color of the dot represents the standard deviation of the sequencing of the three biological replicates (n=3), with blue representing lowest standard deviation (0.00) and orange representing the highest standard deviation (0.20). The x-axis indicates if the RNA species is transcribed by RNA pol-III or RNA pol-II. The y-axis shows the average number of adenosines that are added to a transcript for a particular RNA species. The boxplot shows the minimum, 25th percentile, median, 75th percentile, and maximum of the data. The median adenosine added per transcript was 0.47 for RNA pol-III transcripts, and 0.04 for RNA pol-II transcripts. The IQR was 0.49 for RNA pol-III transcripts and 0.09 for RNA pol-II transcripts. B) The median adenosine added per transcript for species terminating in a genomically encoded adenosine was 0.03 with an IQR 0.04); for uridine the median was 0.19 with an IQR of 0.48; for cytidine the median was 0.09 with an IQR of 0.26; and for guanine the median was 0.04 with an IQR of 0.08. The Kolmogorov–Smirnov test gave a p-value of 0.16 for the comparison of uridine-terminating transcripts to adenosine-terminating transcripts; a p-value of 0.66 for the comparison of cytidine-terminated transcripts to adenosine-terminated transcripts; and 0.99 for guanosine-terminated transcripts to adenosine-terminated transcripts.


A
Adenylation of sncRNA Species Transcribed by RNA Pol III and II
Average Adenosine Added per Transcript
0.0
0.2
0.4
0.6
0.8
1.0
Standard Deviation of Replicates
0.0
0.2
7SL2
Y1
7SL1
RNase MRP
7SK
Y3
7SL3
Y4
U6atac
RNase P
5S rRNA
U6
Y5
p = 0.001
U11
U5F-1
U2, U4atac, U13
U7
U5B-1, U5D-1,
U4-1, U1, U5E-1
U4-2, U5A-1,
TERC, U3
U2-2P, U12
RNA Pol III Transcripts
RNA Pol II Transcripts
RNA Species Transcription

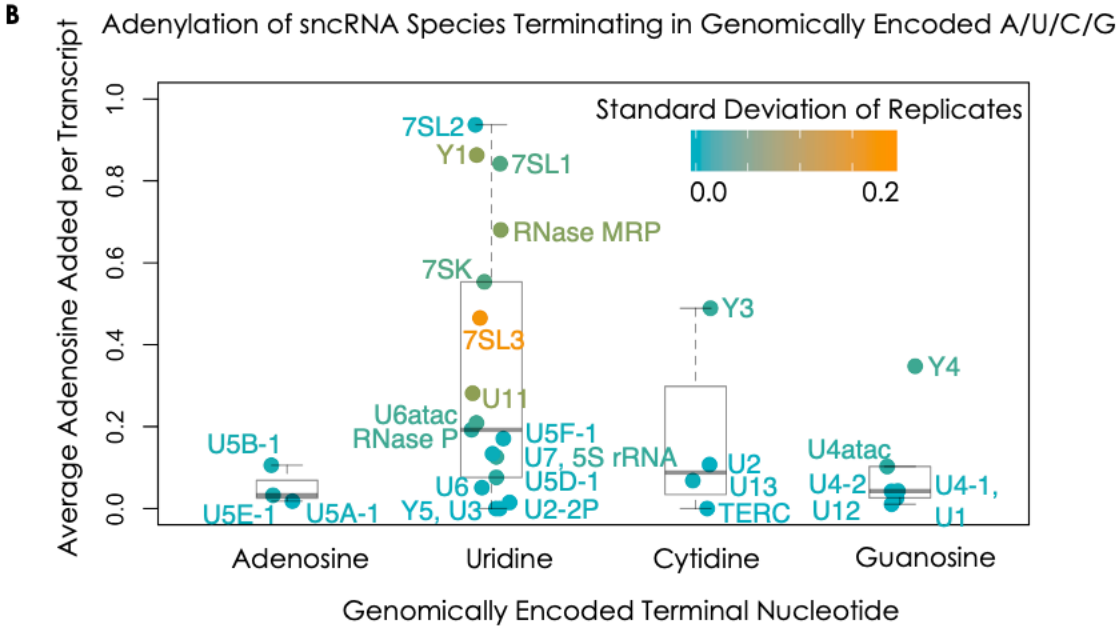

B
Adenylation of sncRNA Species Terminating in Genomically Encoded A/U/C/G
Average Adenosine Added per Transcript
0.0
0.2
0.4
0.6
0.8
1.0
Standard Deviation of Replicates
0.0
0.2
7SL2
Y1
7SL1
RNase MRP
7SK
7SL3
U11
U6atac
RNase P
U5F-1
U7, 5S rRNA
U5D-1
U6
Y5, U3
U2-2P
U5B-1
U5E-1
U5A-1
Y3
U2
U13
TERC
Y4
U4atac
U4-2
U12
U4-1,
U1
Adenosine
Uridine
Cytidine
Guanosine
Genomically Encoded Terminal Nucleotide

Figure 2. Adenylated RNA pol-III transcribed sncRNA species are modified at a mature length and have both mono and oligoadenylation. Plots show wildtype 3' end RNA-seq data. A) Each plot graphically represents for each species the percent of transcripts that were modified with the four RNA nucleotides (A, G, C, or U) for four post-transcriptional added 3' end nucleotide positions 1 to 4. B) Cumulative plot analysis of the ten most adenylated RNA pol-III transcribed species from Figure 1A. Shown is the average of sequencing data from three biological wildtype samples. Solid lines represent the 3' ends of sequenced RNA transcripts including any post-transcriptional nucleotides. Dotted lines represent predicted genomically encoded 3' ends, with the shading between the solid and dotted lines representing post-transcriptional nucleotides. The x-axis shows the position of the 3' end, with "0" representing the mature genomically encoded length of the RNA species.

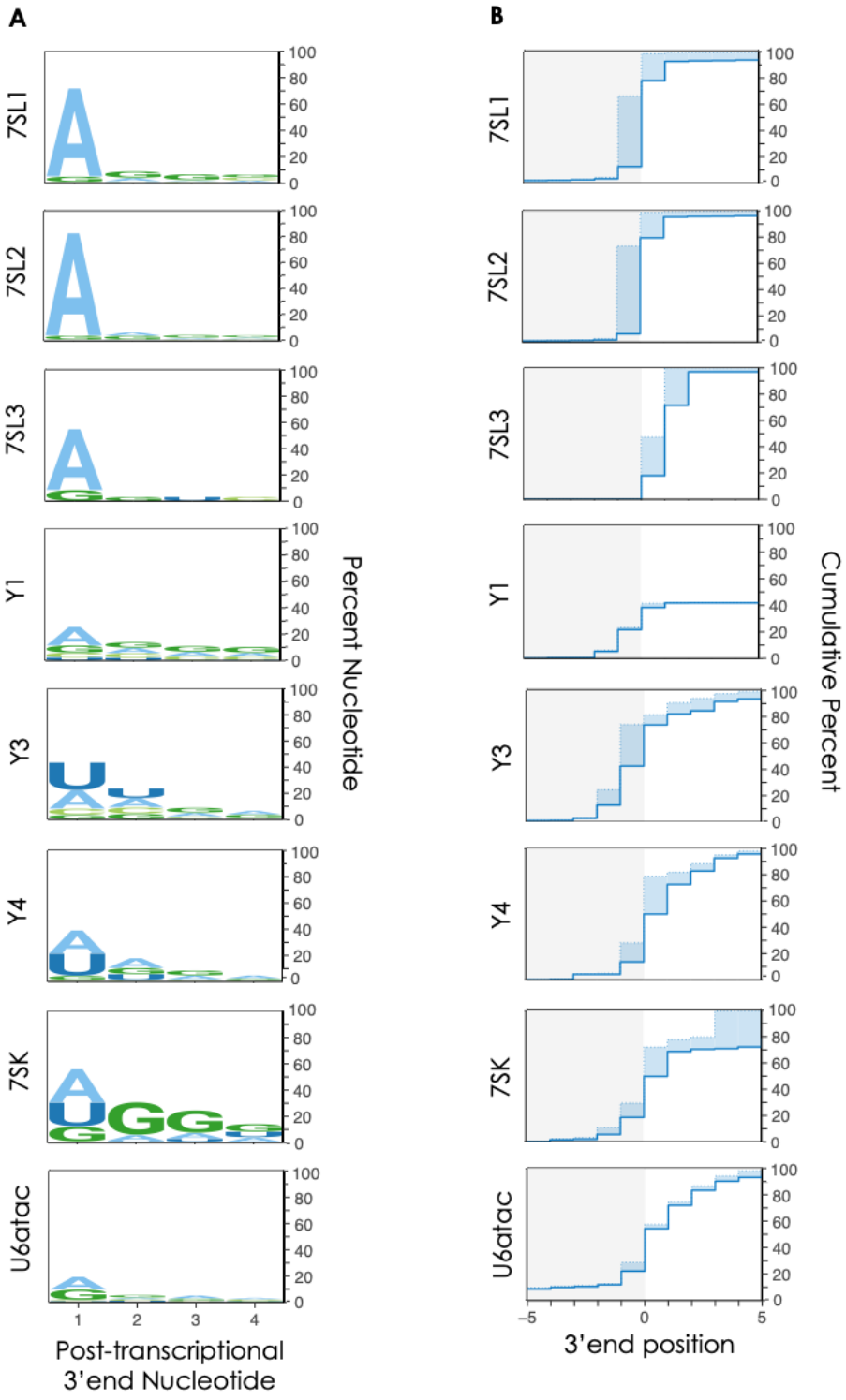
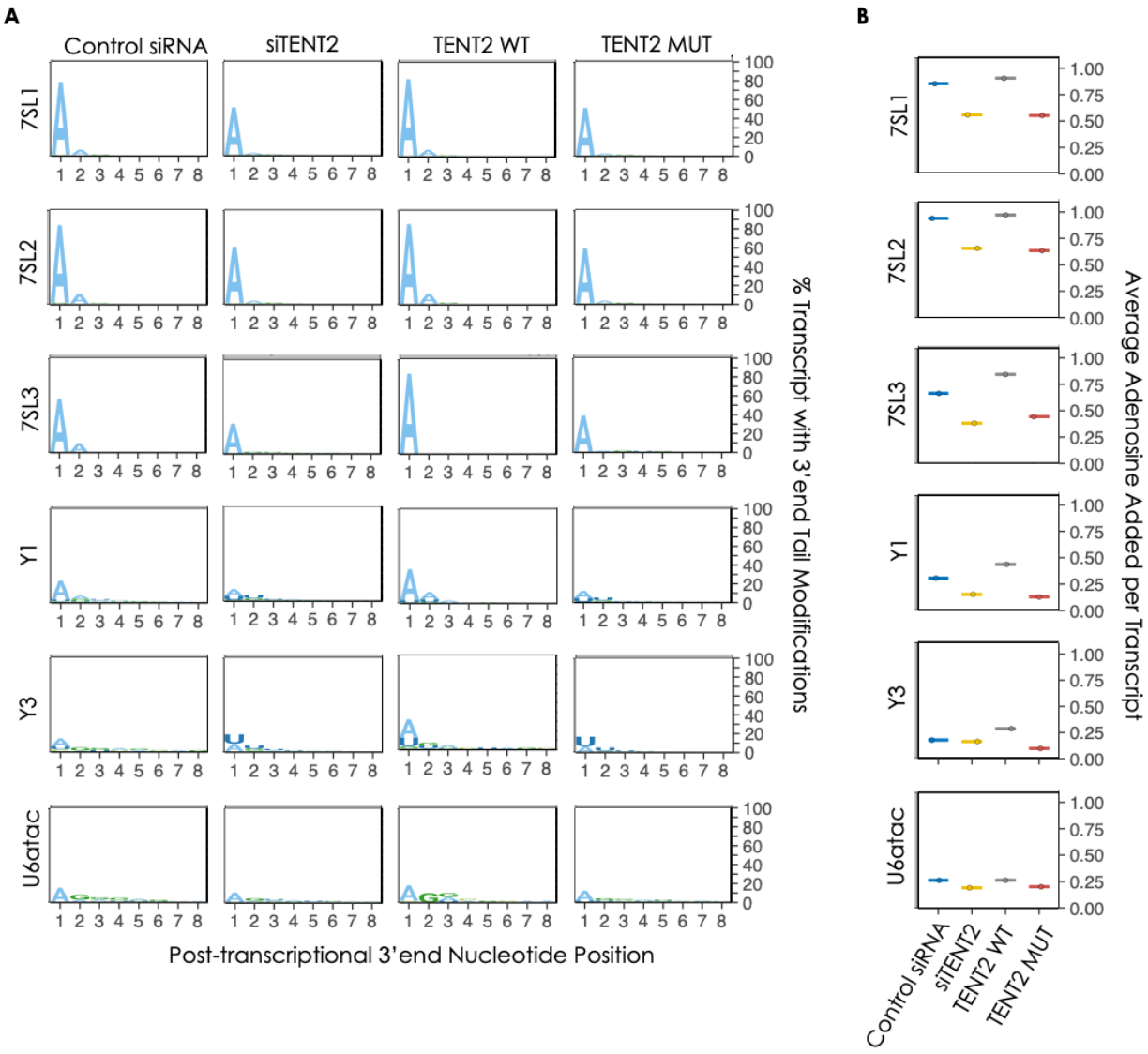

A
B
7SL1
7SL2
7SL3
Y1
Y3
Y4
7SK
U6atac
Percent Nucleotide
Cumulative Percent
Post-transcriptional
3'end Nucleotide
3'end position

Table 3. SncRNAs with significantly changed 3' ends upon TENT2 inactivation

RNA Name	P-Value, Control siRNA vs siTENT2	P-Value, TENT2 WT vs. TENT2 mutant
7SL1	2.99×10^{-134}	0.00
7SL2	9.09×10^{-83}	2.19×10^{-281}
7SL3	4.75×10^{-5}	0.00011829
Y1	6.55×10^{-12}	2.56×10^{-8}
Y3	2.32×10^{-27}	0.01884629
U6atac	8.38×10^{-16}	1.97×10^{-5}

Figure 3. Seqlogo analysis of candidate sncRNA transcripts shows decrease in post-transcriptional 3' end adenylation when functional TENT2 is absent.
These graphs were generated from 3' end sequencing data of sncRNA transcripts isolated from a TENT2 knockdown experiment. The four conditions of the knockdown experiment are 1) a control siRNA condition, 2) an siTENT2 condition, 3) an siTENT2 condition in cell line that expresses an extra wildtype TENT2 mRNA unaffected by siTENT2, and 4) an siTENT2 condition in cell line that expresses an extra mutant TENT2 mRNA both unaffected by siTENT2 and that translates into a non-functional TENT2. The sncRNA transcripts isolated and sequenced from each condition were then analyzed to identify candidate transcripts that showed significant (p-value ≤ 0.05 from Kolmogorov–Smirnov test) changes in adenylation between the conditions with functional TENT2 present (first and third conditions) and functional TENT2 absent (second and fourth conditions). A) Each plot graphically represents for each candidate transcript the percent of transcripts that were modified with the four RNA nucleotides (A, G, C, or U) for four post-transcriptional added 3' end nucleotide positions. B) Plots represent the level of adenylation for each species by the average adenosine added per transcript for each species in each condition.


A
Control siRNA
siTENT2
TENT2 WT
TENT2 MUT
7SL1
7SL2
7SL3
Y1
Y3
U6atac
% Transcript with 3'end Tail Modifications
Post-transcriptional 3'end Nucleotide Position
B
Average Adenosine Added per Transcript
Control siRNA
siTENT2
TENT2 WT
TENT2 MUT

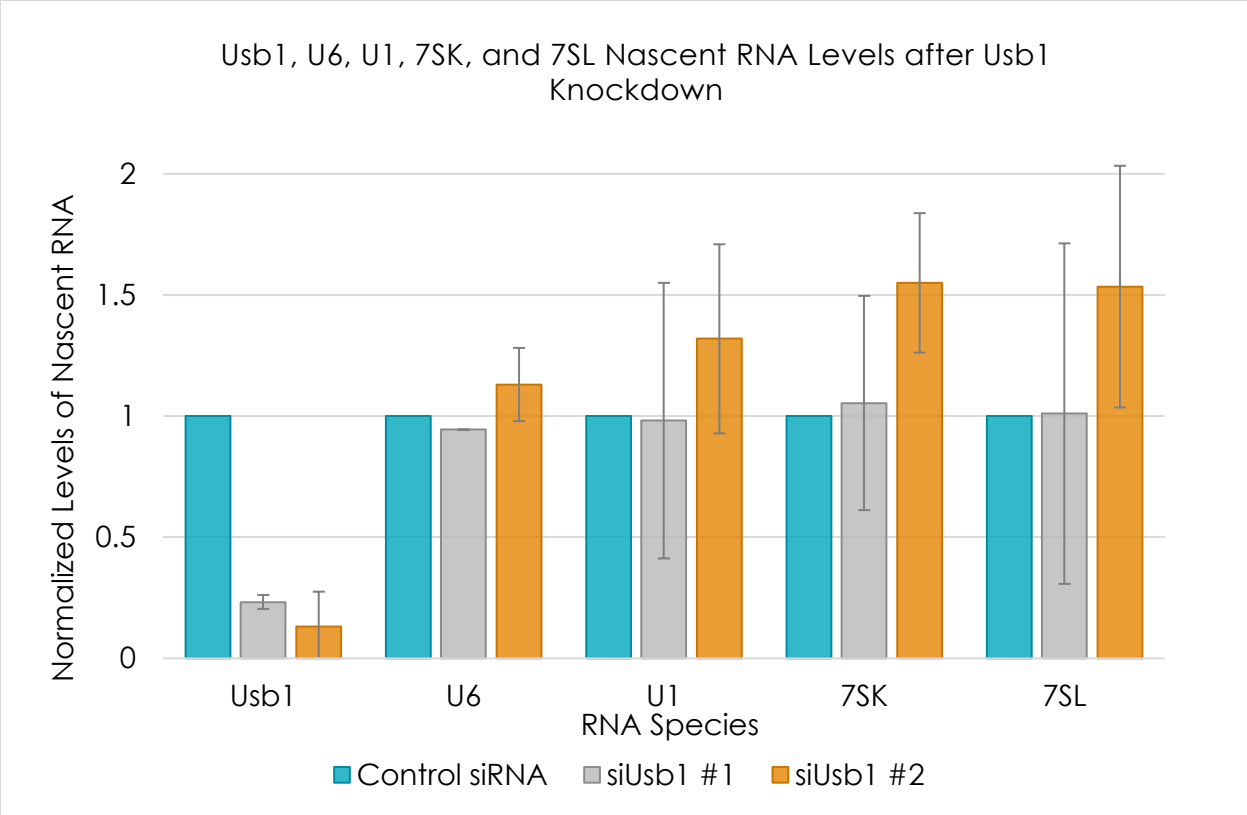


Figure 4. Nascent Usb1, U6, U1, 7SK, and 7SL RNA transcript levels show reduced Usb1 levels. Nascent RNA extracted from 293-tREX cells after Usb1 knockdown by the control siRNA siLuciferase, siUsb1 #1, siUsb1 #2 and analyzed using qPCR. RNA levels normalized to the average of levels of GAPDH, RPS16, and aldolase mRNAs. The blue bar shows the control siRNA conditions, the grey bar shows the siUsb1 #1 conditions, and the orange bar shows the siUsb1 #2 conditions. Error bars represent the standard deviation of three biological replicates (n=3).

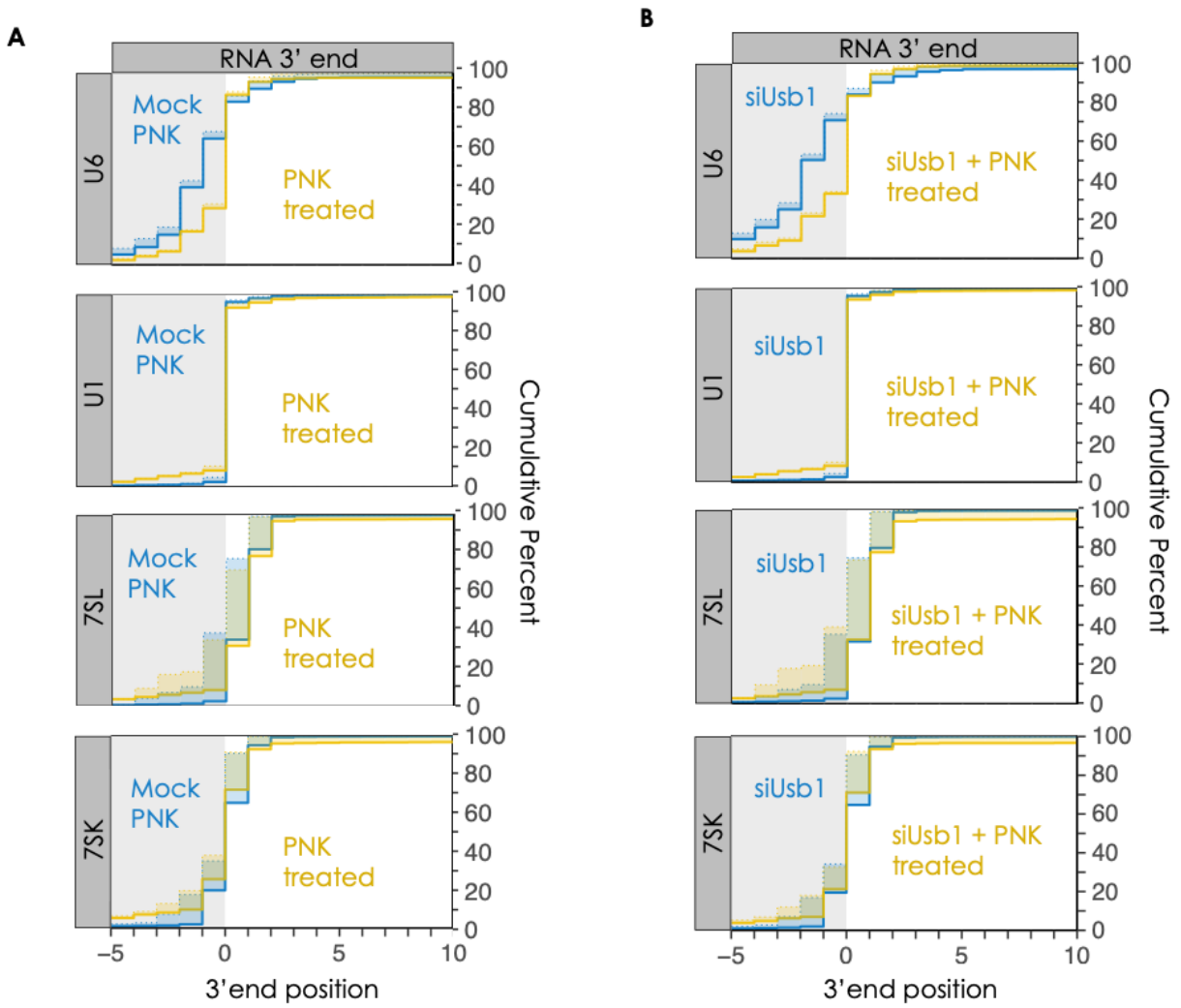


Figure 5. Transcripts 7SL, 7SK, and U1 do not show evidence of 3' end cyclic phosphorylation by Usb1. A) Cumulative plot analysis of U1, 7SL, and 7SK of gene-specific 3' end RNA sequencing data for T4 PNK treated conditions and mock T4 PNK treated conditions. Shown is the average of sequencing data from three biological replicates. Solid lines represent the 3' ends of sequenced RNA transcripts including any post-transcriptional nucleotides. Dotted lines represent predicted genomically encoded 3' ends, with the shading between the solid and dotted lines representing post-transcriptional nucleotides. The x-axis shows the position of the 3' end, with "0" representing the mature genomically encoded length of the RNA species. B) Cumulative plot analysis of U1, 7SL, and 7SK of gene-specific 3' end RNA sequencing data for siUsb1 knockdown with T4 PNK treated conditions and siUsb1 knockdown with mock T4 PNK treated conditions.

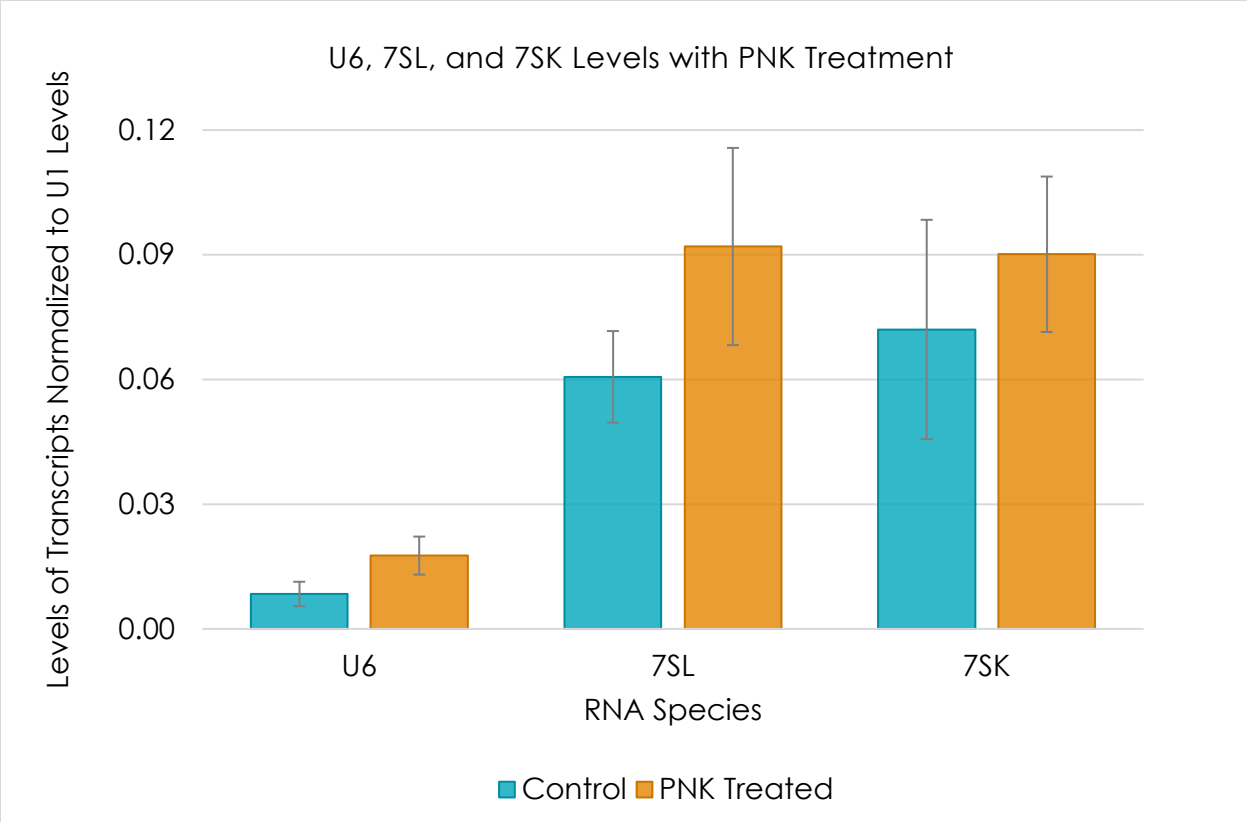


Figure 6. U6 transcripts have a 3' end cyclic phosphate while 7SK, and 7SL RNA transcripts do not have a 3' end cyclic phosphate. Total RNA extracted from 293-tREX cells after control siRNA treatment. RNA was treated with PNK to remove 3' end cyclic phosphates or a control PNK treatment. Transcripts without 3' end cyclic phosphate will be amplified by qPCR, while transcripts with 3' end cyclic phosphate will not be amplified. RNA levels normalized to the average of levels U1. The blue bar shows the control PNK conditions and the orange bar shows the PNK treated conditions. Error bars represent the standard deviation of three biological replicates.

Discussion

My findings demonstrate that many sncRNA species that are transcribed by RNA pol-III are post-transcriptionally adenylated. The extent of adenylation varies by species but, as a group, is significantly more extensive that observed for RNA pol-II transcripts. This adenylation is particularly striking for 7SL RNAs, which showed adenylation, primarily consisting of a single adenosine, at the mature 3' end of about 60 to 90% of the RNA population at steady state depending on the species (Figure 2). Y RNAs were also primarily monoadenylated, but adenylation was seen for only 10-20% of the population, and uridylation was observed for a subset of these RNAs as well (Figure 2). 7SK and U6atac RNAs also accumulated primarily monoadenylated tails on a fraction of the population (Figure 2). Depletion or inactivation of TENT2 caused lowered levels of adenylation of 7SL RNAs, Y1 and Y3 RNAs, and U6atac RNA, demonstrating that TENT2 is involved in the adenylation of at least this subset of RNA pol-III transcripts (Figure 3). Investigating whether another 3' end modification, a 3' end cyclic phosphate, is present in pol-III transcripts other than U6 snRNA, I did not find evidence that 7SL or 7SK RNAs contain such a modification, as no changes were observed in 3' end sequencing or ligation-mediated qRT-PCR assays in the absence or presence of PNK, which removes 3' end cyclic phosphates (Figures 4-6).

As monoadenylation of RNA pol-III transcripts appears to be a common modification, it is important to consider what it's purpose may be. The extent of adenylation at steady state varies widely between species, with the most highly adenylated species in the analysis, 7SL2 RNA, showing over 80% adenylation, and others, such as U6atac, 7SK and Y RNAs with about 20% or less of transcripts adenylated

(Figure 2). This suggests that only a subset of each species becomes adenylated during their life-time, or that 3' end adenosines are transient and can be removed at later stages in the RNA life cycles.

As previous research shows that adenylation by TENT2 can stabilize transcripts (D'Ambrogio et al. 2012) (Katoh et al. 2009), it may be the case that adenylation, particularly of highly adenylated species like 7SL RNAs, also promotes stability, for example by protecting the pol-III transcripts, which generally terminate in uridines, from uridine-specific 3' to 5' exonucleases (Munoz-Tello et al. 2015). As some sncRNA species such as 7SL RNAs have a multitude of pseudogenes (Ullu and Weiner 1984), it is possible that post-transcriptional adenylation is a way to differentiate functional from non-functional transcripts. Another possibility is that adenylation plays a role in pol-III transcript processing and/or assembly into ribonucleoprotein complexes, or even in the function of the mature RNA. However, it is less likely that this adenylation plays a central role in the functioning of those RNA species with low steady-state adenylation levels. Future studies should distinguish between these possibilities to clarify the purpose of this observed monoadenylation.

As my findings demonstrated that RNA pol-III sncRNA species are adenylated to a much higher extent than RNA pol-II species, the question of how pol-III sncRNAs are selectively adenylated is raised. A clue may lie with U11 snRNA, which was the only RNA pol-II transcript that showed high levels of adenylation. All RNA pol-III species have a 3' end oligouridine tail that is genomically encoded as this is the signal for RNA pol-III to terminate transcription (Dergai and Hernandez 2019). Even though U11 snRNA is transcribed by RNA pol-II, which terminates transcription by a different mechanism, U11 snRNA still terminates in two uridines (Table 1) and is the only analyzed RNA pol-II

transcript with this characteristic. It is possible that this oligouridine is involved with signaling the transcript for adenylation, although it cannot be the only signal as not all RNA pol-III transcripts are adenylated.

Finally, although my findings did not indicate the presence of 3' end cyclic phosphates on 7SL or 7SK transcripts, future studies should globally examine the presence of cyclic phosphates on sncRNAs. PNK assay combined with global RNA-seq could allow for the bioinformatic search of candidates for the addition of 3' end cyclic phosphates. One potential issue with the PNK assay used here is that it only indirectly measures the presence of cyclic phosphates, which increases the possibility that other factors are influencing the results. It may be useful to perform future assays that monitor the presence of 3' end cyclic phosphates by using an RNA ligase specific to transcripts with a cyclic phosphate, such as a tRNA ligase. This would help provide more direct evidence for 3' end cyclic phosphorylation.

Materials and Methods

DIS3L2 analysis

Sequencing data retrieval

RNA-sequencing data from wild type HEK293 cells was retrieved in the form of FASTQ files from the NCBI GEO repository with the accession number GSE82336 (samples GSM2190619, GSM2190620, GSM2190621) (Łabno et al. 2016).

List of genes

Gene sequences are from the ENSEMBL database. In order to account for any mis-annotations of the 3' ends of genes on ENSEMBL, in-house python scripts were used to identify where the majority of sequenced transcripts terminated (not including any post-transcriptional modifications) and the 3' end of all genes was assigned to that nucleotide. This resulted in deviations from ENSEMBL's annotation of the 3' end (Table 2).

Analysis of 3' end data

The sequencing fastq files were aligned with STAR to the human genome version GRCh38.p13 from GENCODE. The resulting SAM files from this alignment were run through in-house python scripts that use gene sequences from ENSEMBL to generate files containing the quantification of tail nucleotide modifications. Two custom R scripts were then used to plot the data from these files to generate scatter plot graphs showing the levels of adenylation on transcripts transcribed by RNA pol-II vs RNA pol-III,

and the levels of adenylation on transcripts with a genomically-encoded terminal adenosine, uridine, cytidine, or guanine.

TENT2 knockdown and analysis of target candidates

Plasmid synthesis and creation of stable cell lines (completed by Cody Ocheltree)

The open reading frame of the human *TENT2* gene was inserted into the pcDNA5/FRT/TO vector with a 3xFLAG sequence 5' of the ORF. Stable cell lines were generated using human HEK FLp-In T-REx-293 cells. For the failed rescue mutant cell line, base pair mutations (Martin et al. 2008) were created with Q5 site-directed mutagenesis kit (NEB), where amino acids 213 and 215 were changed from aspartate to alanine. HEK FLp-In T-REx-293 cells were also created from pcDNA5/FRT/TO-3xFLAGalone.

Tissue culture (completed by Cody Ocheltree)

HEK FLp-In T-REx-293 cells were grown at 37° C with 5% CO2 and passaged before 100% confluency. Cells were grown in DMEM (Gibco) with 10% FBS (Gibco) and penicillin/streptomycin (Gibco). Experiments were performed between the third and tenth passage after thawing from -80° C.

siRNA knockdown (completed by Cody Ocheltree)

Cells were plated at ~8% confluency for siRNA treatment in 6 well plates. siRNA knockdowns were performed with siLentFect (Bio-rad) according to the manufacturer's recommendations. Two sequential knockdowns were performed with 20 nM final

concentration of siLuciferase or siTENT2 (Supplementary Table 1) at 72 hours and 24 hours prior to cell harvest. To induce expression of 3xFLAG, 3xFLAG-TENT2 WT, or 3xFLAG-TENT2 DADA, tetracycline was added 24 hours prior to harvest to media at a final concentration of 50 ng/ml.

RNA isolation and library preparation for sequencing (completed by Cody Ocheltree)

RNA was isolated using TRIzol (ThermoFisher) following the manufacturer's protocol. To create libraries for sequencing, 1 ug of RNA was treated with a half-reaction of Ribo-Zero (Illumina) to deplete rRNA. 3' adapters with randomers bases (Supplementary Table 1) were ligated using T4 RNA ligase 1 (NEB). RNAs between 100 and 500 bp were size-selected on a 9% denaturing polyacrylamide gel to enrich for sncRNAs and exclude tRNAs. cDNA was made using SuperScript III (Thermo) with the AR17 primer (Supplementary Table 1). The 5' linker 3Tr3 (Supplementary Table 1) was ligated using RNA ligase (NEB). Libraries were amplified via PCR with TruSeq CD indices (Illumina) (Supplementary Table 1) and Q5 polymerase (NEB), then size-selected again from an agarose gel for 175-600 bases. Libraries were cleaned using Ampure beads (Beckman Coulter).

Sequencing of sncRNA libraries (completed by Cody Ocheltree)

Sequencing was performed on a MiSeq system (Illumina) using a 300 cycle Reagent Kit V2. 200 cycles were dedicated to the 3' end of RNAs, and 100 cycles were dedicated to the 5'end. Data was recovered from the sequencer in the form of FASTQ files.

Analysis of 3' end data

The sequencing fastq files were aligned with STAR to the human genome version GRCh38.p13 from GENCODE. The resulting SAM files from this alignment were run through in-house python scripts that use gene sequences from ENSEMBL to generate files containing the quantification of tail nucleotide modifications. Kolmogorov–Smirnov tests comparing the post-transcriptional adenylation for each gene in the control siRNA and siTENT2 datasets, and for each gene in the siTENT2 + TENT2 rescue and siTENT2 + TENT2 failed rescue were performed. Genes that had significant differences (p-value of less than or equal to 0.05) in the adenylation of their tails in both the siTENT2 + TENT2 rescue and siTENT2 + TENT2 failed rescue comparisons were considered to be potential candidates for TENT2 targets. For each candidate, graphs of the average adenosine added to the RNA transcripts as well as plots that graphically display the DNA sequence logo weighted by the number of each nucleotide present at each position (seqlogo plots) were generated using in-house R scripts.

TENT4A/B cell line creation

Plasmid synthesis and creation of stable cell lines

Open reading frames of TENT4A and TENT4B were inserted into pcDNA5/FRT/TO-3xFLAG from pcDNA3-3xFLAG. Stable cell lines for the TENT4A and TENT4B exogenous insertions were generated using human FLp-In T-REx-293 cells. To create catalytic null mutants of TENT4A and TENT4B, base pair mutations were created using the Q5 site-

directed mutagenesis kit (NEB) following the manufacturer instructions. For the TENT4A failed rescue the primers (Supplementary Table 1) were used to change amino acids 277 and 279 aspartic acids into alanines (GAC ATC GAC → GCC ATC GCC). For the TENT4B failed rescue, the primers (Supplementary Table 1) were used to similarly change amino acids 177 and 179 aspartic acids into alanines (GAC ATA GAC → GCC ATA GCC). The failed-rescue mutant and rescue plasmids were first transformed into DH5alpha e. coli, then transfected into blasticidin, zeocin-resistant parental FLp-In T-REx-293cells on 6-well plates with DMEM (Gibco). After transformation, cells were split and plated at 25% confluency on 10cm plates with DMEM with 5ug/ml blasticidin and 200 ug/ml hygromycin antibiotic concentrations. Cell colonies were then grown up as single colonies, picked, grown up in 6-well plates, then frozen in cryogenic tubes in liquid nitrogen.

Usb1 knockdown and analysis

Tissue culture

FLp-In T-REx-293 cells were grown at 37° C with 5% CO2 and passaged before 100% confluency. Cells were grown in DMEM (Gibco) with 10% FBS (Gibco) and penicillin/streptomycin (Gibco). Experiments were performed between the second and tenth passage after thawing from -80° C.

siRNA knockdown

Cells were plated at ~10% confluency for siRNA treatment in 6 well plates. siRNA knockdowns were performed with siLentFect (Bio-rad) according to the manufacturer's recommendations. Two sequential knockdowns were performed with 20 nM final concentration of siLuciferase siUsb1 #1 or siUsb1 #2 (Supplementary Table 1) at 72 hours and 24 hours prior to cell harvest. To tag nascently transcribed RNA transcripts, 5-Ethynyl Uridine from the Click-iT Nascent RNA Capture Kit (ThermoFisher) was added to media 8 hours before cell harvest for a final concentration of 200uM.

Total RNA isolation and nascent RNA capture

RNA was isolated using TRIzol (ThermoFisher) following the manufacturer's protocol. 10 ug of linear acrylamide (VWR) was used during RNA precipitation. To isolate nascent RNA, 1.5 ug of RNA was treated with the Click-iT Nascent RNA Capture Kit (ThermoFisher) following the manufacturer's protocol.

cDNA Synthesis and qPCR

250 ng of nascently-transcribed RNA was used to synthesize cDNA using SuperScript III First-Strand Synthesis System (ThermoFisher). qPCR was performed in 10ul total reaction on 96-well plate on a QuantStudio3 machine (ThermoFisher) using Fast SYBR Green Reagent (ThermoFisher) with 0.5uM concentrations of both forward and reverse primers for U1, Usb1, 7SL, 7SK, U6, GAPDH, RPS16, and aldolase (Supplementary Table 1). Ct values were normalized to the average Ct values of the housekeepers (GAPDH, RPS16, and aldolase) using the formula 2^(housekeeper Ct - experimental Ct),

then normalized to the control siRNA (siLuciferase) by dividing the siUsb1 condition by the control condition.

T4 Polynucleotide Kinase (PNK) treatment

500ng of total RNA from control siRNA and siUsb1 #2 conditions were used in the following PNK (BioLabs) treatment and mock PNK treatment. PNK and mock PNK master mixes were prepared with final concentrations of 0.002 U/uL TURBO DNase (ThermoFisher), 1mM RNase Out (ThermoFisher), 1 mM DTT (ThermoFisher), and 700 U/uL T4 PNK (BioLabs) or the equivalent volume in water for the mock PNK treatments. These were prepared in a buffer with final concentrations of 70mM Tris-HCl pH 6.5 (company), 10mM $MgCl_2$ (company), and 5mM DTT (ThermoFisher). The master mixes were added to the RNA for a total 100ul reaction volume. The samples were then incubated at 37C for 20 minutes, then purified with a phenol-chloroform treatment. PNK/mock PNK treated RNA were ligated with 3' adapters with randomers bases (Supplementary Table 1) using T4 RNA ligase 1 (NEB). RNA was again purified with phenol-chloroform, and cDNA was made using SuperScript III (ThermoFisher) with the AR17 primer (Supplementary Table 1).

qPCR of cDNA from PNK treatment

qPCR was performed with diluted cDNA in 10ul total reaction on 96-well plate on an QuantStudio3 machine (ThermoFisher) using the Fast SYBR Green Reagent (ThermoFisher) with 0.5uM concentrations of both forward and reverse primers for U1,

Usb1 7SL, 7SK, U6 (Supplementary Table 1). Ct values were normalized to the Ct values of U1 using the formula 2^(U1 Ct - experimental Ct). Then the formula 1-(mock PNK conditions/PNK conditions) was used to find the fraction of transcripts with a 3' end cyclic phosphate.

Library preparation for sequencing of cDNA from PNK treatment

To create libraries for sequencing, cDNA synthesized after PNK treatment was amplified via PCR using the Q5® High-Fidelity DNA Polymerase (NEB) with AR17 primer and barcoded gene-specific primers for U1, U6, 7SL, and 7SK. and then a second time with TruSeq CD indices (Illumina) (Supplementary Table 1). Libraries were cleaned using Ampure beads (Beckman Coulter).

Sequencing of sncRNA libraries

Sequencing was performed on a MiSeq system (Illumina) using a 300 cycle Reagent Kit V2. 200 cycles were dedicated to the 3' end of RNAs, and 100 cycles were dedicated to the 5'end. Data was recovered from the sequencer in the form of FASTQ files.

Analysis of 3' end data

This sequencing data was run through in-house python scripts in order to generate three datasets that aligned the sequences to a database of U1, U6, 7SL, and 7SK sequences obtained from ENSEMBL. These generated datasets contained the

aligned gene from the database, the sequence from the sequencer, the number of times that sequence aligned, the 3' end location and length, and the sequence of any post-transcriptional tail on the transcript. For each of the four genes, plots were generated using in-house scripts that graphically display the DNA sequence logo weighted by the number of each nucleotide present at each position (seqlogo plots), as well as plots that display the percent of transcripts that terminate at each nucleotide position from -5 bases and +10 bases relative to the 3' end (cumulative plots).

References

Andreazzoli M, Gerbi SA. 1991. Changes in 7SL RNA conformation during the signal recognition particle cycle. *EMBO J* **10**: 767–777.

Arimbasseri AG, Rijal K, Maraia RJ. 2013. Transcription termination by the eukaryotic RNA polymerase III. *Biochim Biophys Acta* **1829**: 318–330.

Baumgartner M, Olthof AM, Aquino GS, Hyatt KC, Lemoine C, Drake K, Sturrock N, Nguyen N, Al Seesi S, Kanadia RN. 2018. Minor spliceosome inactivation causes microcephaly, owing to cell cycle defects and death of self-amplifying radial glial cells. *Development* **145**: dev166322.

Berndt H, Harnisch C, Rammelt C, Stöhr N, Zirkel A, Dohm JC, Himmelbauer H, Tavanez J-P, Hüttelmaier S, Wahle E. 2012. Maturation of mammalian H/ACA box snoRNAs: PAPD5-dependent adenylation and PARN-dependent trimming. *RNA* **18**: 958–972.

Brogie JE, Price DH. 2017. Reconstitution of a functional 7SK snRNP. *Nucleic Acids Research* **45**: 6864–6880.

Bunch H. 2018. Gene regulation of mammalian long non-coding RNA. *Mol Genet Genomics* **293**: 1–15.

Chang DD, Clayton DA. 1987. A mammalian mitochondrial RNA processing activity contains nucleus-encoded RNA. *Science* **235**: 1178–1184.

Chen X, Wolin SL. 2004. The Ro 60 kDa autoantigen: insights into cellular function and role in autoimmunity. *J Mol Med (Berl)* **82**: 232–239.

Chung CZ, Jo DHS, Heinemann IU. 2016. Nucleotide specificity of the human terminal nucleotidyltransferase Gld2 (TUT2). *RNA* **22**: 1239–1249.

Core L, Adelman K. 2019. Promoter-proximal pausing of RNA polymerase II: a nexus of gene regulation. *Genes Dev*. http://genesdev.cshlp.org/content/early/2019/05/23/gad.325142.119 (Accessed August 5, 2021).

Costa FF. 2005. Non-coding RNAs: new players in eukaryotic biology. *Gene* **357**: 83–94.

Croce CM. 2009. Causes and consequences of microRNA dysregulation in cancer. *Nat Rev Genet* **10**: 704–714.

D'Ambrogio A, Gu W, Udagawa T, Mello CC, Richter JD. 2012. Specific miRNA stabilization by Gld2-catalyzed monoadenylation. *Cell Rep* **2**: 1537–1545.

Dergai O, Hernandez N. 2019. How to Recruit the Correct RNA Polymerase? Lessons from snRNA Genes. *Trends in Genetics* **35**: 457–469.

Didychuk AL, Butcher SE, Brow DA. 2018. The life of U6 small nuclear RNA, from cradle to grave. *RNA* **24**: 437–460.

Didychuk AL, Montemayor EJ, Carrocci TJ, DeLaitsch AT, Lucarelli SE, Westler WM, Brow DA, Hoskins AA, Butcher SE. 2017. Usb1 controls U6 snRNP assembly through evolutionarily divergent cyclic phosphodiesterase activities. *Nat Commun* **8**: 497.

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**: 15–21.

Eddy SR. 2001. Non-coding RNA genes and the modern RNA world. *Nat Rev Genet* **2**: 919–929.

Esakova O, Krasilnikov AS. 2010. Of proteins and RNA: The RNase P/MRP family. *RNA* **16**: 1725–1747.

Esteller M. 2011. Non-coding RNAs in human disease. *Nat Rev Genet* **12**: 861–874.

Guerrier-Takada C, Gardiner K, Marsh T, Pace N, Altman S. 1983. The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* **35**: 849–857.

Hilcenko C, Simpson PJ, Finch AJ, Bowler FR, Churcher MJ, Jin L, Packman LC, Shlien A, Campbell P, Kirwan M, Dokal I, Warren AJ. 2013. Aberrant 3' oligoadenylation of

spliceosomal U6 small nuclear RNA in poikiloderma with neutropenia. *Blood* **121**: 1028–1038.

Hojo H, Yashiro Y, Noda Y, Ogami K, Yamagishi R, Okada S, Hoshino S, Suzuki T. 2020. The RNA-binding protein QKI-7 recruits the poly(A) polymerase GLD-2 for 3′ adenylation and selective stabilization of microRNA-122. *J Biol Chem* **295**: 390–402.

Hombach S, Kretz M. 2016. Non-coding RNAs: Classification, Biology and Functioning. *Adv Exp Med Biol* **937**: 3–17.

Jarrous N. 2017. Roles of RNase P and Its Subunits. *Trends Genet* **33**: 594–603.

Jawdekar GW, Henry RW. 2008. Transcriptional regulation of human small nuclear RNA genes. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* **1779**: 295–305.

Ji X, Zhou Y, Pandit S, Huang J, Li H, Lin CY, Xiao R, Burge CB, Fu X-D. 2013. SR proteins collaborate with 7SK and promoter-associated nascent RNA to release paused polymerase. *Cell* **153**: 855–868.

Katoh T, Sakaguchi Y, Miyauchi K, Suzuki T, Kashiwabara S, Baba T, Suzuki T. 2009. Selective stabilization of mammalian microRNAs by 3′ adenylation mediated by the cytoplasmic poly(A) polymerase GLD-2. *Genes Dev* **23**: 433–438.

Kowalski MP, Krude T. 2015. Functional roles of non-coding Y RNAs. *Int J Biochem Cell Biol* **66**: 20–29.

Kwak JE, Wang L, Ballantyne S, Kimble J, Wickens M. 2004. Mammalian GLD-2 homologs are poly(A) polymerases. *Proc Natl Acad Sci U S A* **101**: 4407–4412.

Łabno A, Warkocki Z, Kuliński T, Krawczyk PS, Bijata K, Tomecki R, Dziembowski A. 2016. Perlman syndrome nuclease DIS3L2 controls cytoplasmic non-coding RNAs and provides surveillance pathway for maturing snRNAs. *Nucleic Acids Research* **44**: 10437–10453.

LaCava J, Houseley J, Saveanu C, Petfalski E, Thompson E, Jacquier A, Tollervey D. 2005. RNA degradation by the exosome is promoted by a nuclear polyadenylation complex. *Cell* **121**: 713–724.

Lubas M, Christensen MS, Kristiansen MS, Domanski M, Falkenby LG, Lykke-Andersen S, Andersen JS, Dziembowski A, Jensen TH. 2011. Interaction profiling identifies the human nuclear exosome targeting complex. *Mol Cell* **43**: 624–637.

Martin G, Doublié S, Keller W. 2008. Determinants of substrate specificity in RNA-dependent nucleotidyl transferases. *Biochim Biophys Acta* **1779**: 206–216.

Matera AG, Wang Z. 2014. A day in the life of the spliceosome. *Nat Rev Mol Cell Biol* **15**: 108–121.

Mattick JS, Makunin IV. 2006. Non-coding RNA. *Human Molecular Genetics* **15**: R17–R29.

Mattijssen S, Welting TJM, Pruijn GJM. 2010. RNase MRP and disease. *WIREs RNA* **1**: 102–116.

Menezes MR, Balzeau J, Hagan JP. 2018. 3′ RNA Uridylation in Epitranscriptomics, Gene Regulation, and Disease. *Front Mol Biosci* **0**. https://www.frontiersin.org/articles/10.3389/fmolb.2018.00061/full (Accessed August 5, 2021).

Meyer S, Temme C, Wahle E. 2004. Messenger RNA turnover in eukaryotes: pathways and enzymes. *Crit Rev Biochem Mol Biol* **39**: 197–216.

Mroczek S, Dziembowski A. 2013. U6 RNA biogenesis and disease association. *Wiley Interdiscip Rev RNA* **4**: 581–592.

Mroczek S, Krwawicz J, Kutner J, Lazniewski M, Kuciński I, Ginalski K, Dziembowski A. 2012. C16orf57, a gene mutated in poikiloderma with neutropenia, encodes a putative phosphodiesterase responsible for the U6 snRNA 3′ end modification. *Genes Dev* **26**: 1911–1925.

Munoz-Tello P, Rajappa L, Coquille S, Thore S. 2015. Polyuridylation in Eukaryotes: A 3'-End Modification Regulating RNA Life. *Biomed Res Int* **2015**: 968127.

Nomura Y, Montemayor EJ, Virta JM, Hayes SM, Butcher SE. 2019. Structural basis for the evolution of cyclic phosphodiesterase activity in the U6 snRNA exoribonuclease Usb1. *Nucleic Acids Research* **48**. https://www.osti.gov/pages/biblio/1603143 (Accessed August 5, 2021).

Nomura Y, Roston D, Montemayor EJ, Cui Q, Butcher SE. 2018. Structural and mechanistic basis for preferential deadenylation of U6 snRNA by Usb1. *Nucleic Acids Research* **46**: 11488–11501.

Ogami K, Cho R, Hoshino S. 2013. Molecular cloning and characterization of a novel isoform of the non-canonical poly(A) polymerase PAPD7. *Biochem Biophys Res Commun* **432**: 135–140.

Park J-L, Lee Y-S, Kunkeaw N, Kim S-Y, Kim I-H, Lee YS. 2017. Epigenetic regulation of noncoding RNA transcription by mammalian RNA polymerase III. *Epigenomics* **9**: 171–187.

Pessa HKJ, Will CL, Meng X, Schneider C, Watkins NJ, Perälä N, Nymark M, Turunen JJ, Lührmann R, Frilander MJ. 2008. Minor spliceosome components are predominantly localized in the nucleus. *Proc Natl Acad Sci U S A* **105**: 8655–8660.

Rammelt C, Bilen B, Zavolan M, Keller W. 2011. PAPD5, a noncanonical poly(A) polymerase with an unusual RNA-binding motif. *RNA* **17**: 1737–1746.

Reiner R, Ben-Asouli Y, Krilovetzky I, Jarrous N. 2006. A role for the catalytic ribonucleoprotein RNase P in RNA polymerase III transcription. *Genes Dev* **20**:1621–1635.

Schmidt M-J, Norbury CJ. 2010. Polyadenylation and beyond: emerging roles for noncanonical poly(A) polymerases. *Wiley Interdiscip Rev RNA* **1**: 142–151.

Sinha KM, Gu J, Chen Y, Reddy R. 1998. Adenylation of Small RNAs in Human Cells: DEVELOPMENT OF A CELL-FREE SYSTEM FOR ACCURATE ADENYLATION ON THE 3′-END OF HUMAN SIGNAL RECOGNITION PARTICLE RNA *. *Journal of Biological Chemistry* **273**: 6853–6859.

Song J, Song J, Mo B, Chen X. 2015. Uridylation and adenylation of RNAs. *Sci China Life Sci* **58**: 1057–1066.

Täuber H, Hüttelmaier S, Köhn M. 2019. POLIII-derived non-coding RNAs acting as scaffolds and decoys. *J Mol Cell Biol* **11**: 880–885.

Turunen JJ, Niemelä EH, Verma B, Frilander MJ. 2013. The significant other: splicing by the minor spliceosome. *WIREs RNA* **4**: 61–76.

Ullu E, Weiner AM. 1984. Human genes and pseudogenes for the 7SL RNA component of signal recognition particle. *EMBO J* **3**: 3303–3310.

Walter P, Blobel G. 1982. Signal recognition particle contains a 7S RNA essential for protein translocation across the endoplasmic reticulum. *Nature* **299**: 691–698.

Wang L, Eckmann CR, Kadyk LC, Wickens M, Kimble J. 2002. A regulatory cytoplasmic poly(A) polymerase in Caenorhabditis elegans. *Nature* **419**: 312–316.

Willis IM, Moir RD. 2018. Signaling to and from the RNA Polymerase III Transcription and Processing Machinery. *Annual Review of Biochemistry* **87**: 75–100.

Wlotzka W, Kudla G, Granneman S, Tollervey D. 2011. The nuclear RNA polymerase II surveillance system targets polymerase III transcripts. *EMBO J* **30**: 1790–1803.

Zhao Y, Ransom JF, Li A, Vedantham V, von Drehle M, Muth AN, Tsuchihashi T, McManus MT, Schwartz RJ, Srivastava D. 2007. Dysregulation of cardiogenesis, cardiac conduction, and cell cycle in mice lacking miRNA-1-2. *Cell* **129**: 303–317.