Effect of disrupting the mRNA turnover on the genome-wide RNA 3'-end structure

Jurković, Borna

UNIVERSITY OF RIJEKA
DEPARTMENT OF BIOTECHNOLOGY
Graduate university programme
Biotechnology in Medicine

Borna Jurković

Effect of disrupting the mRNA turnover on the genome-wide RNA 3'-end structure

Master´s thesis

Rijeka, 2023

UNIVERSITY OF RIJEKA
DEPARTMENT OF BIOTECHNOLOGY
Graduate university programme
Biotechnology in Medicine

Borna Jurković

Effect of disrupting the mRNA turnover on the genome-wide RNA 3'-end structure

Master´s thesis

Rijeka, 2023

Mentor: Michał Małecki, PhD
Co-mentor: Christian A. Reynolds, PhD

SVEUČILIŠTE U RIJECI
ODJEL ZA BIOTEHNOLOGIJU
Diplomski sveučilišni studij
Biotehnologija u medicini

Borna Jurković

Utjecaj poremećaja u raspadu mRNA na strukturu RNA 3´kraja na razini genoma

Diplomski rad

Rijeka, 2023.

Mentor: dr.sc. Michał Małecki
Co-mentor: dr. sc. Christian A. Reynolds

Acknowledgements

The experimental part of this thesis was done at the Institute of Genetics and Biotechnology, University of Warsaw, Warsaw, Poland as a part of ERASMUS+ mobility.

I would like to give my warmest thanks to my mentor dr. sc. Michał Małecki. I am extremely grateful for the support and guidance provided by you for conducting and writing the thesis.

I would also like to thank dr. sc. Lidia Lipińska-Zubrycka, dr.sc. Konrad Kosicki and dr.sc. Karolina Łabędzka-Dmoch for welcoming me to the institute and being really kind both profesionally and personally.

I would like to extend my appreciation to the PhD student Maciej Grochowski for showing me around in the lab. No matter how many stupid questions I had, you always answered them without judgement. I enjoyed our discussions and working with you.

This thesis has 52 pages, 8 tables, 15 figures and 64 citations.

ABSTRACT

Uridylation is a common widespread 3'-end modification in eukaryotes; however, its impact on mRNA fate is not clear. The mRNA decay mechanism in *S. pombe* is quite similar to that in humans, where the first step usually involves uridylation of the 3´-end of mRNA by terminal uridyltransferases (TUTases). The aim of this project was to investigate the effects on mRNA decay and mRNA tails when gene expression levels of the factors involved in mRNA decay are manipulated. We decided to delete the enhancer of mRNA decapping gene, *edc1*, and the exonuclease gene *xrn1*, while overexpressing the terminal uridyltransferases genes, *cid1* and *cid16*. The transformations were based on the cellular intake of a linear cassette and homologous recombination of the same cassette. The transformed strains were verified using colony-PCR, and overexpression strains were additionally verified using western blotting. Subsequently, mRNA was isolated and purified, and NGS libraries were prepared using a custom GW-3´RACE (TAIL-seq) protocol. The samples were then sent for paired-end Illumina sequencing. The sequencing data was analyzed using bioinformatic tools, and included read trimming, filtering and genome alignment. Although the libraries for strains overexpressing cid1 and cid16 did not work, the cid16 expression dependent strain showed a growth defect. Other main findings were as follows: deletion of *edc1* or *xrn1* resulted in the shortening of mRNA tail length, increased mRNA uridylation, and accumulation of non-tailed mRNAs. In other words, the deletion of *edc1* or *xrn1* produced the same results, namely the accumulation of mRNA decay intermediaries. Additionally, the differential gene expression analysis revealed six commonly up-regulated genes between the mutants, with the most up-regulated gene being *ubi4*, which encodes for ubiquitin.

Keywords: mRNA decay, urydilation, *S. pombe*

SAŽETAK

Uridilacija je česta modifikacija 3'-kraja u eukariotima; međutim, njen utjecaj na sudbinu mRNA nije jasan. Mehanizam raspada mRNA kod *S. pombe* prilično je sličan onom kod ljudi, gdje prvi korak obično uključuje uridilaciju 3´-kraja mRNA terminalnim uridiltranferazama (TUTazama). Cilj ovog projekta bio je istražiti učinke na raspad mRNA i repove mRNA kada se manipuliraju ekspresije gena faktora uključenih u raspad mRNA. Odlučili smo izbrisati gen *edc1*, koji je pojačavač dekapiranja mRNA, i gen *xrn1*, koji kodira egzonukleazu, dok smo prekomjerno izražavali gene terminalnih uridiltranferaza, *cid1* i *cid16*. Transformacije su se temeljile na staničnom unosu linearnog kasetnog elementa i homologne rekombinacije iste kasete. Transformirani sojevi su potvrđeni korištenjem colony-PCR-a, a sojevi s prekomjernim izražavanjem su dodatno potvrđeni korištenjem western blot analize. Nakon toga je mRNA izolirana i pročišćena, a NGS libraries su pripremljene korištenjem prilagođenog GW-3´RACE (TAIL-seq) protokola. Uzorci su zatim poslani na paired-end sekvenciranje Illumina tehnologijom. Podaci sekvenciranja su analizirani korištenjem bioinformatičkih alata, uključujući podrezivanje očitanja, filtriranje i poravnanje s referentnim genomom. Iako biblioteke za sojeve s prekomjernim izražavanjem *cid1* i *cid16* nisu funkcionirale, soj s prekomjernim izražavanjem *cid16* je pokazao defekt rasta. Ostali glavni rezultati su sljedeći: brisanje *edc1* ili *xrn1* rezultiralo je skraćivanjem duljine repova mRNA, povećanom uridilacijom mRNA i nakupljanjem mRNA bez repova. Drugim riječima, brisanje *edc1* ili *xrn1* proizvelo je iste rezultate, odnosno nakupljanje intermediarnih produkata mRNA raspada. Također, analiza diferencijalnog izraza gena je otkrila šest zajednički up-reguliranih gena između mutanata, pri čemu je najviše up-regulirani gen bio *ubi4*, koji kodira ubiquitin.

Ključne riječi: raspad mRNA, uridilacija, *S. pombe*

Table of Contents

1)INTRODUCTION

1.1. Cytoplasmic mRNA degradation

To meet cellular needs during the cell cycle and respond to environmental changes, RNA transcription and degradation are regulated through several mechanisms that modify the transcriptome (1). mRNA degradation is the process by which messenger RNA (mRNA) molecules are broken down and eliminated from the cell. Alongside transcription, precursor mRNA processing, and mRNA transport mechanisms, this process is crucial for maintaining the correct level of transcripts within a cell and preventing the accumulation of unwanted or damaged mRNA molecules.

In eukaryotic cells, mRNA degradation is a complex process and can occur through several pathways, the most prominent being 5' to 3' and 3' to 5' exonucleolytic decay. To regulate the activity of these pathways, eukaryotic mRNAs require both the 5' 7-methylguanosine cap and the 3' poly(A) tail structures for stability. The decay process is triggered by the loss of either structure, or alternatively, it can be initiated by internal endonucleolytic cleavage. The major events in mRNA decay are therefore shortening or removal of the poly(A) tail (deadenylation) and the hydrolysis of the 5' mG-cap (decapping), followed by ribonucleolytic digestion.

Deadenylation is often the rate-limiting step in mRNA decay and is catalyzed by deadenylases. In eukaryotes, there are two conserved deadenylase complexes: Ccr4-Not and Pan2-Pan3 (2,3). Interestingly, Pan2-Pan3's activity is dependent on poly(A)-binding protein (PABP), while CCR4-NOT activity is inhibited by PABP (4).

Both of these deadenylases work together, and it has been suggested that the Pan2-Pan3 complex removes adenosines until the poly(A) tail is 12-25 nucleotides long, after which Ccr4-Not removes the remaining tail in a process referred to as terminal deadenylation (5,6). Decapping is the second important event in mRNA decay and is neccessary if mRNA is to be degraded in 5´→ 3´direction. It can follow deadenylation but deadenylation dependent decapping was also demonstrated. In *S. pombe* decapping is

mainly done by the decapping complex Dcp1-Dcp2 and is facilitated by the Lsm1-7 complex. The Lsm1-7 complex is a complex of seven Sm-like proteins which binds to the 3' UTR of target mRNAs to recruit decapping complex Dcp1-Dcp2 (7). Other important decapping factors which regulate decapping activity are enhancers of decapping which work by boosting the removal rate of the 5′7-mG cap (8). There is evidence decapping can be inhibited by the translation initiation factors (9). The last step in every mRNA decay pathway is the ribonucleolytic digestion. In 5′→ 3′ decay pathways, digestion is entirely done by the exonuclease Xrn1. On the other hand, in 3′→ 5′decay pathway, RNA can be digested by the exosome. There is also another 3′→ 5′decay pathway discovered which is done by the exoribonuclease Dis3L2. In *S. pombe*, Dis3L2 has a preferece for degrading urydilated RNAs while functioning independently of the exosome (10). In humans, mutations in Dis3L2 are linked with the Perlman syndrome of fetal overgrowth and a higher supcetibility to a wide range of cancers (11,12).

mRNA decay pathways can either be: 1) deadenylation-dependent, 2) deadenylation-independent or 3) endonucleases-dependent. The deadenylation-dependent pathway is the most dominant one in eukaryotes and starts with the deadenylation. Following deadenylation, mRNA can further be degraded in one of two possible directions; either 5′→ 3′ which involves decapping and digestion by the Xrn1 exoribonuclease or 3′ → 5′ decay by the exosome.

On the other hand, some mRNAs are degraded via the deadenylation-independent pathway. In this pathway, mRNA is not deadenylated; rather, it is decapped and degraded in the 5' → 3' direction (13).

The degradation process of mRNA through endonuclease-mediated mRNA decay begins by the internal cleavage, creating two fragments both containing a single unprotected end. These fragments are then broken down by Xrn1 and the exosome. This pathway is found to be common in *D. melanogaste*r but no so much in other eukaryotes (13,14).

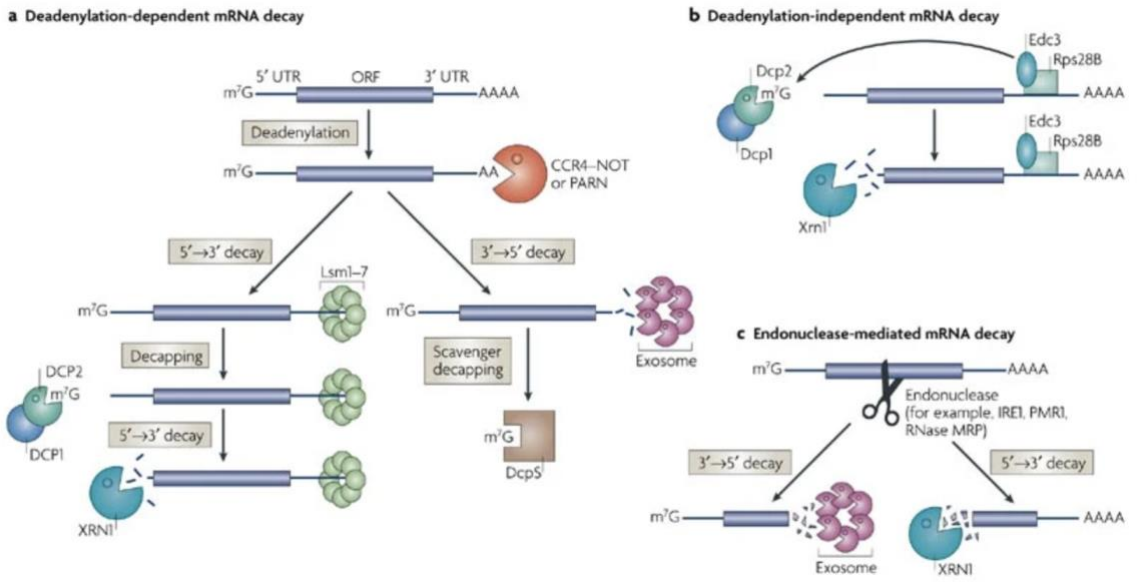

Figure 1. **A graphical representation of mRNA decay pathways.** Taken from Garneau et al. 2007

1.2. Uridylation in mRNA decay

Uridylation is a posttranscriptional modification that targets all categories of eukaryotic RNA. The uridylation of polyadenylated mRNAs is conserved across eukaryotes, with the exception of *S. cerevisiae* (15,16). In mammals, more than 85% of mRNAs undergo terminal uridylation at a frequency higher than 1%, as demonstrated in various cell lines such as NIH 3T3 and HeLa cells, indicating the widespread nature of this modification (15).

While the consequences of RNA uridylation depend on the RNA type and species, in *S. pombe*, 3' RNA uridylation can trigger both 5' to 3' and 3' to 5' RNA decay, specifically sRNA or mRNA decay (17).

Uridylation primarily targets completely or partially deadenylated mRNAs, but it has also been shown to mark mRNAs with long poly(A) tails.

In humans, it has been suggested that oligo-U tails are more effective in promoting decay compared to mono-U tails. However, in *S. pombe*, it is still unclear whether the number of added uridines plays a role (18–20).

In the uridylation-dependent 5' to 3' pathway, the LSm1-7 complex recognizes and binds to the uridylated tail of the mRNA, promoting decapping by the Dcp1/2 enzymes. Removal of the 5' protective cap renders the mRNA vulnerable to digestion by the 5' to 3' exonuclease Xrn1 (21).

Interestingly, the Lsm1-7 complex is also capable of binding to oligo(U) tails of mRNA, and it has been demonstrated that the addition of a uridine residue at the 3' end of the RNA promotes decapping when compared to RNA without a terminal uridine residue (22).

Deadenylated and uridylated mRNAs can also be degraded through the 3' to 5' pathway, which involves digestion by the uridylation-independent exosome or the uridylation-dependent exonuclease Dis3L2. Dis3L2 exhibits a strong preference for oligouridylated RNA and it has been shown adenylated RNAs completely inhibit DIS3L2 (10).

1.3. Selected factors involved in mRNA decay in S. pombe

1.3.1. Terminal uridyltransferases

The enzymes responsible for uridylating RNA are called terminal uridyltransferases (TUTases), and they catalyze the untemplated addition of uridines at the 3' termini of RNAs. TUTases belong to subgroup 2 of the DNA polymerase β-like nucleotidyltransferase superfamily, and *in S. pombe*, there are two known TUTases: Cid1 and Cid16. The core catalytic domain of TUTases consists of a Pol β-like nucleotidyltransferase motif and a poly(A) polymerase-associated motif, which has evolved to bind uridine 5'-triphosphate (UTP) rather than adenosine 5'-triphosphate (ATP) (17). Among the two *S. pombe* TUTases, Cid1 is primarily responsible for uridylation of mRNA. Deletion of the cid1 gene has been shown to result in a significant reduction (61.4%) in mRNA uridylation, while the deletion of Cid16 does not significantly impact the uridylation rate. Interestingly, a single or double TUTase deletion doesn´t have a significant impact on *S. pombe* growth under standard conditions (23). The UTP selectivity in Cid1 is attributed to conserved aspartate and glutamate residues, as well as a

crucial histidine residue (His336). Substituting His336 with another amino acid renders Cid1 unable to add UTP, instead adding ATP (24,25). Cid1 was initially identified in a study of cell cycle checkpoint control. When a combination of hydroxyurea (a DNA synthesis inhibitor) and caffeine (an inhibitor of a protein kinase involved in S-phase checkpoint control) is present, it can lead to S-phase arrest in *S. pombe* cells. Deletion of Cid1 in this study was shown to increase sensitivity to the combination of HU and caffeine, leading to the name Caffeine-induced death suppressor 1 (26). Although Cid16 is not the major factor in mRNA uridylation, it has been shown to be the primary uridyl-transferase for uridylation of Argonaute-bound sRNAs. Deletion of the cid16 gene results in the absence of 3' end uridylation of sRNAs, suggesting that Cid16 is the sole sRNA uridyl-transferase (27).

1.3.2. EDC1

The activity of the Dcp1-Dcp2 mRNA-decapping complex is heavily regulated by additional decapping factors. In *S. pombe*, these factors include Edc1, Edc3, and Pdc1 (known as edc4 in humans).

Edc1, also known as Dbs1 or enhancer of decapping 1, is a 33.35 kDa protein that functions by binding to Dcp1-Dcp2 and activating the decapping complex (8,28). The Edc1 LXPX domain interacts with the Dcp1 EVH1 domain, and the Edc1 YAG motif strengthens the binding between the RNA and the decapping complex. This Edc1 region is found across different *Schizosaccharomyces* species (29). Moreover, the binding of Edc1 alone does not induce structural conformational changes in the Dcp1-Dcp2 complex, but the binding of mRNA to the Edc1-Dcp1-Dcp2 complex does (30). In *S. pombe*, Δ*edc1* cells are significantly longer (19.9 μm vs 14.5 μm) compared to the wild type (31).

Compared to *S. pombe*, there is much more research evidence about the edc family of proteins in *S. cerevisiae*. In *S. cerevisiae*, the decapping factors are Edc1, Edc2, Edc3, and Pby1, which interact with the Dcp1-Dcp2

complex. In the Δ*DCP x SKI* deletion strain, overexpression of *EDC1* and *EDC2* suppresses the mRNA decapping defects. Furthermore, Edc1 is shown not to be the rate-limiting factor for mRNA decay, and the Δ*EDC1* and Δ*EDC1* x *EDC2* strains do not show any growth defects under normal conditions (32). On the other hand, Δ*EDC1*, Δ*EDC2*, or Δ*EDC1* x *EDC2* show impaired growth under stress conditions, specifically during heat stress. Heat stress has the greatest effect on the growth of the Δ*EDC1* x *EDC2* double mutant. There is also some evidence suggesting that Edc1 and Edc2 might play a role in protein translation during heat stress (33).

The deletion of *EDC1* and *EDC2* affects the rate of mRNA decapping. In the presence of only Edc1, the decapping activity was increased 140-fold compared to the double deletion mutant (28). Edc1 and Edc2 enhance decapping by promoting a tighter enzyme-substrate complex, possibly by providing an additional RNA binding surface (34).

1.3.3. XRN1

Xrn1, also known as Exo2, is a 5' to 3' exonuclease located in the cytoplasm that degrades decapped mRNAs. This 152.59 kDa protein belongs to a large family of conserved exonucleases (35,36). It has been shown that Xrn1 interacts with DCP1/DCP2, decapping enhancers, and LSM1-7 proteins, and that Xrn1's proline-rich motif binds to the aromatic groove of Dcp1 (29,37,38). Xrn1 prefers degrading substrates containing a 5'-monophosphate end compared to those with a 5'-triphosphate end, and its activity is blocked by the cap structure of the mRNA (39,40). Xrn1 also degrades a special class of lncRNA called XRN1-sensitive Unstable Transcripts (XUTs) (41).

In S. cerevisiae, the Δ*XRN1* strain had a 1.5 to 1.8-fold larger average cell volume, and its generation time was 2.0-fold longer compared to the wild type (42,43). Moreover, the Δ*XRN1* strain increases the frequency of autophagosome formation, and in mammalian cells, this deletion enhances autophagy and virus infection. Both of these pieces of evidence show that Xrn1 plays a role in autophagy (44). RNA of some viruses (such as dengue

virus and zika virus) is resistant to Xrn1 degradation due to their folded structure (45,46).

There is also some evidence showing that Xrn1 acts as a transcription factor. Xrn1, along with other decay factors (dcp2, lsm1), binds approximately 30 nucleotides upstream of transcription start sites. It has also been shown that these three decay factors tend to bind to gene promoters that code for mRNAs with a high turnover rate (47).

In a recent study, *XRN1* was identified as a potential therapeutic target in cancer research. Silencing *XRN1* alone did not suppress tumor growth, but it potentiated the efficacy of cancer immunotherapy and suppressed tumor growth in mice. RNA accumulation induced by *XRN1* silencing triggered the IFN signaling pathway, which stimulates an immune response (48).

1.4. Methods used for profiling of mRNA tails

1.4.1. The beginnings of mRNA tails profiling

In order to better study mRNA tails and its poly(A) tail length, methods of profiling mRNA tails were developed. All of the methods that first emerged were focused to determine poly(A) length. The first method emerged in 1974 and compares on gel motility of mRNA with to mRNA without poly(A) tail. The name of this method is RNAse H/oligo(dT) assay and it uses oligo(dT) nucleotides to bind to the complementary poly(A) tail. Two samples of the same mRNA are used and only in one sample oligo(dT) nucleotides are added.

Afterwards, RNAse H is added and it specifically cleaves RNA in DNA-RNA hybrids. Since in the second samples there is no DNA-RNA hybrids due to no oligo(dT) nucleotides added, there is also no ribonucleolytic digestion.

The both samples are then analysed using Northern Blot and the difference in electrophoretic motility is used to calculate the poly(A) tail length (49).

The next technique used was Rapid Amplification of cDNA Ends-poly(A) Test (RACE-PAT) assay. This is a type of RACE assay that rapidly amplifies reversely transcribed cDNA with a poly(A) tail and determines mRNAs

poly(A) tail length. In this assay, mRNAs are mixed with oligo(dT)-sequence linked to a GC rich anchor (oligo[dT]-anchor) which later serves as a primer for reverse transcription. Hibridization of oligo(dT)-anchor and mRNA can take place at any position throughout the entire length of the tail, therefore the end result after reverse transcription is a heterogeneous pool of cDNA primed at all possible positions along the poly(A) tail. Next step is amplification using PCR and interestingly, the role of the anchor sequence on the oligo(dT) is to prevent the progressive shortening of the poly(A) tail during PCR. The PCR products are then analysed using gel electrophoresis (50,51).

The ligase-mediated-poly(A) Test (LM-PAT) is a modification to the standard RACE-PAT assay and it appears to be more sensitive in detecting poly(A) tail length. The only difference from RACE-PAT assay is that in LM-PAT isolated RNA is first mixed with phosphorylated oligo(dT) nucleotides and T4 DNA ligase. When this nucleotides reach the 3´end of the poly(A) tail there is a small poly(A) region (<10 nucleotides) which remains unpaired. Only then oligo(dT)-sequence linked to a GC rich anchor (oligo[dT]-anchor) is added. This method enables specific binding oligo(dT)-anchor to the 3`termini of the poly(A) tail (51,52).

These two are not the only two PAT methods; there are similar PAT methods such as ePAT, sPAT, HIRE-PAT, each having its own modifications.

Every method has its own drawbacks and so do the methods explained above. The information they provided was limited to individual transcripts and did not specify the length of the mRNA poly(A) tail.

Before NGS there were some approaches at studying transcript termini at a transcriptome-wide scale using microarray analysis. Although these methods are better at then PAT assays, they still fail to determine precise transcript poly(A) length. With the development of NGS and ultimately RNA-seq a major breakthrough was made in mRNA tail profiling. RNA-Seq is better at detecting transcripts with low abudance, has a broader dynamic range when compared to microarrays. Additionally, with RNA-Seq the

drawbacks of microarrays are bypassed such as cross-hybridization and non-specific hybridization (53).

1.4.2. mRNA tail profiling using sequencing technologies

1.4.2.1. TAIL-seq

TAIL-seq, alongside PAL-seq, is one of the first methods to precisely measure mRNA poly(A). It is a type of RNA-seq method that involves the depletion of noncoding RNA, ligation of a 3' adaptor prior to RNA fragmentation, partial digestion of RNAs while preserving poly(A) tails, incorporation of a biotin residue in the 3' adaptor for easier purification, and the use of 15 degenerate sequences to improve sequencing performance. TAIL-seq employs paired-end sequencing with Illumina Seq and utilizes a specialized algorithm for detecting signals from long T stretches. This approach overcomes the challenge of low sequencing quality in NGS sequencing with long homopolymeric stretches by detecting the transition from long homopolymeric stretches (cDNA poly[T], complementary to mRNA poly[A]) to the 3' UTR. Compared to previously mentioned methods, TAIL-seq is more accurate, as demonstrated by sequencing synthetic poly(A) spike-ins, which showed high accuracy. When compared to the high-resolution poly(A) tail (Hire-PAT) assay, TAIL-seq exhibited a highly similar pattern for five spike-ins and ten tested mRNAs.
The use of TAIL-seq has led to several discoveries, including finding that poly(A) tails are approximately 100 nucleotides shorter than previously believed, the widespread occurrence of uridylation downstream of the poly(A) tail, and the presence of guanylation downstream of longer (>40 nt) poly(A) tails (15).
Two years after the introduction of TAIL-seq, a modified version called mRNA TAIL-Seq (mTAIL-Seq) was developed. Compared to its predecessor, mTAIL-Seq offers approximately 1000-fold enhanced sequencing depth for mRNAs and detects an average of approximately 5-fold more genes per million reads. Additionally, mTAIL-Seq requires approximately 100-fold less RNA for library preparation. Instead of the ligation of a linear 3' adaptor,

mTAIL-Seq uses splint-ligation with a hairpin 3' adaptor, enabling specific capture of poly(A) tails and eliminating the rRNA depletion step. While mTAIL-Seq seems superior to classical TAIL-seq in various aspects, it has limitations in capturing poly(A) tails shorter than 8 nucleotides and those with 3' modifications. Even with the use of special hairpin adapters containing one or two adenosines at the overhang, uridylated tails were still detected at a lower frequency (54).

1.4.2.2. Third-generation sequencing

There are several other methods for studying transcriptome termini based on Illumina sequencing, including TED-Seq, Poly(A)-Seq, EnD-Seq, 3'RACE-Seq, and circTAIL-Seq. In recent years, the use of third-generation sequencing (TGS) has been implemented in transcriptome profiling. TGS has the ability to overcome the drawbacks of NGS, such as short sequencing read length, better analysis of homopolymeric stretches, and providing information beyond the 3' UTR of mRNA.

One TGS method, called Full-length poly(A) and mRNA sequencing (FLAM-Seq), utilizes PacBio sequencing and has demonstrated the production of full-length mRNA sequences with high accuracy (error rate lower than ~0.8%). Similar to TAIL-Seq, FLAM-Seq can effectively detect non-A nucleotides in mRNA tails and identify those nucleotides (55).

Another PacBio-based method is Poly(A)-inclusive RNA isoform sequencing (PAIso-seq). It shares the advantages of FLAM-Seq but faces a similar limitation in precisely measuring the frequency of non-A modifications. A common limitation of all PacBio methods is the high cost required to achieve high coverage. PacBio is not the only TGS technology used to determine poly(A) tail length. A method developed by Geralde et al. utilizes Nanopore sequencing with minimal library preparation. It avoids cDNA synthesis and PCR amplification, thereby eliminating RT and PCR bias, making it superior in this aspect to other sequencing methods. The library preparation steps include splint ligation with a 3' adapter, optional RT, and ligation with a

sequencing adapter. However, this method specifically captures poly(A) mRNAs only. Additionally, it has a higher error rate compared to NGS methods. To address this, RT can be performed to create an RNA-cDNA hybrid, which may reduce intramolecular secondary RNA structure (56). Interestingly, the length of the poly(A) tail cannot be directly determined through sequencing with this method, and the use of a bioinformatic pipeline such as Nanopolish-polyA or TailFindR is required (57,58).

2) AIMS OF THE THESIS

In this project we aimed to understand the impact certain factors such as Edc1, Xrn1, Cid1 and Cid16 have on mRNA uridylation and on mRNA decay. We hypothesized the deletions of *edc1* and *xrn1* will slow the rate of mRNA decay and the overexpressions of *cid1* and *cid16* will increase the mRNA uridylation rate and therefore the rate of mRNA decay.

We aimed to:

1) Create *edc1* and *xrn1* deletion strains, and *cid1* and *cid16* overexpression strains and verify them
2) Cultivate newly created strains
3) Isolate RNA of newly created strains
4) Prepare NGS libraries
5) Analyse NGS data

3) MATERIALS AND METHODS

3.1. Yeast strains and growth conditions

S. pombe strains were grown in Yeast Supplemented Media (YES) (5 g/l of BioShop yeast extract, 30 g/l of glucose, 225 mg/l of adenine, histidine, leucine, uracil, lysine) to an optical density of OD_{600} = 0.5. To induce the overexpression of certain genes in transformed strains, cells were grown in Formedium Edinburgh Minimal Media broth (EMM) (3.11 g/l potassium hydrogen phthalate, 0.28 g/l di-sodium hydrogen phosphate, 5.18 g/l ammonium chloride, 20.44 g/l glucose, 1.21 g/l di-potassium hydrogen phosphate, 0.81 g/l sodium chloride, 0.50 g/l magnesium chloride, 0.40 g/l magnesium sulphate, 16 mg/l calcium chloride, 42 mg/l di-sodium sulphate, 3.23 g/l EMM broth premix). Cells were always incubated and shaked at 30°C, 150 RPM and the OD_{600} was measured using Biochrom Mltrospec 10 Cell Density Meter. Two biological replicates for each strain were grown.

3.2. Growth assay

Exponentially growing cultures of cells were grown at 30°C in 96-well plates in 100 µl volume of starting OD_{600} of 0.10. Optical density was measured every 15 minutes for 48 hours using micro-bioreactor Bioscreen C. Afterwards, the growth curves were visualised using a custom shiny app.

3.3. Plasmid quantification

To quantify the plasmid pFa6a-kanMX6-P3nmt1-3HA competent *E. Coli* strain *DH5*α was transformed. Bacteria were first thawed on ice for 15 minutes. 20 ng of plasmid was added and bacteria were incubated on ice for 30 minutes. Afterwards they were heatshocked for 90 seconds at 42°C. 1 ml of SOD was added and bacteria were incubated on 37°C for 1 hour. Bacteria were plated on 50 ug/ml ampicillin LB plates. The following day bacteria were inoculated to liquid LB media, the day after plasmids were

isolated using ThermoScientific GeneJet Plasmid MiniPrep Kit and the concentrations were measured using Nanodrop 2000.

3.4. PCR

All the primers (for PCR and colony-PCR) were designed using http://bahlerweb.cs.ucl.ac.uk/cgi-bin/PPPP/pppp_n_term.pl.

The mix consisted of 25 µl of 2X ThermoFisher Phusion Mastermix, 1 µl of 10 uM forward primer, 1 µl of 10 uM reverse primer, 1 µl of 20 ng/µl plasmide (for the gene deletions pFA6a-hphMX6 was used and for the gene overexpression pFa6a-kanMX6-P3nmt1-3HA was used) and 22 µl of Milli-Q H_2O.

The following program was run:

1) 98°C for 00:02:00
2) 98°C for 00:00:30
3) 46°C for 00:00:30
4) 72°C for 00:02:30
5) Goto step 2 for 5 times
6) 98°C for 00:00:30
7) 52°C for 00:00:30
8) 72°C for 00:02:30
9) Goto step 6 for 31 times
10) 72°C for 00:07:00
11) 4°C forever

Table 1. **Primer sequences for gene amplification using plasmid as a template.** Table shows which primers were used for each gene for adequate plasmid needed for amplifying DNA fragment later used in transformation.

Plasmid	Gene	Forward primer/reverse primer
pFA6a-hphMX6	Δ*edc1*	5′-TATCCTTAACCAACATCAATCACTAAGAATACAGAATCCATGATCATTAGGGATTACTCGAATTTTGACTATAGTAAAACGGATCCCCGGGTTAATTAA- 3′ / 5′-AAAAAAGTGGAATAGTTAGAACAGAACACGAGGCTTAAAAAGGTGATATGTCCTTAGACATAAATAAGCATCCAATGAAAGAATTCGAGCTCGTTTAAAC- 3′
	Δ*xrn1*	5′-TGTTGACATTTGTTCTCTACTAACAACAATACTTTCTAAATTTATTCCTATTTAACTGTTAATGAGAAAGTTCGTTCAGTCGGATCCCCGGGTTAATTAA- 3′ / 5′-TGTTGTCTTAAAACCTTCTATAAAGTTAAAGAATTCTATGCTAATAAAATTTATTATCACACCACTAAAATACAATTGGGGAATTCGAGCTCGTTTAAAC- 3′
pFa6a-kanMX6-P3nmt1-3HA	*cid1*	5′-AAGACCATTAAACAGTGAATAATTACTCAAAGAGAAAGTGAATAAGTTTGCCACTTTCACCTTTTGAGCAGTAGACGAATGAATTCGAGCTCGTTTAAAC- 3′/ 5′-TGTAAATTTTTGTGAATTTCTGCCTCAATCTCTTCAACTGTGTGAACACCAGGAATAAATTGTGCAGAAGAAATGTTCATGCACTGAGCAGCGTAATCTG- 3′
	cid16	5′-TTATTTATTCTGTGATCCTTAATTAAAAGCAAAGCCTAAAAAGAATTTTTACTTTTTTATTTAAAAATATCACTCGTGAA

		GAATTCGAGCTCGTTTAAAC- 3´/ 5´- ATTGAAATTTCTTTGCTCTTGGCCTTTGCGTAAGCAAAC AAAATTTGTACAGGTTTTAACAATAATTTGGCAAATAGC ATGCACTGAGCAGCGTAATCTG- 3´

3.5. Transformation

S. pombe wildtype strain JB22 was first woken up from the cell bank and plated on YES agar plate on day 1. On days 2, 3 and 4, the preculture 1, 2 (4.5 mL YES) and the culture (40 mL YES) were inoculated to the starting OD_{600} = 0.1.

The 40 mL culture was grown to an OD_{600} = 0.5 and centrifuged 3000 RCF, 3 minutes, room temperature (RT). Supernatant was removed, pelet was resuspended in 1 mL of buffer 1 (10% 1M LiAc, 10% 10x Te, 80% Milli-Q H_2O) and transferred to the 1.5 mL Eppendorf tube. The samples were centrifuged again 3000 RCF, 3 minutes, RT and supernatant was removed. 100 µl of buffer 1 was added again and the pelet was resuspended. 100 µl of the resuspended cells were transferred to a 2 mL Eppendorf tube and mixed with 10 µl of PCR product and 10 µl of UltraPure™ Salmon Sperm DNA solution which was preheated to 95°C in thermal mixer beforehand. After the 10 minutes of incubation on RT, samples were mixed with 280 µl of buffer 2 (10% 1M LiAc, 10% 10x Te, 80% PEG) and then incubated for 1 hour on 30°C. After the incubation, samples were mixed with 43 µl of DMSO and heatshocked for 7 minutes at 42°C. 500 µl of YES media was then added and the samples were centrifuged 2000 RCF, 3 min, RT. After removing supernatant, the pelets were again resuspended in 500 µl of YES media, transfered to glass tubes containing 4,5 mL of YES media and incubated overnight. The following day samples were centrifuged 3000 RCF, 3 minutes, RT. The most of the supernatant was removed, the pelet was resuspended in a little bit of left over media and plated to antibiotic plates; 100 ug/ml hygromicin b YES plates for deletion strains or 100 ug/ml

G418 YES plates for overexpression strains. Three days after eight colonies were picked and plated again on new antibiotic YES plate. After the colony PCR and verification of the transformed colonies, two transformed colonies were transferred to glass tubes with 4.5 ml of YES media and incubated overnight. The following day 1.2 ml of the culture was centrifuged, 600 µl of supernatant was removed and the remaining 600 µl was mixed with 600 µl of glycerol so it could be stored in the cell bank at -70°C.

3.6. Colony-PCR

Small amounts of eight colonies plated on antibiotic plates were picked with a pipette tip and mixed with 10 µl of 0.2% SDS. They were heated in the thermal cycler to 95°C for 10 minutes. PCR tubes were spinned down and 10 µl of Milli-Q water was added. The samples were spinned down for 10 seconds. The PCR mix consisted of 10 µl of VWR Red Taq DNA polymerase mix 2x, 1 µl of 10 uM forward primer, 1 µl of 10 uM reverse primer, 1 µl of the supernatant of the sample and 7 µl of Milli-Q water.
The following program was run:
1) 95°C for 00:03:00
2) 95°C for 00:00:30
3) 48°C for 00:00:30
4) 72°C for 00:01:15
5) Goto step 2 for 29 times
6) 72°C for 00:05:00
7) 4°C forever

Table 2. **Primer sequences used in colony PCR.** Table shows which primers were used in assessing whether gene incorporations in transformations were successful.

Gene	Forward primer/reverse primer
Δ*edc1_check*	5′-ATCTTTGTGCCGAATCTTTGTT- 3′ / 5′-TAAAAACCAAAAGCGAAGGAAA- 3′
Δ*xrn1_check*	5′-AACTAAACTCCTTCCTTCGCCT- 3′/ 5′-TGTTGATGTTCCATTTGATGGT- 3′
cid1_check	5′-TGATTAAGGCTTTCAACGGATT- 3′/ 5′-GTTTAAACGAGCTCGAATTC- 3′
cid16_check	5′-GCATAGGTTGCTTCATGCTCTA- 3′/ 5′-GTTTAAACGAGCTCGAATTC- 3′

*for *cid1_check* and *cid16_check* universal primer was used

3.7. Agarose gel and electrophoresis

1% gel was made mixing agarose with 1x TAE buffer and microwaved until the agarose was completely dissolved. After cooling down to around 50°C ethidium bromide was added to the concentration of 10 ng/ml. After the gel solidified it was put into the gel box for electrophoresis. The box was also filled with 1x TAE buffer. The ladder used is ThermoScientific GeneRµler 1k DNA Ladder Mix 0.5 ug/µl. 1 µl of PCR products for transformations were mixed with 5 µl of Thermofisher 6x Orange DNA Loading Dye before loading onto the gel. The volume of ladder and loaded samples was 5 µl. In the colony PCR, 9 µl of samples was loaded. The constant voltage of 105V/125V was used respectively for 50 mL/100mL of gel for 25 minutes. The visualisation of the gel was done using Quantum ST4 Xpress.

3.8. Cell collecting

On days 1, 2 and 3, the precultures 1, 2 (4.5 mL YES/EMM) and the cultures (15 mL YES/EMM – for the RNA isolation, 20 mL YES/EMM – for the western blot) were inoculated to the starting OD_{600} = 0.1. Wild type and deletion

strains were grown in liquid YES media while overexpression strains were grown in liquid EMM media. The strain overexpressing *cid16* gene was grown in liquid EMM media supplemented with thiamine; low and high concentrations of thiamine. (LT and HT concentrations were 16 ng/ml and 5 ug/ml respectively).

The 15/20 mL (RNA isolation/western blot) cultures grown to an OD_{600} = 0.5 were centrifuged at 3000 RCF, 3 minutes, 4°C. Supernatant was removed and the pelet was resuspended in 1 mL of 4C Milli-Q water. The sample was transferred to a 2 ml Eppendorf tube and centrifuged max RCF, 30 seconds, 4°C. Supernatant was removed again and the pelet with the Eppendorf tube was transferred to liquid nitrogen for 1 minute and then stored in -70°C for the later use.

3.9. Protein isolation

The stored pelets were taken out of -70°C and resuspended in 300 µl of lysis buffer (HEPES 50 mM, 150 mM NaCl, 5 mM EDTA pH 8.0, 1% TRITON-X100, 1 mM PMSF). The suspension was transferred to another Eppendorf tube with 250 µl of Roth 0,5 mm glass beads. The samples were lysed by vortexing them on high speed 1 minute and then putting them on ice for 1 minute. This cycle was repeated 10 times. Afterwards, the bottom of the Eppendorf tube was pierced using a hot needle and the Eppendorf tube was placed in a 40 mL Falcon tube following centrifuge 2500 RCF, 2 minutes, 4°C. Lysate was collected and transferred into new Eppendorf tube, centrifuged max RCF, 10 minutes, 4°C. 100 µl of supernatant was collected and transferred into new Eppendorf tube. 1 µl of the samples were mixed with 49 µl of PBS and 950 µl of (5x dilluted) Bio-rad Bradford reagent. Concentration of proteins were measured using NanoDrop 2000. 20 µl of 6x Lemmliego (300 mM TRIS-HCl pH 6.8, 10% SDS, 50% glycerol, 0.05% Bromophenol Blue) was added to the 100 µl of the samples and the samples were boiled on thermoblock on 95°C for 8 minutes.

3.10. SDS Page

5 mL of the 5% stacking gel (0.83 ml of 30% of acrylamide mix, 0.63 ml of 1.0 M of Tris pH 6.8, 0.05 ml of 10% SDS, 0.05 ml of 10% APS, 0.005 ml of TEMED and 3.4 ml of Milli-Q water) and 30 ml of the 10% separating gel (10.0 ml of 30% of acrylamide mix, 7.5 ml of 1.5 M of Tris pH 8.8, 0.3 ml of 10% SDS, 0.3 ml of 10% APS, 0.012 ml of TEMED and 11.9 ml of Milli-Q water) were used. 5 µl of the ladder VWR peqGOLD Protein Marker IV and 20 ug of each protein sample were loaded onto the gel. First part of electrophoresis was performed at 20 mA for 15 minutes and the second part was 30 mA until the visible bands reached the bottom of the gel.

3.11. Western blot

After the SDS-PAGE, the proteins were transferred to a Whatman nitrocellµlose membrane. The gel, papers and the membrane were soaked in transfer buffer (50 mM Tris BASE, 380 mM glycine, 0.1% SDS, 20% methanol). Transfer was performed in a Bio-Rad Mini Trans-Blot® apparatus at a voltage of 105 V and the current of 280 mA for 2 hours. The membrane was then washed with 1x TBS buffer (50 mM Tris-Cl of pH 7.5, 150 mM NaCl) and stained with PonceauS for 2 minutes to visualize the proteins and therefore transfer quality. After washing with Milli-Q water a couple of times, the membrane was blocked in 5% milk dissolved in 1x TBS buffer for 1 hour. The next step was incubation with 10 ng/µl of Invitrogen anti-HA mouse antibody (catalogue number: 5B1D10) dissolved in 5% milk in 1x TBS overnight at 4°C. After washing the membrane three times for 10 minutes in 1x TBS buffer, the membrane was incubated with 1 ng/µl of Sigma-Aldrich anti-Mouse IgG goat antibody (catalog number: MFCD00162638) for 1 hour. After incubation the membrane was washed again like in the previous step, 1.5 ml of Bio-Rad Clarity™ Western ECL Substrate Kit was added and the image was developed using AGFA CP1000.

3.12. Library preparation (TAIL-SEQ protocol)

1) RNA isolation

600 µl of TES buffer was added to the previously frozen cell pelet and quickly resuspended. 600 µl of the resuspended cells was transferred to an Eppendorf tube with 600 µl of phenol. The samples were incubated on thermoblock on 65°C while shaking 550 RPM. The sample was then centrifuged 15k RCF, 10 minutes, RT. 550 µl of the top fraction was transferred to an Eppendorf tube containing 550 µl of phenol. The sample was vortexed for 5 seconds high speed following centrifuge 15k RCF, 10 minutes, RT. 500 µl of the top layer was transferred to an Eppendorf tube with 500 µl of chloroform. The sample was again vortexed for 5 seconds high speed following centrifuge 15k RCF, 5 minutes, RT. The top layer was transferred to an Eppendorf tube containing 40 µl of 3M NaAc. 1.2 mL of 100% EtOH was added and mixed and the sample was vortexed and incubated on -20°C for 1h. The next step was centrifugation max RCF, 10 min, 4°C. After getting rid of supernatant, 500 µl of 70% EtOH was carefµlly added without mixing and the sample was centrifuged max RCF, 5 minutes, RT. Getting rid of supernatant again, adding 500 µl 70% EtOH and centrifuge max RCF, 5 minutes, RT. Getting rid of supernatant again, short spinning on centrifuge and removing whole supernatant. After letting the sample dry for 15 minutes on RT, 100 µl of nuclease free water (NFW) was added and the sample was resuspended. The sample was left to incubate on RT for couple of minutes, was vortexed and then put on ice. The concentration of the RNA was measured using Nanodrop 2000.

2) RNA cleaning

Cleaning was done according to protocol of RNeasy Mini Kit by QIAGEN using 100 µg of RNA and concentration measuring using Nanodrop 2000.

3) rRNA depletion

Depletion was done according to protocol of riboPOOL by siTOOLS using 5 µg of RNA and concentration measuring using Nanodrop 2000.

4) Clean-up

Clean-up was done according to protocol of RNeasy MinElute Cleanup Kit by QIAGEN to the final elution volume of 20 µl.

5) Adapter ligation

The volume of the RNA suspension was reduced by evaporating water using RVC 2-25 cd plus for 10 minutes in 50°C (there shoµld be around 8 µl left in the Eppendorf tube). 8 µl of RNA was mixed with 1 µl of 2 uM nextera adapter and the RNA was denatured on thermoblock on 65°C for 5 minutes. 9 µl of RNA was mixed with 11 µl of the following ligation mix: 1.5 µl of NEB ligation buffer 10x, 0.45 µl of 33 mM DTT, 0.75 µl of NEB T4 RNA ligase 2 truncated, 0.5 µl of Thermofisher RNase OUT, 7.5 µl of 50% PEG.

The mix was put in a thermal cycler and the following program was run:

1) 30°C for 06:00:00
2) 65°C for 00:20:00 to deactivate the enzyme
3) 4°C forever.

6) Clean up like in step 4).

7) Reverse transcription

The volume of the RNA suspension was reduced by evaporating water using speedvac centrifugal concentrator *RVC 2-25 cd plus* for 10 minutes in 50°C (there shoµld be around 8 µl left in the Eppendorf tube). 8 µl of RNA was mixed with 1 µl of 50 uM RT_Nextera and 1 µl of 10 uM dNTPs and the RNA was denatured on thermoblock on 65°C for 5 minutes. 10 µl of RNA was mixed with 10 µl of the following mix: 4 µl of Thermofisher FS buffer, 1 µl of Thermofisher Superscript III, 1 µl of Thermofisher RNase OUT, 3 µl of NFW.

The mix was put in a thermal cycler and the following program was run:

1) 55°C for 00:45:00
2) 70°C for 00:15:00

3) 4°C forever.

8) Clean-up with BeckmanCoμlter AMPure XP beads (1:1 ratio; added 12 μl of NFW; collected 10 μl)

Equal volume of resuspended AMPure XP beads was added to the ssDNA (in this case 20 μl) and incubated 5 minutes on RT. The probe was put on magnetic rack and incubated for another 3 minutes. The supernatant was carefully discarded and probes were washed twice with 180 μl of 80% EtOH. After second wash entire EtOH was collected. The beads were let to dry to a point when cracks start to appear at their surface. After removing the probe from the rack, beads were resuspended in 12 μl of NFW and incubated for 2 minutes. The tube was again placed on magnetic rack for 1 minute and 10 μl of the ssDNA suspension was transferred to a new Eppendorf tube.

9) Second strand synthesis

10 μl of the ssDNA was mixed with 10 μl of the following mix: 1 μl of NEB RNase H, 2.5 μl of NEB E. coli DNA polymerase, 0.5 μl of NEB E. coli DNA ligase I, 0.5 μl of 10 uM dNTPs, 2 μl of NEB Next Second Strand Synthesis Buffer, 3.5 μl of NFW.

The mix was put in a thermal cycler and the following program was run:

1) 16°C for 02:30:00

2) 4°C forever.

10) Clean-up with BeckmanCoμlter AMPure XP beads (1:1 ratio; added 7 μl of NFW; collected 6 μl).

11) Tagmentase Tn5 transposase oligos annealing and loading them onto Tn5

Tn5-rev and Tn5-A were mixed in a 1:1 ratio, put in a thermal cycler and the following program was run:

1) 95°C for 00:05:00

2) Slow ramping to 65°C
3) 65°C for 00:05:00
4) Slow ramping to 4°C
5) 4°C forever.

Annealed oligos were then dilluted to 35 uM (10 µl of oligos + 4.3 µl of NFW). 0.5 µl of diluted oligos were mixed with 10 µl of 2 mg/ml Tn5. The mix was then incubated in termomixer at 23°C, 30 minutes, 350 RPM.

12) Library tagmentation
Tn5 buffer was prepared (50% DMF, 50 mM TAPS pH 8.5, 25 mM MgCl2) and 2 µl were mixed with 6 µl of dsDNA and 2 µl of loaded Tn5. The thermal cycler was first heated to 55°C for 5 minutes and then the mix was put in a thermal cycler and the following program was run:
1) 55°C for 00:15:00
2) 80°C for 00:05:00
3) 4°C forever.

13) 12 µl of NFW was added and clean-up with BeckmanCoµlter AMPure XP (1:1 ratio; added 12 µl of NFW; collected 10 µl).

14) PCR
10 µl of the tagmentation product was mixed with 0.5 µl of each NEXTERA indeks, 4 µl of ThermoFisher HF buffer, 0.5 µl of 10 uM dNTPs, 0.7 µl of DMSO, 0.2 µl of ThermoFisher Phusion polymerase and 4.1 µl of NFW.
The mix was put in a thermal cycler and the following program was run:
1) 95°C for 00:03:00
2) 98°C for 00:00:20
3) 62°C for 00:00:15
4) 72°C for 00:00:30
5) Goto step 2 for 19 times
6) 72°C for 00:03:00

7) 4°C forever.

15) Clean-up with BeckmanCoμlter AMPure XP (0.9:1 ratio; added 22 μl of NFW; collected 20 μl).

16) Clean-up with BeckmanCoμlter AMPure XP (0.9:1 ratio; added 11 μl of NFW; collected 10 μl).

17) Concentration measurement and quality control using 2100 Bioanalyzer instrument and DNA high sensitivity kit.

Table 3. **Oligonucleotide sequences used in preparing libraries**

Name	Sequence (5'->3')
NEXTERA_adapter	/5rApp/CTGACNNNNNNCTGTCTCTTATACACATCTCCGAGCCCACGAGAC/3ddC/
RT_Nextera	GTCTCGTGGGCTCGGAGATG
Nextera i7 index primer	CAAGCAGAAGACGGCATACGAGAT[i7]GTCTCGTGGGCTCGG
Nextera i5 index primer	AATGATACGGCGACCACCGAGATCTACAC[i5]TCGTCGGCAGCGTC

3.13. Bioinformatic analysis

Raw sequencing data (fastq files) was analysed using a pipeline made by dr. Lidia Lipińska-Zubrycka (https://github.com/lidkalee/tailseq_complete_analysis).
S. pombe reference genome (assembly name ASM294v2) was obtained from https://www.ebi.ac.uk/ena/browser/view/GCA_000002945.2 and the used packages were fastq-pair (1.0), fastp (0.22.0), star (2.7.10b), samtools (1.3.1), bedtools (2.30.0) and ucsc-bedgraphtobigwig (357). Figures were created using program Rstudio (2022.07.0 Build 548) and the differential gene expression was done using DeSeq2 (1.38.2).

4) RESULTS

4.1. Yeast strain preparations

To study changes in uridylation, we decided to manipulate the expression levels of genes involved in mRNA decay in the cytoplasm. Specifically, we either deleted these genes or increased their expression levels to observe how it affected uridylation. Two types of yeast strains were prepared using linear inserts (PCR products) that were incorporated into the yeast genome through transformation and homologous recombination. To enable homologous recombination at the desired genomic location, overhangs homologous to specific genomic loci were added to the linear DNA fragments during PCR. This procedure follows a standardized method using the pFA6A plasmid family as a template for producing linear DNA fragments (59).

4.1.1. Gene deletion strains

According to the transformation of the *S. pombe* wild-type strain into deletion or overexpression strains, a specific template (plasmid) was used in PCR. Gene deletion strains were obtained by replacing the gene of interest with a linear insert (a gene for hygromycin B resistance) using the plasmid pFA6a-hphMX6 as the template during PCR. The transformed cells no longer contained the gene of interest (in this case, either *edc1* or *xrn1*), but they acquired resistance to hygromycin B, allowing them to grow on hygromycin B YE plates.

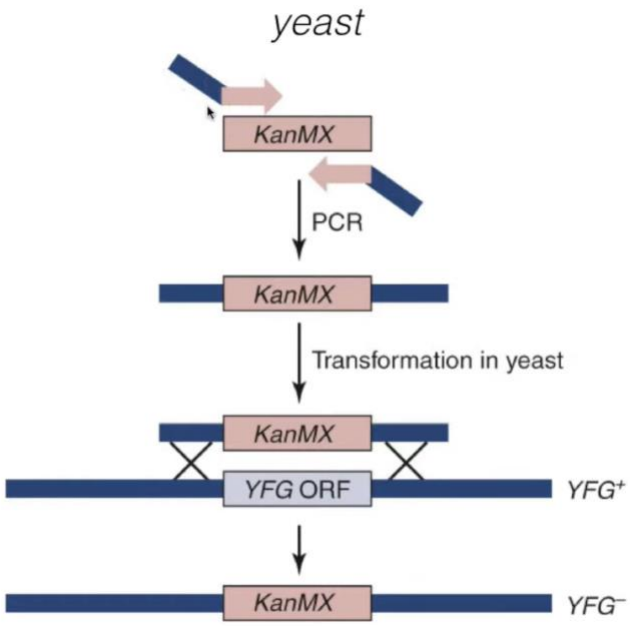

Figure 2. **Example of a graphical representation of gene deletion in yeast using homologous recombination and kanamycin resistance gene**

4.1.2. Gene overexpression strains

For gene overexpression strains, the plasmid pFa6a-kanMX6-P3nmt1-3HA was used as the template during PCR. This enabled the production of a linear insert consisting of the gene for kanamycin resistance, the nmt1 promoter, and an HA tag. The transformed cells acquired resistance to kanamycin and were able to grow on kanamycin YE plates. The use of the nmt1 promoter allowed for the induction of overexpression only when the cells were grown in EMM medium. EMM medium does not contain thiamine, which triggers transcription controlled by the nmt1 promoter, leading to the overexpression of the gene (in this case, either *cid1* or *cid16*). The HA tag was utilized in western blotting as an additional means of confirming the transformation.

4.2. Transformation verification

To confirm the success of transformations in both deletion and overexpression strains, colony-PCR was conducted (Figure 3). Successful transformation indicated that the PCR construct integrated into the correct position in the *S. pombe* genome. In addition, a western blot was performed to verify the overexpression of genes *cid1* and *cid16* in the respective strains. Since the overexpression constructs included HA tags, a mouse anti-HA antibody was used as the primary antibody, followed by an anti-mouse antibody as the secondary antibody.

Bands were observed for both repetitions of the *cid1* overexpression strain, but none were detected for the *cid16* overexpression strain (Figure 4).

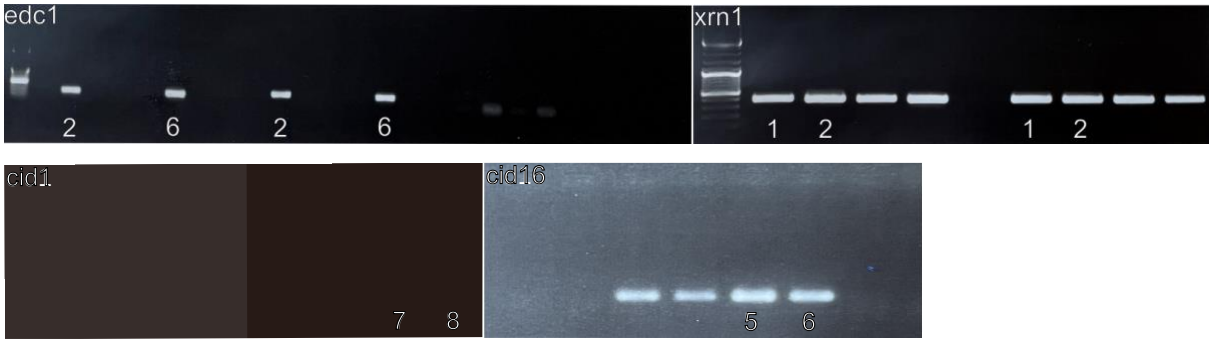

Figure 3. **Agarose gel after colony-PCR of deletion (top) and overexpression (bottom) strains.** For deletion strains, both 5´and 3´ends were checked and for overexpression strains only 5´ was checked. Numbers show chosen colonies which were later inoculated and saved to cell bank.

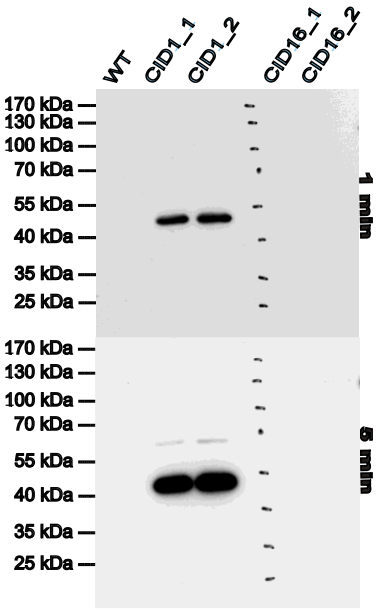

Figure 4. **Western blot of overexpressed cid1 and cid16 constructs containing HA tag after 1 minute of detection (top) and 5 minute of detection (bottom).** Constructs were detected using mouse anti-HA primary antibody and secondary anti-mouse antibody. Wild type strain was used a negative control.

4.3. Overexpression of *cid16* induces growth defects

Despite lack of signal on the western blot, strain overexpressing *cid16* exhibited growth defect that was expression dependent, suggesting production of functional protein. Deletion strains for Δ*edc1* and Δ*xrn1* exhibited a minor growth defect (data not shown). To further investigate the growth behavior of the cid16 overexpression strain, we performed Bioscreen experiments using EMM medium with varying thiamine concentrations (no thiamine, low thiamine, and high thiamine).

As expected, the wild type strain displayed similar growth curves regardless of the thiamine concentration. The exponential phase was consistent across all conditions, but the wild type strain grown in the absence of thiamine showed a slight decline during the late stationary phase.

In contrast, the strain overexpressing *cid16* demonstrated distinct growth patterns in EMM medium with different thiamine concentrations. When the thiamine concentration was high, it exhibited a growth curve similar to that

of the wild type strain, indicating that the nmt1 promoter was not active and *cid16* was not overexpressed. However, the mutant strain exhibited slower growth in EMM medium with low thiamine and even slower growth in EMM medium without thiamine, indicating an inverse relationship between the activity of the nmt1 promoter and the thiamine concentration. Interestingly, mutant grown in low thiamine experienced weaker expontential phase than the one grown in no thiamine (Figure 5).

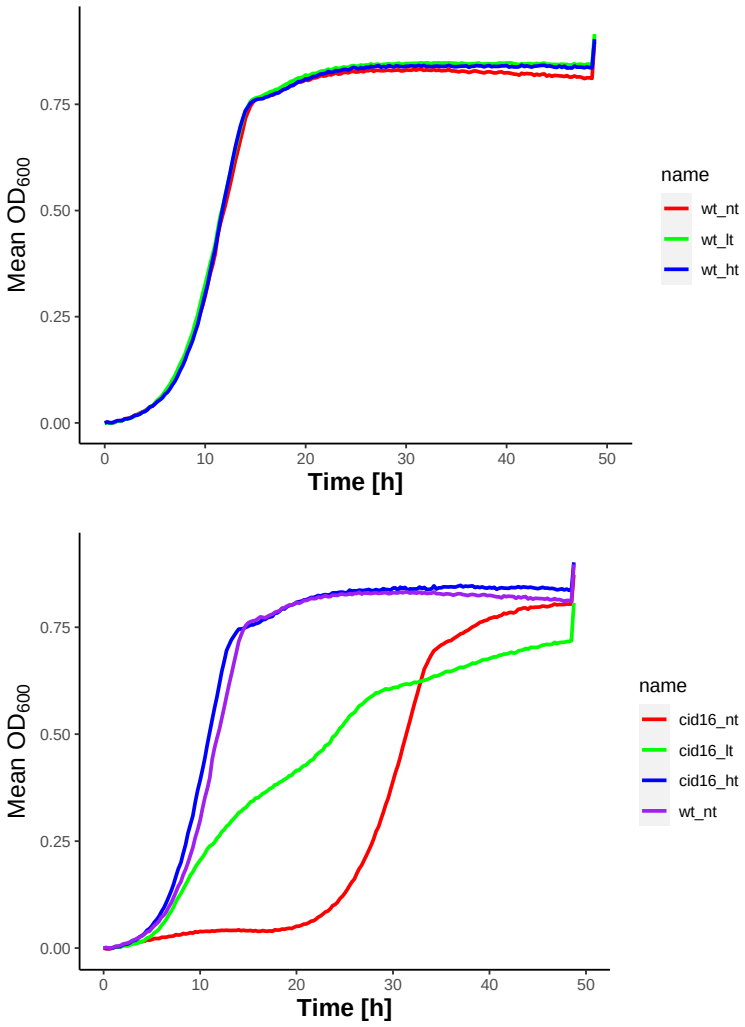

Figure 5. **Overexpression of cid16 induces growth defects.** Growth curves of wild type strain (top) and strain overexpressing cid16 (bottom). Strains were grown in EMM medium with no (nt), low (lt) and high (ht) concentrations of thiamine and growth curves were generated using Bioscreen data over the time span of fifty hours.

4.4. Sequencing data

To gain insight into the molecular phenotypes resulting from the deletion or overexpression of mRNA decay factors, we employed a custom gw-3'RACE protocol to generate sequencing libraries. This method offers superior accuracy in profiling mRNA tails compared to any other existing methods and is capable of detecting various mRNA tail modifications in addition to poly(A) modifications (15).

Following library preparation, quality control using the 2100 Bioanalyzer revealed that the libraries from strains overexpressing cid1 and cid16 did not yield satisfactory results. However, library preparation was successful for the wild type (wt), Δ*edc1*, and Δ*xrn1* strains, which were subsequently subjected to sequencing.

The Illumina sequencing generated paired-end reads (forward and reverse; R1 and R2) in fastq format. Prior to the bioinformatic analysis, MultiQC analysis was performed. The average quality scores of all samples exceeded 30 on the Phred Score scale, indicating a base call accuracy of 99.9%. Notably, all samples exhibited a sharp decline in Phred Score at base position 74. Additionally, the samples contained a significant number of duplicated sequences, which is commonly observed in RNA data analysis and suggests the presence of highly abundant transcripts (Figure 6).

The fastq files displayed noticeable differences in read counts, with edc1_2 having the highest number of reads and wt_2 having the lowest. After trimming adapters and low-quality reads, aligning the reads to the reference genome, performing data fixing, and filtering out data shared with wild type 1 and 2, sample edc1_1 retained the highest number of reads, while wt_1 retained the least (Figure 7). The result of the bioinformatic analysis was data organized in rows in columns (csv files) and such can be seen in Table 4.

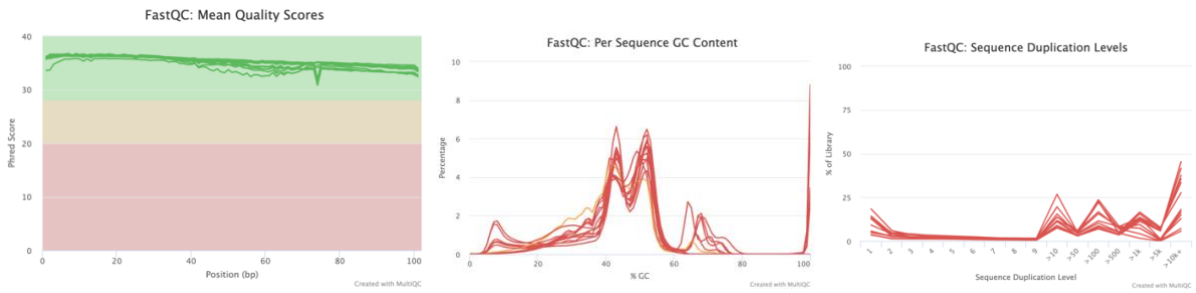

Figure 6. **Mean Quality Scores (left), Per Sequence GC Content (middle) and Sequence Duplication Levels (right) of sequenced samples wt_1, wt_2, edc1_1, edc1_2, xrn1_1 and xrn1_2 using FastQC analysis tool.** FastQC analysis was done before any bioinformatic processing.

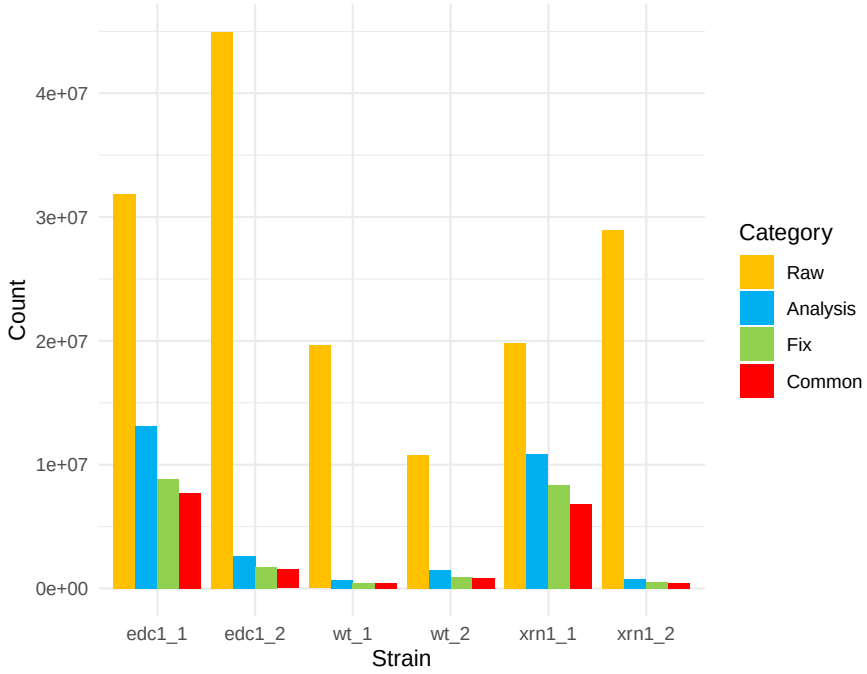

Figure 7. **Read counts of raw sequenced files prior to bioinformatic analysis (Raw), files after the bioinformatic analysis (Analysis), files after the data fix (Fix) and files after the gene filtering to match common genes to both wild type repetitions (Common).**

Table 4. **Preview of edc1_1 (edc1 repetition 1) file in program R after bioinformatic processing.** Table has 7 729 122 rows (reads) and 14 columns: chr (chromosome), start_R1 (annotated start of the forward read), strand_R1 (plus or minus strand of the forward read), gene_start (annotated start site of the gene), gene_stop (annotated stop site of the gene), gene (gene ID), coord_R2 (coordinate of the reverse read), RNA_type (type of the RNA), tail_from (method tail was acquired according to bioinformatic tools), tail_len (tail length in nucleotides), stop_R2 (annotated stop site of the reverse read), distance_to_TES (the distance to transcription end site).

	chr	start_R1	stop_R1	strand_R1	gene_start	gene_stop	gene	coord_R2	RNA_type	tail_from	tail_len	tail_type	stop_R2	distance_to_TES
1	II	2726392	2726493	-	2726356	2727959	SPBC4F6.18c	2726376	mRNA	cigar	0	no_tail	2726375	-19
2	I	987692	987773	-	987569	987825	SPSNRNA.03	987693	snRNA	cigar	0	no_tail	987692	-123
3	II	2190199	2190300	+	2189653	2190349	SPBC17G9.10	0	mRNA	grep	51	polyA	NA	NA
4	II	3958832	3958928	-	3958831	3959086	SPSNRNA.07	3958833	snRNA	cigar	0	no_tail	3958832	-1
5	II	2411250	2411351	+	2411189	2411361	SPSNORNA.34	2411272	snoRNA	cigar	0	no_tail	2411361	0
6	I	987745	987846	-	987569	987825	SPSNRNA.03	987693	snRNA	cigar	0	no_tail	987692	-123
7	I	4869996	4870083	+	4869146	4870212	SPNCRNA.254	4869997	ncRNA	cigar	0	no_tail	4870083	-129
8	I	4869996	4870083	+	4869465	4870169	SPNCRNA.1032	4869997	ncRNA	cigar	0	no_tail	4870083	-86
9	I	3395068	3395169	+	3394874	3395200	SPSNORNA.20	3395112	snoRNA	cigar	0	no_tail	3395201	1

Showing 1 to 10 of 7,729,122 entries, 14 total columns

RNA tails were characterized into no, mixed, poly(A), poly(AU) and oligo(U) tails. In this study, we aimed to investigate the prevalence and distribution of these various tail types across different RNA classes and especially in mRNAs.

Tailed reads were predominantly identified on mRNA, which aligns with our understanding of the 3' ends of various RNA classes. Surprisingly, a small fraction of uridylated tails was observed in mitochondrial RNA across all samples. Other RNA classes exhibited a minor fraction of tailed reads as well: non-coding RNA (ncRNA), small nucleolar RNA (snoRNA), and transfer RNA (tRNA) displayed poly(A) tails. Notably, ncRNA, along with mRNA, exhibited a slight fraction of poly(AU) tails.

Among the different RNA types, mRNA showed the smallest fraction of non-tailed reads (Figure 8).

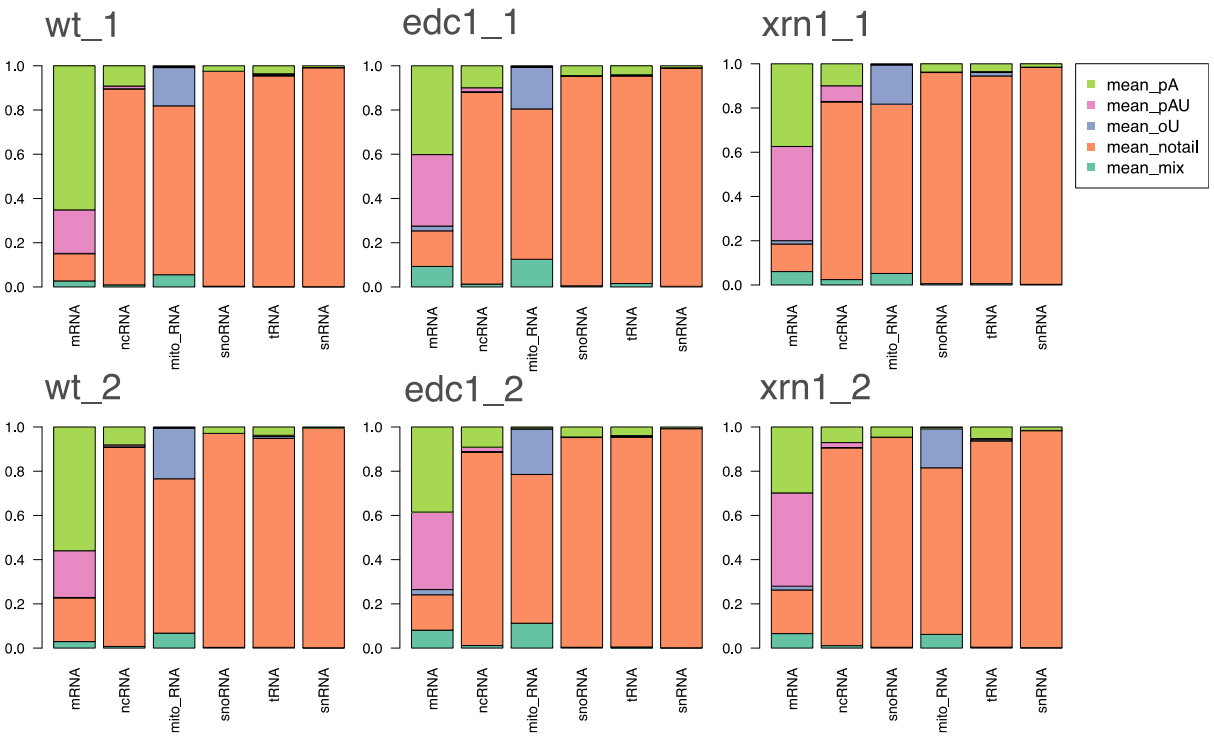

Figure 8. **Only a small fraction of mRNA is non tailed.** The characterisation of RNA tails in various RNA types, n > 20 where n = number of reads for each gene.

4.5. *Edc1* and *xrn1* deletion result in shortening of mRNA tail length

GW-3´-RACE enables reliable estimation of mRNA tail lengths. Using GW-3´-RACE, we obtained valuable insights into the distribution of mRNA tail lengths. The analysis of Δ*edc1* repetitions revealed a consistent pattern, with median tail lengths of 17 and 16. The second quartile ranged from 12 to 15 for the first repetition and from 11 to 24 for the second repetition. In contrast, the wild type repetitions displayed a larger variance in tail lengths, with median values of 26 and 22. The second quartile spanned from 18 to 37 for the first repetition and from 16 to 30 for the second repetition. Similarly, Δ*xrn1* repetitions exhibited medians of 17 and 15, with the second quartile ranging from 12 to 26 for the first repetition and from 11 to 22 for the second repetition. Remarkably, the minimum and maximum tail lengths across all samples were found to be 1 and 90, respectively.

Notably, wild type repetition 1 demonstrated a substantially wider second quartile range compared to the other samples, indicating a greater diversity in mRNA tail lengths within this particular repetition. On the whole, the mutants exhibited shorter mRNA tail lengths and a pronounced accumulation of mRNA molecules with shorter tails (Figure 9).

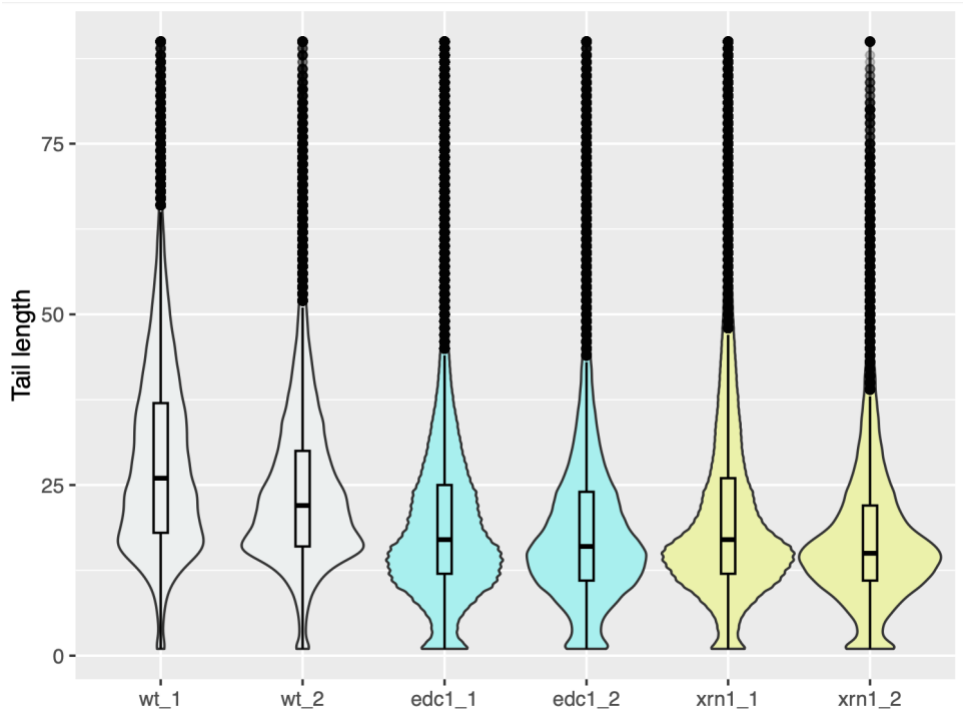

Figure 9. ***Edc1* and *xrn1* deletion result in shortening of mRNA tail length.** Violin plot for mRNA tail length. The boxes show the 75th (upper horizontal line), median (middle bold horizontal line) and the 25th (lower horizontal line) percentiles of the distribution, n > 20 where n = number of reads for each gene.

4.6. *Edc1* and *xrn1* deletion result in increased mRNA uridylation

The mRNA tails of both wild type repetitions predominantly consist of poly(A) tails, accounting for 77% in wild type repetition 1 and 73% in wild type repetition 2. The oligo(U) content is relatively small, representing only 1% in wild type repetition 1 and 2% in wild type repetition 2. The remaining

portion consists of poly(AU) tails, accounting for 22% in wild type repetition 1 and 25% in wild type repetition 2.

In comparison, Δ*edc1* exhibits the highest content of oligo(U) tails, with repetition 1 at 5% and repetition 2 at 4.8%. Its poly(AU) content is also higher than wild type, comprising 39% in repetition 1 and 42.8% in repetition 2. Although the oligo(U) content is lower for Δ*xrn1* compared to Δ*edc1*, with 2.8% in repetition 1 and 3.7% in repetition 2, their poly(AU) content is the highest among all samples, reaching 49.9% in repetition 1 and 50.7% in repetition 2 (Figure 10).

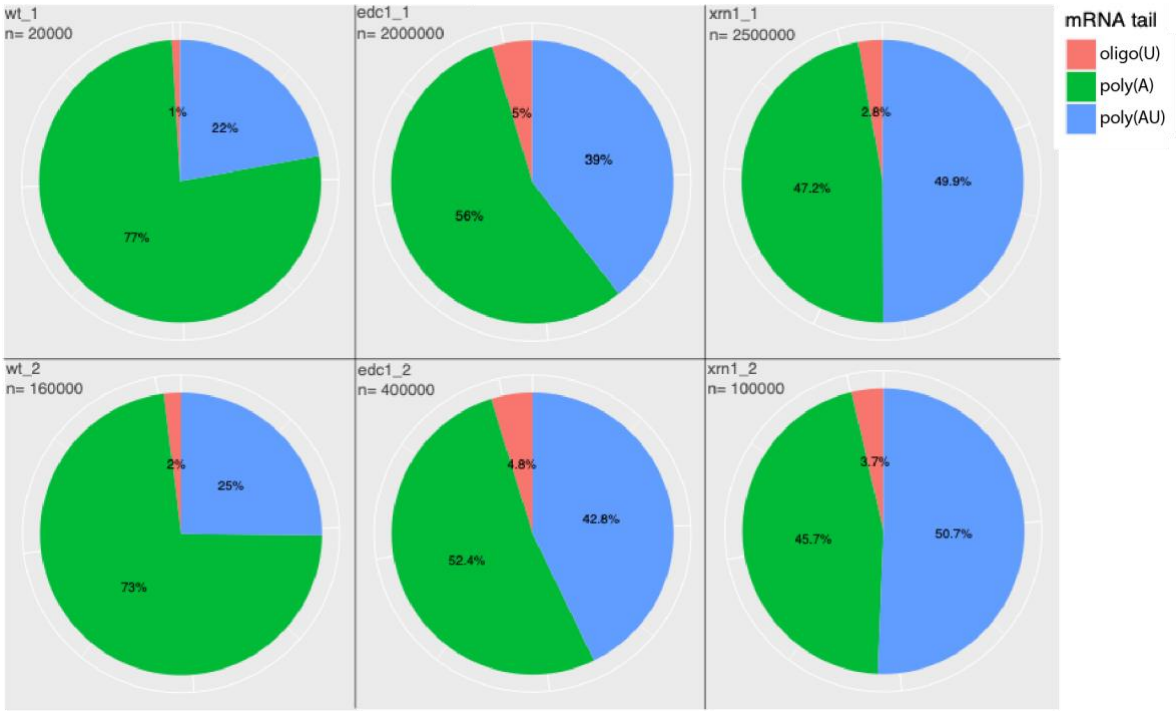

Figure 10. **Δ*Edc1* and Δ*xrn1* show greater poly(AU) and oligo(U) mRNA content compared to wild type.** Pie graph showing percentage ratios of each catergory of mRNA tails, n > 20 where n = number of reads for each gene.

The majority of mRNA reads in wild type repetition 1 have tails ranging from 12 to 61 nucleotides in length. In wild type repetition 2, there is a narrower distribution of both poly(AU) and poly(A) tails, with the majority falling between 12 and 40 nucleotides in length.

The most common tail lengths for poly(AU) and poly(A) tails are approximately 19 and 26 nucleotides, respectively, across both repetitions. The oligo(U) tails in wild type range from 1 to 3 nucleotides in length, with 1 nt tails being the most prevalent and 3 nt tails being the least represented.

In the case of Δ*edc1*, the majority of counts are distributed between 6 and 33 tails in length. The highest count for poly(A) tails occurs at a length of 26 nucleotides, while the highest count for poly(AU) tails falls within the range of 12 to 19 nucleotides. There is also a significant increase in 1 nt length poly(A) tails. Δ*xrn1* exhibits a similar pattern to Δ*edc1*, but with the difference of lacking the small peak for both poly(A) and poly(AU) tails at 26 nucleotides in length (Figure 11).

Both repetitions of Δ*edc1* and Δ*xrn1* show a broader distribution of oligo(U) tails across tail lengths. Δ*edc1* repetition 1 and 2 have ranges of 1 to 14 nt and 1 to 10 nt, respectively. Δ*xrn1* repetition 1 has the widest range of oligo(U) tail lengths, spanning from 1 to 19 nt, while repetition 2 ranges from 1 to 6 nt.

All samples exhibit the highest count of oligo(U) tails at a length of 1 nt, and there is a trend of decreasing counts as tail length increases (Figure 12).

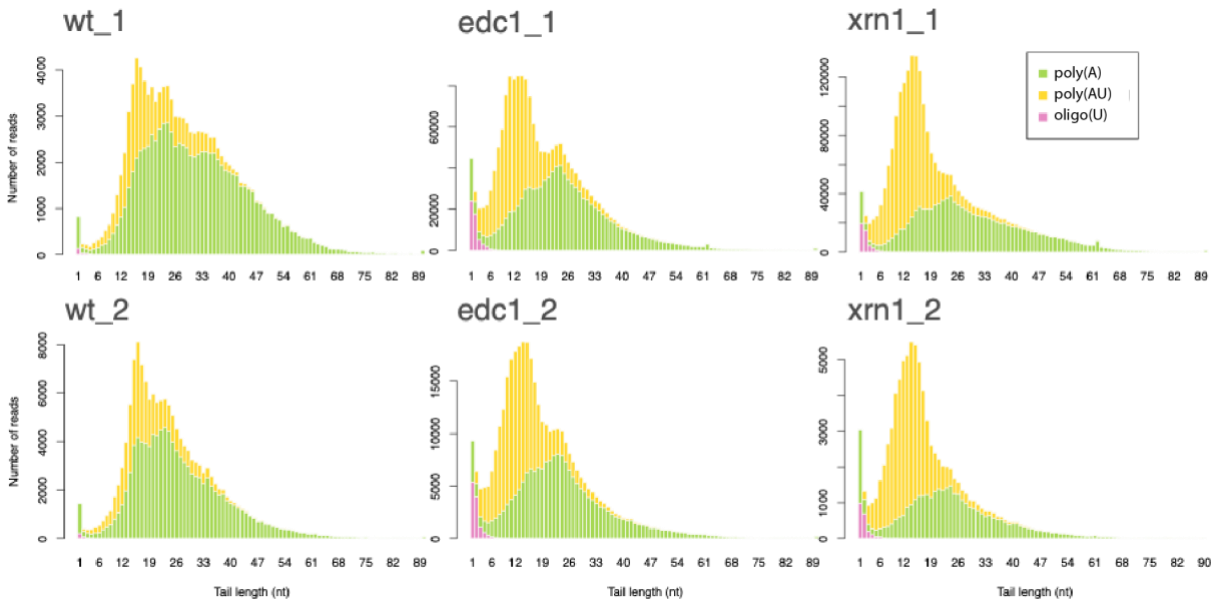

Figure 11. **Oligo(U) tails of Δ*edc1* and Δ*xrn1* strains are shorter than six nucleotides.** Counts of reads of each type of mRNA tail per tail length, n > 20 where n = number of reads for each gene.

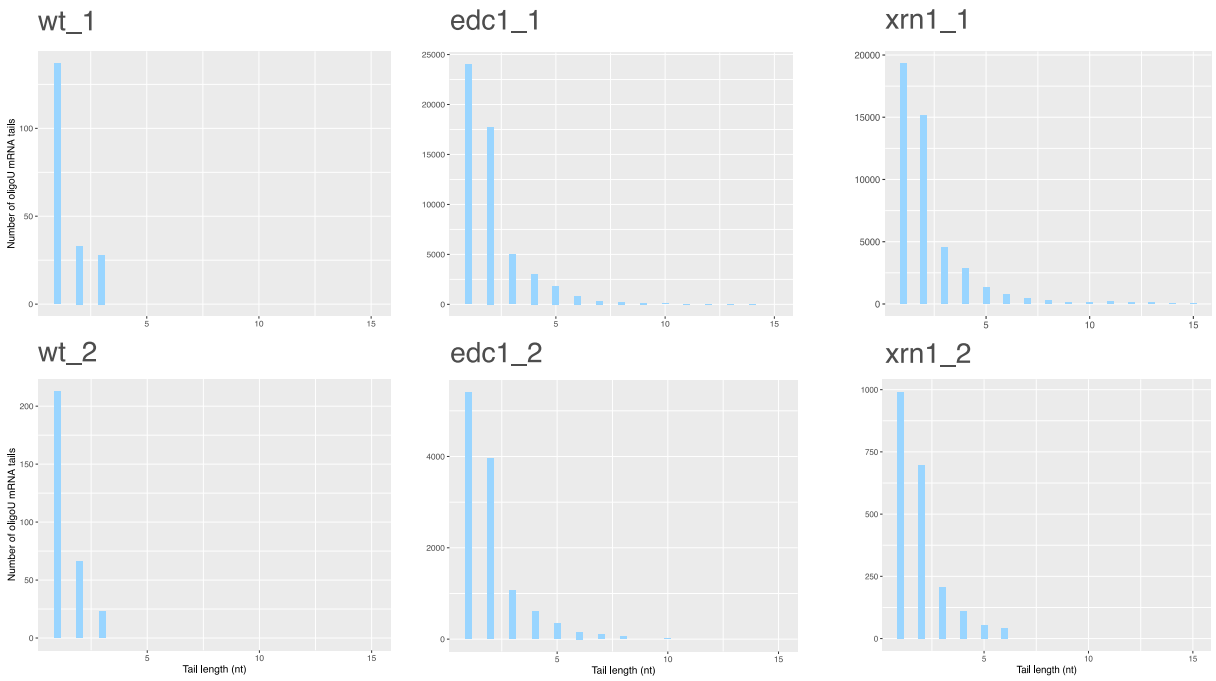

Figure 12. **Oligo(U) mRNA tails are mostly one or two uridines long.** Bar plot showing the frequency of each oligo(U) mRNA tail length, n, m > 20 where n = number of reads for each gene and m = number of reads for each tail length.

4.7. *Edc1* and *xrn1* deletion result in accumulation of non-tailed mRNAs

Figure 13 indicates that the large peak in all samples corresponds to reads mapping to annotated transcription end sites (TES). Wild type samples exhibit better coverage compared to the mutants, with a higher number of reads annotated across the entire range of -1000 to 250 kb of TES. In contrast, the mutants show sharper peaks closer to 0 kb to TES. (Figure 13).

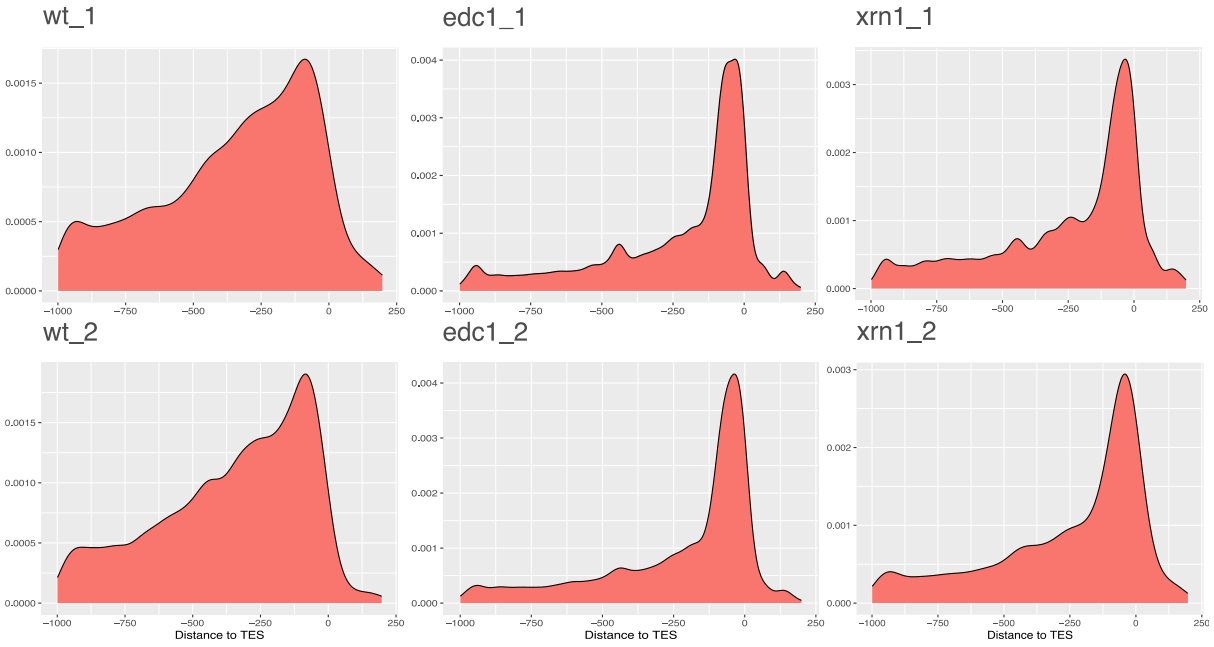

Figure 13. **Δ*edc1* and Δ*xrn1* samples possess a similary density pattern ranging from -1000 to 250 kb of annotated transcription end site.** Density plot ranging -1000 to 250 kb of annotated transcription end site, reads of genes SPAC6F6.03c, SPAC10F6.03c, SPCC663.10, SPCC1235.08c, SPBC1734.10c, SPAC6G10.11c, SPAC13G6.02c are excluded, n > 20 where n = number of reads for each gene.

4.8. Differentialy expressed genes

Out of all repetitions, the ones of Δ*xrn1* have the most variances (Figure 14). After performing differential gene analysis on 648 genes using DeSeq2 (adjusted p-value < 0.05, log2fold > 1), the results for Δ*edc1* samples indicate that 15 genes (2.30%) are upregulated (Table 6). The results for Δ*xrn1* samples suggest that 24 genes (3.70%) are upregulated (Table 7), while 11 genes (1.7%) are downregulated (Table 8). Both mutants share six common upregulated genes: *stg1* (SM22/transgelin-like actin modulating protein Sm22), *atd1* (aldehyde dehydrogenase), *gfa1* (glutamine-fructose-6,6-phosphate transaminase Gfa1), *rlp7* (ribosomal protein L7-like Rlp7 involved in ribosome biogenesis), *cmt1* (O-methyltransferase), and *ubi4* (ubiquitin). Performing the hypergeometric test the calculated p-value for gene overlap is less than 1e-4 indicating he statistical significance of the gene overlap between mutants. Upon examining the functions of the six common upregulated genes, it was found that they are involved in actin binding during mitosis, metabolism, ribosome biogenesis, transferase activity, and ubiquitin-related processes. The gene *ubi4*, encoding for ubiquitin, showed the highest level of upregulation in both mutants, with log2fold values of 2.60 for Δ*edc1* and 2.70 for Δ*xrn1*, respectively. Ubi4 is a stress-response protein involved in protein degradation, and research has shown its direct connection to RNA decay and the TOR signaling pathway. Specifically, the suppression of Ubi4 by Tor2 stabilizes Pir1, a component of the RNA elimination network that promotes RNA decay. Ubi4 is required for Pir1 degradation, and therefore, the upregulated levels of *ubi4* could be an indication of a lower rate of RNA degradation (60).

The most plausible explanation for the upregulation of ubi4 is that when mRNA degradation rates decrease, there is an increase in mRNA abundance, leading to higher protein synthesis and turnover, ultimately resulting in increased ubiquitination. Furthermore, the upregulation of *ubi4* also suggests a stress response due to the gene deletions in the mutants.

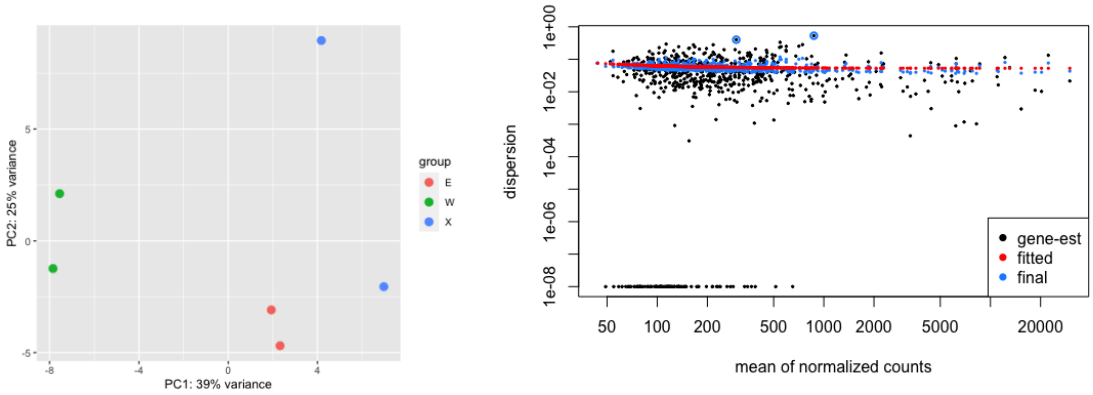

Figure 14. **PCA plot (left) and dispersion estimate (right) plot of differentialy expressed genes of deletion strains Δ*edc1* (E) and Δ*xrn1* (X) compared to wildtype (W).**

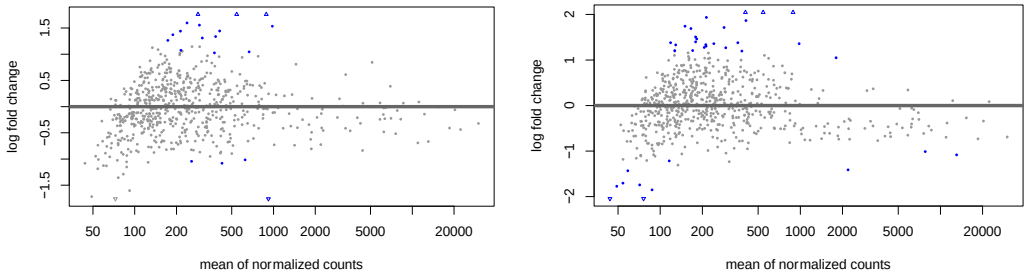

Figure 15. **MA plot of differentialy expressed genes of deletion strains Δ*edc1* (left) and Δ*xrn1* (right) compared to wildtype, adjusted p value < 0.05.**

Table 5. **Up-regulated genes in Δ*edc1* compared to wild type, log2 fold change ascending, adjusted p value < 0.05.**

	Gene ID	log2FoldChange	Gene	Gene product
1	SPBC1D7.01	1.03	pfd1	prefoldin subunit Pfd1
2	SPAC821.10c	1.04	sod1	superoxide dismutase Sod 1
3	SPBC713.08	1.07	mim1	mitochondrial MIM complex subunit Mim1
4	SPAC17H9.09c	1.26	ras1	GTPase Ras1
5	SPAC23H3.02c	1.31	ini1	RING finger-like protein Ini1
6	SPAC4F8.10c	1.34	stg1	SM22/transgelin-like actin modulating protein Stg1
7	SPAC959.03c	1.37	utp7	U3 snoRNP-associated protein Utp7
8	SPAC9E9.09c	1.44	atd1	aldehyde dehydrogenase
9	SPBC12C2.11	1.44	gfa1	glutamine-fructose-6-6-phosphate transaminase Gfa1
10	SPAC664.06	1.53	rlp7	ribosomal protein L7-like Rlp7 involved in ribosome biogenesis
11	SPAC637.06	1.55	gmh5	Gmh5 alpha-1,2-galactosyltransferase
12	SPAPB24D3.08c	1.60	-	NADP-dependent oxidoreductase
13	SPAC26F1.07	2.32	-	NADPH-dependent aldo-keto reductase
14	SPBC119.03	2.52	cmt1	O-methyltransferase, human COMT catechol homolog 1
15	SPBC337.08c	2.60	ubi4	protein modifier, ubiquitin

Table 6. **Down-regulated genes in Δ*edc1* compared to wild type, log2 fold change descending, adjusted p value < 0.05.**

	Gene	log2FoldChange	Gene	Gene product
1	SPBC14F5.04c	-1.02	pgk1	phosphoglycerate kinase Pgk1
2	SPAC2F3.11	-1.04	ppx1	exopolyphosphatase, prune Ppx1
3	SPAC6B12.18	-1.08	gon7	EKC/KEOPS complex subunit Gon7
4	SPBC32F12.11	-1.97	tdh1	glyceraldehyde-3-phosphate dehydrogenase Tdh1

Table 7. **Up-regulated genes in Δ*xrn1* compared to wild type, log2 fold change ascending, adjusted p value < 0.05.**

	Gene ID	log2FoldChange	Gene	Gene product
1	SPAC27E2.11c	1.05	-	uncharacterised
2	SPAC4F8.10c	1.20	stg1	SM22/transgelin-like actin modulating protein Stg1
3	SPCC1682.06	1.21	-	uncharacterised
4	SPCC576.09	1.21	rps20	40S ribosomal protein S20
5	SPBC28F2.02	1.27	mep33	translation machinery associated protein Mep33
6	SPAC8C9.03	1.28	cgs1	cAMP-dependent protein kinase regulatory subunit Cgs1
7	SPBC244.02c	1.30	utp6	U3 snoRNP-associated protein Utp6
8	SPBC23G7.05	1.33	sui1	translation initiation factor eIF1
9	SPCC1223.11	1.34	ptc22	serine/threonine protein phosphatase PP2C catalytic subunit Ptc2
10	SPCC1223.08c	1.36	dfr1	dihydrofolate reductase/lysophospholipase fusion protein Dfr1
11	SPAC664.06	1.36	rlp7	ribosomal protein L7-like Rlp7 involved in ribosome biogenesis
12	SPCC1739.13	1.38	ssa2	Hsp70 family heat shock protein Ssa2
13	SPAC1D4.13	1.38	byr1	MAP kinase kinase Byr1
14	SPBC21C3.06	1.40	-	uncharacterised
15	SPAC25H1.04	1.46	mug105	K48-linkage specific deubiquitinase
16	SPAC17D4.01	1.51	pex7	peroxin-7
17	SPBC725.07	1.69	pex5	peroxisomal targeting signal receptor Pex5
18	SPBC119.03	1.71	cmt1	O-methyltransferase, human COMT catechol homolog 1
19	SPCC965.05c	1.74	thp1	uracil DNA N-glycolase Thp1
20	SPBC12C2.11	1.87	gfa1	glutamine-fructose-6-phosphate transaminase Gfa1
21	SPAC9E9.09c	1.94	atd1	aldehyde dehydrogenase
22	SPCC330.06c	2.32	pmp20	thioredoxin-related chaperone Pmp20
23	SPBC337.08c	2.70	ubi4	protein modifier, ubiquitin
24	SPAC26F1.07	3.03	-	NADPH-dependent aldo-keto reductase

Table 8. **Down-regulated genes in Δ*xrn1* compared to wild type, log2 fold change descending, adjusted p value < 0.05.**

	Gene ID	log2FoldChange	Gene	Gene product
1	SPAC3G6.13c	-1.01	rpl4104	60S ribosomal protein L41
2	SPAC4F10.14c	-1.08	btf3	nascent polypeptide-associated complex beta subunit
3	SPBC19C7.07c	-1.22	sen34	tRNA-splicing endonuclease catalytic subunit Sen34
4	SPAC521.05	-1.41	rps802	40S ribosomal protein S8
5	SPBC1709.02c	-1.43	vrs1	cytoplasmic valine-tRNA ligase Vrs1/Vas1
6	SPAC15E1.08	-1.70	naa10	NatA N-acetyltransferase complex catalytic subunit Naa10
7	SPCC553.07c	-1.74	kpa1	DinB translesion DNA repair polymerase, pol kappa
8	SPCC297.04c	-1.77	set7	histone lysine H3-K37 methyltransferase Set7
9	SPBC146.08c	-1.85	tif1102	translation initiation factor eIF1A-like
10	SPBC3H7.07c	-2.07	ser2	phosphoserine phosphatase Ser2
11	SPBC4B4.01c	-2.15	ptk1	fumble family pantothenate kinase

5) DISCUSSION

5.1. Cid16 phenotype

Although the successful transformation of the strain overexpressing Cid16 terminal uridyltransferase was confirmed through colony-PCR, the absence of visible bands on the western blot polyacrylamide gel does not negate the success of the transformation. Moreover, the strain exhibited a growth defect, indicating a phenotype resulting from the overexpression of *cid16*. Due to this growth defect, we decided to cultivate the strain in EMM medium with low and high thiamine concentrations in order to decrease the activity of the nmt1 promoter and assess its impact on growth. Our findings revealed that lower promoter activity resulted in improved growth of the mutant. What is interesting is that strain overexpressing *cid16* grown in no thiamine has stronger exponential growth phase than the one grown in low thiamine concentration, indicating a mutation in one of these strains.

One possible explanation for the absence of Cid16 on the gel is that when Cid16 is overexpressed, it tends to aggregate, and this protein aggregation might be causing the growth defect. In other words, the aggregation of Cid16 could potentially be detrimental to the yeast cells, leading to their demise. This hypothesis could be investigated by analyzing the non-soluble fraction of the cells; however, due to time constraints, I was unable to perform this experiment.

Another explanation could be attributed to protein instability caused by the 3xHA tag. It has been observed that some yeast proteins tagged with a 3xHA tag experience reduced functional activity and protein instability compared to when they are tagged with other tags. The specific sequence of amino acids that separate the protein from the 3xHA tag is a critical factor influencing the stability of the protein (61).

Lastly, the growth defect observed could be attributed to increased Cid16 activity. Although it is possible that the end result of increased Cid16 activity is enhanced mRNA decay, the absence of this phenotype in Cid1, which is the primary mRNA TUTase in *S. pombe*, challenges this notion.

With increased Cid16 activity, there is also increased decay of Argonaute-bound small RNAs, which could potentially be the cause of the observed phenotype.

5.2. Cid1 and cid16 libraries

The unsuccessful libraries for strains overexpressing *cid1* and *cid16* are highly likely to be the result of human error. It is possible that errors occurred during the handling of yeast cells or the preparation of the libraries. Although the same reagents were used for successful libraries with deletion strains, it is unlikely that the issue with failed libraries for strains overexpressing *cid1* and *cid16* was caused by the reagents. However, we cannot exclude the possibility of other factors such as inconsistencies in experimental conditions or variations in the behavior of the strains being studied. Therefore, it is crucial to thoroughly assess all aspects of the experimental design and execution to identify the underlying causes of the failed libraries. This will help prevent the repetition of the same errors in future experiments. Unfortunately, due to time constraints, I was unable to repeat these libraries.

5.3. Variances in repetitions

While Figure 13 suggests that deletion strains have longer oligouridylated residues compared to two nucleotides, it is essential to consider the overall read counts. Edc1_1, edc1_2, and xrn1_1 have significantly higher read counts than the wild type repetitions. Only xrn1_2, among the deletion samples, has similar read counts to the wild type repetitions. Additionally, a previous study observed longer oligouridylated residues when decapping was inhibited (19).

Nevertheless, it is important to acknowledge other variations observed between repetitions. These include differences in tail length between wild type samples in Figure 9, variations in wild type repetitions in Figure 11, and discrepancies in gene expression patterns among Δ*xrn1* samples in

Figure 15. To mitigate such variations in future experiments, it is advisable to replicate each strain in three repetitions. This will provide more robust and reliable data, reducing the impact of experimental variability.

5.4. Main findings

The most significant findings of this experiment were the observed phenotype of strain overexpressing *cid16* and the similarities between the Δ*edc1* and Δ*xrn1* deletion strains. Both deletions led to an increase in mRNA with shorter tails, an elevated frequency of mRNA uridylation, and the accumulation of non-tailed mRNAs. These characteristics are indicative of mRNA degradation intermediates, specifically deadenylated mRNAs (mRNAs with shorter or no tail) and uridylated mRNAs. The increased uridylation in both deletion strains suggests that uridylation may be a prerequisite for both decapping and 5' to 3' decay in *S. pombe*.

Although these deletions resulted in the accumulation of mRNA intermediates and posed challenges for mRNA degradation, both mutants remained viable. Both deletion strains exhibit a reduced growth rate when grown at 30°C (data not shown) and severely reduced growth rates at lower temperatures (31,62).

Therefore, in the cases of both mutants, 5' to 3' decay is inhibited. The molecular phenotype observed was tail shortening and extensive uridylation of both poly(A) and deadenylated mRNAs. This outcome proves that in mRNA decay, uridylation is upstream of decapping and 5' to 3' decay by Xrn1. In such cases, the Lsm1-7 complex serves as the element linking uridylation to 5' to 3' decay. Additionally, in both mutants, I observed the accumulation of deadenylated mRNAs whose 3'-ends map close to the transcription end site. Notably, mRNA decay in both mutants relies on the 3' to 5' decay pathway. The accumulation of deadenylated mRNAs suggests that this decay pathway is not as efficient in picking up deadenylated mRNA substrates, resulting in their accumulation. The oligouridylation of deadenylated mRNAs observed in the studied mutants may lead to their decay in the 3' to 5' decay pathway by Dis3l2. However, based on my data,

it cannot be excluded that these species may also serve as substrates for the Lsm1-7 complex and subsequent 5' to 3' decay. These results are consistent with findings in *S. pombe* as well as other organisms. Mutants in decapping and deadenylation pathways in *S. pombe* also exhibit the accumulation of oligouridylated mRNAs (19). In humans, suppression of mRNA decay factors in both the 5' to 3' (Xrn1, Dcp1, Lsm1) and 3' to 5' (Dis3L2, Rrp41) pathways leads to the accumulation of oligouridylated mRNAs (20). In *S. cerevisiae*, knocking out factors in either the 5' to 3' or 3' to 5' mRNA decay pathway has minimal effects (63,64).

6)CONCLUSION

The decay of mRNA plays a crucial role in regulating gene expression and maintaining cellular homeostasis. The gene expression levels of factors involved in mRNA decay have a direct impact on the kinetics of mRNA decay. In the case of the *edc1* and *xrn1* deletions, there is a defect in the mRNA decay pathway, resulting in the accumulation of mRNA decay intermediates characterized by shorter tails, increased uridylation, and the accumulation of non-tailed mRNAs. Despite these defects, the deletion strains remain viable without major growth defects. Interestingly, the most significantly up-regulated gene was *ubi4*, which encodes for ubiquitin, indicating a stress response due to the gene deletions.

On the other hand, the overexpression of *cid16* resulted in a growth defect. This defect could be attributed to various factors, including increased Cid16 activity, Cid16 aggregation, or the presence of the HA tag. Further research is needed to investigate and better understand this issue. Notably, the overexpression of *cid1* did not have an impact on growth or mRNA decay, suggesting that uridylation may not be the rate-limiting step in this context. These results strongly suggest that uridylation is a prerequisite for the 5' to 3' decay pathway which has not been discovered before. Furthermore, it suggests that the 3' to 5' decay pathway is unable to adequately compensate for the inhibition of the 5' to 3' decay pathway.

To improve the reliability of experimental results and minimize the impact of variations between repetitions, it is recommended to replicate each sample in at least three independent repetitions. This will help ensure the robustness of the findings and provide more reliable data for analysis.

7) LITERATURE

1. CHEADLE C. Stability Regulation of mRNA and the Control of Gene Expression. Ann N Y Acad Sci. 2005 Nov 1;1058(1):196–204.
2. Wahle E, Winkler GS. RNA decay machines: Deadenylation by the Ccr4–Not and Pan2–Pan3 complexes. Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms. 2013 Jun;1829(6–7):561–70.
3. Wolf J, Passmore LA. mRNA deadenylation by Pan2–Pan3. Biochem Soc Trans. 2014 Feb 1;42(1):184–7.
4. Tucker M, Staples RR, Valencia-Sanchez MA, Muhlrad D, Parker R. Ccr4p is the catalytic subunit of a Ccr4p/Pop2p/Notp mRNA deadenylase complex in *Saccharomyces cerevisiae*. EMBO J. 2002 Mar 15;21(6):1427–36.
5. Tucker M, Valencia-Sanchez MA, Staples RR, Chen J, Denis CL, Parker R. The Transcription Factor Associated Ccr4 and Caf1 Proteins Are Components of the Major Cytoplasmic mRNA Deadenylase in Saccharomyces cerevisiae. Cell. 2001 Feb;104(3):377–86.
6. Yamashita A, Chang TC, Yamashita Y, Zhu W, Zhong Z, Chen CYA, et al. Concerted action of poly(A) nucleases and decapping enzyme in mammalian mRNA turnover. Nat Struct Mol Biol. 2005 Dec 13;12(12):1054–63.
7. Tharun S, Parker R. Targeting an mRNA for Decapping. Mol Cell. 2001 Nov;8(5):1075–83.
8. Vidya E, Duchaine TF. Eukaryotic mRNA Decapping Activation. Front Genet. 2022 Mar 23;13.
9. Schwartz DC, Parker R. Mutations in Translation Initiation Factors Lead to Increased Rates of Deadenylation and Decapping of mRNAs in *Saccharomyces cerevisiae*. Mol Cell Biol. 1999 Aug 1;19(8):5247–56.
10. Malecki M, Viegas SC, Carneiro T, Golik P, Dressaire C, Ferreira MG, et al. The exoribonuclease Dis3L2 defines a novel eukaryotic RNA degradation pathway. EMBO J. 2013 Mar 15;32(13):1842–54.
11. Viegas SC, Silva IJ, Apura P, Matos RG, Arraiano CM. Surprises in the 3'-end: 'U' can decide too! FEBS Journal. 2015 Sep;282(18):3489–99.
12. Chang HM, Triboulet R, Thornton JE, Gregory RI. A role for the Perlman syndrome exonuclease Dis3l2 in the Lin28–let-7 pathway. Nature. 2013 May 17;497(7448):244–8.
13. Garneau NL, Wilusz J, Wilusz CJ. The highways and byways of mRNA decay. Nat Rev Mol Cell Biol. 2007 Feb;8(2):113–26.
14. Gatfield D, Izaurralde E. Nonsense-mediated messenger RNA decay is initiated by endonucleolytic cleavage in Drosophila. Nature. 2004 Jun;429(6991):575–8.
15. Chang H, Lim J, Ha M, Kim VN. TAIL-seq: Genome-wide Determination of Poly(A) Tail Length and 3' End Modifications. Mol Cell. 2014 Mar;53(6):1044–52.
16. Sement FM, Ferrier E, Zuber H, Merret R, Alioua M, Deragon JM, et al. Uridylation prevents 3' trimming of oligoadenylated mRNAs. Nucleic Acids Res. 2013 Aug 1;41(14):7115–27.
17. De Almeida C, Scheer H, Zuber H, Gagliardi D. RNA uridylation: a key posttranscriptional modification shaping the coding and noncoding transcriptome. WIREs RNA. 2018 Jan 5;9(1).

18. Rissland OS, Mikulasova A, Norbury CJ. Efficient RNA Polyuridylation by Noncanonical Poly(A) Polymerases. Mol Cell Biol. 2007 May 1;27(10):3612–24.
19. Rissland OS, Norbury CJ. Decapping is preceded by 3′ uridylation in a novel pathway of bulk mRNA turnover. Nat Struct Mol Biol. 2009 Jun 10;16(6):616–23.
20. Lim J, Ha M, Chang H, Kwon SC, Simanshu DK, Patel DJ, et al. Uridylation by TUT4 and TUT7 Marks mRNA for Degradation. Cell. 2014 Dec;159(6):1365–76.
21. Nagarajan VK, Jones CI, Newbury SF, Green PJ. XRN 5′→3′ exoribonucleases: Structure, mechanisms and functions. Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms. 2013 Jun;1829(6–7):590–603.
22. Song MG, Kiledjian M. 3′ Terminal oligo U-tract-mediated stimulation of decapping. RNA. 2007 Dec;13(12):2356–65.
23. Lipińska-Zubrycka L, Grochowski M, Bähler J, Małecki M. Pervasive mRNA uridylation in fission yeast is catalysed by both Cid1 and Cid16 terminal uridyltransferases. PLoS One. 2023 May 23;18(5):e0285576.
24. Yates LA, Durrant BP, Fleurdépine S, Harlos K, Norbury CJ, Gilbert RJC. Structural plasticity of Cid1 provides a basis for its distributive RNA terminal uridylyl transferase activity. Nucleic Acids Res. 2015 Mar 11;43(5):2968–79.
25. Yates LA, Fleurdépine S, Rissland OS, De Colibus L, Harlos K, Norbury CJ, et al. Structural basis for the activity of a cytoplasmic RNA terminal uridylyl transferase. Nat Struct Mol Biol. 2012 Aug 1;19(8):782–7.
26. Wang SW, Norbury C, Harris AL, Toda T. Caffeine can override the S-M checkpoint in fission yeast. J Cell Sci. 1999 Mar 15;112(6):927–37.
27. Pisacane P, Halic M. Tailing and degradation of Argonaute-bound small RNAs protect the genome from uncontrolled RNAi. Nat Commun. 2017 May 25;8(1):15332.
28. Borja MS, Piotukh K, Freund C, Gross JD. Dcp1 links coactivators of mRNA decapping to Dcp2 by proline recognition. RNA. 2011 Feb;17(2):278–90.
29. Wurm JP, Overbeck J, Sprangers R. The *S. pombe* mRNA decapping complex recruits cofactors and an Edc1-like activator through a single dynamic surface. RNA. 2016 Sep;22(9):1360–72.
30. Wurm JP, Holdermann I, Overbeck JH, Mayer PHO, Sprangers R. Changes in conformational equilibria regulate the activity of the Dcp2 decapping enzyme. Proceedings of the National Academy of Sciences. 2017 Jun 6;114(23):6034–9.
31. Chen JS, Beckley JR, Ren L, Feoktistova A, Jensen MA, Rhind N, et al. Discovery of genes involved in mitosis, cell division, cell wall integrity and chromosome segregation through construction of *Schizosaccharomyces pombe* deletion strains. Yeast. 2016 Sep;33(9):507–17.
32. Dunckley T, Tucker M, Parker R. Two Related Proteins, Edc1p and Edc2p, Stimulate mRNA Decapping in *Saccharomyces cerevisiae*. Genetics. 2001 Jan 1;157(1):27–37.
33. Neef DW, Thiele DJ. Enhancer of decapping proteins 1 and 2 are important for translation during heat stress in *Saccharomyces cerevisiae*. Mol Microbiol. 2009 Sep;73(6):1032–42.
34. SCHWARTZ D, DECKER CJ, PARKER R. The enhancer of decapping proteins, Edc1p and Edc2p, bind RNA and stimulate the activity of the decapping enzyme. RNA. 2003 Feb;9(2):239–51.
35. Szankasi P, Smith GR. A single-stranded DNA exonuclease from Schizosaccharomyces pombe. Biochemistry. 1992 Jul;31(29):6769–73.

36. Parker R, Song H. The enzymes and control of eukaryotic mRNA turnover. Nat Struct Mol Biol. 2004 Feb;11(2):121–7.
37. Tharun S, He W, Mayes AE, Lennertz P, Beggs JD, Parker R. Yeast Sm-like proteins function in mRNA decapping and decay. Nature. 2000 Mar;404(6777):515–8.
38. Nissan T, Rajyaguru P, She M, Song H, Parker R. Decapping Activators in Saccharomyces cerevisiae Act by Multiple Mechanisms. Mol Cell. 2010 Sep;39(5):773–83.
39. Hsu CL, Stevens A. Yeast cells lacking 5'-->3' exoribonuclease 1 contain mRNA species that are poly(A) deficient and partially lack the 5' cap structure. Mol Cell Biol. 1993 Aug;13(8):4826–35.
40. Stevens A, Maupin MK. A 5' → 3' exoribonuclease of Saccharomyces cerevisiae: Size and novel substrate specificity. Arch Biochem Biophys. 1987 Feb;252(2):339–47.
41. van Dijk EL, Chen CL, d'Aubenton-Carafa Y, Gourvennec S, Kwapisz M, Roche V, et al. XUTs are a class of Xrn1-sensitive antisense regulatory non-coding RNA in yeast. Nature. 2011 Jul 22;475(7354):114–7.
42. Larimer FW, Hsu CL, Maupin MK, Stevens A. Characterization of the XRN1 gene encoding a 5'→3' exoribonuclease: sequence data and analysis of disparate protein and mRNA levels of gene-disrupted yeast cells. Gene. 1992 Oct;120(1):51–7.
43. Larimer FW, Stevens A. Disruption of the gene XRN1, coding for a 5'→3' exoribonuclease, restricts yeast cell growth. Gene. 1990 Oct;95(1):85–90.
44. Delorme-Axford E, Abernathy E, Lennemann NJ, Bernard A, Ariosa A, Coyne CB, et al. The exoribonuclease Xrn1 is a post-transcriptional negative regulator of autophagy. Autophagy. 2018 May 4;14(5):898–912.
45. Chapman EG, Moon SL, Wilusz J, Kieft JS. RNA structures that resist degradation by Xrn1 produce a pathogenic Dengue virus RNA. Elife. 2014 Apr 1;3.
46. Akiyama BM, Laurence HM, Massey AR, Costantino DA, Xie X, Yang Y, et al. Zika virus produces noncoding RNAs using a multi-pseudoknot structure that confounds a cellular exonuclease. Science (1979). 2016 Dec 2;354(6316):1148–52.
47. Haimovich G, Medina DA, Causse SZ, Garber M, Millán-Zambrano G, Barkai O, et al. Gene Expression Is Circular: Factors for mRNA Degradation Also Foster mRNA Synthesis. Cell. 2013 May;153(5):1000–11.
48. Ran XB, Ding LW, Sun QY, Yang H, Said JW, Zhentang L, et al. Targeting RNA exonuclease XRN1 potentiates efficacy of cancer immunotherapy. Cancer Res. 2023 Jan 13;
49. Sippel AE, Stavrianopoulos JG, Schutz G, Feigelson P. Translational Properties of Rabbit Globin mRNA after Specific Removal of Poly(A) with Ribonuclease H. Proceedings of the National Academy of Sciences. 1974 Nov;71(11):4635–9.
50. Sallés FJ, Darrow AL, O'Connell ML, Strickland S. Isolation of novel murine maternal mRNAs regulated by cytoplasmic polyadenylation. Genes Dev. 1992 Jul;6(7):1202–12.
51. Sallés FJ, Richards WG, Strickland S. Assaying the Polyadenylation State of mRNAs. Methods. 1999 Jan;17(1):38–45.
52. Sallés FJ, Strickland S. Analysis of Poly(A) Tail Lengths by PCR: The PAT Assay. In: RNA-Protein Interaction Protocols. New Jersey: Humana Press; p. 441–8.
53. Zhao S, Fung-Leung WP, Bittner A, Ngo K, Liu X. Comparison of RNA-Seq and Microarray in Transcriptome Profiling of Activated T Cells. PLoS One. 2014 Jan 16;9(1):e78644.

54. Lim J, Lee M, Son A, Chang H, Kim VN. mTAIL-seq reveals dynamic poly(A) tail regulation in oocyte-to-embryo development. Genes Dev. 2016 Jul 15;30(14):1671–82.
55. Legnini I, Alles J, Karaiskos N, Ayoub S, Rajewsky N. FLAM-seq: full-length mRNA sequencing reveals principles of poly(A) tail length control. Nat Methods. 2019 Sep 5;16(9):879–86.
56. Garalde DR, Snell EA, Jachimowicz D, Sipos B, Lloyd JH, Bruce M, et al. Highly parallel direct RNA sequencing on an array of nanopores. Nat Methods. 2018 Mar 15;15(3):201–6.
57. Workman RE, Tang AD, Tang PS, Jain M, Tyson JR, Razaghi R, et al. Nanopore native RNA sequencing of a human poly(A) transcriptome. Nat Methods. 2019 Dec 18;16(12):1297–305.
58. Krause M, Niazi AM, Labun K, Torres Cleuren YN, Müller FS, Valen E. *tailfindr* : alignment-free poly(A) length measurement for Oxford Nanopore RNA and DNA sequencing. RNA. 2019 Oct;25(10):1229–41.
59. Bähler J, Wu JQ, Longtine MS, Shah NG, Mckenzie III A, Steever AB, et al. Heterologous modules for efficient and versatile PCR-based gene targeting inSchizosaccharomyces pombe. Yeast. 1998 Jul;14(10):943–51.
60. Wei Y, Lee NN, Pan L, Dhakshnamoorthy J, Sun LL, Zofall M, et al. TOR targets an RNA processing network to regulate facultative heterochromatin, developmental gene expression and cell proliferation. Nat Cell Biol. 2021 Mar 11;23(3):243–56.
61. Saiz-Baggetto S, Méndez E, Quilis I, Igual JC, Bañó MC. Chimeric proteins tagged with specific 3xHA cassettes may present instability and functional problems. PLoS One. 2017 Aug 11;12(8):e0183067.
62. Szankasi P, Smith GR. Requirement of S. pombe exonuclease II, a homologue of S. cerevisiae Sep1, for normal mitotic growth and viability. Curr Genet. 1996 Sep 26;30(4):284–93.
63. He F, Li X, Spatrick P, Casillo R, Dong S, Jacobson A. Genome-Wide Analysis of mRNAs Regulated by the Nonsense-Mediated and 5′ to 3′ mRNA Decay Pathways in Yeast. Mol Cell. 2003 Dec;12(6):1439–52.
64. Houalla R, Devaux F, Fatica A, Kufel J, Barrass D, Torchet C, et al. Microarray detection of novel nuclear RNA substrates for the exosome. Yeast. 2006 Apr 30;23(6):439–54.