Simplified Nonsense: New Methods for Interrogating NMD

Hanna Alalam

Department of Chemistry and Molecular Biology
Faculty of Science
University of Gothenburg

UNIVERSITY OF GOTHENBURG

Gothenburg 2023

Cover illustration: “The nonsense goober” by Karl Persson

Simplified Nonsense: New Methods for Interrogating NMD

Hanna.alalam@gu.se

ISBN 978-91-8069-163-5 (PRINT)
ISBN 978-91-8069-164-2 (PDF)

Printed in Borås, Sweden 2023
Printed by Stema Specialtryck AB

To those who light the candles of my path

ABSTRACT

Nonsense mediated decay (NMD) is a pathway that regulates RNA turnover. Since its discovery, this pathway has been implicated in a variety of cellular processes ranging from differentiation to the restriction of viral replication. While NMD has been heavily studied since its discovery, the understanding of how the pathway carries out its function has been a long and convoluted process, where the current cornerstones that establish our present understanding of the regulatory mechanisms are continuously challenged.

In this thesis, new methods were explored with the goal to provide tools that would simplify investigating the NMD pathway and potentially other pathways regulating RNA. We studied the use of nucleotide conversion methods and their applicability to yeast. Additionally, we designed a set of reporters that allow *in vivo* monitoring of NMD with an easy-to-read phenotype as an output. Moreover, we modified a reporter that was developed during the construction of the NMD reporters to also be applicable for alternative studies. In this particular case, we adapted one of our reporters to the study of the SARS-CoV-2 major protease (NSP5). Overall, simplified methods to interrogate cellular NMD were successfully constructed, in addition to establishing a sensitive yeast-based system for the detection of anti-viral compounds.

Keywords: mRNA degradation, metabolic labeling, nonsense-mediated decay, Major protease, SARS-CoV-2

SAMMANFATTNING

Nonsensmedierad nedbrytning (NMD) är en signaleringsväg som reglerar omsättningen av RNA. Alltsedan upptäckten av denna signaleringsväg har många olika cellulära processer kopplats till NMD-reglering, såsom celldifferentiering och begränsandet av virusreplikation. Fastän NMD har studerats flitigt sedan det upptäcktes, har den mekanistiska förståelsen av NMD varit en lång och komplicerad process där de nuvarande grundläggande hörnstenarna än idag ifrågasätts.

I denna avhandling har nya metoder utforskats för att tillhandahålla verktyg som förenklar undersökandet av NMD och potentiellt även andra RNA-reglerande signaleringsvägar. Vi har utforskat hur man kan använda metoder för konversion av nukleotider och hur dessa kan tillämpas på jäst. Vi har också designat en uppsättning reportrar som tillåter *in vivo*-avläsning av NMD med en tydlig fenotyp som resultat. Dessutom modifierade vi några av de koncept som utvecklades under konstruktionen av NMD-reportrarna för att även vara tillämpliga på alternativa studier. I det här specifika fallet anpassade vi ett av våra koncept till att studera Nsp5-proteaset (major protease) från SARS-CoV-2. På detta sätt konstruerades därmed förenklade metoder för att undersöka cellulär NMD och ett känsligt jästbaserat system för detektion av antivirala föreningar.

Nyckelord: mRNA-nedbrytning, metabolisk inmärkning, nonsens-medierad nedbrytning, Major protease, SARS-CoV-2

LIST OF PAPERS

This thesis is based on the following studies, referred to in the text by their Roman numerals.

I. **Alalam, H.**, Zepeda-Martínez, J.A. and Sunnerhagen, P. (2022). Global SLAM-Seq for accurate mRNA decay determination and identification of NMD targets. RNA 28, 905–915.

II. **Alalam, H.** and Sunnerhagen, P. (2023). New reporters for monitoring cellular NMD. Manuscript.

III. **Alalam, H.**, Sigurdardóttir, S., Bourgard, C., Tiukova, I., King, R.D., Grøtli, M. and Sunnerhagen, P. (2021). A genetic trap in yeast for inhibitors of SARS-CoV-2 main protease. mSystems 6, e01087-21.

Not included in this thesis:

Alalam, H., Graf, F.E., Palm, M., Abadikhah, M., Zackrisson, M., Boström, J., Fransson, A., Hadjineophytou, C., Persson, L., Stenberg, S., Mattsson, M., Ghiaci, P., Sunnerhagen, P., Warringer, J. and Farewell, A., 2020. A High-Throughput Method for Screening for Genes Controlling Bacterial Conjugation of Antibiotic Resistance. mSystems 5, e01226-20.

AUTHOR CONTRIBUTION

I. Conceived the project, prepared all the sequencing samples, carried out all the analysis, prepared all the figures and co-wrote the manuscript.

II. Conceived the project, carried out all the experiments and analysis, prepared all the figures and co-wrote the manuscript.

III. Conceived the project, constructed all the strains, carried out the initial characterization of the system, prepared all the figures, co-analyzed the data and co-wrote the manuscript.

CONTENT

AIM OF THE THESIS

This thesis primarily deals with nonsense mediated decay (NMD) using *Saccharomyces cerevisiae* (hereafter referred to as yeast) as a model organism. We aimed to create new methods that simplify quantifying NMD, either at the RNA half-life level or through reporters. It is important to note that while all experiments were done in yeast, it is very difficult to discuss NMD in a yeast-isolated context. Especially given that the way information was accumulated on this pathway was through the integration of knowledge coalescing from the study of NMD from multiple different model organisms.

While NMD is generally a well-conserved process, important differences exist between experimental systems. Throughout this thesis, efforts have been made to point out the differences relative to other organisms (primarily human) when applicable. Moreover, it must be noted that NMD is a highly complex pathway and no published work to date has determined a unified functioning model on how the pathway works. On the upside, this leaves plenty of room for further research and model refinement to be built on top of the vast wealth of information that has been amassed on NMD since 1979.

Unfortunately, during the writing of this thesis the world had the displeasure of dealing with the SARS-COV-2 pandemic. The pandemic was a trying time for all of us and poised a unique challenge, especially for the scientific community. I believe it is absolutely essential that we remember the lessons we learned from this bitter experience which for me was the importance of adaptability in facing new challenges. In light of this, some of the reporters originally intended for NMD assessment were adapted for a viral protein study and were incorporated in this thesis that was subsequently split into two parts to simplify reading the thesis. The aims of each paper included in this thesis are summarized as follows:

I. In paper I, we aimed to adapt a new metabolic labeling method to determine RNA half-life in a simplified manner to yeast.
II. In paper II, we aimed to create reporters that reflect the NMD status of the cell and characterize them in the deletion collection.

III. In paper III, we aimed to modify a concept that was developed in paper II to monitor the suppression of a viral protein by chemical inhibitors.

I thoroughly enjoyed writing this thesis. I have gained a great appreciation of all the hard work that has gone into the dissection of this pathway from the past 44 years and I am proud to present the culmination of the work that has been carried out during my PhD studies. In the end, I really hope you enjoy reading this work as much as I have had writing it.

INTRODUCTION – NMD

Eukaryotic cellular homeostasis is maintained through complex interweaved systems of regulation that are divided into multiple layers. The amount of a given protein can be regulated through transcriptional upregulation or downregulation in response to stimuli, the transcript itself can be subject to post-transcriptional regulation, and the translated protein can be further modified to alter its function. An additional aspect for consideration is the control of the half-life on both the RNA level and protein level. The half-lives of transcripts and proteins vary between gene products and this can also be regulated to maintain cellular homeostasis. Various previous works have focused on different parts of these regulatory networks; however, RNA stability has been less explored relative to transcription and translation due to the nature of the experiments required to capture this parameter and the low reproducibility of measured stabilities between labs using the older techniques.

RNA molecules exist for a finite period of time. The time required for the initial observed amount to be reduced to half is referred to as half-life and is used as a measure of stability of a given RNA. Under certain circumstances it becomes crucial for the cells to degrade a given transcript quickly to circumvent the accumulation of non-productive or even toxic transcripts. Quality control systems exist within eukaryotic cells to ensure timely degradation of aberrant transcripts.

Nonsense mediated decay (NMD) is the best studied example of a RNA quality control system and functions in both regulation of wild-type transcripts and the removal of aberrant transcripts [1]. The first evidence of NMD emerged from the studies of mutated form of the yeast *URA3* locus harboring premature termination codons (PTCs) at various positions [2]. It was observed that *URA3* transcripts originating from mutants carrying PTCs had a reduced half-life relative to the wild-type transcripts and the strength of reduction was dependent on the location of the PTC within the open reading frame of the transcript [2]. Within a short time of the previous work, a study on a mutant PTC containing allele of the β-globin gene from a thalassemia patient found that the mutated transcript had a reduced half-life relative to the wild-type allele [3]. The authors of the study correctly speculated that the effect is caused by RNA instability due to a nonsense (termination codon) mutation. The next set of experimental evidence came from the studies of

alleles of *unc*-54, *lin*-29 and *tra*-2 from *Caenorhabditis elegans* for which extragenic suppressors could be found [4]. They identified suppressors for these mutations and named them Suppressors of Morphogenesis in Genitalia 1-6 (SMG1-6) and correctly postulated that the identified suppressors affect RNA stability [4]. Their results indicated that the currently unidentified process (later to be known as NMD) was acting through trans-acting proteins. Orthologs of some of these proteins were then identified in yeast by studying the changes in *his4-38* mRNA stability and were termed Up frameshift proteins (UPF1-3) [5], [6]. The term NMD was coined in 1993 [7]. The authors studied the effect of having a PTC in a normally stable mRNA in yeast. Combining the knowledge from previous studies they showed that addition of PTCs rendered the *PGK1* transcript unstable and that the extent of the destabilization is dependent on the location of the PTC within the open reading frame (with earlier occurring PTCs being more destabilizing) and that the process was controlled by the trans-acting factor Upf1. This process was termed NMD due to the fact that all observed substrates up to that point have been PTC-containing transcripts.

NMD PROTEINS

The NMD pathway is now known to not only target PTC-containing transcripts but also to regulate wild type transcripts. The functionality of the pathway depends on three core factors (UPF1, UPF2 and UPF3), which have been shown to be conserved in a large number of eukaryotic cells [8]. Disruption of any of these three factors in yeast has been shown to stabilize known NMD targets in addition to a significant portion of the transcriptome [9].

UPF1 is an ATP-dependent RNA helicase and is the most conserved out of the core proteins [8]. UPF1 is a highly modular protein that is extensively regulated. It mainly consists of a cysteine–histidine-rich (CH) located towards the N-terminal part of the protein, an RNA helicase region located centrally in the protein and a C-terminal serine and glutamine rich (SQ) domain [10]. The SQ domain and the N-terminal region have a high density of S/T–Q dipeptides, which are the target of the protein kinase SMG-1 in higher eukaryotes hence conferring a phosphorylation based regulation of UPF1 by SMG-1 [11], [12]. It is of note that the *S. cerevisiae* homolog of UPF1 has a reduced density of S/T–Q dipeptides motifs and this is a consequence of the loss of SMG-1 in yeast evolution [8]. Although UPF1 is still a phosphoprotein in yeast, the identity of the kinase responsible for the phosphorylation is not known [13] and a clear requirement for phosphorylation in yeast has not been established. UPF1 contains a highly processive helicase domain [14]. The helicase domain is divided into multiple sub-domains: RecA1 and RecA2 are the catalytic subdomains while 1B and 1C are regulatory subdomains [15]–[17]. Finally, the CH domain is required for interaction with UPF2 and the auto-inhibition of the helicase domain of UPF1 [18]–[20]. It has been shown that in mammalian cells the concentration of UPF1 surpasses that of UPF2 and UPF3 indicating that UPF1 may serve functions beyond its role in the NMD pathways. This is indeed the case as it has been demonstrated that UPF1 participates in additional decay pathways besides NMD [21]. Moreover, UPF1 possesses a E3-ubiquitin ligase activity [22] that promotes the degradation of the MYOD protein, which functions in the regulation of myogenesis [23].

The UPF2 protein can be thought of as a bridging component interacting with both UPF1 and UPF3 [24] and this interaction is required for NMD in yeast. UPF2 interacts with the UPF1 CH domain to prevent its inhibition of

the UPF1 helicase domain. There are some studies that specifically used a UPF2 knockout to examine the effects on spermatogenesis and liver function and development [25], [26], however, in both of these cases the effects are confined to UPF2 function on NMD rather than an alternative function.

UPF3 is the least conserved of the three core factors [8]. Yeast UPF3 is required for NMD since deletion of any of the core proteins leads to the upregulation of a set of 907 overlapping transcripts [9]. The UPF3 protein contains nuclear localization signals and nuclear export signals and hence can shuttle between the nuclear and cytoplasmic compartments [27]–[30]. It also contains an RNA recognition motif domain (RRM) that interacts with UPF2 instead of RNA [19], [31], [32], although there is evidence that UPF3B (see below) can bind RNA [33]. In humans, UPF3 is unique among NMD core proteins due to the presence of two sister proteins UPF3A and UPF3B, of which UPF3B is considered the primary NMD factor with UPF3A acting as backup or as an antagonist in certain situations [34]–[36]. Of interest is the recent observation of UPF3A involvement in the genetic compensation response in zebrafish that offers an explanation for phenotypic discrepancies between knockouts vs. knockdowns[37]–[39]. Specifically, a PTC-containing transcript of Capn3, a calcium-dependent cysteine protease, was able to trigger compensatory expression of Capn8 and Capn12 while a knockdown of Capn3 was not, and this effect was mediated through the association of UPF3A with COMPASS (complex of proteins associated with Set1) [37].

Additional proteins other than the core proteins have been identified in multicellular organisms and are referred to as SMG-1 and SMG5-9. SMG-1 is the previously discussed kinase that is required for UPF1 phosphorylation (discussed in more details in NMD models below), and no equivalent of it exists in yeast. SMG-8 and SMG-9 regulate the kinase activity of SMG-1 [40] and also do not have a yeast equivalent. The last set of proteins (SMG-5, SMG-6, SMG-7) function after PTC recognition with SMG-5 and SMG-7 acting as a heterodimer that recruits the CCR4-NOT complex hence allowing deadenylation, decapping and exonucleolytic degradation of target RNA [41]. This represent the first branch of transcript degradation while the second branch is carried out by SMG-6 through the activation of its endonuclease activity after its interaction with the SMG-5/SMG-7 complex [42], [43]. The equivalent of SMG-5/SMG-7 in yeast is EBS1 and the equivalent of SMG-6 is NMD4, although the yeast equivalent has lost its endonucleolytic activity

[44]. However, the roles of the yeast equivalents are less clear than for their mammalian counterparts.

NMD SUBSTRATES

NMD targets both natural transcripts and faulty transcripts. Faulty transcripts can arise from mutations on the DNA level, transcriptional errors, PTC-introducing recombination events (*e.g.* VDJ recombination), and splicing errors. On the other hand, natural transcripts that are targeted by NMD are seemingly normal transcripts that harbor specific NMD triggering features. Examples of such features include: upstream open reading frames (uORFs), atypically long 3'-untranslated regions (3′-UTRs) and translation frameshifting sequences. Below examples of NMD substrates in yeast are presented.

Inefficiently spliced introns. The *RPL28* (*CYH2*) gene carries an inefficiently spliced intron [45], [46]. This inefficient splicing causes part of the cytoplasmic transcripts to contain the intron when it should be removed during the maturation process, and translation of these intron-containing transcripts triggers NMD due to the presence of PTCs originating from the intron. The intron-containing substrate of *RPL28* has been long used as an initial test for NMD efficiency due to the large accumulation of the intron-containing isoform when NMD is inactivated [47]. When assessing NMD efficiency, the ratio of the intron-containing *RPL28* transcript vs. mature form is used as a first pass test before carrying out the more technically challenging RNA half-life measurements [48]–[50].

uORFs. *ALR1* and *CPA1* are the best known examples of uORF-based NMD regulated transcripts. *CPA1* encodes the small subunit of carbamoyl phosphate synthetase, which is required for the synthesis of carbamoyl phosphate. The 5' untranslated region (5′-UTR) of the *CPA1* transcript contains a single uORF called the arginine attenuator peptide, which is responsive to the level of arginine in the media. The presence of arginine causes ribosomes to stall at the uORF stop codon, consequently activating NMD to degrade the *CPA1* transcript [51]. This regulation was utilized to identify an extragenic suppressor of the arginine mediated negative regulation of *CPA1* called "*CPAR*" as *UPF1* [51] (also detailed in paper II). *ALR1* encodes a plasma membrane magnesium transporter. The *ALR1* 5'-UTR contains three uORFs of which uORF 3 is primarily responsible for the NMD-mediated degradation of the *ALR1* transcript [52]. Discovery of this regulation of *ALR1* offers an explanation for the observed level of increased PTC read-through in yeast cells mutated in the core NMD proteins [53]–[56],

which is partly due to increased intracellular magnesium concentration as a consequence of the stabilization of the *ALR1* transcript [52].

Atypically long 3'-UTR. 3'-UTRs are typically short in yeast with a range of 50-200 nucleotides [57]. A number of transcripts with a 3'-UTR length of around 350 or higher have been shown to be regulated by NMD [49], [58]. Sometimes the RNAs harboring such long UTRs can be produced with UTR species of variant lengths due to alternative end processing [59], and not all the different UTR length species are targeted by NMD. The *MAK31* RNA for example can have a 3'-UTR of a length of either 200 nucleotides or 920 nucleotides [49], and only the latter is downregulated by NMD. On the other hand, the *MPA43* RNA can have a UTR length of either 300 or 600 nucleotides and both are regulated by NMD [49]. Finally, other transcripts like *PGA1* only has one identified major length of 750 [49] and seems to be constitutively targeted by NMD under standard growth conditions.

Translation frameshifting sequences. *EST3* is the best example of a gene in yeast that undergoes +1 ribosomal frameshifting. The *EST3* gene is arranged into two ORFs. The 5'-most ORF is 276 nucleotides in length and is overlapped with the downstream 270 base pair ORF by a heptameric sequence (CUU-AGU-U) [60]. +1 frameshifting at the heptameric site causes the sequence to be read as CUU-GUU (leucine and valine) while skipping the central A nucleotide [60]. Transcripts that fail to undergo the frameshift will encounter an internal stop codon instead, hence rendering them NMD substrates.

The above are just few examples of the many substrates of NMD in yeast. While the previously mentioned features are a good starting point to assess whether a prospective transcript is NMD regulated or not, the presence of any of these features does not absolutely indicate that a given transcript is NMD regulated. As an example, the *YAP1* and *GCN4* transcripts both contain uORFs in their 5'-UTR. However, both are resistant to NMD due to the presence of a stability element in the transcript that binds Pub1 and circumvents NMD activation [61]. Moreover, the *SSY5* transcript carries a long 3'-UTR [49] and should theoretically be targeted for NMD degradation. However, the transcripts evades NMD by a still undefined mechanism [62].

NMD MODELS

The most important question that needs to be asked is: how are PTCs recognized by the NMD machinery? Two widely accepted models are prevalent to explain NMD substrate selection. The first model is the SURF model [63] and the second model is the faux 3′-UTR model [64].

The SURF model is an exon junction complex (EJC)-dependent model suitable for explaining NMD on spliced transcripts. EJC is a multimeric protein complex of which the core consists of eIF4A3, Y14/MAGOH heterodimer and the BTZ protein [65], [66]. The EJC is deposited on a given transcript in a strictly splicing-dependent manner [65] with the majority being deposited 20-24 nucleotides upstream of exon-exon junctions [67]–[69]. UPF3 is capable of binding to the EJC either in the nucleus or cytoplasm [66] and recruits UPF2 to the complex, hence the EJC serves as an anchor point for NMD proteins [70]. Upon encountering a stop codon upstream of an EJC, the NMD machinery would recognize it as a PTC since normal termination codons are in the last exon, hence would not normally have a downstream EJC. Recognition of the PTC by the NMD machinery depends on the action of the SURF complex, which consists of SMG-1 (complexed with SMG-8 and SMG-9), UPF1, ERF1 and ERF3 [63], [71]. The association of the SURF complex with the PTC is thought to delay translation termination [72] and allows the detection of this aberrant messenger ribonucleoprotein particle (mRNP), hence allowing for the bridging of the SURF complex with the EJC [73] through NMD core proteins interaction. Subsequently, UPF1 is phosphorylated by SMG-1, and the SMG-1 complex along with eRFs dissociate and the ribosome is recycled [63]. Phosphorylated UPF1 helicase activity is further stimulated by the interaction with UPF2 [74]. Furthermore, the phosphorylation of UPF1 causes it to interact with a complex consisting of SMG5, SMG7, SMG6, and protein phosphatase 2A [75]. This complex then dephosphorylates the UPF1 protein. The phosphorylation-dephosphorylation cycle leads to activation of NMD on the transcript and recruitment of enzymes involved in degrading the free 5‘ and 3’ ends [76], [77] that are generated by SMG6 endoribonuclease activity at the PTC [78].

The SURF model can explain how NMD is triggered on spliced transcripts; however, it lacks the ability to explain how non-spliced transcripts are targeted by NMD and is less relevant to yeast because it contains few introns. The faux 3’-UTR model is an alternative model which was derived from

yeast experimental data. It suggests an alternative mechanism for NMD activation based on the length of the 3'-UTR rather than proximity to the EJC. The model hinges on the conjecture that translation termination at a PTC is slow and inefficient compared to a normal termination codon [64], [79], which allows for the assembly of the UPF1 complex. The recognition of PTCs in this model is based on the distance to the poly-A binding protein (PAB1). Normally, PAB1 would bind ERF1 and ERF3 to stimulate translational termination; however, in this case the long distance between PTCs and PAB1 would compromise the ability of PAB1 to promote termination. Instead UPF1 would associate with ERF1 and ERF3 and trigger NMD on the transcript [64]. This simpler model is able to explain NMD on non-spliced transcripts and is the relevant model for explaining NMD in yeast. A condensed visual summary of both models is depicted in Figure 1.

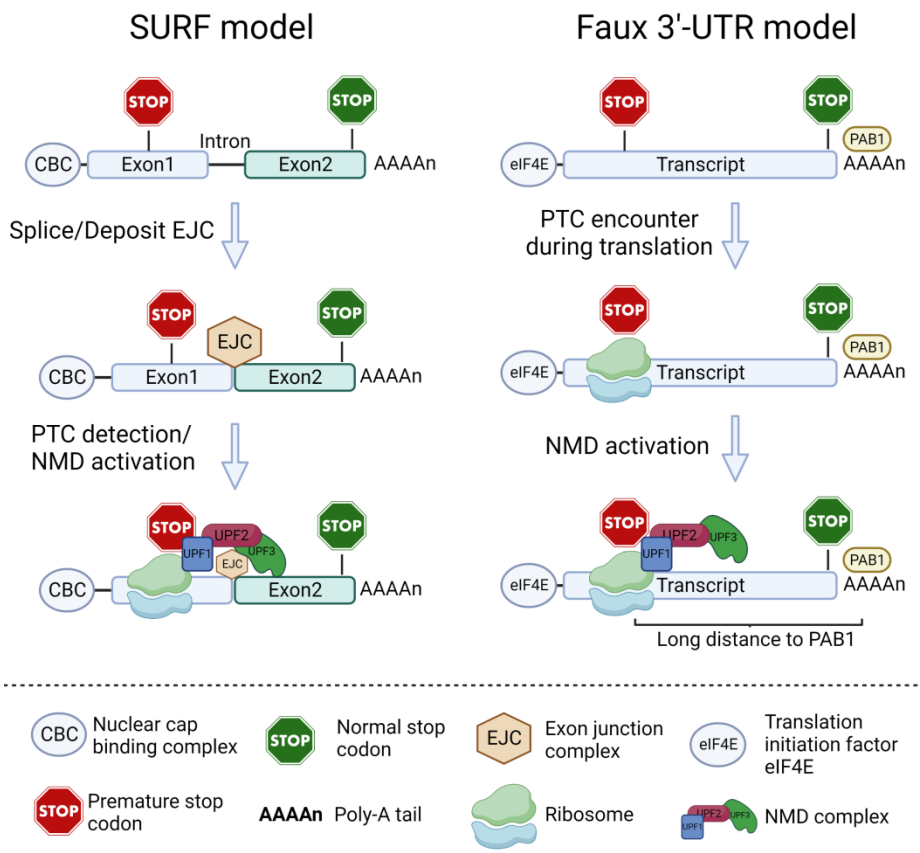

Figure 1 - NMD models. The SURF model is an EJC-dependent model that hinges on anchoring of the NMD factors to the EJC and subsequent NMD activation, should a premature stop codon be encountered. The faux 3'-UTR model proposes that NMD is activated due to the long distance that would separate the premature termination codon from the PAB1 protein.

CRITIQUE TO CURRENT NMD MODELS

The models described for NMD have endured for a relatively long period since their publication. However, new data relating to the hypothesized slow termination, release factor binding to UPF1, PAB1 requirement for PTC discrimination and the transcript resistance to NMD post EJC clearing, has strongly challenged the notion that NMD transcript recognition can be fully captured by the simplistic models above. None of these critiques should be viewed as completely discrediting the models. Rather, they call for expansion and refinement of what is currently known, especially considering that the models above give a reasonable starting point for studying NMD activation.

Slow termination at PTCs. Both models hypothesize that termination at PTC is slow, which could be a cue for NMD activation since it would be detected as an aberrant termination. However, recent data indicated that no difference in ribosomal occupancy was detected between an NMD sensitive 3'-UTR vs. a non-NMD sensitive 3'-UTR [80]. Hence, NMD activation is not dependent on ribosomal stalling. As this result was produced using human cells lysates, it is not clear whether the same holds true for yeast. The ribosomal stalling at the *CPA1* uORF has been shown to be important for NMD activation on that particular transcript at least [81], [82].

Release factors binding to UPF1. Experiments using purified proteins showed that in yeast, UPF1 can bind both ERF1 and ERF3 while UPF2 and UPF3 can bind ERF3 [83], [84]. On the other hand, a study that uses human UPFs found no such interaction for UPF1. Instead an interaction between ERF3a and UPF3B was found [85], which raises questions about the UPF1 recruitment as a part of the SURF complex being mediated by ERFs.

PAB1 requirement for PTC discrimination. Experiments using reporters that lack a poly-A tail, which is a binding site of PAB1, have shown that the poly-A itself had no effect on NMD discrimination [86]. This was done using PTC-containing reporters that had a ribozyme attached downstream of the stop codon causing the transcript to lose its poly-A tails. Then, the RNA half-life of the PTC-containing reporter is compared to a PTC-less equivalent in a yeast strain mutated for *SKI7* to prevent exosomal degradation of the transcripts [86]. The PTC-containing transcript degraded faster in an NMD-dependent manner indicating that the presence of the poly-A tails is not a requirement for NMD recognition of the PTC [86]. While these results

indirectly show that PAB1 is not required for NMD, a direct effect of PAB1 that is independent from its binding to the poly-A tail cannot be ruled out. Hence, another experiment was done in a strain lacking *PAB1* and *RRP6* (an *rrp6* deletion is required to suppress the inviability of *pab1* deletion strains), and still PTC-containing reporters were degraded in an NMD dependent manner. This indicates that PAB1 does not play a role in NMD target recognition [86]. Furthermore, additional data showed that the interaction of UPF1 and ERF3 in yeast is not affect by PAB1. This rules out a situation in which PAB1 outcompetes UPF1 for ERF3 binding, therefore suppressing NMD [87].

Transcript resistance to NMD post EJC clearing. A previously held view on EJC-based NMD was that NMD can only occur during the first round of translation, called the pioneer round, due to the subsequent displacement of the EJC complexes by the elongating ribosomes [88]. Transcripts are initially bound at their cap with nuclear cap binding complex (CBC) consisting of CBP20 and CBP80, which later is exchanged for eIF4E. The pioneering translation round would occur on the CBC-bound form of the transcript, and would cause the triggering of NMD in case of PTCs occurring in proximity of an EJC. However, after this round EJCs are stripped from the transcript by the ribosomes and since the SURF model heavily hinges on an EJC requirement for NMD activation further activation of NMD on a transcript should be theoretically impossible even if it contains a PTC. This has been proven incorrect since eIF4E-bound transcripts have been found to undergo NMD as well [89], [90]. Additional support for the possibility of NMD being triggered in any round of translation came from a yeast study that placed NMD core proteins under an inducible promoter. This study showed that post induction, pre-existing NMD targets got destabilized again, indicating that transcripts do not acquire immunity to NMD after completing the first round of translation [91].

There are attempts at proposing models that unify data from various systems to create a more general framework of PTC recognition. An “extended” faux 3’-UTR proposes that a long 3’-UTR is the primary determinant for NMD and that the EJC involvement is an evolutionary adaptation in mammalian cells rather than being a general NMD recognition mechanism in all organisms [92], explaining how EJC-cleared transcripts can still be targeted for NMD. The model also proposes that it is the physical distance between PAB1 and PTC due to RNA folding is what determines NMD activation

rather than the absolute length of UTR downstream of the PTC [92]. Even with all this data and models, there is currently no model that explains the exact steps required for NMD target recognition that is universal across organisms.

METHODS FOR STUDYING NMD

NMD has been studied using different methods varying from early screens aimed at identifying NMD core proteins using a phenotype that only becomes apparent when NMD is inactivated, such as the *his4-38* mRNA [93], [94], biochemical techniques to determine NMD protein interaction partners [48], RNA-seq to measure difference in transcript steady-state in strains mutated in NMD core proteins vs. wild type strains [95], and RNA half-life measurements in NMD inactivated strains compared to wild type strains [96]. This section will primarily focus on methods that are utilized for the study of RNA half-life as they are most relevant for this thesis. Not all of the methods that will be discussed below have been explicitly used for studying RNA half-life in an NMD context, however, description of them will be added as there is no real technical obstacle for their application for studying NMD. Moreover, example works using the methods in NMD studies have been added when applicable.

RNA half-life measurement is considered the gold standard for identifying NMD substrates since direct measuring of half-lives can unmask indirect effects on RNA steady-state levels due to the true NMD target being upstream in a pathway. The main classes of techniques to study RNA half-life are: transcriptional inhibition either using temperature sensitive alleles of genes involved in RNA metabolism or chemical inhibition; gene control using modified promoters under exogenous control; and metabolic labeling using modified nucleotide analogs.

TRANSCRIPTIONAL INHIBITION

The core concept behind transcriptional inhibition studies is the prevention of formation of new RNA transcripts upon introducing the cells to specific cues, be it a non-permissive condition for normal RNA synthesis/export or chemical inhibitors. This causes the already existing RNA molecules to decay in a time dependent fashion following first order decay kinetics. The amount of RNA remaining at a given time point can be calculated using the following equation: $\text{abundance}(\text{time}) = \text{abundance}_{\text{SteadyState}} * e^{(-\text{kdecay} * \text{time})}$, where kdecay is the decay rate [97]. Fitting the data from a transcriptional inhibition study using nonlinear least squares allows the calculation of a transcript's half-life using the equation half-life = ln(2)/kdecay [97].

An early example in yeast of using transcriptional inhibition used the temperature sensitive *rna1*-ts136 allele [98]–[100], which is required for cytoplasmic RNA export [101]. This allele was created using 1-Methyl-3-nitro-1-nitrosoguanidine (MNNG) mutagenesis and selection for colonies that grew at 23°C but failed to grow at 36°C [98]. Subsequently, RNA and protein accumulation were checked using radioactive labeling. Both showed a sharp decline upon switching to a non-permissive temperature [99]. The sharp decrease in RNA accumulation was likely due to the lack of export of transcripts encoding proteins required for transcription, rather than a direct effect on RNA synthesis. Nonetheless, this allele was used to indirectly infer RNA half-lives by measuring the decline in protein synthesis post temperature shift, since no new transcripts are exported to the cytoplasm for protein synthesis [100]. Now we know that this is unlikely to truly reflect RNA half-life given that both RNA and proteins can have different modifications that would cause the half-lives of these two molecules not to correlate and this is even disregarding the effect of the temperature shift would in itself have on translation. The method based on the *rna1*-ts136 allele was not widely adopted due to its indirect nature and the difficultly in applying it genome-wide.

The temperature sensitive allele of *RPB1* (*rpb1-1*), which encodes the largest subunit of RNA polymerase II, is an alternative to the *rna1*-ts136 allele. The allele was constructed by hydroxylamine hydrochloride mutagenesis of *RPB1* on a plasmid and then reintroduction into yeast while selecting for cells that grow at 24°C but not at 36°C [102]. Unlike the *rna1*-ts136 allele, the effect of the *rpb1-1* allele is directly on RNA synthesis, hence sampling time points

post the temperature shift allows the measurement of RNA half-life directly since no new transcripts are produced and existing transcripts start decaying in a time-dependent manner. Pairing of microarray or RNA-seq data from strains carrying the *rpb1-1* allele allows the monitoring of RNA decay genome-wide. Therefore, due to its direct effect and ease of adaptability to genome-wide studies, using the *rpb1-1* allele became a popular way of studying RNA stability. Chemical inhibitors can also substitute for the *rpb1-1* allele since conceptually they follow the same experimental rationale, but allow the use of any strain rather than temperature-sensitive strains, hence greatly simplifying handling. Thiolutin and 1,10-phenanthroline are two examples of compounds that inhibit RNA polymerase II [103], [104] that are used for RNA stability studies in yeast, while actinomycin D has mostly been used with mammalian cell lines [105]. When using RNA polymerase inhibitors, the shift to a non-permissive temperature that is employed when using *rpb1-1* is exchanged for a chemical addition step to inhibit transcription and then similarly *rpb1-1* a sample time series is used to determine RNA stability. A visual presentation of an example transcriptional inhibition experiment is presented in Figure 2.

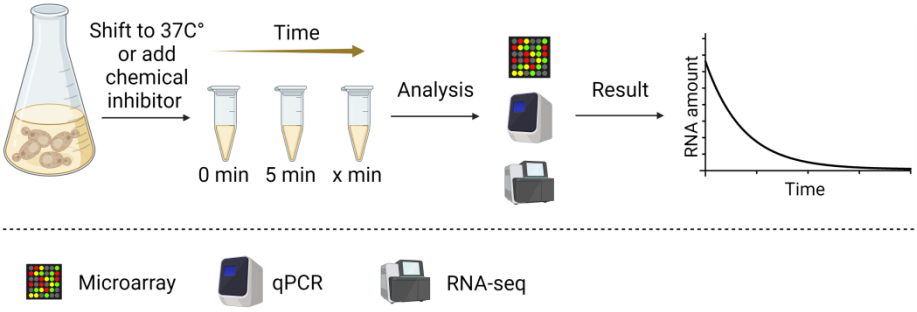

Figure 2 – Example of a transcriptional inhibition experiment. Log phase cells are shifted to a condition that is not permissive to the synthesis of new RNA hence causing the preexisting RNA to decay. Sampling after the shift and analysis using a variety of techniques allows the RNA decay to be measured.

Transcriptional inhibition has been used in multiple studies to determine RNA half-life in yeast. Wang et al used the *rpb1-1* allele in combination with microarrays to determine RNA half-life [106]. Grigull *et al.* used a number of transcriptional inhibitors that included thiolutin and 1,10-phenanthroline and compared the half-lives measurements using these inhibitors vs. *rpb1-1* allele measurements with the study, concluding that thiolutin and

1,10-phenanthroline treatments results were the most similar to *rpb1-1* [107]. Geisberg *et al.* combined RNA-seq with a modified form of transcriptional inhibition termed "anchor away" [108], which depends on the nuclear depletion of a FRB-tagged Rpb1 and anchoring it to FKBP-tagged Rpl13A upon the addition of rapamycin [109]. Lastly, Guan *et al.* used thiolutin in combination with microarrays to measure the RNA half-lives with both a wild type strain and a *upf1* deletion strain [110].

Transcriptional inhibition studies have been widely applied, however, they are not without disadvantages. Comparison of the data derived using the *rpb1-1* allele vs. an alternative method of RNA stability analysis revealed that even at 30^{o}C the *rpb1-1* allele strain suffered from a global decrease in RNA synthesis rate by 2.7 fold [111]. Moreover, half-lives calculated using *rpb1-1* correlated with data sets that were measured under heat shock and osmotic stress rather than unstressed cells [111]. This indicates that the *rpb1-1* method is not suited to measuring RNA half-life that reflects cells' physiological condition, as the method itself causes an activation of the stress response pathway and alters RNA metabolism even under the permissive condition. Similarly, thiolutin and 1,10-phenanthroline interfere with various cellular signaling pathways including: target of rapamycin complex (TORC1), high osmolarity glycerol response (HOG1), and protein kinase C (PKC), as well as causes the induction of processing bodies (P-bodies) specifically by thiolutin [112]. All of these issues offer an explanation to the low correlation seen between various transcriptional inhibition studies and the discrepancies in RNA half-lives when comparing to the values obtained from more recent data sets derived using newer techniques. Hence, the current recommendation is to avoid using transcriptional inhibition as a method to study RNA stability whether using chemical inhibitors or allelic inhibitors.

GENE CONTROL

Gene control is essentially transcriptional inhibition scaled down to a single transcript. The transcript of interest is placed under an exogenously controlled promoter, and a time series is sampled post addition of the transcriptional inhibitor that is specific to the promoter used. To date, the glucose-sensitive *GAL1* promoter and tetracycline repressor (TetR) fusion protein-based promoters have been used for this purpose. This technique has not been used for global assessment of RNA stability due to its poor scalability; however, it does offer the advantage of being less disruptive than global inhibition of RNA synthesis.

The yeast *GAL1* promoter is highly sensitive to the carbon source in the media, being induced approximately 1000-fold in galactose media relative to growth in the presence of glucose [113] and was initially utilized for gene control. Constructs carrying the transcript of interest are fused to the *GAL1* promoter and grown in galactose media followed by the addition of glucose to inhibit the transcription of new transcripts of the target, and samples are taken in a time series following glucose addition. The addition of glucose serves as the equivalent of temperature shift/inhibitor addition discussed above and allows for the quantitation of the degradation of the preexisting transcript prior to glucose addition. *GAL1*-based constructs have been used to assess the half-life of a small set of transcripts [114]–[116], although it must be pointed out that the authors in the previous studies chose to combine gene control with the *rpb1-1* allele to avoid artifacts arising from the switch in the carbon source. A *GAL1*-based gene control system was also used to assess the nuclear escape rate of the *PTC7* intron and calculate the contribution of NMD to this process by using a *upf2* deletion strain [117]. The use of this system is not always desirable as it limits the choice of carbon source to use, moreover, wash steps to remove the previous carbon source adds an additional layer of technical difficulty to those assays especially when handling a larger number of samples. Newer studies tend to move away from *GAL1*-based systems and opt for orthogonal systems.

Newer forms of gene control abandon the use of endogenous promoter in favor of orthogonal systems. TetR fusion proteins are most widely used system for this purpose. TetR is a bacterial transcriptional repressor that binds to an operator named *tetO* with high affinity and specificity to prevent the expression of the tetracycline exporter (*tetA*) in the absence of

tetracycline [118]. Uncontrolled expression of TetA has a significant fitness cost for the carrying bacteria [119], hence TetR is highly efficient in preventing the expression of TetA and only dissociates to allow expression when tetracycline is present. Given these properties, *tetR* from the Tn10 transposon was chosen for the initial use in eukaryotic systems. Gossen *et al.* constructed the first synthetic transcription system for use in mammalian cells line by fusing the *tetR* to a 127 C-terminal fragment of the herpes simplex virus VP16 transcriptional activator and called it a "tetracycline-controlled transactivator" (tTA) [120]. tTA was paired with a minimalistic cytomegalovirus promoter (CMV) in which *tetO* sites were implanted and this gave rise to a system in which the absence of tetracycline (or its analogs) would cause the TetR moity from tTA to bind its operator, effectively recruiting the VP16 activation domain to the minimal promoter and allowing transcription [120]. In the presence of tetracycline, TetR would dissociate from *tetO* and shut down expression causing the expression levels to be controlled by the amount of tetracycline added to the media [120].

The previous is considered an example of a "tet-off" system since the presence of tetracycline causes the expression to be turned off. Following the success of the tet-off system the same group developed a reverse system that used a reverse tTA (rtTA) [121]. The concept is identical to the tet-off system but instead of the wild type *tetR*, a mutant form was created through random mutagenesis that binds to its operator in the presence of tetracycline rather than its absence (rtetR) hence the expression/repression is reversed relative to the tet-off system therefore it is called a "tet-on" system [121]. Since the conception of these two core systems, additional components that fuse TetR to alternative effectors such as transcriptional repressors, *e.g.* fusion of TetR to KRAB transcription repression domain from the human kox1 protein, created a system that causes transcriptional repression in the absence of tetracycline when paired with a fully functional CMV promoter that has been modified with *tetO* sites [122].

The tTA based systems has been adapted to yeast by using a minimal *CYC1* promoter instead of the minimal CMV promoter [123]. The transcriptional repressor fusions (either to tetR or rtetR) were also adapted for yeast by exchanging the mammalian transcriptional repressors for general transcriptional repressor from yeast *e.g. SSN1*, *TUP1* and *SUM1* [124], [125]. Given the simplicity of the systems described, their adaptation for use in RNA decay studies is not surprising since the only requirement is the

addition of tetracycline at the point the transcriptional shutoff is required. An excellent example of use of tetracycline-based systems is the work of Baudrimont *et al.* who took advantage of a rtetR/*SUM1* fusion to create substitutional promoters that retain all wild-type regulation by careful design of the insertion position of the *tetO* site in wild type promoter, thus allowing measuring decay in near wild type contexts [124]. Another example is the work of Dehecq *et al.* that used a tet-off system to show the effect of the deletion of *NMD4* in combination with *EBS1* on the decay on an NMD transcript [48]. The orthogonal systems described are easy to use and are well suited when studying a single transcript; however, they do require time-consuming construction prior to the beginning of the experiment.

METABOLIC LABELING

Metabolic labeling is based on the usage of modified nucleosides that are incorporated into cellular RNA upon exposure. Earlier studies used a radioactive nucleotide like [^{3}H]-adenine or radioactive [^{32}P]-phosphate [126], [127]. This kind of labeling did not get widely adopted for RNA stability studies both due to the effect of radioactive nucleotides in eliciting cellular damage and the laborious experimental protocols involved in using radioactive labels. Radioactive labeling is now largely superseded by non-radioactive modified nucleosides. The most common example of non-radioactive nucleotides used to metabolic labeling is uracil modified with a thiol group such as 4-thiouracil (4-tU) and 4-thiouridine (4-sU).

4-tU is usually used for metabolic labeling in yeast while 4-sU is used in mammalian cell lines. This is due to yeast being able to transport 4-tU efficiently but not 4-sU [128]. 4-tU is activated intracellularly in yeast for RNA incorporation through the pyrimidine salvage pathway by the uracil phosphoribosyltransferase (Fur1) enzyme. On the other hand, mammalian cells do not have an active pyrimidine salvage pathway [129], hence cannot activate 4-tU for RNA incorporation. However, in both cases RNA labeled with the thiol-modified bases can be separated from non-labeled RNA through biotinylation of the thiol group followed by streptavidin separation [130], [131]. In a recent advancement, chemical conversions schemes were conceived that alleviate the need for laborious streptavidin separation [132], [133]. SLAM-seq (thiol(SH)-linked alkylation for the metabolic sequencing of RNA) utilizes iodoacetamide to alkylate the 4-tU or 4-sU leading to formation of derivatives that base pair with guanine during reverse transcription, hence causing that position to be detected as a thymine to cytosine conversion in the RNA-seq output [132]. The thymine to cytosine conversion rate can then be used to calculate RNA half-life if samples are taken in a time series (discussed in paper I). Similarly to SLAM-seq, TUC-seq (ThioUridine-to-Cytidine-Sequencing) causes thymine to cytosine conversion by utilizing an osmium tetroxide (OsO_4) solution in an ammonium chloride buffer (NH_4Cl) buffer that causes the conversion of 4-sU into cytosine [133].

Unlike transcriptional inhibition methods, metabolic labeling allows for more flexibility in experimental design pertaining to how cells are exposed to the label. There are two major schemes that have been utilized for metabolic

labeling namely pulse-chase and approach to equilibrium. A pulse-chase scheme consists of adding the modified nucleotide (label) for a short period of time followed by replacing the labeling media with media containing an unmodified nucleotide followed by time series sampling. The disappearance of the labeled RNA from the samples originating from the chase can then be used to calculate the RNA half-life by using similar first order decay equations to the ones described for transcriptional inhibition. In contrast to pulse-chase, approach to equilibrium follows the increase the labeled fraction following the exposure to the label. A time series sampling is carried out after the introduction of the label and labeled fraction is isolated. The increase in the labeled fraction is dependent on the decay rate and is described in the equation:

$$\text{abundance}(\text{time}) = \text{abundance}_{\text{SteadyState}} * (1 - e^{\wedge[-(\text{kdecay} + \text{kgrowth}) * \text{time}]})$$

where kdecay is the degradation rate and kgrowth is the growth rate that is used to compensate for the dilution of RNA due to active growth [97]. Hence the half-life can be calculated using half-life = ln(2)/(kdecay). An example experiment using metabolic labeling using either a pulse-chase or approach to equilibrium is depicted in Figure 3.

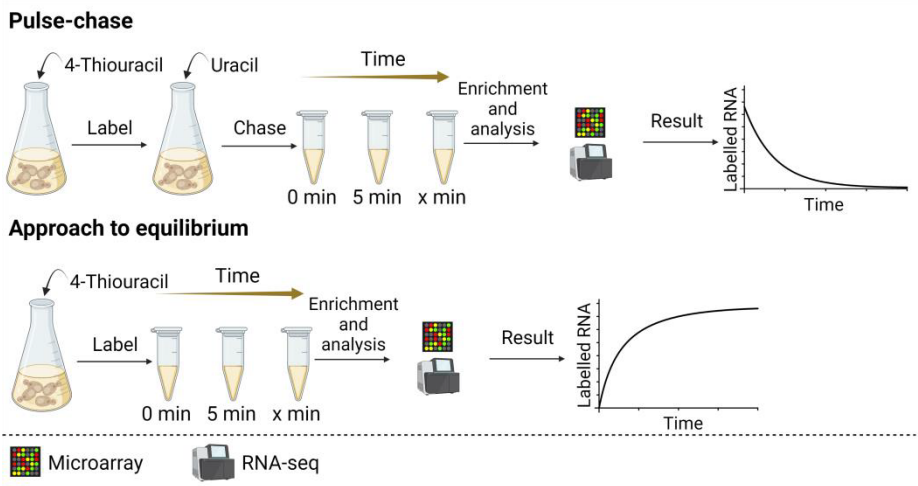

Figure 3 – Example of a metabolic labeling experiment. In a pulse-chase scheme, log phase cells are labeled by 4-thiouracil for a set period followed by chasing by uracil. A sample time series is taken during the chase and analyzed to measure RNA kinetics. In this scheme, the chase causes a reduction in the fraction of labelled RNA. In an approach to equilibrium scheme, the chase step is omitted and instead a time series sample is taken after the addition of the label and analyzed to measure RNA kinetics. In this scheme, the fraction of labelled RNA increases rather than decreases.

The majority of recent RNA half-life datasets from yeast uses some form of metabolic labeling. Munchel *et al.* used 4-tU and streptavidin separation with a pulse-chase scheme and RNA-seq to calculate the decay rate of mRNA at a genome-wide scale under multiple conditions including: replete glucose, glucose starvation, galactose growth, and glucose-to-galactose shift [134]. Neymotin *et al.* used 4-tU and streptavidin separation with an approach to equilibrium scheme and RNA-seq to determine the kinetics of coding and non-coding transcripts [135]. Lastly, our lab wished to avoid the using the laborious streptavidin separation and adapted SLAM-seq using 4-tU for use in yeast in order to compare the RNA half-lives in a wild type strain compared to a NMD mutant strain [136].

Metabolic labeling is currently the most common technique for measuring RNA half-life by virtue of its ease of use and its minimally perturbing nature and scalability. However, the currently used labels are not completely innocuous since 4-tU has been found to modestly induce P-bodies [112]. Nevertheless, we expect that metabolic labeling will continue being the dominant method for RNA half-life in the future especially with the advent of the simplified chemical conversion methods.

NMD AND HUMAN HEALTH

NMD has been heavily studied due to its intimate association with human health. Among others, NMD has important roles in normal developmental processes, lymphocyte development, human disease arising from PTC and resistance to viral disease. NMD involvement in these various processes has prompted the search for NMD-inhibiting drugs that represent a viable therapeutic avenue for exploration depending on the underlying condition. Therefore, methods that facilitate NMD studies are important for studying the processes above.

NMD IN DEVELOPMENTAL PROCESSES

NMD functions in a highly regulated manner rather than being a simple on-or-off switch, making it suitable for processes that require coordinated changes in the stability of a target subset of transcripts such as differentiation processes [137]. Indeed a decrease in NMD efficiency seems to be important for the differentiation of various cell types as was found in neuronal differentiation [138], adipocyte differentiation [139], myogenesis [140] and endoderm differentiation [141]. This section is limited to the two well-studied examples of NMD involvement in differentiation of neuronal cells and myogenesis.

An early hint of the involvement of NMD in neuronal development came from the observation that mutations in *UPF3B* caused intellectual disability in humans [142]. Further evidence came from the identification of a micro-RNA (miR-128) that suppresses a subset of factors involved in NMD, *e.g. UPF1* during mouse brain development [143]. Additional support for the requirement of NMD downregulation for neural cell differentiation was derived from the observation that forced expression of *UPF1* when it should be downregulated inhibited this differentiation [138]. Collectively, these pieces of evidence point to the possibility that particular NMD target transcripts that are required for neuronal differentiation are downregulated when the NMD efficiency in the cell is high. Hence, when differentiation is required, the inhibition of NMD factors by miR-128 causes a drop in the NMD efficiency which allows the normally degraded pro-differentiation factors to accumulate and initiate the process. In line with this hypothesis, *SMAD7*, a differentiation factor for neurons, was found to be an NMD target and that its knockdown prevented the differentiation of neuronal cells [138]. However, the results indicating that NMD is required for neuronal differentiation were challenged by the observation of reduced differentiation ability of neural progenitor cells derived from older mouse embryos when *UPF3B* is downregulated [144]. These results do not have to be necessarily conflicting and could rather mean that the oscillation of NMD efficiency from low to high is required to first commit the cells to the differentiation process and then allows the completion of the process [137]. However, such an explanation needs to be treated with care especially since two different NMD factors were depleted in each study.

Similar to neuronal differentiation, NMD downregulation was shown to be important for myogenesis, and this is dependent on the alteration between the activity ratio of two different RNA decay pathways, NMD and Staufen-mediated decay (SMD) [140]. SMD is a pathway somewhat analogous to NMD and is dependent on UPF1 and STAU-1 [145], [146]. In the SMD pathway, a STAU-1 binding site (SBS) positioned downstream of a stop codon causes the decay of the SBS-carrying transcript if STAU-1 binds to it, and this decay occurs in a UPF1-dependent manner [145], [146]. The SMD and NMD pathways are in competition due to overlap between UPF2 and STAU-1 binding sites on UPF1. Consequently, binding of one will exclude the binding of the other and reduce the efficiency of the corresponding pathway [140]. A study in C2C12 myoblasts (MBs) showed that the efficiency of SMD increases relative to NMD during differentiation to myotubes (MTs) [140]. This has important implications for differentiation because the decay of the *PAX3* mRNA, which is an SMD target, promotes myogenesis [147], while concurrent downregulation of NMD allows the *MYOG* transcript, which encodes a transcription activator required for differentiation [148], to stabilize. This promotes differentiation in a single coordinated action through altering the relative activity of SMD to be greater than NMD [140]. *MYOG* is not the only UPF1-sensitive factor involved in the differentiation. A subsequent study showed additional more complex roles that UPF1 can carry out during this differentiation [149]. The MYOD protein, another important factor for differentiation, is prototypically degraded via the E3-ubiquitin ligase activity of UPF1, while the *MYOD* RNA levels are not affected [149]. The decrease of UPF1 during differentiation [140] allows for the accumulation of MYOD as well as MYOG to promote myogenesis. The previous two examples paint an intricate picture on how NMD can shape developmental processes and it would not be surprising if additional examples of such interactions are identified in other pathways.

V(D)J recombination is a recombination process that occurs in developing lymphocytes and is important for the production of a diverse repertoire of T-cell receptors and antibodies to allow the recognition of antigens originating from diverse pathogens [150]. V(D)J recombination causes the joining of heavy and light chains of antigen receptors, termed variable (V), diversity (D), and joining (J) in a variety of configurations in order to combat pathogens. However, this process does not always lead to a productive arrangement. In fact, two thirds of such events lead to introduction of PTCs through frameshifts, and such PTCs are recognized by NMD [151]. It is

important that the PTC-containing transcripts originating from such non-productive recombination be downregulated as they are expected to impact the efficiency of the immune response through the production of dominant-negative truncated proteins [152]. This is supported by the observation that NMD efficiency is very high in activated B-cells leading to almost complete removal (~95 %) of PTC-containing immunoglobulin transcripts [153]. This function of NMD exemplifies how NMD can act as a "vacuum cleaner" to protect functional proteins from the dominant effects of truncated proteins that arise from propensity of biological systems to have a level of inaccuracy in their function whether it is a process like transcription or as in this case recombination.

HUMAN DISEASES

Single base polymorphisms (SNPs) that lead to the formation of stop codons are estimated to account for ~20 % of disease-related SNPs [154]. This is logical considering that mutations that lead to the alteration of amino acid identity can still produce a sufficiently functional protein while this is less likely for truncated proteins. Considering that NMD is important for the recognition and subsequent degradation of PTCs, it has emerged as an important modifier of disease manifestation. NMD in this situation can act as a double-edged sword. In some cases, it is important that NMD degrades a given transcript to prevent truncated proteins from affecting the healthy form of the proteins hence causing either asymptomatic or mild clinical manifestation of the disease. In other cases, NMD mediated degradation reduces the amount of a truncated protein that still retains some beneficial activity, hence causing a more severe manifestation of the disease. There are many examples of diseases where NMD is involved [155] but, this section will be limited to two examples where NMD has opposing effects on the manifestation of disease.

The typical example discussed in the context of NMD protection of heterozygotes is that of β-thalassemia. Hemoglobin, the oxygen-carrying protein in red blood cells, is composed of two chains of α-globin and two chains of β-globin. β-thalassemia major manifests when two copies of the β-globin gene are mutated with PTCs causing the inability to produce hemoglobin, which results in severe anemia requiring lifelong blood transfusions for survival [156]. In comparison, heterozygote carriers are asymptomatic, indicating that one copy is sufficient to support adequate hemoglobin production. Thus, the disease is considered recessive [156]. When a dominant form of the disease in heterozygous carriers (β-thalassemia intermedia), which displays milder symptoms than β-thalassemia major, was identified, it prompted an investigation into the molecular mechanism of this form of the disease [157]. It was found that in these patients a PTC in the last exon of the β-globin gene was responsible for the phenotype [157]. This PTC cannot be detected by NMD due to its position late in the transcript. This allows for the accumulation of truncated β-globin, which in turn places an increased requirement for proteolytic decay to remove the excess protein leading to reduction in the efficiency of erythropoiesis [157]. This situation exemplifies the importance of NMD recognition of PTCs in the protection of heterozygotes from pathogenesis.

NMD triggered due to PTCs is not always beneficial in a disease context. Unlike β-thalassemia, Duchenne muscular dystrophy (DMD), a disease causing progressive weakness in the muscles due to mutations in the dystrophin gene [158], [159], is worsened when NMD is activated on the DMD transcript [160], [161]. This is supported by the observation that patients carrying a rare NMD-insensitive mutation in the transcript suffer from a milder form of the disease called Becker muscular dystrophy due to the truncated peptide retaining some activity [155], [161]. This situation exemplifies a situation where NMD suppression is beneficial for treatment. In line with this, the recommended treatment for DMD is a small molecule called Ataluren that causes stop codon read-through, hence suppressing NMD [162].

RESISTANCE TO VIRAL DISEASE

NMD serves a role in the restriction of viral RNA replication [163], [164] due to genomic configurations, unique in evolution, that viruses have to adopt. These may resemble NMD-triggering features. Viral genomes are usually compact due to size constraints of packing genomic material into viral capsids. As a consequence, strategies such as alternative splicing and multicistronic expression are utilized by viruses. Such strategies lead to the production of RNAs that have long 3'-UTRs hence causing them to be targeted by NMD [165]. Viral NMD targeting will force the viruses to evolve specific mechanisms to evade NMD and ensure their replication. These mechanisms are broadly categorized into *cis* and *trans* mechanisms. Again, examples of such evasion mechanisms are many; therefore, the section will be limited to one example per category. A more complete list can be found in a recent review [165].

The *cis* mechanisms consist of specific sequences that prevent the function of NMD. The Colorado tick fever virus (CTFV) is a member of the *Reoviridae* virus family and is the causative agent of the Colorado tick fever [166]. CTFV protein VP9 is a 36 kDa product of RNA segment 9 with can be extended into longer form through a stop codon read-through [167]. The location of the stop should normally activate NMD, however, an RNA hairpin structure in the vicinity prevents NMD activation by promoting stop codon readthrough, although the exact details of the mechanism are not known [166], [168].

The *trans* mechanisms are mediated through virally encoded proteins that globally downregulate NMD by targeting NMD proteins rather than shielding specific viral transcripts. The human T-cell lymphotropic virus type 1 (HTLV1) is a positive-sense RNA virus implicated in adult T-cell leukemia and HTLV-1-associated myelopathy (reviewed in [169]). The HTLV-1 tax protein has been shown to prevent NMD degradation of the HTLV-1 transcripts as well as endogenous NMD targets [170], [171]. This action was mediated by the binding of tax to UPF1 and both lowering its binding affinity for RNA and inhibiting its ATPase activity [171]. Viral protein-mediated inhibition of NMD as well as *cis* mechanisms represent an interesting avenue of research that can clarify the regulation of the NMD pathway as well as shed a light on virus-host interactions that may assist in identifying additional therapies.

CONCLUDING REMARKS ON NMD

NMD is a highly complex pathway and much effort has gone into dissecting it. The previous sections discussed NMD from its discovery to the latest advancements in our understanding of the pathway supplemented with various examples. It is without a doubt that many other aspects of NMD regulation remain to be elucidated. I for one look forward to how the intricacies of the NMD pathway will unravel in the future and what new therapies are going to be developed based on an expanded understanding of NMD.

INTRODUCTION – SARS-COV-2

Coronaviruses are human pathogens that cause respiratory symptoms. The first human coronavirus (HCoV) was called strain B814 (after the sample number) and it was isolated from the nasal washing from a male child suffering from the common cold by culturing it using human embryo tracheal tissue [172]. After the isolation of strain B814, multiple other strains were isolated, of which strains 229E and OC43 became the best characterized [173]. The official designation of the name "coronaviruses" came from a letter sent by eight virologists describing the work that had been done in isolating novel respiratory viruses and rationalizing the naming based of images acquired of the viruses with electron microscopy using negative staining [174], [175]. They described the coronaviruses as "…has a characteristic electron microscopic appearance resembling, but distinct from, that of myxoviruses. Particles are more or less rounded in profile; although there is a certain amount of polymorphism, there is also a characteristic "fringe" of projections 200 Å long, which are rounded or petal shaped, rather than sharp or pointed, as in the myxoviruses. This appearance, recalling the solar corona, is shared by mouse hepatitis virus and several viruses recently recovered from man, namely strain B814, 229E and several others." [175].

The discovered coronaviruses were generally understudied due to the mild nature of the illness they cause and the similarity of their symptoms to other viruses. The first highly pathogenic coronavirus was severe acute respiratory syndrome coronavirus (SARS-CoV), also referred to as SARS-CoV-1, which originated from southern China [176]. The SARS-CoV virus caused the first pandemic of the 21st century between 2002 and 2004 [177] and resulted in over 8000 infections with a case fatality ratio of ~10 % [176]. Two additional human coronaviruses (HCoVs) were identified after the initial SARS-CoV pandemic, one being identified from a 7-month infant suffering from bronchiolitis in 2004, and the other from a patient suffering from pneumonia in 2005. These viruses were called HCoV-NL63 (NetherLand 63) and HCoV-HKU1 (Hong Kong University 1), respectively, and both caused mild symptoms in immunocompetent patients [178], [179]. Eight years after the end of the SARS-CoV pandemic, another highly pathogenic coronavirus was isolated from a 60 year old man who died from acute pneumonia in Saudi Arabia [180]. The newly isolated virus was named Middle East respiratory syndrome–related coronavirus (MERS-CoV) and led to an additional 2062

infections (from April 2012 to December 2022) with a case fatality ratio of 36 % [181]. The latest coronavirus to be isolated is severe acute respiratory syndrome coronavirus 2 (SARS‑CoV‑2), which is the causative agent of the recent devastating coronavirus disease 2019 pandemic (COVID-19).

SARS‑CoV‑2 was originally identified in patients suffering from pneumonia caused by an unknown etiology in Wuhan, China [182]. Sequencing of the patients' samples lead to the identification of a viral particle that showed the highest similarity to a bat SARS-like CoV (bat-SL-CoVZC45) that was found in all the tested samples [182]. Subsequent isolation of the virus in mammalian cell lines further strengthened the possibility that a new coronavirus, distinct from previous coronaviruses that caused severe disease in humans had been isolated [182]. Further reconstruction of the viral genome from the sequenced samples solidified the conclusion that a new betacoronavirus from the *Sarbecovirus* subgenus of the *Coronaviridae* family has been identified and is the cause of the pneumonia that the patients in Wuhan were suffering from [182]. In the following sections, a general overview of the SARS-CoV-2 virus will be presented. Topics such as variant tracking and vaccines are considered beyond the scope of what is relevant for this thesis and will not be included.

PROTEINS

SARS-CoV-2 carries one of the largest genome for RNA viruses at ~ 30 kb [182], and its genome encodes 26 proteins divided into three classes: nonstructural proteins (NSPs), structural proteins and accessory proteins [183]. The transcription and translation processes of SARS-COV-2 are complex and many aspects of those processes are still not clear. To simplify the discussion, a brief summary for each protein from the non-structural and structural protein classes is presented first, as they are better characterized than the accessory proteins, before the replication cycle of the virus is described.

NSPS

NSPs range in size from 13 to 1299 amino acids and carry out the functions required for the production of new viral particles. There are a total of 16 NSPs, with many mapped interactions with human proteins [184]. Below a brief summary of the NSPs is presented.

NSP1. NSP1 is the protein encoded closest to the 5'-proximal end of the SARS-COV-2 genome and is expressed early during the infection cycle [185]–[188]. Studies in SARS-COV-1 showed that NSP1 binds the ribosomal 40S subunit and occludes efficient translation of the host's transcripts, while allowing the translation of SARS-COV-1 specific transcripts [188]. Additionally, a template-dependent RNA cleavage function was identified for NSP1, hence representing a two-pronged approach of inhibiting the host translation while promoting the translation of viral transcripts [188]. NSP1 from SARS-COV-2 showed similar activities to those demonstrated for NSP1 from SARS-COV-1. Specifically, SARS-COV-2 NSP1 was shown to bind to the 40S ribosomal subunit to prevent the translation of cellular mRNA [189]. Interestingly, it was shown that SARS-COV-2 transcripts themselves do not completely escape the translation inhibition imposed by NSP1, although the transcripts that carry the viral 5'-UTR are more efficiently translated [189]. This led to the hypothesis that NSP1 expression imposed a limitation on the number of active ribosomes that are capable of translation and this limited pool of active ribosomes would be shifted to the translation of viral transcripts due the translation-enhancing function of their 5'-UTR [189]. Additional work has also shown that SARS-COV-2 NSP1 can further prevent host mRNA translation through preventing nuclear export via its interaction with the mRNA export receptor heterodimer NXF1-NXT1 [190].

NSP2. NSP2 was found to be important for evading cellular immunity. The evasion function of NSP2 is mediated through stabilizing the GIGYF2/4EHP translational repressor complex and causing increased translational repression of type I interferon β [191], [192]. This was further supported by the observation that knockdown of either GIGYF2 or 4EHP strongly reduced viral replication in infected cell lines and increased the detectible amount of type I interferon β [191].

NSP3, NSP4, and NSP6. NSP3, NSP4, and NSP6 are required for the formation of double membrane vesicles (DMVs), in which the viral replicates while being shielded from the immune response [193], [194]. Additionally, NSP3 participates in other functions by virtue of its multi-domain nature, which is not surprising considering that it is the largest protein in SARS-CoV2 with a size of ~ 214 kDa (Mariano et al., 2020). Specifically, NSP3 contains a papain-like protease domain (PLpro) required for the maturation of NSP1, NSP2 and NSP3 [195]. PLpro activity is not limited to the maturation of viral protein but also for immune evasion through the cleavage of the Ubiquitin and Ubiquitin-like Interferon-stimulated Gene 15 protein (ISG15) [194].

NSP5. NSP5 is usually referred to by the name major protease (or 3CLpro). The protease has 11 cleavage sites and is required for the maturation of the vast majority of the NSPs (hence the name major protease) [196]. The major protease is a highly attractive drug target due to the lack of human homologs combined with its high conservation in coronaviruses, hence giving the possibility of developing pan-coronal antivirals [197], [198]. The only currently approved oral anti-viral is called Paxlovid, which was developed by Pfizer. Paxlovid is a combination of nirmatrelvir and ritonavir and is used to target the major protease.

NSP7, NSP8, NSP12 and NSP13. NSP12 in association with NSP7 and NSP8 constitute the viral polymerase complex, also called the replication transcription complex (RTC) [199]. NSP12 is the RNA-dependent RNA polymerase (RdRp) and it possesses some activity without its main co-factors [200]. However, NSP7 and NSP8 are required for full activity [201]. NSP8 functions in the production of RNA primers required for NSP12, while NSP7 acts to stabilize NSP8 [202], [203]. NSP13 is an RNA helicase [204] that interacts with the core polymerase complex (NSP7,NSP8 and NSP12), likely to facilitate copying the RNA sequences that contain secondary structures [183]. Additionally, NSP13 is involved in viral RNA capping due to its RNA 5′-triphosphatase activity [204].

NSP9. NSP9 is an RNA binding protein with an unknown function [205]. It could be speculated that NSP9 has a role in RNA synthesis due to its interaction with NSP8 [206].

NSP10, NSP14 and NSP16. SARS-COV-2 transcripts are capped through four reactions: initially NSP13 produces a pp-RNA through its RNA 5′-triphosphatase activity followed by the addition of guanosine monophosphate to the pp-RNA through an unknown enzyme; finally two methylation steps produce cap0 and cap1 through the action of N^7-methyltransferase (N^7-MTase) activity of NSP14 and 2′-O-methyltransferase (2′-O-MTase) activity of NSP16, respectively [207], [208], [194]. NSP10 is a co-factor required the 2′-O-MTase activity of NSP16 [208]. Additionally, NSP14 participates in maintaining viral genome integrity via its 3’ to 5’ exonuclease activity [209].

NSP15. During the course of SARS-COV-2 replication, double-stranded RNA intermediates are produced due the production of a negative sense copy of the positive sense genome that contains a poly-U stretch [210]. The poly-U stretch is recognized by the MDA5 and subsequently activates the type I interferon (IFN) response [211], [210]. NSP15 is an endoribonuclease that cleaves at uridine residues, hence it functions in immune evasion by reducing the IFN response through limiting the length and abundance of poly-U stretches in the negative sense RNA [210].

STRUCTURAL PROTEINS

Structural proteins constitute the components required for the physical shell of the viral particles as well as to arrange the genome within it. There are four structural proteins in SARS-CoV-2 and they are encoded by the most 3' third of the genome. Below a brief description of each is presented.

Spike protein. The spike protein (S protein) is a transmembrane glycoprotein consisting of two subunits named S1 and S2 [212]. The S protein functions as a trimer and is required for binding the human ACE2 receptor and mediating the fusion of viral particle with the host cells [213], [214]. The S1 subunit contains an N-terminal domain, the receptor-binding domain (RBD), and two subdomains (SD1 and SD2) [213]; the RBD domain is responsible for binding to the ACE2 receptor [214]. The S2 subunit spans the viral membrane and consists of a fusion peptide (FP), two heptad-repeat regions (HR1 and HR2), a transmembrane domain (TM), and a cytoplasmic tail (CT) [213], [214]. The FP and HR1 and HR2 are essential for viral fusion processes [214], which is initiated through the cleavage of the S protein through host proteases, e.g Transmembrane Protease Serine 2 (TMPRSS2) and furin [215].

Nucleocapsid protein. The nucleocapsid protein (N protein) is an RNA binding protein that is highly similar to the N protein from SARS-CoV-1. The primary function of the N protein is binding the viral genome to form a ribonucleoprotein complex that is required for the assembly of new virus and may also facilitate replication and transcription [216]. Additional functions besides the packing function have also been attributed to the N protein. *e.g.* prevention of viral protein degradation and antagonization of IFN [217], [218]. Interestingly the N protein from the mouse hepatitis virus has been shown to inhibit NMD [219] and the SARS-CoV-2 N protein has an interaction with UPF1 [184] hinting at the possibility that it also can inhibit NMD. However, no evidence of this has been found thus far.

Envelope protein. The envelope protein (E protein) is a small protein with a primary function in viral assembly and release [220]. The interactions of the E protein with the M protein and NSP2 and NSP3 facilitates the production of spherical viral particles by inducing the curvature of the ER membrane [221].

Membrane protein. The membrane protein (M protein) like the E protein is primarily involved in viral assembly and release [220]. The M protein also plays a part in viral immune evasion by inhibiting the activation of Nuclear Factor Kappa B [222].

REPLICATION CYCLE

The SARS-COV-2 large RNA genome necessitates coordinated replication and translation to ensure both adequate production of viral particles and the assembly of new viral particles. Significant efforts have gone into dissecting the replication of coronaviruses, which is fortunate considering that these pathways are shared with SARS-COV-2. Here, a brief description of the replication is presented based on a recent review that incorporates currently known aspects of coronavirus replication [223]. It is helpful to refer to the protein description section above while reading this section.

The viral genome is a positive-sense RNA, which allows it to be directly translated upon S protein-mediated entry into the cells. Initially two large poly-protein peptides are translated from the genome, namely pp1a and pp1ab, which encode the NSPs. NSPs originating from pp1a are more abundant than from pp1ab due to the requirement of -1 ribosomal shift to produce pp1ab. The maturation of the NSPs from this pseudo polycistronic arrangement (referred to as pseudo since there are no stop codons between the NSPs as opposed to true polycistronic arrangements from prokaryotes) is carried out by proteolytic cleavage of the poly-protein by NSP3 and NSP5, which auto-catalyze their own excision. Once the initial sets of NSPs are produced, the next step of the viral replication cycle aims at creating a cellular environment conducive to the production of new viral particles. This environment preparation step is multifaceted and requires the coordinated actions of multiple NSPs. NSP1 directs translation toward viral transcripts while NSP2, NSP3 and NSP15 impairs the immune response. During this time, the DMVs that are formed by the action of NSP3, NSP4 and NSP5 are converted into what is called a replication organelle, which are characteristic of coronavirus infection, and function as a site for viral RNA synthesis that is shielded from cellular immune surveillance. All the steps above are crucial for establishing the RTC, which is perhaps the most important complex for the virus. The RTC is responsible for generating negative sense copies of the viral genome, which is then used to create more positive copies of the genome to produce additional NSPs or new viral genomes.

An important aspect to consider is how the structural proteins are expressed. Unlike the NSPs they are not expressed as one large polyprotein and each of them carries its own stop codon, which makes the synthesis of the downstream protein difficult. Coronaviruses utilize a discontinuous

transcription strategy to overcome this problem by generating translation-competent sub-genomic RNA fragments (sgRNAs). The last third of the SARS-COV-2 genome encodes the structural proteins and contains sequences called transcription regulatory sequences (TRSs). Once the RTC reaches such a sequence, it can stop transcription and skip over to another TRSs located at the beginning of the genome called the TRS leader (TRS-L). This skipping action creates a truncated negative sense fragment of the genome that carries the coding sequence of a protein (or multiple proteins) and the 5'-UTR of the genome, which is most likely done to prevent the down regulation of the translation of such fragments by NSP1 once they are converted into a positive sense sgRNAs. It is of note that while these sgRNAs can code for more than one protein, they are expected to act as a mono-cistron due to the presence of stop codons in the structural proteins. Utilizing the trick highlighted here, SARS-CoV-2 can express its full protein repertoire, which leaves one last remaining step: to assemble and release new viral particles. The viral release step is less well studied. However, the structural proteins play a role in assembling new viral particles, which are then released either through exocytosis or lysosomal trafficking. The structure of the viral genome and aspects of viral protein production during replication are presented in Figure 4.

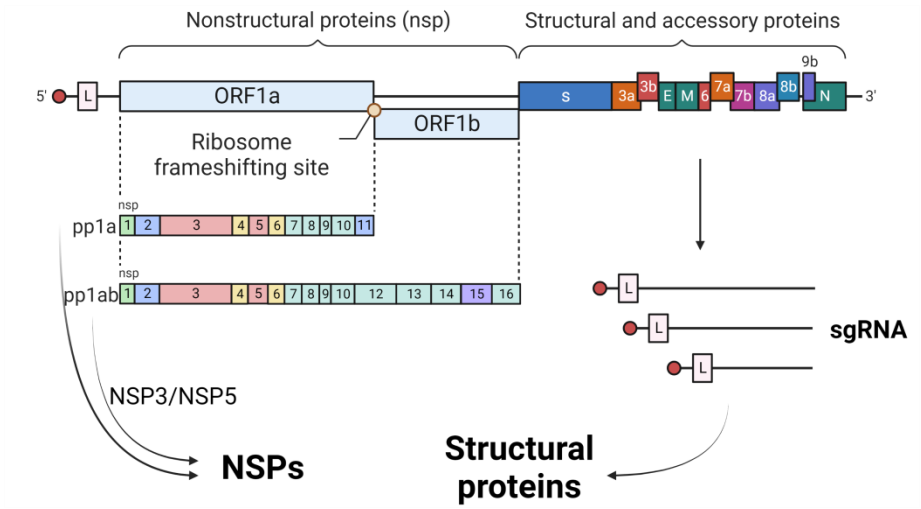

Figure 4 – Genome of SARS-CoV-2 and protein production. A densely coding genome is directly translated into two large poly-proteins via ribosomal frameshifting. The poly-proteins are processed into mature NSPs through the action of NSP3 and NSP5. On the other hand, structural and accessory proteins are produced via sub-genomic RNA.

CORONAVIRUSES AND YEAST

The pandemic caused by SARS-CoV-2 had severe consequences world-wide. It also caused multiple nuisances such as lockdowns and travel restrictions. The scientific community was faced with a unique situation of lack of lab supplies since the majority of the available resources were channeled towards combating SARS-CoV-2. This was particularly disastrous for RNA labs such as ours, since SARS-CoV-2 studies utilized similar reagents to what is used for the study of RNA in other organisms. We decided that instead of completely halting our scientific activity, we would also join the efforts of combating SARS-CoV-2. We reasoned that yeast can easily be used to identify possible inhibitors for NSP5. The past success with expressing active PLpro from SARS-CoV in yeast encouraged us to pursue this further [224]. Prior to initiating this concept, we had been developing reporters for NMD in yeast cells. Some of those reporters were based on an earlier concept that aimed to define changes in NMD based on differential toxicity. However, we quickly abandoned those designs due to their high toxicity in the cells. Nonetheless, it turns out that with minor modifications to that NMD reporter (from thesis paper II) we could turn it into a cleavage reporter for NSP5 instead, which allowed detection of anti-viral compounds with a high sensitivity (detailed in thesis paper III). Subsequent to our initial report of NSP5 expression in yeast [225], other groups expressed NSP5 in yeast and found a unique niche for using yeast in mapping the effect of substituting every position in NSP5 with all possible amino acids [226] or identifying mutations that would cause resistance to nirmatrelvir (the active ingredient of Paxlovid) [227]. Recently, another genetic system in yeast was developed to assess the viral capping protein (NSP14) and was able to identify important functional residues [228]. It is our hope that additional systems will continue being developed using yeast, as the safety and ease of handling would allow faster dissemination of these experimental systems to address key aspects of SARS-CoV-2 biology as well as lead to the identification of new antivirals.

CONCLUDING REMARKS

In conclusion, the pandemic has been a difficult time for us all. It is important to realize that this represents a wakeup call from nature not to neglect observing our surroundings until the situation is severe. While many people consider the pandemic to be over, this is not true. New infections are occurring every day and new variants are likely to emerge, hence, a new severe pandemic is not a matter of "if" rather a matter of "when". On a more positive note, it was great seeing what science can accomplish when resources and minds are aligned to a singular goal.

SUMMARY AND FUTURE PERSPECTIVES

PAPER I

RNA half-life study methodologies vary. In paper I, we show that the methodologies used for RNA half-life studies are not equally reliable and some methods have little biological relevance. We also show that methods based on nucleotide conversion can easily be adapted to yeast and the resulting half-lives calculated from such methods are highly correlated to more recent and robust methods. The output of paper I is summarized as follows:

- SLAM-seq was adapted for use for yeast as an alternative for the laborious pull-down based approaches.
- RNA half-lives were calculated on a genome-wide scale in a wild-type strain and an NMD knockout strain and 225 new transcripts targeted by NMD were identified.
- We show that a pulse-chase approach gives highly correlated half-life values to those obtained with the approach to equilibrium scheme, only when the samples are analyzed by SLAM-seq and not by pull-down approaches.

In the future, we envision that SLAM-seq will be useful for any kind of biological question that aims to investigate aspects of RNA metabolism in yeast. We are aiming to produce additional SLAM-seq data sets from different stress conditions and assess how different stressors affect RNA dynamics. However, additional experiments to refine the method further will be carried out. In our work, we matched our label concentration to another work that used the pulse-chase approach. However, other groups use different label concentrations and it is not clear what effect it would have on our application. Moreover, we used strains that are auxotrophic for uracil, therefore necessitating adding a small amount of uracil during the labeling phase, and the effect of using strains that are prototrophic for uracil on the labeling efficiency has not been investigated. The above points will be addressed experimentally. However, we speculate that an increase in labeling efficiency will be observed both with increasing the label concentration and using a prototrophic strain while omitting uracil from the labeling media. This is likely to require the use of longer read lengths compared to what we

used in our publication to compensate for increased misalignment that stems from the increase of number of mismatches expected for each transcript.

PAPER II

NMD reporters that were used in early studies of NMD depended on rare mutant transcript forms that render them NMD sensitive. In paper II, we constructed NMD sensitive reporters in a systematic manner and identified a design that is able to detect changes in NMD even with strains mutated in NMD factors that have a relatively mild effect on NMD. Additionally, we characterized two of our designed constructs using the yeast deletion library in order to identify other factors that possibly influence NMD efficiency. The output of paper II is summarized as follows:

- Multiple NMD reporters were designed and tested in an NMD proficient background and an NMD defective background. In these reporters, growth was used as the phenotype that is coupled to NMD as this will allow the transfer of the reporter to other laboratories without the requirement of any specialized equipment to read the output.
- We show that the requirement for the amount of different auxotrophic markers to support growth varies. This allows for additional reporter design to be made with a prior expectation of how much background growth is to be predicted.
- The results of characterization of the reporters in the yeast deletion library show that the histidine marker-based reporter performs better than the leucine marker-based reporter in detecting weak NMD factors. Additionally, we provide support to the recent observation that *TMA20* influences NMD and we identify *VPS21* as an influencing factor to the efficiency of NMD only under the condition where NMD is challenged by strongly expressing an NMD-targeted transcript.

Currently, we are investigating how the deletion of *VPS21* might be influencing NMD through analyzing downstream targets such as VMA proteins (check paper II for details). The designed constructs offer the possibility to investigate how NMD is affected by different conditions such as environmental stress or *trans*-acting viral proteins. We speculate that our simple system should facilitate such efforts; moreover, we plan to expand the repertoire of our reporters to include an intron-containing reporter as the

current system cannot detect possible NMD effects that are coupled to splicing. One critique to the current system is the dependence on amino acid drop-out condition, which in itself might have an unexpected effect on NMD. In order to address this we started preparing vectors that carry GFP hence circumventing the need to use auxotrophic markers.

PAPER III

When the pandemic first began, *in vitro* methods were used to identify anti-virals that target the SARS-CoV-2 major protease. However, as knowledge accumulated on the major protease it became clear that these methodologies were not optimized and this resulted in large fraction of false positive compounds that did not have any effect on the major protease *in vivo* (check paper III for details). In paper III, we created a sensitive *in vivo* yeast-based system that allows the identification of anti-virals more reliably and allows compounds that are expected to be cytotoxic to be ruled out earlier as compared to *in vitro* methods. The output of paper III is summarized as follows:

- We show that functional SARS-CoV-2 major protease can be expressed in yeast and phenotypically causes slow growth most likely due to cleavage of yeast proteins.

- We optimized the parameters of a sensitive readout positive selection system for inhibitors of the SARS-CoV-2 major protease in yeast based on a bacterial RNA toxin. The system allows the detection of anti-viral activity of various potential inhibitors by inhibiting the toxicity that results from coupling the major protease to activation of the toxin.

We expect the design of this system to be compatible with other proteases that represent viable therapeutic targets from other pathogens; hence we plan to build similar systems intended for different proteases. Furthermore, the use of fluorescent markers as the direct protease substrate is being considered to replace the toxin. Finally, work-flows to completely convert the system into an automated system are being explored.

ACKNOWLEDGEMENTS

I cannot believe that my PhD journey is at its end. The Lundberg lab has literally been my home (my extra tooth brush is still at the office). During this journey I fluctuated from living like a normal human to living like a bat in the dark to do my experiments. During my time here, I met many people that I will be forever grateful for and I genuinely believe that I neither would be here nor be the person I am today without them.

Per, I will forever be grateful for you being my supervisor. You allowed me to be a "free-range" PhD student and pursue all the experiments that I wanted to do. I believe I grew as a scientist in your lab and this would not have been possible in any other lab. Over the years I grew to appreciate your "punny" sense of humor more than I care to admit. I believe that I definitely knew I wanted to work with you when one day you made a comment about me devouring an entire bar of blue marzipan for breakfast after an overnight experiment. Your reaction was to simply disregard that I looked like hell and have not slept for 24 hours and say "is that bar made out of smurfs, hahahah".

Anne, I remember my first interaction with you was complaining about my grade in your course. I did not imagine that throughout the year you would become one of my best friends in the world. I really appreciate every minute we spent either watching brain-less shows or talking about random topics. I really enjoyed hanging out with **Jonas** as well; you are a kind and gentle soul and an amazing chef. I appreciate all the buckets of food that I got after hanging out with the both of you. Watching Chewie being goofy while hanging out with you two is one my fondest memories.

Markus. Thank you for being my examiner. Your level-headedness and diplomatic approach has really helped me during my studies. I appreciate every moment you took to help me resolve any issues that came up.

Jonas and **Maria,** I really appreciate you being my co-supervisors. While unfortunately the pandemic got in the way of us interacting as much as I would like to, I am happy that you were always one e-mail away when I needed you. Also, a special thanks to all PIs for always being available for random chats. Specifically, I will always remember **Marc** randomly pulling me into his office to tell me new cool discoveries his group made about

lipids. Same goes for **Julie** when we would chat about everything we have done over the weekend.

To my fellow PhD students we went from being complete strangers to very close friends, everything that I have done here would not have been possible without you keeping me sane. **Michelle**, you cackle like a deranged demon and I have grown to recognize that noise as the comfort of having you around and that everything will be alright. I really appreciate all the help and support you offered throughout the years and I will always remember all the silly stuff we did in the lab and I promise we will go see the moo moos again. **Sansan,** you are the warmest soul I ever met and I knew I was going to like you from the day we met. I appreciate everything you have done for me and I will be waiting for that home-made baked bread you promised. **Sunniva**, my lab mate, we have been working side by side closely for years. I really like complaining about everything with you and always appreciated all the support that you have given me throughout the years while I was stressed. I promise we will play the VR all day when this is over. **Alfred**, you are basically a big teddy. We fluctuated between talking about really serious topics to completely meaningless ones and in either case, it was great. It is important that we never barbeque with each other again, since we can only get away with not burning down the cabin once. **Martin**, you have been my office mate for years and it has been great (and yes it is "our office" comrade). It is unfortunate that I could not corrupt you with my bad habits but I am not going to give up. **Emma**, it took a while to take you out of your shell but now look at you, you are as weird as the rest of us, especially with your baking of excessive amounts of pies. **Stefanie**, every time you showed up, I always popped by for a chat, talking to you is such a comfort even when everything is not going ok and I look forward to visiting you and Matthias. **Joana**, I will miss your sassy personality. I am very impressed by the amounts of jokes that you single handedly managed to introduce to the lab (especially those targeting Martin). **Karl**, you were my teaching partner and it was the most fun I had during my teaching, you have a dark sense of humor that I appreciate very much, never ever change it. **Simon**, the cult guru, I am still amazed at the amount of stuff that I learned from you and every time we talked I had to go down a Wikipedia rabbit-hole. **Davide**, you are the definition of fun, I knew that every time you came around nerdy activities will be planned. I really miss having you in Sweden. **Katharina**, you are an amazing baker and I really appreciate how much thought went into making stuff for me since my hate for chocolate makes everything difficult. I really

liked all the gatherings you planned and will miss playing with Logan. **Johanna**, you have always looked out for me and I like when you dropped in every now and then to check up on me. I will miss our drunken discussions about Swedish politics. **Silvana**, you were among my favorite students and I am very happy you started your own PhD journey. I really expect you will do great. **Paula**, you just started in our lab and I really like all your peculiarities. The more I knew about your hobbies (like singing opera) the more I appreciated you. I just wished we had more time to hang out before I finished.

I would like to thank everyone else that I have interacted with during the years and made my stay at Lundberg more enjoyable. I would like to thank **Carmen, James, Johan, Mariana, Ulrika, Bruno, Lars, Peter D, Valida, Elisa, Elena, Kexin, Jessie, Jeanette, Marie** and all the students I have taught throughout the years.

Finally, a very big thank you to my family for your unwavering support and love, we have been through a lot these past years and I know me being away did not make things easier. Now I can finally say that after all the hardships we have been through, it was worth it and we will meet soon.

REFERENCES

[1] A. Nickless, J. M. Bailis, and Z. You, "Control of gene expression through the nonsense-mediated RNA decay pathway," *Cell & Bioscience*, vol. 7, no. 1, p. 26, May 2017, doi: 10.1186/s13578-017-0153-7.

[2] R. Losson and F. Lacroute, "Interference of nonsense mutations with eukaryotic messenger RNA stability.," *Proc Natl Acad Sci U S A*, vol. 76, no. 10, pp. 5134–5137, Oct. 1979.

[3] L. E. Maquat, A. J. Kinniburgh, E. A. Rachmilewitz, and J. Ross, "Unstable β-globin mRNA in mRNA-deficient β0 thalassemia," *Cell*, vol. 27, no. 3, Part 2, pp. 543–553, Dec. 1981, doi: 10.1016/0092-8674(81)90396-2.

[4] J. Hodgkin, A. Papp, R. Pulak, V. Ambros, and P. Anderson, "A new kind of informational suppression in the nematode Caenorhabditis elegans," *Genetics*, vol. 123, no. 2, pp. 301–313, Oct. 1989, doi: 10.1093/genetics/123.2.301.

[5] P. Leeds, S. W. Peltz, A. Jacobson, and M. R. Culbertson, "The product of the yeast UPF1 gene is required for rapid turnover of mRNAs containing a premature translational termination codon.," *Genes Dev.*, vol. 5, no. 12a, pp. 2303–2314, Dec. 1991, doi: 10.1101/gad.5.12a.2303.

[6] P. Leeds, J. M. Wood, B. S. Lee, and M. R. Culbertson, "Gene products that promote mRNA turnover in Saccharomyces cerevisiae.," *Mol Cell Biol*, vol. 12, no. 5, pp. 2165–2177, May 1992.

[7] S. W. Peltz, A. H. Brown, and A. Jacobson, "mRNA destabilization triggered by premature translational termination depends on at least three cis-acting sequence elements and one trans-acting factor.," *Genes Dev.*, vol. 7, no. 9, pp. 1737–1754, Sep. 1993, doi: 10.1101/gad.7.9.1737.

[8] B. Causier, Z. Li, R. De Smet, J. P. B. Lloyd, Y. Van de Peer, and B. Davies, "Conservation of Nonsense-Mediated mRNA Decay Complex Components Throughout Eukaryotic Evolution," *Sci Rep*, vol. 7, no. 1, Art. no. 1, Nov. 2017, doi: 10.1038/s41598-017-16942-w.

[9] A. Celik, R. Baker, F. He, and A. Jacobson, "High-resolution profiling of NMD targets in yeast reveals translational fidelity as a basis for substrate selection," *RNA*, vol. 23, no. 5, pp. 735–748, May 2017, doi: 10.1261/rna.060541.116.

[10] D. Lavysh and G. Neu-Yilik, "UPF1-Mediated RNA Decay—Danse Macabre in a Cloud," *Biomolecules*, vol. 10, no. 7, p. 999, Jul. 2020, doi: 10.3390/biom10070999.

[11] A. Yamashita, T. Ohnishi, I. Kashima, Y. Taya, and S. Ohno, "Human SMG-1, a novel phosphatidylinositol 3-kinase-related protein kinase, associates with components of the mRNA surveillance complex and is

involved in the regulation of nonsense-mediated mRNA decay," *Genes Dev*, vol. 15, no. 17, pp. 2215–2228, Sep. 2001, doi: 10.1101/gad.913001.

[12] A. Grimson, S. O'Connor, C. L. Newman, and P. Anderson, "SMG-1 Is a Phosphatidylinositol Kinase-Related Protein Kinase Required for Nonsense-Mediated mRNA Decay in Caenorhabditis elegans," *Mol Cell Biol*, vol. 24, no. 17, pp. 7483–7490, Sep. 2004, doi: 10.1128/MCB.24.17.7483-7490.2004.

[13] C. Lasalde *et al.*, "Identification and functional analysis of novel phosphorylation sites in the RNA surveillance protein Upf1," *Nucleic Acids Research*, vol. 42, no. 3, pp. 1916–1929, Feb. 2014, doi: 10.1093/nar/gkt1049.

[14] F. Fiorini, D. Bagchi, H. Le Hir, and V. Croquette, "Human Upf1 is a highly processive RNA helicase and translocase with RNP remodelling activities," *Nat Commun*, vol. 6, no. 1, Art. no. 1, Jul. 2015, doi: 10.1038/ncomms8581.

[15] S. Chakrabarti *et al.*, "Molecular mechanisms for the RNA-dependent ATPase activity of Upf1 and its regulation by Upf2," *Mol Cell*, vol. 41, no. 6, pp. 693–703, Mar. 2011, doi: 10.1016/j.molcel.2011.02.010.

[16] Z. Cheng, D. Muhlrad, M. K. Lim, R. Parker, and H. Song, "Structural and functional insights into the human Upf1 helicase core," *EMBO J*, vol. 26, no. 1, pp. 253–264, Jan. 2007, doi: 10.1038/sj.emboj.7601464.

[17] M. Clerici *et al.*, "Unusual bipartite mode of interaction between the nonsense-mediated decay factors, UPF1 and UPF2," *EMBO J*, vol. 28, no. 15, pp. 2293–2306, Aug. 2009, doi: 10.1038/emboj.2009.175.

[18] F. Fiorini, M. Boudvillain, and H. Le Hir, "Tight intramolecular regulation of the human Upf1 helicase by its N- and C-terminal domains," *Nucleic Acids Res*, vol. 41, no. 4, pp. 2404–2415, Feb. 2013, doi: 10.1093/nar/gks1320.

[19] H. Chamieh, L. Ballut, F. Bonneau, and H. Le Hir, "NMD factors UPF2 and UPF3 bridge UPF1 to the exon junction complex and stimulate its RNA helicase activity," *Nat Struct Mol Biol*, vol. 15, no. 1, pp. 85–93, Jan. 2008, doi: 10.1038/nsmb1330.

[20] J. Kadlec, D. Guilligay, R. B. Ravelli, and S. Cusack, "Crystal structure of the UPF2-interacting domain of nonsense-mediated mRNA decay factor UPF1," *RNA*, vol. 12, no. 10, pp. 1817–1824, Oct. 2006, doi: 10.1261/rna.177606.

[21] Y. K. Kim and L. E. Maquat, "UPFront and center in RNA decay: UPF1 in nonsense-mediated mRNA decay and beyond," *RNA*, vol. 25, no. 4, pp. 407–422, Apr. 2019, doi: 10.1261/rna.070136.118.

[22] S. Takahashi *et al.*, "Upf1 potentially serves as a RING-related E3 ubiquitin ligase via its association with Upf3 in yeast," *RNA*, vol. 14, no. 9, pp. 1950–1958, Sep. 2008, doi: 10.1261/rna.536308.

[23] Q. Feng, S. Jagannathan, and R. K. Bradley, "The RNA surveillance factor UPF1 represses myogenesis via its E3 ubiquitin ligase activity," *Mol Cell*, vol. 67, no. 2, pp. 239-251.e6, Jul. 2017, doi: 10.1016/j.molcel.2017.05.034.

[24] F. He, A. H. Brown, and A. Jacobson, "Upf1p, Nmd2p, and Upf3p are interacting components of the yeast nonsense-mediated mRNA decay pathway," *Mol Cell Biol*, vol. 17, no. 3, pp. 1580–1594, Mar. 1997, doi: 10.1128/MCB.17.3.1580.

[25] J. Bao *et al.*, "UPF2-Dependent Nonsense-Mediated mRNA Decay Pathway Is Essential for Spermatogenesis by Selectively Eliminating Longer 3'UTR Transcripts," *PLoS Genet*, vol. 12, no. 5, p. e1005863, May 2016, doi: 10.1371/journal.pgen.1005863.

[26] L. A. Thoren *et al.*, "UPF2 Is a Critical Regulator of Liver Development, Function and Regeneration," *PLoS One*, vol. 5, no. 7, p. e11650, Jul. 2010, doi: 10.1371/journal.pone.0011650.

[27] B. S. Lee and M. R. Culbertson, "Identification of an additional gene required for eukaryotic nonsense mRNA turnover," *Proc Natl Acad Sci U S A*, vol. 92, no. 22, pp. 10354–10358, Oct. 1995, doi: 10.1073/pnas.92.22.10354.

[28] J. Lykke-Andersen, M. D. Shu, and J. A. Steitz, "Human Upf proteins target an mRNA for nonsense-mediated decay when bound downstream of a termination codon," *Cell*, vol. 103, no. 7, pp. 1121–1131, Dec. 2000, doi: 10.1016/s0092-8674(00)00214-2.

[29] G. Serin, A. Gersappe, J. D. Black, R. Aronoff, and L. E. Maquat, "Identification and Characterization of Human Orthologues to Saccharomyces cerevisiae Upf2 Protein and Upf3 Protein (Caenorhabditis elegans SMG-4)," *Molecular and Cellular Biology*, vol. 21, no. 1, pp. 209–223, Jan. 2001, doi: 10.1128/MCB.21.1.209-223.2001.

[30] R. L. Shirley, M. J. Lelivelt, L. R. Schenkman, J. N. Dahlseid, and M. R. Culbertson, "A factor required for nonsense-mediated mRNA decay in yeast is exported from the nucleus to the cytoplasm by a nuclear export signal sequence," *J Cell Sci*, vol. 111 (Pt 21), pp. 3129–3143, Nov. 1998, doi: 10.1242/jcs.111.21.3129.

[31] N. H. Gehring, G. Neu-Yilik, T. Schell, M. W. Hentze, and A. E. Kulozik, "Y14 and hUpf3b form an NMD-activating complex," *Mol Cell*, vol. 11, no. 4, pp. 939–949, Apr. 2003, doi: 10.1016/s1097-2765(03)00142-4.

[32] E. D. Karousis, S. Nasif, and O. Mühlemann, "Nonsense-mediated mRNA decay: novel mechanistic insights and biological impact," *Wiley Interdiscip Rev RNA*, vol. 7, no. 5, pp. 661–682, Sep. 2016, doi: 10.1002/wrna.1357.

[33] J. Kadlec, E. Izaurralde, and S. Cusack, "The structural basis for the interaction between nonsense-mediated mRNA decay factors UPF2

and UPF3," *Nat Struct Mol Biol*, vol. 11, no. 4, pp. 330–337, Apr. 2004, doi: 10.1038/nsmb741.

[34] W.-K. Chan *et al.*, "A UPF3-mediated regulatory switch that maintains RNA surveillance," *Nat Struct Mol Biol*, vol. 16, no. 7, pp. 747–753, Jul. 2009, doi: 10.1038/nsmb.1612.

[35] J. B. Kunz, G. Neu-Yilik, M. W. Hentze, A. E. Kulozik, and N. H. Gehring, "Functions of hUpf3a and hUpf3b in nonsense-mediated mRNA decay and translation," *RNA*, vol. 12, no. 6, pp. 1015–1022, Jun. 2006, doi: 10.1261/rna.12506.

[36] E. Y. Shum *et al.*, "The Antagonistic Gene Paralogs Upf3a and Upf3b Govern Nonsense-Mediated RNA Decay," *Cell*, vol. 165, no. 2, pp. 382–395, Apr. 2016, doi: 10.1016/j.cell.2016.02.046.

[37] Z. Ma *et al.*, "PTC-bearing mRNA elicits a genetic compensation response via Upf3a and COMPASS components," *Nature*, vol. 568, no. 7751, pp. 259–263, Apr. 2019, doi: 10.1038/s41586-019-1057-y.

[38] A. Rossi *et al.*, "Genetic compensation induced by deleterious mutations but not gene knockdowns," *Nature*, vol. 524, no. 7564, Art. no. 7564, Aug. 2015, doi: 10.1038/nature14580.

[39] P. Zhu *et al.*, "Short body length phenotype is compensated by the upregulation of nidogen family members in a deleterious nid1a mutation of zebrafish," *J Genet Genomics*, vol. 44, no. 11, pp. 553–556, Nov. 2017, doi: 10.1016/j.jgg.2017.09.011.

[40] A. Yamashita *et al.*, "SMG-8 and SMG-9, two novel subunits of the SMG-1 complex, regulate remodeling of the mRNA surveillance complex during nonsense-mediated mRNA decay," *Genes Dev*, vol. 23, no. 9, pp. 1091–1105, May 2009, doi: 10.1101/gad.1767209.

[41] B. Loh, S. Jonas, and E. Izaurralde, "The SMG5-SMG7 heterodimer directly recruits the CCR4-NOT deadenylase complex to mRNAs containing nonsense codons via interaction with POP2," *Genes Dev*, vol. 27, no. 19, pp. 2125–2138, Oct. 2013, doi: 10.1101/gad.226951.113.

[42] V. Boehm *et al.*, "SMG5-SMG7 authorize nonsense-mediated mRNA decay by enabling SMG6 endonucleolytic activity," *Nat Commun*, vol. 12, no. 1, Art. no. 1, Jun. 2021, doi: 10.1038/s41467-021-24046-3.

[43] E. Huntzinger, I. Kashima, M. Fauser, J. Saulière, and E. Izaurralde, "SMG6 is the catalytic endonuclease that cleaves mRNAs containing nonsense codons in metazoan," *RNA*, vol. 14, no. 12, pp. 2609–2617, Dec. 2008, doi: 10.1261/rna.1386208.

[44] A. Gilbert and C. Saveanu, "Unusual SMG suspects recruit degradation enzymes in nonsense-mediated mRNA decay," *BioEssays*, vol. 44, no. 5, p. 2100296, 2022, doi: 10.1002/bies.202100296.

[45] U. Swida, E. Thüroff, and N. F. Käufer, "Intron mutations that affect the splicing efficiency of the CYH2 gene of Saccharomyces

cerevisiae," *Mol Gen Genet*, vol. 203, no. 2, pp. 300–304, May 1986, doi: 10.1007/BF00333970.

[46] U. Swida, E. Thüroff, E. Steinert, and N. F. Käufer, "A non-conserved sequence in the 5'region of the CYH2 intron from Saccharomyces cerevisiae controls splicing efficiency of the pre-mRNA," *Yeast*, vol. 4, no. 3, pp. 209–217, Sep. 1988, doi: 10.1002/yea.320040306.

[47] F. He, S. W. Peltz, J. L. Donahue, M. Rosbash, and A. Jacobson, "Stabilization and ribosome association of unspliced pre-mRNAs in a yeast upf1- mutant.," *Proc Natl Acad Sci U S A*, vol. 90, no. 15, pp. 7034–7038, Aug. 1993.

[48] M. Dehecq *et al.*, "Nonsense-mediated mRNA decay involves two distinct Upf1-bound complexes," *EMBO J*, vol. 37, no. 21, p. e99278, Nov. 2018, doi: 10.15252/embj.201899278.

[49] B. W. Kebaara and A. L. Atkin, "Long 3′-UTRs target wild-type mRNAs for nonsense-mediated mRNA decay in Saccharomyces cerevisiae," *Nucleic Acids Res*, vol. 37, no. 9, pp. 2771–2778, May 2009, doi: 10.1093/nar/gkp146.

[50] B. Luke, C. M. Azzalin, N. Hug, A. Deplazes, M. Peter, and J. Lingner, "Saccharomyces cerevisiae Ebs1p is a putative ortholog of human Smg7 and promotes nonsense-mediated mRNA decay," *Nucleic Acids Res*, vol. 35, no. 22, pp. 7688–7697, 2007, doi: 10.1093/nar/gkm912.

[51] F. Messenguy, F. Vierendeels, A. Piérard, and P. Delbecq, "Role of RNA surveillance proteins Upf1/CpaR, Upf2 and Upf3 in the translational regulation of yeast CPA1 gene," *Curr Genet*, vol. 41, no. 4, pp. 224–231, Jul. 2002, doi: 10.1007/s00294-002-0300-4.

[52] M. J. O. Johansson and A. Jacobson, "Nonsense-mediated mRNA decay maintains translational fidelity by limiting magnesium uptake," *Genes Dev*, vol. 24, no. 14, pp. 1491–1495, Jul. 2010, doi: 10.1101/gad.1930710.

[53] K. M. Keeling *et al.*, "Leaky termination at premature stop codons antagonizes nonsense-mediated mRNA decay in S. cerevisiae," *RNA*, vol. 10, no. 4, pp. 691–703, Apr. 2004, doi: 10.1261/rna.5147804.

[54] A. B. Maderazo, F. He, D. A. Mangus, and A. Jacobson, "Upf1p control of nonsense mRNA translation is regulated by Nmd2p and Upf3p," *Mol Cell Biol*, vol. 20, no. 13, pp. 4591–4603, Jul. 2000, doi: 10.1128/MCB.20.13.4591-4603.2000.

[55] W. Wang, K. Czaplinski, Y. Rao, and S. W. Peltz, "The role of Upf proteins in modulating the translation read-through of nonsense-containing transcripts," *EMBO J*, vol. 20, no. 4, pp. 880–890, Feb. 2001, doi: 10.1093/emboj/20.4.880.

[56] Y. Weng, K. Czaplinski, and S. W. Peltz, "Genetic and biochemical characterization of mutations in the ATPase and helicase regions of the Upf1 protein," *Mol Cell Biol*, vol. 16, no. 10, pp. 5477–5490, Oct. 1996, doi: 10.1128/MCB.16.10.5477.

[57] J. H. Graber, G. D. McAllister, and T. F. Smith, "Probabilistic prediction of Saccharomyces cerevisiae mRNA 3′-processing sites," *Nucleic Acids Res*, vol. 30, no. 8, pp. 1851–1858, Apr. 2002.

[58] R. Deliz-Aguirre, A. L. Atkin, and B. W. Kebaara, "Copper tolerance of Saccharomyces cerevisiae nonsense-mediated mRNA decay mutants," *Curr Genet*, vol. 57, no. 6, pp. 421–430, Dec. 2011, doi: 10.1007/s00294-011-0356-0.

[59] K. S. Kim Guisbert, H. Li, and C. Guthrie, "Alternative 3′ Pre-mRNA Processing in Saccharomyces cerevisiae Is Modulated by Nab4/Hrp1 In Vivo," *PLoS Biol*, vol. 5, no. 1, p. e6, Jan. 2007, doi: 10.1371/journal.pbio.0050006.

[60] D. Taliaferro and P. J. Farabaugh, "An mRNA sequence derived from the yeast EST3 gene stimulates programmed +1 translational frameshifting," *RNA*, vol. 13, no. 4, pp. 606–613, Apr. 2007, doi: 10.1261/rna.412707.

[61] M. J. Ruiz-Echevarría and S. W. Peltz, "The RNA binding protein Pub1 modulates the stability of transcripts containing upstream open reading frames," *Cell*, vol. 101, no. 7, pp. 741–751, Jun. 2000, doi: 10.1016/s0092-8674(00)80886-7.

[62] J. Obenoskey, D. R. Lane, A. L. Atkin, and B. W. Kebaara, "Immunity of the Saccharomyces cerevisiae SSY5 mRNA to nonsense-mediated mRNA decay," *Frontiers in Molecular Biosciences*, vol. 1, 2014, Accessed: Jan. 04, 2023. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fmolb.2014.00025

[63] I. Kashima *et al.*, "Binding of a novel SMG-1–Upf1–eRF1–eRF3 complex (SURF) to the exon junction complex triggers Upf1 phosphorylation and nonsense-mediated mRNA decay," *Genes Dev*, vol. 20, no. 3, pp. 355–367, Feb. 2006, doi: 10.1101/gad.1389006.

[64] N. Amrani, R. Ganesan, S. Kervestin, D. A. Mangus, S. Ghosh, and A. Jacobson, "A faux 3'-UTR promotes aberrant termination and triggers nonsense-mediated mRNA decay," *Nature*, vol. 432, no. 7013, pp. 112–118, Nov. 2004, doi: 10.1038/nature03060.

[65] N. H. Gehring, S. Lamprinaki, M. W. Hentze, and A. E. Kulozik, "The Hierarchy of Exon-Junction Complex Assembly by the Spliceosome Explains Key Features of Mammalian Nonsense-Mediated mRNA Decay," *PLOS Biology*, vol. 7, no. 5, p. e1000120, May 2009, doi: 10.1371/journal.pbio.1000120.

[66] H. L. Hir, J. Saulière, and Z. Wang, "The exon junction complex as a node of post-transcriptional networks," *Nat Rev Mol Cell Biol*, vol. 17, no. 1, Art. no. 1, Jan. 2016, doi: 10.1038/nrm.2015.7.

[67] C. Hauer *et al.*, "Exon Junction Complexes Show a Distributional Bias toward Alternatively Spliced mRNAs and against mRNAs Coding for Ribosomal Proteins," *Cell Reports*, vol. 16, no. 6, pp. 1588–1603, Aug. 2016, doi: 10.1016/j.celrep.2016.06.096.

[68] J. Saulière *et al.*, "CLIP-seq of eIF4AIII reveals transcriptome-wide mapping of the human exon junction complex," *Nat Struct Mol Biol*, vol. 19, no. 11, Art. no. 11, Nov. 2012, doi: 10.1038/nsmb.2420.

[69] G. Singh *et al.*, "The Cellular EJC Interactome Reveals Higher-Order mRNP Structure and an EJC-SR Protein Nexus," *Cell*, vol. 151, no. 4, pp. 750–764, Nov. 2012, doi: 10.1016/j.cell.2012.10.007.

[70] I. Aznarez *et al.*, "Mechanism of Nonsense-Mediated mRNA Decay Stimulation by Splicing Factor SRSF1," *Cell Reports*, vol. 23, no. 7, pp. 2186–2198, May 2018, doi: 10.1016/j.celrep.2018.04.039.

[71] A. Yamashita *et al.*, "SMG-8 and SMG-9, two novel subunits of the SMG-1 complex, regulate remodeling of the mRNA surveillance complex during nonsense-mediated mRNA decay," *Genes Dev*, vol. 23, no. 9, pp. 1091–1105, May 2009, doi: 10.1101/gad.1767209.

[72] D. Lavysh and G. Neu-Yilik, "UPF1-Mediated RNA Decay—Danse Macabre in a Cloud," *Biomolecules*, vol. 10, no. 7, p. 999, Jul. 2020, doi: 10.3390/biom10070999.

[73] H. Chamieh, L. Ballut, F. Bonneau, and H. Le Hir, "NMD factors UPF2 and UPF3 bridge UPF1 to the exon junction complex and stimulate its RNA helicase activity," *Nat Struct Mol Biol*, vol. 15, no. 1, pp. 85–93, Jan. 2008, doi: 10.1038/nsmb1330.

[74] T. M. Franks, G. Singh, and J. Lykke-Andersen, "Upf1 ATPase-dependent mRNP disassembly is required for completion of nonsense-mediated mRNA decay," *Cell*, vol. 143, no. 6, pp. 938–950, Dec. 2010, doi: 10.1016/j.cell.2010.11.043.

[75] F. Lejeune, "Nonsense-Mediated mRNA Decay, a Finely Regulated Mechanism," *Biomedicines*, vol. 10, no. 1, p. 141, Jan. 2022, doi: 10.3390/biomedicines10010141.

[76] F. Lejeune, X. Li, and L. E. Maquat, "Nonsense-mediated mRNA decay in mammalian cells involves decapping, deadenylating, and exonucleolytic activities," *Mol Cell*, vol. 12, no. 3, pp. 675–687, Sep. 2003, doi: 10.1016/s1097-2765(03)00349-6.

[77] B. Loh, S. Jonas, and E. Izaurralde, "The SMG5-SMG7 heterodimer directly recruits the CCR4-NOT deadenylase complex to mRNAs containing nonsense codons via interaction with POP2," *Genes Dev*, vol. 27, no. 19, pp. 2125–2138, Oct. 2013, doi: 10.1101/gad.226951.113.

[78] E. Huntzinger, I. Kashima, M. Fauser, J. Saulière, and E. Izaurralde, "SMG6 is the catalytic endonuclease that cleaves mRNAs containing nonsense codons in metazoan," *RNA*, vol. 14, no. 12, pp. 2609–2617, Dec. 2008, doi: 10.1261/rna.1386208.

[79] N. Amrani *et al.*, "Aberrant termination triggers nonsense-mediated mRNA decay," *Biochem Soc Trans*, vol. 34, no. Pt 1, pp. 39–42, Feb. 2006, doi: 10.1042/BST20060039.

[80] E. D. Karousis, L.-A. Gurzeler, G. Annibaldis, R. Dreos, and O. Mühlemann, "Human NMD ensues independently of stable ribosome

stalling," *Nat Commun*, vol. 11, no. 1, Art. no. 1, Aug. 2020, doi: 10.1038/s41467-020-17974-z.

[81] A. Gaba, A. Jacobson, and M. S. Sachs, "Ribosome occupancy of the yeast CPA1 upstream open reading frame termination codon modulates nonsense-mediated mRNA decay," *Mol Cell*, vol. 20, no. 3, pp. 449–460, Nov. 2005, doi: 10.1016/j.molcel.2005.09.019.

[82] Z. Wang, A. Gaba, and M. S. Sachs, "A highly conserved mechanism of regulated ribosome stalling mediated by fungal arginine attenuator peptides that appears independent of the charging status of arginyl-tRNAs," *J Biol Chem*, vol. 274, no. 53, pp. 37565–37574, Dec. 1999, doi: 10.1074/jbc.274.53.37565.

[83] K. Czaplinski *et al.*, "The surveillance complex interacts with the translation release factors to enhance termination and degrade aberrant mRNAs," *Genes Dev*, vol. 12, no. 11, pp. 1665–1677, Jun. 1998.

[84] W. Wang, K. Czaplinski, Y. Rao, and S. W. Peltz, "The role of Upf proteins in modulating the translation read-through of nonsense-containing transcripts," *EMBO J*, vol. 20, no. 4, pp. 880–890, Feb. 2001, doi: 10.1093/emboj/20.4.880.

[85] G. Neu-Yilik *et al.*, "Dual function of UPF3B in early and late translation termination," *EMBO J*, vol. 36, no. 20, pp. 2968–2986, Oct. 2017, doi: 10.15252/embj.201797079.

[86] S. Meaux, A. van Hoof, and K. E. Baker, "Nonsense-mediated mRNA decay in yeast does not require PAB1 or a poly(A) tail," *Mol Cell*, vol. 29, no. 1, pp. 134–140, Jan. 2008, doi: 10.1016/j.molcel.2007.10.031.

[87] S. Kervestin, C. Li, R. Buckingham, and A. Jacobson, "Testing the faux-UTR model for NMD: analysis of Upf1p and Pab1p competition for binding to eRF3/Sup35p," *Biochimie*, vol. 94, no. 7, pp. 1560–1571, Jul. 2012, doi: 10.1016/j.biochi.2011.12.021.

[88] O. Isken and L. E. Maquat, "The multiple lives of NMD factors: balancing roles in gene and genome regulation," *Nat Rev Genet*, vol. 9, no. 9, pp. 699–712, Sep. 2008, doi: 10.1038/nrg2402.

[89] S. Durand and J. Lykke-Andersen, "Nonsense-mediated mRNA decay occurs during eIF4F-dependent translation in human cells," *Nat Struct Mol Biol*, vol. 20, no. 6, pp. 702–709, Jun. 2013, doi: 10.1038/nsmb.2575.

[90] S. C. Rufener and O. Mühlemann, "eIF4E-bound mRNPs are substrates for nonsense-mediated mRNA decay in mammalian cells," *Nat Struct Mol Biol*, vol. 20, no. 6, pp. 710–717, Jun. 2013, doi: 10.1038/nsmb.2576.

[91] A. B. Maderazo, J. P. Belk, F. He, and A. Jacobson, "Nonsense-containing mRNAs that accumulate in the absence of a functional nonsense-mediated mRNA decay pathway are destabilized rapidly upon its restitution," *Mol Cell Biol*, vol. 23, no. 3, pp. 842–851, Feb. 2003, doi: 10.1128/MCB.23.3.842-851.2003.

[92] L. Stalder and O. Mühlemann, "The meaning of nonsense," *Trends in Cell Biology*, vol. 18, no. 7, pp. 315–321, Jul. 2008, doi: 10.1016/j.tcb.2008.04.005.

[93] P. Leeds, J. M. Wood, B. S. Lee, and M. R. Culbertson, "Gene products that promote mRNA turnover in Saccharomyces cerevisiae," *Mol Cell Biol*, vol. 12, no. 5, pp. 2165–2177, May 1992, doi: 10.1128/mcb.12.5.2165-2177.1992.

[94] P. Leeds, S. W. Peltz, A. Jacobson, and M. R. Culbertson, "The product of the yeast UPF1 gene is required for rapid turnover of mRNAs containing a premature translational termination codon," *Genes Dev*, vol. 5, no. 12A, pp. 2303–2314, Dec. 1991, doi: 10.1101/gad.5.12a.2303.

[95] A. Celik, R. Baker, F. He, and A. Jacobson, "High-resolution profiling of NMD targets in yeast reveals translational fidelity as a basis for substrate selection," *RNA*, vol. 23, no. 5, pp. 735–748, May 2017, doi: 10.1261/rna.060541.116.

[96] H. Alalam, J. A. Zepeda-Martínez, and P. Sunnerhagen, "Global SLAM-seq for accurate mRNA decay determination and identification of NMD targets," *RNA*, vol. 28, no. 6, pp. 905–915, Jun. 2022, doi: 10.1261/rna.079077.121.

[97] A. Lugowski, B. Nicholson, and O. S. Rissland, "Determining mRNA half-lives on a transcriptome-wide scale," *Methods*, vol. 137, pp. 90–98, Mar. 2018, doi: 10.1016/j.ymeth.2017.12.006.

[98] L. H. Hartwell, "Macromolecule Synthesis in Temperature-sensitive Mutants of Yeast," *J Bacteriol*, vol. 93, no. 5, pp. 1662–1670, May 1967.

[99] H. T. Hutchison, L. H. Hartwell, and C. S. McLaughlin, "Temperature-sensitive Yeast Mutant Defective in Ribonucleic Acid Production," *J Bacteriol*, vol. 99, no. 3, pp. 807–814, Sep. 1969.

[100] H. Koch and J. D. Friesen, "Individual messenger RNA half lives in Saccharomyces cerevisiae," *Mol Gen Genet*, vol. 170, no. 2, pp. 129–135, Feb. 1979, doi: 10.1007/BF00337787.

[101] J. Becker, F. Melchior, V. Gerke, F. R. Bischoff, H. Ponstingl, and A. Wittinghofer, "RNA1 encodes a GTPase-activating protein specific for Gsp1p, the Ran/TC4 homologue of Saccharomyces cerevisiae," *J Biol Chem*, vol. 270, no. 20, pp. 11860–11865, May 1995, doi: 10.1074/jbc.270.20.11860.

[102] M. Nonet, C. Scafe, J. Sexton, and R. Young, "Eucaryotic RNA polymerase conditional mutant that rapidly ceases mRNA synthesis.," *Mol Cell Biol*, vol. 7, no. 5, pp. 1602–1611, May 1987.

[103] A. Jimenez, D. J. Tipper, and J. Davies, "Mode of action of thiolutin, an inhibitor of macromolecular synthesis in Saccharomyces cerevisiae," *Antimicrob Agents Chemother*, vol. 3, no. 6, pp. 729–738, Jun. 1973, doi: 10.1128/AAC.3.6.729.

[104] G. C. Johnston and R. A. Singer, “RNA synthesis and control of cell division in the yeast S. cerevisiae,” *Cell*, vol. 14, no. 4, pp. 951–958, Aug. 1978, doi: 10.1016/0092-8674(78)90349-5.
[105] T. Wada and A. Becskei, “Impact of Methods on the Measurement of mRNA Turnover,” *Int J Mol Sci*, vol. 18, no. 12, p. 2723, Dec. 2017, doi: 10.3390/ijms18122723.
[106] Y. Wang, C. L. Liu, J. D. Storey, R. J. Tibshirani, D. Herschlag, and P. O. Brown, “Precision and functional specificity in mRNA decay,” *Proc Natl Acad Sci U S A*, vol. 99, no. 9, pp. 5860–5865, Apr. 2002, doi: 10.1073/pnas.092538799.
[107] J. Grigull, S. Mnaimneh, J. Pootoolal, M. D. Robinson, and T. R. Hughes, “Genome-wide analysis of mRNA stability using transcription inhibitors and microarrays reveals posttranscriptional control of ribosome biogenesis factors,” *Mol Cell Biol*, vol. 24, no. 12, pp. 5534–5547, Jun. 2004, doi: 10.1128/MCB.24.12.5534-5547.2004.
[108] H. Haruki, J. Nishikawa, and U. K. Laemmli, “The anchor-away technique: rapid, conditional establishment of yeast mutant phenotypes,” *Mol Cell*, vol. 31, no. 6, pp. 925–932, Sep. 2008, doi: 10.1016/j.molcel.2008.07.020.
[109] J. V. Geisberg, Z. Moqtaderi, X. Fan, F. Ozsolak, and K. Struhl, “Global Analysis of mRNA Isoform Half-Lives Reveals Stabilizing and Destabilizing Elements in Yeast,” *Cell*, vol. 156, no. 4, pp. 812–824, Feb. 2014, doi: 10.1016/j.cell.2013.12.026.
[110] Q. Guan *et al.*, “Impact of Nonsense-Mediated mRNA Decay on the Global Expression Profile of Budding Yeast,” *PLOS Genetics*, vol. 2, no. 11, p. e203, Nov. 2006, doi: 10.1371/journal.pgen.0020203.
[111] M. Sun *et al.*, “Comparative dynamic transcriptome analysis (cDTA) reveals mutual feedback between mRNA synthesis and degradation,” *Genome Res*, vol. 22, no. 7, pp. 1350–1359, Jul. 2012, doi: 10.1101/gr.130161.111.
[112] N. Eshleman, X. Luo, A. Capaldi, and J. R. Buchan, “Alterations of signaling pathways in response to chemical perturbations used to measure mRNA decay rates in yeast,” *RNA*, vol. 26, no. 1, pp. 10–18, Jan. 2020, doi: 10.1261/rna.072892.119.
[113] J. S. Flick and M. Johnston, “Two systems of glucose repression of the GAL1 promoter in Saccharomyces cerevisiae.,” *Mol Cell Biol*, vol. 10, no. 9, pp. 4757–4769, Sep. 1990.
[114] C. A. Beelman and R. Parker, “Differential effects of translational inhibition in cis and in trans on the decay of the unstable yeast MFA2 mRNA,” *J Biol Chem*, vol. 269, no. 13, pp. 9687–9692, Apr. 1994.
[115] C. J. Decker and R. Parker, “A turnover pathway for both stable and unstable mRNAs in yeast: evidence for a requirement for deadenylation.,” *Genes Dev.*, vol. 7, no. 8, pp. 1632–1643, Aug. 1993, doi: 10.1101/gad.7.8.1632.

[116] D. Muhlrad, C. J. Decker, and R. Parker, “Deadenylation of the unstable mRNA encoded by the yeast MFA2 gene leads to decapping followed by 5’-->3’ digestion of the transcript.,” *Genes Dev.*, vol. 8, no. 7, pp. 855–866, Apr. 1994, doi: 10.1101/gad.8.7.855.

[117] M. M. Bonde, S. Voegeli, A. Baudrimont, B. Séraphin, and A. Becskei, “Quantification of pre-mRNA escape rate and synergy in splicing,” *Nucleic Acids Res*, vol. 42, no. 20, pp. 12847–12860, Nov. 2014, doi: 10.1093/nar/gku1014.

[118] W. Hillen and C. Berens, “Mechanisms underlying expression of Tn10 encoded tetracycline resistance,” *Annu Rev Microbiol*, vol. 48, pp. 345–369, 1994, doi: 10.1146/annurev.mi.48.100194.002021.

[119] T. N. Nguyen, Q. G. Phan, L. P. Duong, K. P. Bertrand, and R. E. Lenski, “Effects of carriage and expression of the Tn10 tetracycline-resistance operon on the fitness of Escherichia coli K12,” *Mol Biol Evol*, vol. 6, no. 3, pp. 213–225, May 1989, doi: 10.1093/oxfordjournals.molbev.a040545.

[120] M. Gossen and H. Bujard, “Tight control of gene expression in mammalian cells by tetracycline-responsive promoters,” *Proc Natl Acad Sci U S A*, vol. 89, no. 12, pp. 5547–5551, Jun. 1992, doi: 10.1073/pnas.89.12.5547.

[121] M. Gossen, S. Freundlieb, G. Bender, G. Müller, W. Hillen, and H. Bujard, “Transcriptional Activation by Tetracyclines in Mammalian Cells,” *Science*, vol. 268, no. 5218, pp. 1766–1769, Jun. 1995, doi: 10.1126/science.7792603.

[122] U. Deuschle, W. K. Meyer, and H. J. Thiesen, “Tetracycline-reversible silencing of eukaryotic promoters.,” *Mol Cell Biol*, vol. 15, no. 4, pp. 1907–1914, Apr. 1995.

[123] E. Garí, L. Piedrafita, M. Aldea, and E. Herrero, “A set of vectors with a tetracycline-regulatable promoter system for modulated gene expression in Saccharomyces cerevisiae,” *Yeast*, vol. 13, no. 9, pp. 837–848, Jul. 1997, doi: 10.1002/(SICI)1097-0061(199707)13:9<837::AID-YEA145>3.0.CO;2-T.

[124] A. Baudrimont *et al.*, “Multiplexed gene control reveals rapid mRNA turnover,” *Science Advances*, vol. 3, no. 7, p. e1700006, Jul. 2017, doi: 10.1126/sciadv.1700006.

[125] G. Bellí, E. Garí, L. Piedrafita, M. Aldea, and E. Herrero, “An activator/repressor dual system allows tight tetracycline-regulated gene expression in budding yeast,” *Nucleic Acids Research*, vol. 26, no. 4, pp. 942–947, Feb. 1998, doi: 10.1093/nar/26.4.942.

[126] N. E. Hynes and S. L. Phillips, “Turnover of polyadenylate-containing ribonucleic acid in Saccharomyces cerevisiae,” *J Bacteriol*, vol. 125, no. 2, pp. 595–600, Feb. 1976, doi: 10.1128/jb.125.2.595-600.1976.

[127] D. Herrick, R. Parker, and A. Jacobson, “Identification and comparison of stable and unstable mRNAs in Saccharomyces

cerevisiae," *Mol Cell Biol*, vol. 10, no. 5, pp. 2269–2284, May 1990, doi: 10.1128/mcb.10.5.2269-2284.1990.

[128] C. Miller *et al.*, "Dynamic transcriptome analysis measures rates of mRNA synthesis and decay in yeast," *Mol Syst Biol*, vol. 7, p. 458, Jan. 2011, doi: 10.1038/msb.2010.112.

[129] A. C. Ghosh, M. Shimell, E. R. Leof, M. J. Haley, and M. B. O'Connor, "UPRT, a suicide-gene therapy candidate in higher eukaryotes, is required for Drosophila larval growth and normal adult lifespan," *Sci Rep*, vol. 5, p. 13176, Aug. 2015, doi: 10.1038/srep13176.

[130] M. D. Cleary, C. D. Meiering, E. Jan, R. Guymon, and J. C. Boothroyd, "Biosynthetic labeling of RNA with uracil phosphoribosyltransferase allows cell-specific microarray analysis of mRNA synthesis and decay," *Nat Biotechnol*, vol. 23, no. 2, pp. 232–237, Feb. 2005, doi: 10.1038/nbt1061.

[131] L. Dölken *et al.*, "High-resolution gene expression profiling for simultaneous kinetic parameter analysis of RNA synthesis and decay," *RNA*, vol. 14, no. 9, pp. 1959–1972, Sep. 2008, doi: 10.1261/rna.1136108.

[132] V. A. Herzog *et al.*, "Thiol-linked alkylation of RNA to assess expression dynamics," *Nat Methods*, vol. 14, no. 12, Art. no. 12, Dec. 2017, doi: 10.1038/nmeth.4435.

[133] C. Riml, T. Amort, D. Rieder, C. Gasser, A. Lusser, and R. Micura, "Osmium-Mediated Transformation of 4-Thiouridine to Cytidine as Key To Study RNA Dynamics by Sequencing," *Angew Chem Int Ed Engl*, vol. 56, no. 43, pp. 13479–13483, Oct. 2017, doi: 10.1002/anie.201707465.

[134] S. E. Munchel, R. K. Shultzaberger, N. Takizawa, and K. Weis, "Dynamic profiling of mRNA turnover reveals gene-specific and system-wide regulation of mRNA decay," *Mol Biol Cell*, vol. 22, no. 15, pp. 2787–2795, Aug. 2011, doi: 10.1091/mbc.E11-01-0028.

[135] B. Neymotin, R. Athanasiadou, and D. Gresham, "Determination of in vivo RNA kinetics using RATE-seq," *RNA*, vol. 20, no. 10, pp. 1645–1652, Oct. 2014, doi: 10.1261/rna.045104.114.

[136] H. Alalam, J. A. Zepeda-Martínez, and P. Sunnerhagen, "Global SLAM-seq for accurate mRNA decay determination and identification of NMD targets," *RNA*, vol. 28, no. 6, pp. 905–915, Jun. 2022, doi: 10.1261/rna.079077.121.

[137] S. Jaffrey and M. F. Wilkinson, "Nonsense-mediated RNA decay in the brain: Emerging modulator of neural development and disease," *Nat Rev Neurosci*, vol. 19, no. 12, pp. 715–728, Dec. 2018, doi: 10.1038/s41583-018-0079-z.

[138] C. H. Lou *et al.*, "Posttranscriptional control of the stem cell and neurogenic programs by the nonsense-mediated RNA decay

pathway," *Cell Rep*, vol. 6, no. 4, pp. 748–764, Feb. 2014, doi: 10.1016/j.celrep.2014.01.028.

[139] H. Cho *et al.*, "Staufen1-mediated mRNA decay functions in adipogenesis," *Mol Cell*, vol. 46, no. 4, pp. 495–506, May 2012, doi: 10.1016/j.molcel.2012.03.009.

[140] C. Gong, Y. K. Kim, C. F. Woeller, Y. Tang, and L. E. Maquat, "SMD and NMD are competitive pathways that contribute to myogenesis: effects on PAX3 and myogenin mRNAs," *Genes Dev*, vol. 23, no. 1, pp. 54–66, Jan. 2009, doi: 10.1101/gad.1717309.

[141] C.-H. Lou *et al.*, "Nonsense-Mediated RNA Decay Influences Human Embryonic Stem Cell Fate," *Stem Cell Reports*, vol. 6, no. 6, pp. 844–857, Jun. 2016, doi: 10.1016/j.stemcr.2016.05.008.

[142] P. S. Tarpey *et al.*, "Mutations in UPF3B, a member of the nonsense-mediated mRNA decay complex, cause syndromic and nonsyndromic mental retardation," *Nat Genet*, vol. 39, no. 9, pp. 1127–1133, Sep. 2007, doi: 10.1038/ng2100.

[143] I. G. Bruno *et al.*, "Identification of a microRNA that activates gene expression by repressing nonsense-mediated RNA decay," *Mol Cell*, vol. 42, no. 4, pp. 500–510, May 2011, doi: 10.1016/j.molcel.2011.04.018.

[144] L. A. Jolly, C. C. Homan, R. Jacob, S. Barry, and J. Gecz, "The UPF3B gene, implicated in intellectual disability, autism, ADHD and childhood onset schizophrenia regulates neural progenitor cell behaviour and neuronal outgrowth," *Hum Mol Genet*, vol. 22, no. 23, pp. 4673–4687, Dec. 2013, doi: 10.1093/hmg/ddt315.

[145] Y. K. Kim, L. Furic, L. Desgroseillers, and L. E. Maquat, "Mammalian Staufen1 recruits Upf1 to specific mRNA 3'UTRs so as to elicit mRNA decay," *Cell*, vol. 120, no. 2, pp. 195–208, Jan. 2005, doi: 10.1016/j.cell.2004.11.050.

[146] Y. K. Kim, L. Furic, M. Parisien, F. Major, L. DesGroseillers, and L. E. Maquat, "Staufen1 regulates diverse classes of mammalian transcripts," *EMBO J*, vol. 26, no. 11, pp. 2670–2681, Jun. 2007, doi: 10.1038/sj.emboj.7601712.

[147] J. A. Epstein, P. Lam, L. Jepeal, R. L. Maas, and D. N. Shapiro, "Pax3 inhibits myogenic differentiation of cultured myoblast cells," *J Biol Chem*, vol. 270, no. 20, pp. 11719–11722, May 1995, doi: 10.1074/jbc.270.20.11719.

[148] W. E. Wright, D. A. Sassoon, and V. K. Lin, "Myogenin, a factor regulating myogenesis, has a domain homologous to MyoD," *Cell*, vol. 56, no. 4, pp. 607–617, Feb. 1989, doi: 10.1016/0092-8674(89)90583-7.

[149] Q. Feng, S. Jagannathan, and R. K. Bradley, "The RNA surveillance factor UPF1 represses myogenesis via its E3 ubiquitin ligase activity," *Mol Cell*, vol. 67, no. 2, pp. 239-251.e6, Jul. 2017, doi: 10.1016/j.molcel.2017.05.034.

[150] S. M. Christie, C. Fijen, and E. Rothenberg, "V(D)J Recombination: Recent Insights in Formation of the Recombinase Complex and Recruitment of DNA Repair Machinery," *Frontiers in Cell and Developmental Biology*, vol. 10, 2022, Accessed: Jan. 19, 2023. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fcell.2022.886718

[151] S. Li and M. F. Wilkinson, "Nonsense surveillance in lymphocytes?," *Immunity*, vol. 8, no. 2, pp. 135–141, Feb. 1998, doi: 10.1016/s1074-7613(00)80466-5.

[152] J.-M. Lambert, M. O. Ashi, N. Srour, L. Delpy, and J. Saulière, "Mechanisms and Regulation of Nonsense-Mediated mRNA Decay and Nonsense-Associated Altered Splicing in Lymphocytes," *Int J Mol Sci*, vol. 21, no. 4, p. 1335, Feb. 2020, doi: 10.3390/ijms21041335.

[153] A. Tinguely *et al.*, "Cross Talk between Immunoglobulin Heavy-Chain Transcription and RNA Surveillance during B Cell Development," *Mol Cell Biol*, vol. 32, no. 1, pp. 107–117, Jan. 2012, doi: 10.1128/MCB.06138-11.

[154] M. Mort, D. Ivanov, D. N. Cooper, and N. A. Chuzhanova, "A meta-analysis of nonsense mutations causing human genetic disease," *Hum Mutat*, vol. 29, no. 8, pp. 1037–1047, Aug. 2008, doi: 10.1002/humu.20763.

[155] J. N. Miller and D. A. Pearce, "Nonsense-Mediated Decay in Genetic Disease: Friend or Foe?," *Mutat Res Rev Mutat Res*, vol. 0, pp. 52–64, 2014, doi: 10.1016/j.mrrev.2014.05.001.

[156] J. A. Holbrook, G. Neu-Yilik, M. W. Hentze, and A. E. Kulozik, *NMD and Human Disease*. Landes Bioscience, 2013. Accessed: Jan. 19, 2023. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK6093/

[157] G. W. Hall and S. Thein, "Nonsense Codon Mutations in the Terminal Exon of the β-Globin Gene Are Not Associated With a Reduction in β-mRNA Accumulation: A Mechanism for the Phenotype of Dominant β-Thalassemia," *Blood*, vol. 83, no. 8, pp. 2031–2037, Apr. 1994, doi: 10.1182/blood.V83.8.2031.2031.

[158] A. P. Monaco, R. L. Neve, C. Colletti-Feener, C. J. Bertelson, D. M. Kurnit, and L. M. Kunkel, "Isolation of candidate cDNAs for portions of the Duchenne muscular dystrophy gene," *Nature*, vol. 323, no. 6089, Art. no. 6089, Oct. 1986, doi: 10.1038/323646a0.

[159] A. H. Burghes, C. Logan, X. Hu, B. Belfall, R. G. Worton, and P. N. Ray, "A cDNA clone from the Duchenne/Becker muscular dystrophy gene," *Nature*, vol. 328, no. 6129, pp. 434–437, Aug. 1987, doi: 10.1038/328434a0.

[160] D. A. Pillers *et al.*, "Duchenne/Becker muscular dystrophy: correlation of phenotype by electroretinography with sites of

dystrophin mutations," *Hum Genet*, vol. 105, no. 1–2, pp. 2–9, 1999, doi: 10.1007/s004399900111.

[161] T. P. Kerr, C. A. Sewry, S. A. Robb, and R. G. Roberts, "Long mutant dystrophins and variable phenotypes: evasion of nonsense-mediated decay?," *Hum Genet*, vol. 109, no. 4, pp. 402–407, Oct. 2001, doi: 10.1007/s004390100598.

[162] L. Politano, "Read-through approach for stop mutations in Duchenne muscular dystrophy. An update," *Acta Myol*, vol. 40, no. 1, pp. 43–50, Mar. 2021, doi: 10.36185/2532-1900-041.

[163] G. Balistreri *et al.*, "The host nonsense-mediated mRNA decay pathway restricts Mammalian RNA virus replication," *Cell Host Microbe*, vol. 16, no. 3, pp. 403–411, Sep. 2014, doi: 10.1016/j.chom.2014.08.007.

[164] D. Garcia, S. Garcia, and O. Voinnet, "Nonsense-mediated decay serves as a general viral restriction mechanism in plants," *Cell Host Microbe*, vol. 16, no. 3, pp. 391–402, Sep. 2014, doi: 10.1016/j.chom.2014.08.001.

[165] M. W.-L. Popp, H. Cho, and L. E. Maquat, "Viral subversion of nonsense-mediated mRNA decay," *RNA*, vol. 26, no. 11, pp. 1509–1518, Nov. 2020, doi: 10.1261/rna.076687.120.

[166] S. Napthine, C. Yek, M. L. Powell, T. D. K. Brown, and I. Brierley, "Characterization of the stop codon readthrough signal of Colorado tick fever virus segment 9 RNA," *RNA*, vol. 18, no. 2, pp. 241–252, Feb. 2012, doi: 10.1261/rna.030338.111.

[167] F. M. Jaafar, H. Attoui, P. de Micco, and X. de Lamballerie, "Termination and read-through proteins encoded by genome segment 9 of Colorado tick fever virus," *J Gen Virol*, vol. 85, no. Pt 8, pp. 2237–2244, Aug. 2004, doi: 10.1099/vir.0.80019-0.

[168] S. L. Baker and J. R. Hogg, "A system for coordinated analysis of translational readthrough and nonsense-mediated mRNA decay," *PLoS One*, vol. 12, no. 3, p. e0173980, Mar. 2017, doi: 10.1371/journal.pone.0173980.

[169] M. Matsuoka and K.-T. Jeang, "Human T cell leukemia virus type 1 (HTLV-1) and leukemic transformation: viral infectivity, Tax, HBZ, and therapy," *Oncogene*, vol. 30, no. 12, pp. 1379–1389, Mar. 2011, doi: 10.1038/onc.2010.537.

[170] V. Mocquet *et al.*, "The Human T-Lymphotropic Virus Type 1 Tax Protein Inhibits Nonsense-Mediated mRNA Decay by Interacting with INT6/EIF3E and UPF1," *J Virol*, vol. 86, no. 14, pp. 7530–7543, Jul. 2012, doi: 10.1128/JVI.07021-11.

[171] F. Fiorini *et al.*, "HTLV-1 Tax plugs and freezes UPF1 helicase leading to nonsense-mediated mRNA decay inhibition," *Nat Commun*, vol. 9, no. 1, Art. no. 1, Jan. 2018, doi: 10.1038/s41467-017-02793-6.

[172] D. A. Tyrrell and M. L. Bynoe, "CULTIVATION OF A NOVEL TYPE OF COMMON-COLD VIRUS IN ORGAN CULTURES," *Br*

Med J, vol. 1, no. 5448, pp. 1467–1470, Jun. 1965, doi: 10.1136/bmj.1.5448.1467.
[173] S. H. Myint, "Human Coronavirus Infections," in *The Coronaviridae*, S. G. Siddell, Ed. Boston, MA: Springer US, 1995, pp. 389–401. doi: 10.1007/978-1-4899-1531-3_18.
[174] J. D. Almeida and D. A. Tyrrell, "The morphology of three previously uncharacterized human respiratory viruses that grow in organ culture," *J Gen Virol*, vol. 1, no. 2, pp. 175–178, Apr. 1967, doi: 10.1099/0022-1317-1-2-175.
[175] "Virology: Coronaviruses," *Nature*, vol. 220, no. 5168, Art. no. 5168, Nov. 1968, doi: 10.1038/220650b0.
[176] M. CHAN-YEUNG and R. XU, "SARS: epidemiology," *Respirology*, vol. 8, no. Suppl 1, pp. S9–S14, Nov. 2003, doi: 10.1046/j.1440-1843.2003.00518.x.
[177] J. S. M. Peiris *et al.*, "Coronavirus as a possible cause of severe acute respiratory syndrome," *Lancet*, vol. 361, no. 9366, pp. 1319–1325, Apr. 2003, doi: 10.1016/s0140-6736(03)13077-2.
[178] L. van der Hoek *et al.*, "Identification of a new human coronavirus," *Nat Med*, vol. 10, no. 4, pp. 368–373, Apr. 2004, doi: 10.1038/nm1024.
[179] P. C. Y. Woo *et al.*, "Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia," *J Virol*, vol. 79, no. 2, pp. 884–895, Jan. 2005, doi: 10.1128/JVI.79.2.884-895.2005.
[180] A. M. Zaki, S. van Boheemen, T. M. Bestebroer, A. D. M. E. Osterhaus, and R. A. M. Fouchier, "Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia," *N Engl J Med*, vol. 367, no. 19, pp. 1814–1820, Nov. 2012, doi: 10.1056/NEJMoa1211721.
[181] CSR, "MERS outbreaks," *World Health Organization - Regional Office for the Eastern Mediterranean*. http://www.emro.who.int/health-topics/mers-cov/mers-outbreaks.html (accessed Jan. 27, 2023).
[182] N. Zhu *et al.*, "A Novel Coronavirus from Patients with Pneumonia in China, 2019," *N Engl J Med*, vol. 382, no. 8, pp. 727–733, Feb. 2020, doi: 10.1056/NEJMoa2001017.
[183] A. C. Brant, W. Tian, V. Majerciak, W. Yang, and Z.-M. Zheng, "SARS-CoV-2: from its discovery to genome structure, transcription, and replication," *Cell & Bioscience*, vol. 11, no. 1, p. 136, Jul. 2021, doi: 10.1186/s13578-021-00643-z.
[184] D. E. Gordon *et al.*, "A SARS-CoV-2 protein interaction map reveals targets for drug repurposing," *Nature*, vol. 583, no. 7816, Art. no. 7816, Jul. 2020, doi: 10.1038/s41586-020-2286-9.
[185] C. Huang, K. G. Lokugamage, J. M. Rozovics, K. Narayanan, B. L. Semler, and S. Makino, "SARS Coronavirus nsp1 Protein Induces

Template-Dependent Endonucleolytic Cleavage of mRNAs: Viral mRNAs Are Resistant to nsp1-Induced RNA Cleavage," *PLOS Pathogens*, vol. 7, no. 12, p. e1002433, Dec. 2011, doi: 10.1371/journal.ppat.1002433.

[186] W. Kamitani, C. Huang, K. Narayanan, K. G. Lokugamage, and S. Makino, "A two-pronged strategy to suppress host protein synthesis by SARS coronavirus Nsp1 protein," *Nat Struct Mol Biol*, vol. 16, no. 11, Art. no. 11, Nov. 2009, doi: 10.1038/nsmb.1680.

[187] W. Kamitani *et al.*, "Severe acute respiratory syndrome coronavirus nsp1 protein suppresses host gene expression by promoting host mRNA degradation," *Proceedings of the National Academy of Sciences*, vol. 103, no. 34, pp. 12885–12890, Aug. 2006, doi: 10.1073/pnas.0603144103.

[188] K. G. Lokugamage, K. Narayanan, C. Huang, and S. Makino, "Severe Acute Respiratory Syndrome Coronavirus Protein nsp1 Is a Novel Eukaryotic Translation Inhibitor That Represses Multiple Steps of Translation Initiation," *Journal of Virology*, vol. 86, no. 24, pp. 13598–13608, Dec. 2012, doi: 10.1128/JVI.01958-12.

[189] K. Schubert *et al.*, "SARS-CoV-2 Nsp1 binds the ribosomal mRNA channel to inhibit translation," *Nat Struct Mol Biol*, vol. 27, no. 10, Art. no. 10, Oct. 2020, doi: 10.1038/s41594-020-0511-8.

[190] K. Zhang *et al.*, "Nsp1 protein of SARS-CoV-2 disrupts the mRNA export machinery to inhibit host gene expression," *Science Advances*, vol. 7, no. 6, p. eabe7386, Feb. 2021, doi: 10.1126/sciadv.abe7386.

[191] Z. Xu *et al.*, "SARS-CoV-2 impairs interferon production via NSP2-induced repression of mRNA translation," *Proceedings of the National Academy of Sciences*, vol. 119, no. 32, p. e2204539119, Aug. 2022, doi: 10.1073/pnas.2204539119.

[192] L. Zou, C. Moch, M. Graille, and C. Chapat, "The SARS-CoV-2 protein NSP2 impairs the silencing capacity of the human 4EHP-GIGYF2 complex," *iScience*, vol. 25, no. 7, p. 104646, Jul. 2022, doi: 10.1016/j.isci.2022.104646.

[193] M. M. Angelini, M. Akhlaghpour, B. W. Neuman, and M. J. Buchmeier, "Severe acute respiratory syndrome coronavirus nonstructural proteins 3, 4, and 6 induce double-membrane vesicles," *mBio*, vol. 4, no. 4, pp. e00524-13, Aug. 2013, doi: 10.1128/mBio.00524-13.

[194] W. Yan, Y. Zheng, X. Zeng, B. He, and W. Cheng, "Structural biology of SARS-CoV-2: open the door for novel therapies," *Sig Transduct Target Ther*, vol. 7, no. 1, Art. no. 1, Jan. 2022, doi: 10.1038/s41392-022-00884-5.

[195] Y.-S. Han *et al.*, "Papain-like protease 2 (PLP2) from severe acute respiratory syndrome coronavirus (SARS-CoV): expression, purification, characterization, and inhibition," *Biochemistry*, vol. 44, no. 30, pp. 10349–10359, Aug. 2005, doi: 10.1021/bi0504761.

[196] C.-C. Chen *et al.*, "Overview of antiviral drug candidates targeting coronaviral 3C-like main proteases," *The FEBS Journal*, vol. 288, no. 17, pp. 5089–5121, 2021, doi: 10.1111/febs.15696.
[197] H. Yang *et al.*, "Design of wide-spectrum inhibitors targeting coronavirus main proteases," *PLoS Biol*, vol. 3, no. 10, p. e324, Oct. 2005, doi: 10.1371/journal.pbio.0030324.
[198] L. Zhang *et al.*, "Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α-ketoamide inhibitors," *Science*, vol. 368, no. 6489, pp. 409–412, Apr. 2020, doi: 10.1126/science.abb3405.
[199] R. N. Kirchdoerfer and A. B. Ward, "Structure of the SARS-CoV nsp12 polymerase bound to nsp7 and nsp8 co-factors," *Nat Commun*, vol. 10, no. 1, Art. no. 1, May 2019, doi: 10.1038/s41467-019-10280-3.
[200] D.-G. Ahn, J.-K. Choi, D. R. Taylor, and J.-W. Oh, "Biochemical characterization of a recombinant SARS coronavirus nsp12 RNA-dependent RNA polymerase capable of copying viral RNA templates," *Arch Virol*, vol. 157, no. 11, pp. 2095–2104, Nov. 2012, doi: 10.1007/s00705-012-1404-x.
[201] L. Subissi *et al.*, "One severe acute respiratory syndrome coronavirus protein complex integrates processive RNA polymerase and exonuclease activities," *Proceedings of the National Academy of Sciences*, vol. 111, no. 37, pp. E3900–E3909, Sep. 2014, doi: 10.1073/pnas.1323705111.
[202] M. Bartlam, Y. Xu, and Z. Rao, "Structural proteomics of the SARS coronavirus: a model response to emerging infectious diseases," *J Struct Funct Genomics*, vol. 8, no. 2, pp. 85–97, 2007, doi: 10.1007/s10969-007-9024-5.
[203] I. Imbert, R. Ulferts, J. Ziebuhr, and B. Canard, "SARS Coronavirus Replicative Enzymes: Structures and Mechanisms," *Molecular Biology of the SARS-Coronavirus*, pp. 99–114, Jul. 2009, doi: 10.1007/978-3-642-03683-5_7.
[204] K. A. Ivanov and J. Ziebuhr, "Human coronavirus 229E nonstructural protein 13: characterization of duplex-unwinding, nucleoside triphosphatase, and RNA 5'-triphosphatase activities," *J Virol*, vol. 78, no. 14, pp. 7833–7838, Jul. 2004, doi: 10.1128/JVI.78.14.7833-7838.2004.
[205] M.-P. Egloff *et al.*, "The severe acute respiratory syndrome-coronavirus replicative protein nsp9 is a single-stranded RNA-binding subunit unique in the RNA virus world," *Proc Natl Acad Sci U S A*, vol. 101, no. 11, pp. 3792–3796, Mar. 2004, doi: 10.1073/pnas.0307877101.
[206] G. Sutton *et al.*, "The nsp9 Replicase Protein of SARS-Coronavirus, Structure and Functional Insights," *Structure*, vol. 12, no. 2, pp. 341–353, Feb. 2004, doi: 10.1016/j.str.2004.01.016.

[207] Y. Chen *et al.*, "Functional screen reveals SARS coronavirus nonstructural protein nsp14 as a novel cap N7 methyltransferase," *Proc Natl Acad Sci U S A*, vol. 106, no. 9, pp. 3484–3489, Mar. 2009, doi: 10.1073/pnas.0808790106.

[208] M. Bouvet *et al.*, "In Vitro Reconstitution of SARS-Coronavirus mRNA Cap Methylation," *PLoS Pathog*, vol. 6, no. 4, p. e1000863, Apr. 2010, doi: 10.1371/journal.ppat.1000863.

[209] L. D. Eckerle, X. Lu, S. M. Sperry, L. Choi, and M. R. Denison, "High fidelity of murine hepatitis virus replication is decreased in nsp14 exoribonuclease mutants," *J Virol*, vol. 81, no. 22, pp. 12135–12144, Nov. 2007, doi: 10.1128/JVI.01296-07.

[210] M. Hackbart, X. Deng, and S. C. Baker, "Coronavirus endoribonuclease targets viral polyuridine sequences to evade activating host sensors," *Proc Natl Acad Sci U S A*, vol. 117, no. 14, pp. 8094–8103, Apr. 2020, doi: 10.1073/pnas.1921485117.

[211] J. K. Roth-Cross, S. J. Bender, and S. R. Weiss, "Murine Coronavirus Mouse Hepatitis Virus Is Recognized by MDA5 and Induces Type I Interferon in Brain Macrophages/Microglia," *Journal of Virology*, vol. 82, no. 20, p. 9829, Oct. 2008, doi: 10.1128/JVI.01199-08.

[212] A. Mittal, K. Manjunath, R. K. Ranjan, S. Kaushik, S. Kumar, and V. Verma, "COVID-19 pandemic: Insights into structure, function, and hACE2 receptor recognition by SARS-CoV-2," *PLoS Pathog*, vol. 16, no. 8, p. e1008762, Aug. 2020, doi: 10.1371/journal.ppat.1008762.

[213] L. Duan, Q. Zheng, H. Zhang, Y. Niu, Y. Lou, and H. Wang, "The SARS-CoV-2 Spike Glycoprotein Biosynthesis, Structure, Function, and Antigenicity: Implications for the Design of Spike-Based Vaccine Immunogens," *Front Immunol*, vol. 11, p. 576622, 2020, doi: 10.3389/fimmu.2020.576622.

[214] Y. Huang, C. Yang, X.-F. Xu, W. Xu, and S.-W. Liu, "Structural and functional properties of SARS-CoV-2 spike protein: potential antivirus drug development for COVID-19," *Acta Pharmacol Sin*, vol. 41, no. 9, pp. 1141–1149, Sep. 2020, doi: 10.1038/s41401-020-0485-4.

[215] M. Hoffmann *et al.*, "SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor," *Cell*, vol. 181, no. 2, pp. 271-280.e8, Apr. 2020, doi: 10.1016/j.cell.2020.02.052.

[216] Q. Huang *et al.*, "Structure of the N-terminal RNA-binding domain of the SARS CoV nucleocapsid protein," *Biochemistry*, vol. 43, no. 20, pp. 6059–6063, May 2004, doi: 10.1021/bi036155b.

[217] X. Lu, J. Pan, J. Tao, and D. Guo, "SARS-CoV nucleocapsid protein antagonizes IFN-β response by targeting initial step of IFN-β induction pathway, and its C-terminal region is critical for the antagonism," *Virus Genes*, vol. 42, no. 1, pp. 37–45, Feb. 2011, doi: 10.1007/s11262-010-0544-x.

[218] Q. Wang *et al.*, "Interactions of SARS coronavirus nucleocapsid protein with the host cell proteasome subunit p42," *Virol J*, vol. 7, p. 99, May 2010, doi: 10.1186/1743-422X-7-99.
[219] M. Wada, K. G. Lokugamage, K. Nakagawa, K. Narayanan, and S. Makino, "Interplay between coronavirus, a cytoplasmic RNA virus, and nonsense-mediated mRNA decay pathway," *Proc Natl Acad Sci U S A*, vol. 115, no. 43, pp. E10157–E10166, Oct. 2018, doi: 10.1073/pnas.1811675115.
[220] S. Satarker and M. Nampoothiri, "Structural Proteins in Severe Acute Respiratory Syndrome Coronavirus-2," *Arch Med Res*, vol. 51, no. 6, pp. 482–491, Aug. 2020, doi: 10.1016/j.arcmed.2020.05.012.
[221] D. Schoeman and B. C. Fielding, "Coronavirus envelope protein: current knowledge," *Virol J*, vol. 16, no. 1, p. 69, May 2019, doi: 10.1186/s12985-019-1182-0.
[222] X. Fang *et al.*, "The membrane protein of SARS-CoV suppresses NF-kappaB activation," *J Med Virol*, vol. 79, no. 10, pp. 1431–1439, Oct. 2007, doi: 10.1002/jmv.20953.
[223] P. V'kovski, A. Kratzel, S. Steiner, H. Stalder, and V. Thiel, "Coronavirus biology and replication: implications for SARS-CoV-2," *Nat Rev Microbiol*, vol. 19, no. 3, Art. no. 3, Mar. 2021, doi: 10.1038/s41579-020-00468-6.
[224] M. Frieman *et al.*, "Yeast based small molecule screen for inhibitors of SARS-CoV," *PLoS One*, vol. 6, no. 12, p. e28479, 2011, doi: 10.1371/journal.pone.0028479.
[225] H. Alalam *et al.*, "A Genetic Trap in Yeast for Inhibitors of SARS-CoV-2 Main Protease," *mSystems*, vol. 6, no. 6, pp. e01087-21, Nov. 2021, doi: 10.1128/mSystems.01087-21.
[226] J. M. Flynn *et al.*, "Comprehensive fitness landscape of SARS-CoV-2 Mpro reveals insights into viral resistance mechanisms," *eLife*, vol. 11, p. e77433, Jun. 2022, doi: 10.7554/eLife.77433.
[227] J. Ou *et al.*, "A yeast-based system to study SARS-CoV-2 Mpro structure and to identify nirmatrelvir resistant mutations," *Res Sq*, p. rs.3.rs-1942964, Aug. 2022, doi: 10.21203/rs.3.rs-1942964/v1.
[228] M. Y. Ornelas, A. Y. Thomas, L. I. Johnson Rosas, R. O. Scoville, and A. P. Mehta, "Synthetic Platforms for Characterizing and Targeting of SARS-CoV-2 Genome Capping Enzymes," *ACS Synth. Biol.*, vol. 11, no. 11, pp. 3759–3771, Nov. 2022, doi: 10.1021/acssynbio.2c00359.