Seminars in Cell and Developmental Biology 154 (2024) 114–122

Contents lists available at ScienceDirect

Seminars in Cell and Developmental Biology

journal homepage: www.elsevier.com/locate/semcdb

Review

# Making sense of mRNA translational "noise"

Yuanhui Mao, Shu-Bing Qian *

*Division of Nutritional Sciences, Cornell University, Ithaca, NY 14853, USA*

ARTICLE INFO

*Keywords:*
Translation
Ribosome
Alternative initiation
Reinitiation
Frameshifting
Start codon
Stop codon

ABSTRACT

The importance of translation fidelity has been apparent since the discovery of genetic code. It is commonly believed that translation deviating from the main coding region is to be avoided at all times inside cells. However, ribosome profiling and mass spectrometry have revealed pervasive noncanonical translation. Both the scope and origin of translational "noise" are just beginning to be appreciated. Although largely overlooked, those translational "noises" are associated with a wide range of cellular functions, such as producing unannotated protein products. Furthermore, the dynamic nature of translational "noise" is responsive to stress conditions, highlighting the beneficial effect of translational "noise" in stress adaptation. Mechanistic investigation of translational "noise" will provide better insight into the mechanisms of translational regulation. Ultimately, they are not "noise" at all but represent a signature of cellular activities under pathophysiological conditions. Deciphering translational "noise" holds the therapeutic and diagnostic potential in a wide spectrum of human diseases.

## 1. Introduction

A central pillar of biology is the high fidelity of gene expression in all organisms. The textbook knowledge tells us the error rate of DNA replication is on the order of $10^{-9}$, RNA transcription $10^{-6}$, and mRNA translation $10^{-4}$ [1–5]. As translation consumes a lion's share of cellular energy, the relatively poor fidelity of protein synthesis is a bit surprising. On average, there is one error for every ten thousand peptide bonds made during mRNA translation. Moreover, the ribosome does not strictly adhere to the code instruction, resulting in alternative start codon selection [6], ribosome frameshifting [7], as well as stop codon readthrough [8]. A growing body of evidence suggests that decreasing the translation fidelity could be beneficial for biological processes [9]. For example, tRNA misacylation by methionine increases up to 10-fold upon exposing cells to oxidative stress [10]. Intriguingly, the resultant "mis-translation" products are more resistant to oxidative stress. Ribosome frameshifting is deleterious in general as premature termination often leads to non-functional translational products. However, some programmed ribosome frameshifting events give rise to new functional products without changing the mRNA template. The expanded coding potential from limited genomes is typical in viruses, including SARS-CoV-2 [11]. Although flexible translation increases proteome diversity and complexity, neither the scope nor the underlying mechanism is fully understood.

To better understand translational "noise", it is important to revisit what we have learned regarding translation processes in eukaryotic cells. Instead of giving a comprehensive review covering all the translation steps, we focus on ribosome decoding from start to stop codons. Eukaryotic mRNA translation typically starts with the attachment of the small ribosomal subunit to the 5'cap of mRNA, followed by the migration of the 48 S preinitiation complex along the 5' untranslated region (5'UTR) until a start codon is encountered [12]. Once a start codon is selected by the scanning ribosome, the engagement of the initiator tRNA is followed by 60 S joining. The efficiency of start codon recognition can be influenced by the codon context as well as many initiation factors, although the precise mechanism remains elusive. It is now well-established that, in addition to the canonical AUG, many non-AUG codons could serve as potential initiation sites [13]. The 80 S ribosome assembly at a particular start codon primarily defines the subsequence open reading frame (ORF). Prior to elongation commitment, the transition of the assembled 80 S from an initiation complex to an elongation complex remains incompletely understood. Since the initiator tRNA is directly placed in the P site of the ribosome, the initiating 80 S at the start codon has no tRNAs at the E and A sites. Additional quality control mechanisms likely exist to ensure the reading frame fidelity from the start codon. Translation elongation commences once the amino acid-charged tRNA is delivered into the ribosomal A site. Upon the formation of the first peptide bond, the ribosome enters the elongation

* Corresponding author.
*E-mail address:* sq38@cornell.edu (S.-B. Qian).

https://doi.org/10.1016/j.semcdb.2023.03.004
Received 16 March 2022; Received in revised form 9 March 2023; Accepted 9 March 2023
Available online 14 March 2023

cycle by decoding the nucleotide triplets (codon) in a successive and nonoverlapping fashion [14]. During elongation, the ribosome does not move at a constant speed but rather proceeds in a stop-and-go traffic manner. Both *cis* sequence elements and *trans* regulatory factors contribute to the variations of elongation speed [15,16]. When the ribosome reaches a stop codon, termination occurs via the concerted action of release factors [17], leading to peptide release, tRNA dissociation, and ribosome separation. In some cases, the empty ribosome remains associated with mRNA and could start a second round of translation from upstream or downstream start codons, leading to translational events in 3'UTRs [18]. It is clear that multiple ORFs are embedded within a single messenger, forming the foundation of translational "noise".

## 2. Detecting translational "noise"

The human transcriptome contains tens of thousands of potential ORFs besides the annotated ORFs [19]. Many sequence features, such as length, codon usage, and evolutionary conservation, are often used to predict the coding potential of putative ORFs [20–22]. However, positive prediction of ORFs by no means indicates active translation. Some ORFs may undergo selective translation in response to stress conditions or in different cell types. On the other hand, advances in high-throughput sequencing and proteomic technologies reveal myriad translational signals that cannot be attributed to the known ORFs. It is thus crucial to distinguish translational "noise" from technical noise.

### 2.1. Ribosome profiling

Ribosome profiling (Ribo-seq) is a powerful technique that provides a snapshot of global translation by sequencing ribosome-protected mRNA fragments (RPFs) [23]. Ribo-seq has become one of the commonly used methods to identify ORFs with active translation (Fig. 1A). The general procedure involves nuclease digestion, size selection, cDNA library construction, and deep sequencing. A typical ribosome footprint is ~29 nucleotide long and different read length could indicate distinct ribosome status during translation. For example, ribosomes with empty A sites give rise to short footprints of 20–22 nt [24], whereas disome-protected fragments could reach up to 65 nt [25]. Nuclease digestion largely contributes to the length heterogeneity that mainly occurs at the 3' end of RPFs [26], leaving the 5' end nucleotide as a robust marker for inferring ribosome positions. When aligned to ORFs, RPFs show a strong three-nucleotide periodicity. This so-called footprint phasing is often used to gauge the quality of Ribo-seq data sets. However, different Ribo-seq methodology often results in varied 5' end read accuracy. For instance, circularization after reverse transcription is known to introduce untemplated nucleotide addition, whereas direct RNA ligation suffers from sequence-dependent biases. As a result, a substantial number of out-of-frame footprints are attributed to technical artifacts.

RPFs tend to accumulate at start and stop codons, forming the boundary of coding regions (CDS). Individual transcripts, however, are typified by peaks and valleys at various CDS positions, blurring the ORF boundary. As a result, computational tools are often needed to identify ORFs with active translation. For example, a predictor based on the length of RPFs (fragment length organization similarity score, FLOSS) showed a high sensitivity to distinguish 5' UTR with active translation from other non-coding RNAs or regions [27]. RiboTaper employed a multitaper approach to characterize the 3-nt periodicity of footprints [28], revealing > 600 mRNAs with active translation in the 5'UTR. RIBO-TISH [29] and RiboCode [30] directly compared ribosome densities in all three different frames of the putative ORF using a non-parametric Wilcoxon rank-sum test, with the null hypothesis that all three frames have similar footprint reads. Since translation termination

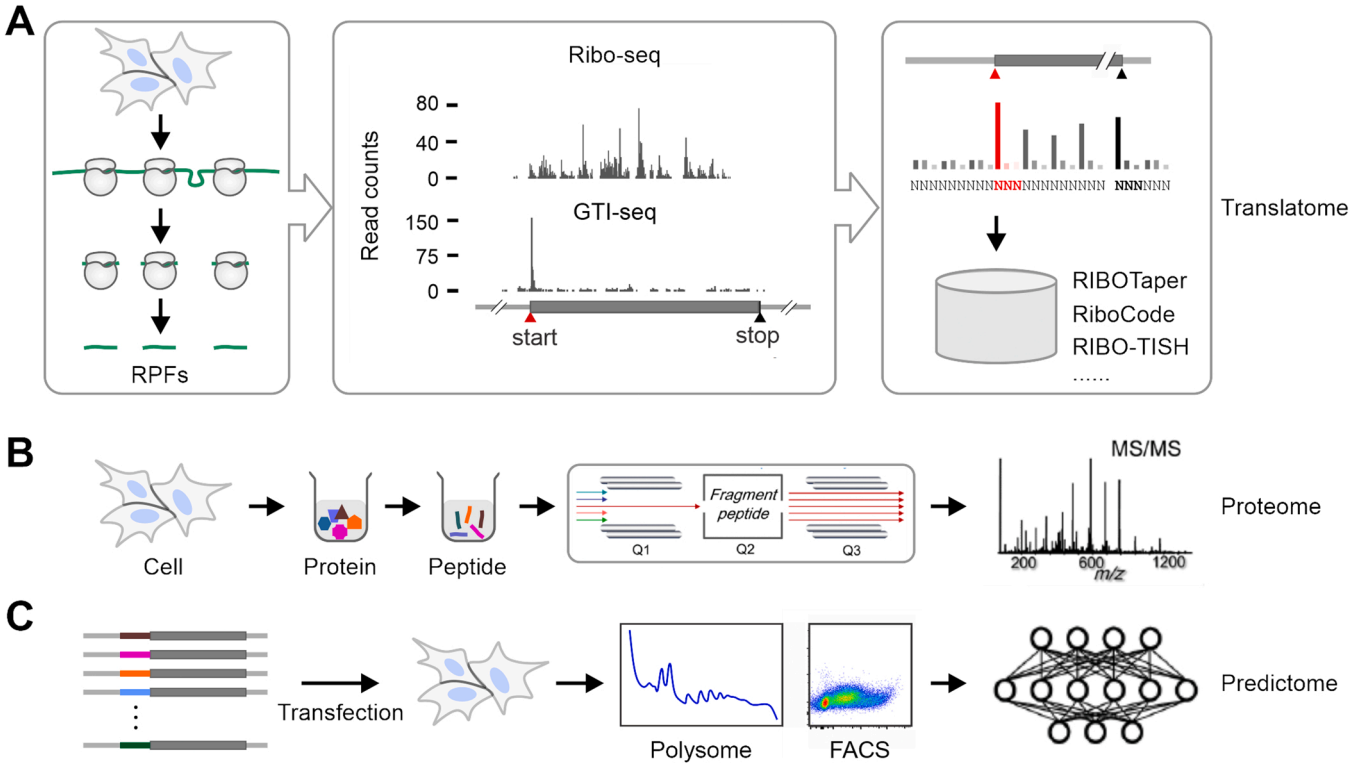


**Fig. 1.** Methods of detecting translational "noise". (A) A comprehensive flowchart integrating ribosome profiling (Ribo-seq) and global translation initiation sequencing (GTI-seq) in identification of alternative ORFs with active translation. The left panel shows the simplified procedure of Ribo-seq. The middle panel shows a typical distribution of footprints on a mRNA generated by Ribo-seq and GTI-seq. The right panel shows computational tools to identify alternative ORFs. (B) A pipeline to identify cellular proteome by mass spectrometry. (C) The deep learning model based on the data from massively paralleled reporter assays to predict the translatability of RNA sequences (predictome).

results in a drop of footprint reads after the stop codon, a ribosome release score helps distinguish ORF translation from the nonspecific background reads [31]. However, such boundary becomes ambiguous for transcripts with overlapping ORFs. Several studies integrated multiple parameters such as footprint length, density, and 3-nt periodicity to achieve ORF classification [32,33]. It appears that the 3-nt periodicity contributes the most to classification efficiency [32].

Despite the success of Ribo-seq in assessing ORF translation, identifying hidden ORFs or neoORFs requires determination of authentic translation initiation sites (TIS-seq). One strategy is to capture initiating ribosomes assembled at start codons. Translation inhibitors like harringtonine act on the first round of peptide bond formation [34]. A short incubation period allows the run-off of elongating ribosomes, thereby specifically halting ribosomes at all possible start sites. Since harringtonine binds to the free 60 S subunit first before it blocks the peptidyltransferase center, it is uncertain whether the presence of this compound affects physiological start codon selection. The compound lactimidomycin, by contrast, binds to the empty E-site of the initiating ribosomes [35], enabling quantitative capture of the 80 S ribosomes assembled at the start codons. Indeed, profiling of RPFs marked by lactimidomycin (GTI-seq) revealed remarkable enrichment of ribosomes at the annotated start codon [36,37]. This strategy has been successfully used to uncover the coding potential of the human cytomegalovirus genome [37], as well as the dynamics of initiating ribosomes in cancer [38]. It is anticipated that profiling of terminating ribosomes, once established, would greatly improve the accuracy of ORF identification.

### 2.2. Mass spectrometry

Ribosome occupancy is associated with but is not a direct measure of translation efficiency (TE). For ORFs with low TE, it is challenging to distinguish active translation from background noise. Mass spectrometry (MS) is the main methodology that enables direct identification of translational products (Fig. 1B). For shotgun proteomics, proteins are digested using proteases and the resulting peptides are separated and identified using tandem mass spectrometry (MS/MS). Surprisingly, on average, 75% of spectra analyzed in an MS experiment cannot be identified [39]. Many of these spectra appear to be of high quality and are likely to have originated from peptides. Since search engines were built upon theoretical spectra derived from user-defined protein sequences, it is possible that the current proteome database is far from complete.

Of course, not all the translational events supported by Ribo-seq could be confirmed by shotgun MS detection, partly because some encoded proteins are short-lived and of low abundance [40,41]. This is particularly true for non-canonical translational events. A promising way to detect those short-lived proteins produced by non-canonical translation is to monitor newly synthesized proteins by incorporating puromycin [42], puromycin analogs such as O-propargyl-puromycin (OPP) [43], or noncanonical amino acid such as *L*-azidohomoalanine (AHA) [44] into the nascent peptides. The metabolically labeled nascent peptides are subsequently biotinylated followed by enrichment via streptavidin beads. To increase the sensitivity of MS detection, protein extraction can be optimized by minimizing degradation mediated by endogenous peptidase and protease activity [45]. Additionally, some enrichment methods can be applied to increase the amount of the low-abundance proteins. For instance, acetic acid precipitation is shown to enrich the small and low abundance proteins [40,46]. Finally, although trypsin is commonly used for most MS samples to digest the proteins, a combination of Lys-C and trypsin could improve the sample quality and increase the number of detected peptides [47].

Besides the sample preparation, MS data analysis is crucial in peptide detection. It is important to compile a reference library comprising all coding potential of the entire genome, generating *in silico* six-frame translation covering both sense and anti-sense strands [48]. Not surprisingly, searching MS data against such inflated library introduces many false positives, and the searching process becomes extremely time-consuming. Incorporating transcriptome from RNA-seq and translatome from Ribo-seq into the analysis pipeline is a common strategy [49,50]. Since MS-based proteomics measures steady-state levels of proteins present in the sample, it is anticipated that the scope of proteome is smaller than that of the translatome. Surprisingly, even when all the reading frames are considered, the identity of millions of spectra remains elusive. One possibility is that many peptides are subjected to unknown post-translational modifications. Another possibility is that our current understanding of translational diversity is still limited. Without knowing the scope of non-canonical translation, how translational "noises" contribute to the proteome landscape remains a fundamental knowledge gap.

### 2.3. Reporter assay

While millions of spectra from MS have no corresponding ribosome occupancy, a large portion of the translation signals from Ribo-seq cannot be validated by MS proteomics. Independent assays are thus required to confirm whether the "noisy" reads represent true translational events. A reporter assay is relatively straightforward albeit with low throughput to validate putative translation signals [45]. Instead of measuring the products of endogenous genes, a reporter assay is more suitable for identification of sequence elements responsible for active translation. For example, a frame-shifting reporter can be constructed by placing a downstream ORF at different reading frames without upstream stop codons [51]. Although luciferase or GFP reporters are commonly used as ORFs, a growing number of short peptides are included in the reporter assays. Typical short ORFs include HiBit, a nanoluciferase peptide, and SIINFEKL, a peptide that can be detected by a monoclonal antibody 25D1 once presented by the mouse MHC class I molecule H-2 $K^b$ on the cell surface [52]. With superior detection sensitivity, those short ORFs greatly increase the versatility of reporter assays in validating translational "noise".

Besides translational validation, massively paralleled reporter assays (MPRAs) can be designed to explore translational "noise" in an unbiased manner (Fig. 1 C). By introducing variation to a particular sequence element, MPRAs contain fully degenerate or endogenous sequence fragments. As a result, the MPRA library size can be orders of magnitude larger than the number of genomic examples. MPRAs have been successfully applied to uncover sequence elements in mediating cap-independent translation [53], start codon selection [54], as well as frameshifting [55]. One limitation of MPRA is the length of the sequence that determines the scope of the sequence coverage. For example, a 10-mer random sequence gives rise to $10^6$, whereas a 50-mer generates $> 10^{30}$ sequence combinations. Given the size and the unbiased nature of a MPRA library, MPRA data is more suitable than endogenous transcript data for deep learning and training predictive models of translational "noise" [56]. Besides random sequences, a multiplex assay containing pooled full-length 5' UTRs uncovered the molecular consequences of 5' UTR mutations in human prostate cancer [57].

## 3. Origins of translational "noise"

Deep analyses of Ribo-seq and MS data sets revealed pervasive translational signals across mRNAs, spanning the 5'UTR to 3'UTR. In this sense, "Un-Translated Region" does not seem to convey the proper functions of these important mRNA segments. Similarly, translational signals exist on many non-coding RNAs, casting doubts on the definition of "non-coding". Nevertheless, the level of translational activity on these non-CDS regions is generally low, and it is challenging to distinguish translational events from background noise. Nevertheless, it is clear that translational signals detected on different mRNA regions rely on distinct mechanisms (Fig. 2).

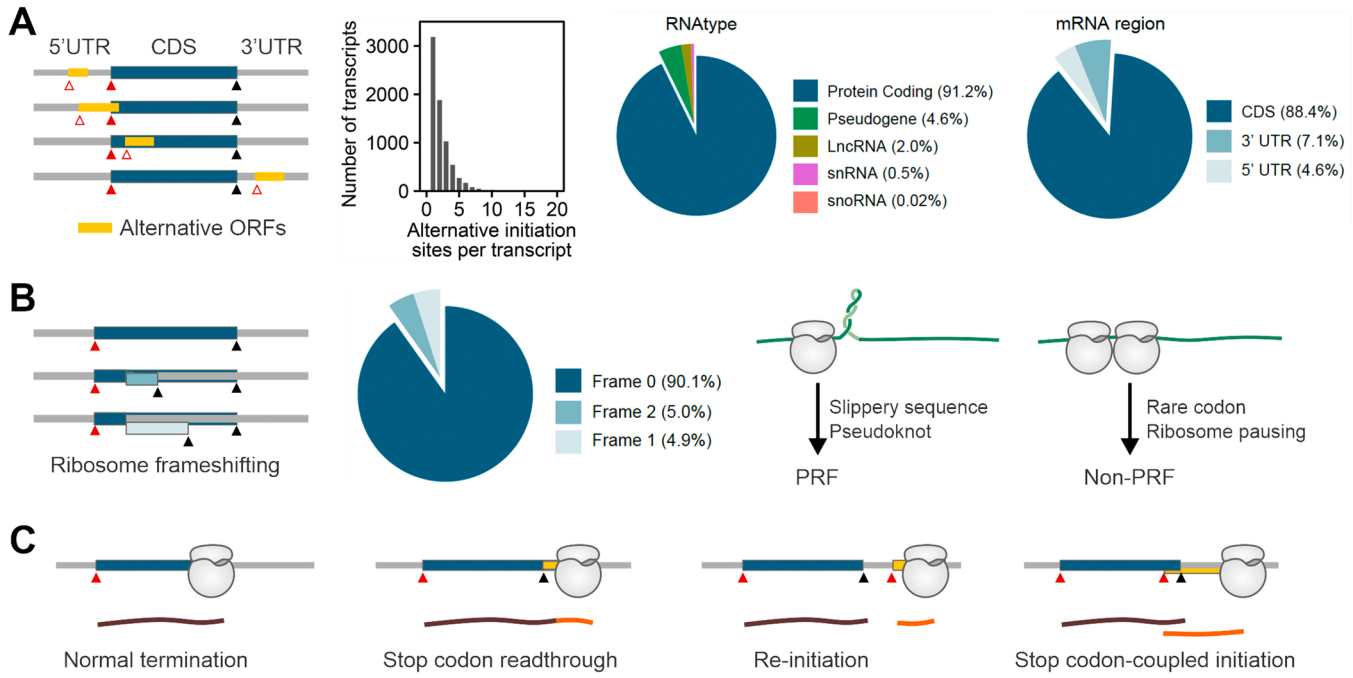


**Fig. 2.** Origins of translational "noise". (A) Alternative ORFs generated by alternative initiation. The left panel show alternative initiation sites in the 5' UTR, CDS or 3' UTR. The authentic and alternative initiation sites are indicated by solid red and empty red triangles, respectively. The main ORFs are indicated by green frames. The histogram shows the frequency of alternative initiation sites on individual mRNAs. The pie chart shows the fraction of footprint reads aligned to different RNA species (left) and mRNA regions (right). Ribosome profiling data is obtained from GSE176058. (B) Alternative ORFs generated by ribosomal frameshifting in the coding region. The left panel shows alternative reading frames (the bottom two) after the frameshifting site. The original frame is shown at top. The middle panel shows the fraction of footprint reads aligned to different reading frames of mRNA. The right panel illustrates the programmed ribosomal frameshifting (PRF) induced by a slippery sequence and pseudoknot, and non-programmed ribosomal frameshifting (Non-PRF) induced by rare codons. (C) Alternative ORFs generated by stop codon readthrough, reinitiation or stop codon-coupled initiation. The start codon and stop codon are indicated by red and black triangles, respectively.

### 3.1. *Alternative initiation*

In eukaryotes, ribosomal scanning is a well-accepted model for start codon selection. It is commonly assumed that the first AUG triplet that the scanning ribosome encounters serves as the start codon for translation. However, one or more potential initiation sites could exist besides the main start codon (Fig. 2 A). Accumulating evidence suggests that many non-AUG triplets, especially near cognate codons that differ from AUG by only one nucleotide, can also serve as start codons [13,58]. Using a series of reporter assays, previous studies uncovered that at least 9 non-AUG triplets capable of translation initiation, with CUG showing the highest initiation efficiency [59]. Analysis of the human transcriptome revealed that more than 90% human mRNAs contain multiple initiators in the 5' UTR [19]. Indeed, the number of uORFs with active translation ranges from several hundred to thousand [6,60–62], depending on the methodology. A single study using high-quality Ribo-seq typically uncovers 1000 to 3000 actively translated uORFs in human cells [63]. By integrating multiple Ribo-seq data sets using various human cell lines with different culture conditions, $>$ 15,000 uORFs spreading across $>$ 7000 mRNAs are listed in the small ORF database sORFs.org [64]. Notably, start sites in 5' UTR are over-represented by non-AUG initiators [19,65], implying that uORF translation is subjected to dynamic regulation. Indeed, nutrient stress has been reported to increase uORF translation. Elevated 5'UTR translation has also been observed during oncogenesis triggered by RAS-MAPK pathway [66].

Besides codon identity, potential start sites are also dependent on the surrounding sequence context. In mammalian cells, the initiator AUG triplet is usually in an optimal context with a purine at position –3 and a guanine at position + 4 (+1 refers to the first position of start codon) [67]. RNA secondary structures [68] and RNA modifications such as $N^6$-methyladenosine [69] may act as barriers to slow down scanning subunits, increasing the probability of alternative initiation in the 5' UTR. In addition to these *cis* sequence elements, the stringency of start codon selection is also subject to regulation by *trans* acting factors such as eIF1[65,66], eIF1A, and eIF5 [70,71]. eIF1 promotes initiation in the 5' UTR when the initiation context is suboptimal [72,73]. The poor initiation context of eIF1 offers an auto-regulatory mechanism to orchestrate eIF1 levels [74]. Not surprisingly, many helicases not only contribute to the scanning process, but also influence start codon selection. Ded1p (DDX3 in human) is required for continuous scanning for structured 5' UTRs [75], and loss-of-function mutations lead to elevated alternative initiation at the near-cognate start codon immediately upstream of RNA structures. eIF4G2 (also named DAP5 or Nat1) is reported to promote uORF bypassing, facilitating translation at the authentic start codon [76].

Inefficient recognition of an initiator codon results in a portion of scanning ribosomes continuing to scan and initiating at a downstream site, in a process known as leaky scanning. Not surprisingly, the number of identified downstream initiation sites is much fewer than the number of upstream initiation sites. Using Ribo-seq approach, a recent study uncovered 2466 upstream initiation sites, but only 13 downstream initiation sites, in human foreskin fibroblasts [47]. The underrepresented downstream initiation is consistent with the directional scanning process [12], but also reflects the challenge of uncovering translational "noise" from the main CDS.

### 3.2. *Frameshifting*

Upon start codon recognition, the 80 S ribosome assembled at the start codon is typically thought to follow the reading frame defined by the start site. Since there is no punctuation between codons, the ribosome is expected to maintain the reading frame for hundreds of codons to ensure proper protein production. Although spontaneous

frameshifting could be deleterious as premature termination often leads to non-functional products, programmed ribosomal frameshifting (PRF) could be advantageous in some circumstances. Many viruses utilize PRF to regulate the production of key enzymes. For example, the frameshifting at the junction between ORF1a and 1b of the SARS-CoV-2 genome is necessary for the synthesis of viral RNA-dependent RNA polymerase and other viral proteins essential for the viral life cycle [11]. Notably, most known PRF events occur at slippery sites and are enhanced by downstream stimulatory elements, such as stem-loop RNA structures or pseudoknots (Fig. 2B) [7,77].

While PRF is pervasive in viruses, there is an increasing list of PRF cases in eukaryotes. Translation of ornithine decarboxylase (ODC) antizyme is one of the best-known examples [78,79]. ODC antizyme synthesis is dependent on the + 1 PRF at the stop codon, which allows ribosomes to translate downstream regions. Intriguingly, the efficiency of + 1 PRF is regulated by the concentration of polyamines, whose biosynthesis is mediated by ODC. Frameshifting can also be induced by rare codons [80,81], suggesting that delayed tRNA delivery into the A-site promotes frameshifting. Supporting this notion, depletion of intracellular tryptophan in melanoma cells leads to ribosome pausing at tryptophan codons and subsequent frameshifting [82,83]. This study highlights the possibility that ribosome pausing promotes frameshifting. Indeed, unresolved ribosome collision is known to promote both + 1 and − 1 frameshifting in yeast and bacteria [25,84,85].

Ribosome pausing is prevalent across the coding region [86], but it is unclear whether all pausing events are associated with certain degree of frameshifting. To identify translational "noise" derived from frameshifting, Ribo-seq data sets with superior resolution will be highly desirable. Since alternative translation also contributes to out-of-frame footprint reads in CDS, the origin of off-track translation is not immediately clear. The main difference between alternative translation and frameshifting lies in the start codon choice. For instance, a sudden appearance of out-of-frame reads in CDS without obvious start codons is likely a result of elongation-associated frameshifting. Given the existence of non-AUG start sites, presence of the initiator $tRNA_i^{Met}$ could be the more definitive evidence for alternative initiation.

### 3.3. Stop codon readthrough

Translation termination occurs upon stop codon recognition, which involves competition between near-cognate tRNAs and release factors in the ribosomal A site [17,87]. Although termination at stop codons is generally efficient, stop codon readthrough occurs in organisms ranging from viruses to mammals (Fig. 2 C) [8]. Stop codon readthrough is especially prevalent in *Drosophila* with hundreds of transcripts undergoing active translation after the stop codon [88–90]. In mammalian cells, the basal level of stop codon readthrough is < 1% [91]. However, the stop codon readthrough can be enhanced in certain contexts, such as the presence of RNA secondary structures or a cytidine immediately after the stop codon. It has been reported that there are 23 human transcripts with the stop codon UGA immediately followed by CUAG, and all of those mRNAs showed elevated stop codon readthrough (>1%), with *OPRL1* mRNA having the highest level of ~31% in HEK293T cells [92,93].

Phylogenetic analysis uncovered a conserved motif around the stop codons associated with high potential of stop codon readthrough [93]. By counting the footprint density in 3' UTRs, Ribo-seq holds the potential of uncovering stop codon readthrough events in human cells [89,94,95]. The readthrough ribosomes are supposed to maintain the same reading frame beyond the stop codon. However, a recent study suggested that ribosomes undergoing stop codon readthrough are subjected to frameshifting [96]. Besides Ribo-seq, a recent study applied a combination of in vitro RNA selection and high-throughput sequencing to characterize RNA features that facilitate stop codon readthrough [97]. This method highlighted the critical role of the cytidine and stem loop structures immediately after the stop codon. By integrating into a machine learning model, it enabled discovery of novel stop codon readthrough events in human endogenous transcripts like *CDKN2B*, *LEPROTL1*, *PVRL3*, and *SFTA2*.

### 3.4. Stop codon-coupled initiation

Successful translation termination involves peptide discharge, ribosome splitting, and mRNA release. In some cases, however, termination at the stop codon is followed by reinitiation in which the ribosome remains bound to the mRNA and resumes scanning for downstream start codons (Fig. 2C) [18]. Reinitiation is often used by viruses to express multiple ORFs from one mRNA. Some cellular mRNAs harboring uORFs are also capable of reinitiation especially under stress conditions. The best characterized example is *GCN4* in yeast and *ATF4* in mammals, whose translation is selectively induced during amino acid deprivation [98]. It has been reported that reinitiation efficiency is inversely correlated with the length of the uORF [99,100]. Beside the uORF length, the distance between the uORF stop codon and the downstream CDS start codon is also crucial. A growing body of evidence suggests that reinitiation relies on certain initiation factors (eIFs) that remain attached to the post-termination ribosomes [101,102]. One such factor is eIF3 complex which has been shown to ensure prolonged retention of post-termination ribosomes on mRNA [103]. A recent study reported that eIF3a undergoes dynamic *O*-GlcNAc modification that influences the 80S-eIF3 complex formation and subsequent translation reinitiation [102].

In addition to scanning-based reinitiation after termination, the post-termination ribosome could migrate bi-directionally to locate start codons in different reading frames. Termination-dependent reinitiation has been described in a number of RNA viruses with overlapping stop and start codons (AUGA or UGAUG). The translation termination-coupled reinitiation mechanism also occurs in mammalian cells, contributing to translational activities in 3'UTRs. As a typical example, *CASQ2* relies on a GAUGAU repeat upstream of the stop codon to enable reinitiation at upstream AUG codons in the + 1 reading frame [104]. It appears that terminating ribosomes are capable of acquiring the initiator tRNA before recycling.

## 4. Functions of translational "noise"

The pervasive translational "noise" outside of and within the canonical coding regions raises an intriguing question concerning the physiological functions of alternative translation events. Based on where, when, and how these translational "noises" are generated, they may exert diverse roles in pathophysiology. Indeed, systematic disruption of translational "noises" can lead to specific transcriptomic and phenotypic changes in human cells.

### 4.1. Regulation of mRNA translation

Alternative translation occurring in the 5' UTR regulates main CDS translation through a variety of mechanisms. Based on the leaky scanning model, uORF translation generally represses the main CDS translation by consuming the scanning ribosomes [105,106]. This is particularly true when the uORF overlaps with the main CDS (Fig. 3A). The inhibitory effect of an uORF can be enhanced by ribosomal pausing in uORF, thereby blocking the leaky scanning. This can be achieved by introducing rare codons into the uORF or by enabling the encoded peptide directly interacting with the ribosome [107–110]. On the other hand, some uORF translation may promote main CDS translation via the reinitiation mechanism when the space between uORF and CDS is optimal [111,112]. One of the well-characterized examples is a group of mRNAs that are involved in the integrated stress response (ISR), including *ATF4*, *GADD34* and *CHOP* [113]. ISR activation is associated with a wide range of human diseases by influencing the translational balance between uORF and CDS [114]. For example, the programmed

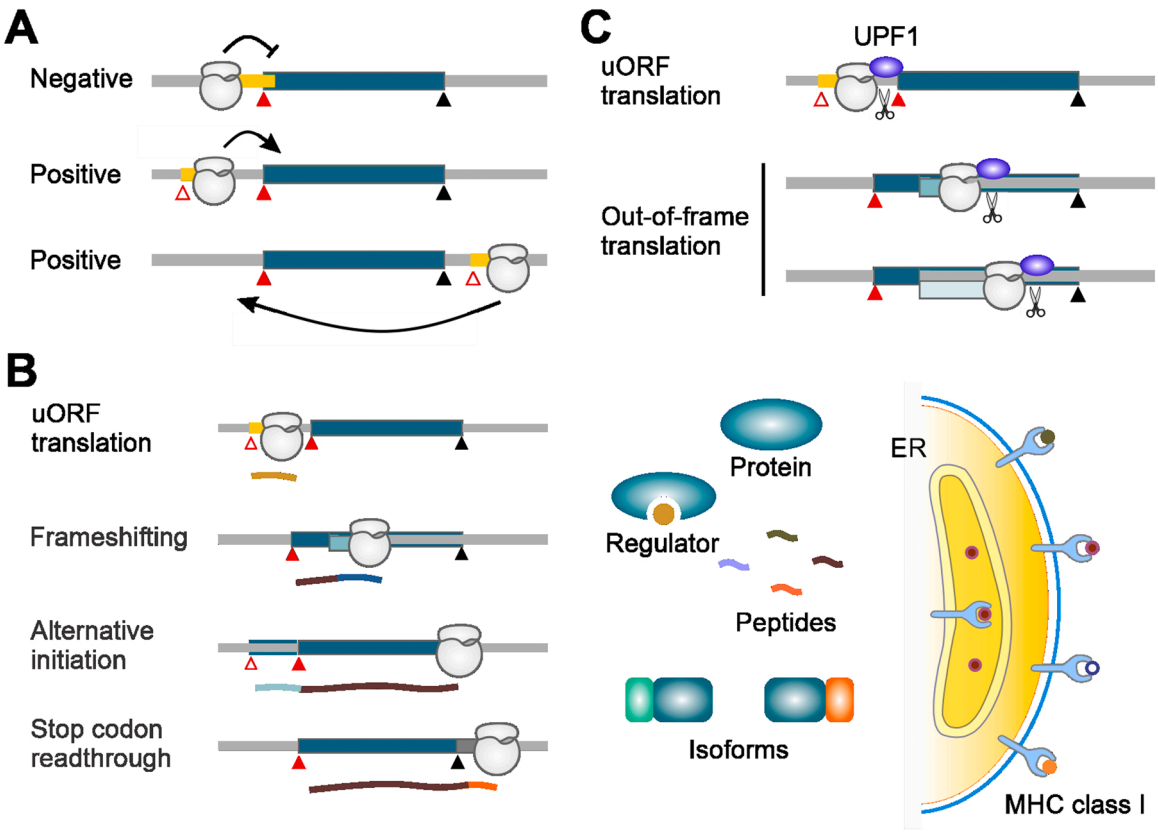


**Fig. 3.** Functions of translational "noise". (A) uORFs can inhibit main ORF translation (top) when overlapping with the main ORF, or promote main ORF translation via reinitiation. Alternative ORFs in 3' UTR may promote main ORF translation (bottom), possibly by facilitating recruitment of initiation factors and ribosomes from 3' UTR to the 5' end of mRNA. (B) Alternative translation products have a wide range of cellular functions. For instance, the uORF products could act as protein regulators, binding partners, and protein isoforms. Many microproteins could also server as precursors for peptides presented by MHC class I molecules. (C) uORF translation or ribosome frameshifting can trigger UPF1-mediated mRNA decay.

death-ligand 1 (PD-L1) is abnormally upregulated on tumor cells and antigen-presenting cells in the tumor microenvironment, resulting in tumor immune escape [115,116]. An uORF driven by CUG overlaps with the CDS of PD-L1 mRNA, thereby inhibiting PD-L1 translation in normal tissues. The uORF translation is bypassed upon ISR activation, contributing to elevated PD-L1 translation in liver and lung cancers.

A recent study reported that translational events in 3' UTR enhances main CDS translation [117]. It has been proposed that a closed mRNA loop may facilitate recruitment of initiation factors and ribosomes from 3' UTR translation to the 5' end of mRNA. Since active translation in 3'UTR is relatively rare, the physiological significance of this mechanism remains to be demonstrated.

### 4.2. *Generation of functional products*

Unlike the annotated main CDS, alternative ORFs typically encode short peptides lacking conservation [118]. An increasing amount of evidence suggests that *de novo* translational products of noncanonical ORFs could have direct cellular functions (Fig. 3B). For instance, small peptides generate by uORFs in fruit fly play crucial roles in development [119]. An uORF found on the tumor-suppressive gene *PTEN* encodes a 31-aa protein, MP31, which regulates lactate metabolism in mitochondria [120]. Interestingly, loss-of-function mutations in MP31 are frequently found in glioblastoma, an aggressive cancer that occurs in the brain or spinal cord. Therefore, the uORF product MP31, like PTEN, acts as a tumor suppressor by regulating lactate metabolism [119]. Another example comes from the uORF on the mRNA encoding the protein kinase C-η (PKC-η), a signaling and antiapoptotic stress kinase involved in cell proliferation, differentiation, and apoptosis. This uORF encodes a 26-aa peptide, uPEP2, which inhibits the kinase activity of PKC-η or other PKCs, thus suppressing proliferation and migration of cancer cells [121].

Besides identification of individual translational products from alternative ORFs, systematic characterization of alternative translational products remains challenging. A recent study used a systematic CRISPR-based screening strategy to knockout hundreds of noncanonical ORFs in human cells [47]. The high throughput screen uncovers 230 uORFs and 2 internal translation that significantly affect cell variability. A close examination of the uORF translation revealed that many uORFs only slightly inhibit main CDS translation (20% ~ 60% reduction in the presence of uORF), suggesting that the main function of those uORFs lies in the encoded micropeptides [47]. Another study expressed 553 noncanonical ORFs in human cells and observed that 73% of ORFs induced significant changes in gene expression [45]. The two high throughput screens uncovered many novel noncanonical ORF-encoded proteins that play critical cellular roles. Interestingly, many noncanonical ORF-encoded peptides can be displayed by MHC-I, contributing to the antigen repertoire [122–126]. A recent study demonstrated that prolonged exposure to IFNγ triggers frameshifting events at the tryptophan codons in melanoma cells [82,83]. The frameshifting resulted in the generation and presentation of out-of-frame peptides at the cell surface, enabling recognition and specific killing of drug-resistant cancer cells by T lymphocytes.

### 4.3. *Regulation of mRNA stability*

It has long been appreciated that efficient translation typically protects mRNA from degradation. However, whether alternative ORF translation influences the mRNA stability remains a fundamental

knowledge gap. Nonsense-mediated decay is an RNA surveillance pathway that eliminates the mRNAs with premature translation termination codons (PTC). Similar to the finding that a longer 3' UTR more likely activates NMD, mRNAs with robust uORF translation are targeted by NMD (Fig. 3C) [127,128]. Using massively parallel reporter assays, a recent study confirmed the negative correlation between uORF translation and mRNA stability. Notably, uORF translation-mediated mRNA decay requires UPF1, a core component of NMD machinery. Besides uORF translation, mRNAs with out-of-frame initiation from the main CDS or ribosomal frameshifting are likely the targets for NMD. Indeed, analysis of NMD targets reveals a higher amount of out-of-frame translation on the mRNAs targeted by NMD [129]. Emerging evidence suggests that the codon optimality is another determinant of mRNA stability [130]. It remains to be determined whether the frameshifting induced by non-optimal codons serves as the underlying mechanism.

## 5. Conclusions

A central tenet of biology is the accurate flow of genetic information from nucleic acids to proteins. Biological noises, however, are often overlooked. In some circumstances, decreasing the fidelity of DNA replication and RNA transcription can be beneficial for biological processes. For example, somatic hypermutation reduces the fidelity of DNA replication by over 1,000-fold and enables B-cells to generate a highly diverse library of receptors [131]. The ~100-fold lower fidelity of retroviral reverse transcriptase enables the generation of a diverse population of retroviruses, some of which can better survive cellular and pharmacological attacks [132]. Translational fidelity is maintained at two fundamental steps: aminoacylation and ribosome decoding. Intriguingly, tRNA misacylation and ribosome recoding have been shown to protect cells from oxidative stress [133]. The rapid development of Ribo-seq technologies over the past decade has provided a unique opportunity for taking a global snapshot of the translatome as well as translational "noise". Understanding the origin of translational "noise" greatly increases the proteome diversity and complexity. From alternative initiation, frameshifting, to stop codon readthrough, the pervasive noncanonical translation could explain thousands of mass spectrometry spectra unannotated from human proteome.

Exploring translational "noises" under different experimental conditions could indicate whether the translational infidelity arises from error, or represents a potential feature conferring an advantage. Indeed, some microproteins derived from noncanonical translation play crucial roles in cell survival. Of note, translational "noises" are typically increased in response to stress conditions. Nutrient stress rapidly alters the proteome landscape via translational reprogramming [134]. Upon amino acid deprivation, general protein synthesis is rapidly suppressed but a subset of mRNAs undergoes selective translation. To support selective protein synthesis, degradative systems are activated to recycle intracellular amino acids. However, what protein sources are preferentially allocated for degradation remains a debatable subject. Previous studies proposed that a ribosome autophagy (ribophagy) pathway supplies internal amino acids during acute nutrient stress [135], but systematic quantitation of ribosome inventory showed minimal ribosome degradation [136]. It is conceivable that stress-induced translational "noises" provide a degradative source for intracellular amino acid recycling, which is essential for selective protein synthesis.

Clearly, combining multiple emerging technologies will paint a multilayered picture of translational "noise" with higher resolution. Future investigation will be required to capture dynamic translational "noise" at single-cell levels with spatial resolution. Another important field is to develop deep learning-based modeling to predict translational "noise" from mRNA sequences. Finally, it is highly desirable if we could fine-tune translational "noises" across transcriptome or towards individual transcripts, which might ultimately lead to the development of new therapeutic strategies for human diseases.

## Conflict of interest

The authors declare no conflicts of interest.

## Acknowledgements

We thank the Qian lab members for critical reading of this manuscript. This work was supported by US National Institutes of Health (R01GM1222814 and DP1GM142101) and HHMI Faculty Scholar (55108556) to S.-B.Q.

## References

[1] Preston B.D., Albertson T.M., Herr A.J. (2010) DNA replication fidelity and cancer. Seminars in cancer biology, (Elsevier), pp 281–293.
[2] J.-F. Gout, W. Li, C. Fritsch, A. Li, S. Haroon, L. Singh, D. Hua, H. Fazelinia, Z. Smith, S. Seeholzer, The landscape of transcription errors in eukaryotic cells, Sci. Adv. 3 (10) (2017), e1701484.
[3] L. Cochella, R. Green, Fidelity in protein synthesis, Curr. Biol. 15 (14) (2005) R536–R540.
[4] T.A. Kunkel, K. Bebenek, DNA replication fidelity, Annu. Rev. Biochem. 69 (1) (2000) 497–529.
[5] H.S. Zaher, R. Green, Fidelity at the molecular level: lessons from protein synthesis, Cell 136 (4) (2009) 746–762.
[6] M.W. Orr, Y. Mao, G. Storz, S.-B. Qian, Alternative ORFs and small ORFs: shedding light on the dark proteome, Nucleic Acids Res. 48 (3) (2020) 1029–1042.
[7] J.D. Dinman, Mechanisms and implications of programmed translational frameshifting, Wiley Inter. Rev. RNA 3 (5) (2012) 661–673.
[8] M. Palma, F. Lejeune, Deciphering the molecular mechanism of stop codon readthrough, Biol. Rev. Camb. Philos. Soc. 96 (1) (2021) 310–329.
[9] K. Mohler, M. Ibba, Translational fidelity and mistranslation in the cellular response to stress, Nat. Microbiol. 2 (9) (2017) 1–9.
[10] N. Netzer, J.M. Goodenbour, A. David, K.A. Dittmar, R.B. Jones, J.R. Schneider, D. Boone, E.M. Eves, M.R. Rosner, J.S. Gibbs, Innate immune and chemically triggered oxidative stress modifies translational fidelity, Nature 462 (7272) (2009) 522–526.
[11] P.R. Bhatt, A. Scaiola, G. Loughran, M. Leibundgut, A. Kratzel, R. Meurs, R. Dreos, K.M. O'Connor, A. McMillan, J.W. Bode, Structural basis of ribosomal frameshifting during translation of the SARS-CoV-2 RNA genome, Science 372 (6548) (2021) 1306–1313.
[12] A.G. Hinnebusch, The scanning mechanism of eukaryotic translation initiation, Annu. Rev. Biochem. 83 (2014) 779–812.
[13] S. Lee, B. Liu, S. Lee, S.-X. Huang, B. Shen, S.-B. Qian, Global mapping of translation initiation sites in mammalian cells at single-nucleotide resolution, Proc. Natl. Acad. Sci. 109 (37) (2012) E2424–E2432.
[14] T.E. Dever, R. Green, The elongation, termination, and recycling phases of translation in eukaryotes, Cold Spring Harb. Perspect. Biol. 4 (7) (2012) a013706.
[15] H. Gingold, Y. Pilpel, Determinants of translation efficiency and accuracy, Mol. Syst. Biol. 7 (1) (2011) 481.
[16] E.M. Novoa, L.R. de Pouplana, Speeding with control: codon usage, tRNAs, and ribosomes, Trends Genet. 28 (11) (2012) 574–581.
[17] C.U. Hellen, Translation termination and ribosome recycling in eukaryotes, Cold Spring Harb. Perspect. Biol. 10 (10) (2018) a032656.
[18] S. Gunišová, V. Hronová, M.P. Mohammad, A.G. Hinnebusch, L.S. Valášek, Please do not recycle! Translation reinitiation in microbes and higher eukaryotes, FEMS Microbiol. Rev. 42 (2) (2018) 165–192.
[19] H. Zhang, Y. Wang, X. Wu, X. Tang, C. Wu, J. Lu, Determinants of genome-wide distribution and evolution of uORFs in eukaryotes, Nat. Commun. 12 (1) (2021) 1–17.
[20] M. Pohl, G. Theiβen, S. Schuster, GC content dependency of open reading frame prediction via stop codon frequencies, Gene 511 (2) (2012) 441–446.
[21] I.T. Rombel, K.F. Sykes, S. Rayner, S.A. Johnston, ORF-FINDER: a vector for high-throughput gene identification, Gene 282 (1–2) (2002) 33–41.
[22] L. Wang, H.J. Park, S. Dasari, S. Wang, J.-P. Kocher, W. Li, CPAT: Coding-Potential Assessment Tool using an alignment-free logistic regression model, Nucleic Acids Res. 41 (6) (2013) e74-e74.
[23] N.T. Ingolia, S. Ghaemmaghami, J.R. Newman, J.S. Weissman, Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling, Science 324 (5924) (2009) 218–223.
[24] C.C.-C. Wu, B. Zinshteyn, K.A. Wehner, R. Green, High-resolution ribosome profiling defines discrete ribosome elongation states and translational regulation during cellular stress, Mol. Cell 73 (5) (2019) 959–970, e955.
[25] S. Meydan, N.R. Guydosh, Disome and trisome profiling reveal genome-wide targets of ribosome quality control, Mol. Cell 79 (4) (2020) 588–602, e586.
[26] A.T. Martens, J. Taylor, V.J. Hilser, Ribosome A and P sites revealed by length analysis of ribosome profiling data, Nucleic Acids Res. 43 (7) (2015) 3680–3687.
[27] N.T. Ingolia, G.A. Brar, N. Stern-Ginossar, M.S. Harris, G.J. Talhouarne, S. E. Jackson, M.R. Wills, J.S. Weissman, Ribosome profiling reveals pervasive translation outside of annotated protein-coding genes, Cell Rep. 8 (5) (2014) 1365–1379.

[28] L. Calviello, N. Mukherjee, E. Wyler, H. Zauber, A. Hirsekorn, M. Selbach, M. Landthaler, B. Obermayer, U. Ohler, Detecting actively translated open reading frames in ribosome profiling data, Nat. Methods 13 (2) (2016) 165–170.
[29] P. Zhang, D. He, Y. Xu, J. Hou, B.-F. Pan, Y. Wang, T. Liu, C.M. Davis, E.A. Ehli, L. Tan, Genome-wide identification and differential analysis of translational initiation, Nat. Commun. 8 (1) (2017) 1–14.
[30] Z. Xiao, R. Huang, X. Xing, Y. Chen, H. Deng, X. Yang, De novo annotation and characterization of the translatome with ribosome profiling data, Nucleic Acids Res. 46 (10) (2018) e61-e61.
[31] M. Guttman, P. Russell, N.T. Ingolia, J.S. Weissman, E.S. Lander, Ribosome profiling provides evidence that large noncoding RNAs do not encode proteins, Cell 154 (1) (2013) 240–251.
[32] Z. Ji, R. Song, A. Regev, K. Struhl, Many lncRNAs, 5'UTRs, and pseudogenes are translated and some are likely to express functional proteins, elife 4 (2015), e08890.
[33] A.P. Fields, E.H. Rodriguez, M. Jovanovic, N. Stern-Ginossar, B.J. Haas, P. Mertins, R. Raychowdhury, N. Hacohen, S.A. Carr, N.T. Ingolia, A regression-based analysis of ribosome-profiling data reveals a conserved complexity to mammalian translation, Mol. Cell 60 (5) (2015) 816–827.
[34] M.-T. Huang, Harringtonine, an inhibitor of initiation of protein biosynthesis, Mol. Pharmacol. 11 (5) (1975) 511–519.
[35] T. Schneider-Poetsch, J. Ju, D.E. Eyler, Y. Dang, S. Bhat, W.C. Merrick, R. Green, B. Shen, J.O. Liu, Inhibition of eukaryotic translation elongation by cycloheximide and lactimidomycin, Nat. Chem. Biol. 6 (3) (2010) 209–217.
[36] X. Gao, J. Wan, B. Liu, M. Ma, B. Shen, S.-B. Qian, Quantitative profiling of initiating ribosomes in vivo, Nat. Methods 12 (2) (2015) 147–153.
[37] N. Stern-Ginossar, B. Weisburd, A. Michalski, V.T.K. Le, M.Y. Hein, S.-X. Huang, M. Ma, B. Shen, S.-B. Qian, H. Hengel, Decoding human cytomegalovirus, Science 338 (6110) (2012) 1088–1093.
[38] L. Dong, Y. Mao, A. Zhou, X.-M. Liu, J. Zhou, J. Wan, S.-B. Qian, Relaxed initiation pausing of ribosomes drives oncogenic translation, Sci. Adv. 7 (8) (2021) eabd6927.
[39] J. Griss, J.M. Foster, H. Hermjakob, J.A. Vizcaino, PRIDE Cluster: building a consensus of proteomics data, Nat. Methods 10 (2) (2013) 95–96.
[40] T. Cardon, F. Hervé, V. Delcourt, X. Roucou, M. Salzet, J. Franck, I. Fournier, Optimized sample preparation workflow for improved identification of ghost proteins, Anal. Chem. 92 (1) (2019) 1122–1129.
[41] B. Fabre, J.-P. Combier, S. Plaza, Recent advances in mass spectrometry–based peptidomics workflows to identify short-open-reading-frame-encoded peptides and explore their functions, Curr. Opin. Chem. Biol. 60 (2021) 122–130.
[42] R. Aviner, T. Geiger, O. Elroy-Stein, Genome-wide identification and quantification of protein synthesis in cultured cells and whole tissues by puromycin-associated nascent chain proteomics (PUNCH-P), Nat. Protoc. 9 (4) (2014) 751–760.
[43] C.M. Forester, Q. Zhao, N.J. Phillips, A. Urisman, R.J. Chalkley, J.A. Oses-Prieto, L. Zhang, D. Ruggero, A.L. Burlingame, Revealing nascent proteomics in signaling pathways and cell differentiation, Proc. Natl. Acad. Sci. USA 115 (10) (2018) 2353–2358.
[44] Y. Ma, D.B. McClatchy, S. Barkallah, W.W. Wood, Yates JR 3rd, Quantitative analysis of newly synthesized proteins, Nat. Protoc. 13 (8) (2018) 1744–1762.
[45] J.R. Prensner, O.M. Enache, V. Luria, K. Krug, K.R. Clauser, J.M. Dempster, A. Karger, L. Wang, K. Stumbraite, V.M. Wang, Noncanonical open reading frames encode functional proteins essential for cancer cell survival, Nat. Biotechnol. 39 (6) (2021) 697–704.
[46] B. Wang, J. Hao, N. Pan, Z. Wang, Y. Chen, C. Wan, Identification and analysis of small proteins and short open reading frame encoded peptides in Hep3B cell, J. Proteom. 230 (2021), 103965.
[47] J. Chen, A.-D. Brunner, J.Z. Cogan, J.K. Nuñez, A.P. Fields, B. Adamson, D. N. Itzhak, J.Y. Li, M. Mann, M.D. Leonetti, Pervasive functional translation of noncanonical human open reading frames, Science 367 (6482) (2020) 1140–1146.
[48] S. Wang, L. Tian, H. Liu, X. Li, J. Zhang, X. Chen, X. Jia, X. Zheng, S. Wu, Y. Chen, Large-scale discovery of non-conventional peptides in maize and Arabidopsis through an integrated peptidogenomic pipeline, Mol. Plant 13 (7) (2020) 1078–1093.
[49] C. Chong, M. Müller, H. Pak, D. Harnett, F. Huber, D. Grun, M. Leleu, A. Auger, M. Arnaud, B.J. Stevenson, Integrated proteogenomic deep sequencing and analytics accurately identify non-canonical peptides in tumor immunopeptidomes, Nat. Commun. 11 (1) (2020) 1–21.
[50] T.F. Martinez, Q. Chu, C. Donaldson, D. Tan, M.N. Shokhirev, A. Saghatelian, Accurate annotation of human protein-coding small open reading frames, Nat. Chem. Biol. 16 (4) (2020) 458–468.
[51] J.W. Harger, J.D. Dinman, An in vivo dual-luciferase assay system for studying translational recoding in the yeast Saccharomyces cerevisiae, Rna 9 (8) (2003) 1019–1024.
[52] Dersh D., Yewdell J.W., Wei J. (2019) A SIINFEKL-based system to measure MHC class I antigen presentation efficiency and kinetics. Antigen Processing, (Springer), pp 109–122.
[53] S. Weingarten-Gabbay, S. Elias-Kirma, R. Nir, A.A. Gritsenko, N. Stern-Ginossar, Z. Yakhini, A. Weinberger, E. Segal, Systematic discovery of cap-independent translation sequences in human and viral genomes, Science 351 (6270) (2016) aad4939.
[54] L. Jia, Y. Mao, Q. Ji, D. Dersh, J.W. Yewdell, S.-B. Qian, Decoding mRNA translatability and stability from the 5′ UTR, Nat. Struct. Mol. Biol. 27 (9) (2020) 814–821.
[55] M. Mikl, Y. Pilpel, E. Segal, High-throughput interrogation of programmed ribosomal frameshifting in human cells, Nat. Commun. 11 (1) (2020) 1–18.
[56] P.J. Sample, B. Wang, D.W. Reid, V. Presnyak, I.J. McFadyen, D.R. Morris, G. Seelig, Human 5′ UTR design and variant effect prediction from a massively parallel translation assay, Nat. Biotechnol. 37 (7) (2019) 803–809.
[57] Y. Lim, S. Arora, S.L. Schuster, L. Corey, M. Fitzgibbon, C.L. Wladyka, X. Wu, I. M. Coleman, J.J. Delrow, E. Corey, L.D. True, P.S. Nelson, et al., Multiplexed functional genomic analysis of 5' untranslated region mutations across the spectrum of prostate cancer, Nat. Commun. 12 (1) (2021) 4217.
[58] N.T. Ingolia, L.F. Lareau, J.S. Weissman, Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes, Cell 147 (4) (2011) 789–802.
[59] M.G. Kearse, J.E. Wilusz, Non-AUG translation: a new start for protein synthesis in eukaryotes, Genes Dev. 31 (17) (2017) 1717–1731.
[60] K. Wethmar, A. Barbosa-Silva, M.A. Andrade-Navarro, A. Leutz, uORFdb—a comprehensive literature database on eukaryotic uORF biology, Nucleic Acids Res. 42 (D1) (2014) D60–D67.
[61] T.E. Dever, I.P. Ivanov, M.S. Sachs, Conserved upstream open reading frame nascent peptides that control translation, Annu. Rev. Genet. 54 (2020) 237–264.
[62] R. Fernandes, L. Romão, Translational regulation by upstream open reading frames and its relevance to human genetic disease, eLS 1 (2020) 408–425.
[63] B. Liu, Y. Han, S.B. Qian, Cotranslational response to proteotoxic stress by elongation pausing of ribosomes, Mol. Cell 49 (3) (2013) 453–463.
[64] V. Olexiouk, W. Van Criekinge, G. Menschaert, An update on sORFs. org: a repository of small ORFs identified by ribosome profiling, Nucleic Acids Res. 46 (D1) (2018) D497–D502.
[65] P. Spealman, A.W. Naik, G.E. May, S. Kuersten, L. Freeberg, R.F. Murphy, J. McManus, Conserved non-AUG uORFs revealed by a novel regression analysis of ribosome profiling data, Genome Res. 28 (2) (2018) 214–222.
[66] A. Sendoel, J.G. Dunn, E.H. Rodriguez, S. Naik, N.C. Gomez, B. Hurwitz, J. Levorse, B.D. Dill, D. Schramek, H. Molina, Translation from unconventional 5′ start sites drives tumour initiation, Nature 541 (7638) (2017) 494–499.
[67] W.L. Noderer, R.J. Flockhart, A. Bhaduri, A.J. Diaz de Arce, J. Zhang, P. A. Khavari, C.L. Wang, Quantitative analysis of mammalian translation initiation sites by FACS-seq, Mol. Syst. Biol. 10 (8) (2014) 748.
[68] A.G. Hinnebusch, Molecular mechanism of scanning and start codon selection in eukaryotes, Microbiol. Mol. Biol. Rev. 75 (3) (2011) 434–467.
[69] J. Zhou, J. Wan, X.E. Shu, Y. Mao, X.-M. Liu, X. Yuan, X. Zhang, M.E. Hess, J. C. Brüning, S.-B. Qian, N6-methyladenosine guides mRNA alternative translation during integrated stress response, Mol. Cell 69 (4) (2018) 636–647, e637.
[70] J.L. Llacer, T. Hussain, A.K. Saini, J.S. Nanda, S. Kaur, Y. Gordiyenko, R. Kumar, A.G. Hinnebusch, J.R. Lorsch, V. Ramakrishnan, Translational initiation factor eIF5 replaces eIF1 on the 40S ribosomal subunit to promote start-codon recognition, Elife (2018) 7.
[71] I.P. Ivanov, J.A. Saba, C.M. Fan, J. Wang, A.E. Firth, C. Cao, R. Green, T.E. Dever, Evolutionarily conserved inhibitory uORFs sensitize Hox mRNA translation to start codon selection stringency, Proc. Natl. Acad. Sci. USA 119 (2022) 9.
[72] J.S. Nanda, Y.-N. Cheung, J.E. Takacs, P. Martin-Marcos, A.K. Saini, A. G. Hinnebusch, J.R. Lorsch, eIF1 controls multiple steps in start codon recognition during eukaryotic translation initiation, J. Mol. Biol. 394 (2) (2009) 268–285.
[73] D. Fijałkowska, S. Verbruggen, E. Ndah, V. Jonckheere, G. Menschaert, P. Van Damme, eIF1 modulates the recognition of suboptimal translation initiation sites and steers gene expression via uORFs, Nucleic Acids Res. 45 (13) (2017) 7997–8013.
[74] I.P. Ivanov, G. Loughran, M.S. Sachs, J.F. Atkins, Initiation context modulates autoregulation of eukaryotic translation initiation factor 1 (eIF1), Proc. Natl. Acad. Sci. 107 (42) (2010) 18056–18060.
[75] U.-P. Guenther, D.E. Weinberg, M.M. Zubradt, F.A. Tedeschi, B.N. Stawicki, L. L. Zagore, G.A. Brar, D.D. Licatalosi, D.P. Bartel, J.S. Weissman, The helicase Ded1p controls use of near-cognate translation initiation codons in 5′ UTRs, Nature 559 (7712) (2018) 130–134.
[76] V.V. Smirnova, E.D. Shestakova, D.S. Nogina, P.A. Mishchenko, T. A. Prikazchikova, T.S. Zatsepin, I.V. Kulakovskiy, I.N. Shatsky, I.M. Terenin, Ribosomal leaky scanning through a translated uORF requires eIF4G2, Nucleic Acids Res. 50 (2) (2022) 1111–1127.
[77] T.E. Dever, J.D. Dinman, R. Green, Translation elongation and recoding in eukaryotes, Cold Spring Harb. Perspect. Biol. 10 (2018) 8.
[78] S. Matsufuji, T. Matsufuji, Y. Miyazaki, Y. Murakami, J.F. Atkins, R.F. Gesteland, S.-i Hayashi, Autoregulatory frameshifting in decoding mammalian ornithine decarboxylase antizyme, Cell 80 (1) (1995) 51–60.
[79] R. Palanimurugan, H. Scheel, K. Hofmann, R. Jürgen Dohmen, Polyamines regulate their synthesis by inducing expression and blocking degradation of ODC antizyme, EMBO J. 23 (24) (2004) 4857–4867.
[80] K. Kawakami, S. Pande, B. Faiola, D.P. Moore, J.D. Boeke, P.J. Farabaugh, J. N. Strathern, Y. Nakamura, D.J. Garfinkel, A rare tRNA-Arg(CCU) that regulates Ty1 element ribosomal frameshifting is essential for Ty1 retrotransposition in Saccharomyces cerevisiae, Genetics 135 (2) (1993) 309–320.
[81] R. Temperley, R. Richter, S. Dennerlein, R.N. Lightowlers, Z.M. Chrzanowska-Lightowlers, Hungry codons promote frameshifting in human mitochondrial ribosomes, Science 327 (5963) (2010), 301-301.
[82] O. Bartok, A. Pataskar, R. Nagel, M. Laos, E. Goldfarb, D. Hayoun, R. Levy, P.-R. Körner, I.Z. Kreuger, J. Champagne, Anti-tumour immunity induces aberrant peptide presentation in melanoma, Nature 590 (7845) (2021) 332–337.

[83] J. Champagne, A. Pataskar, N. Blommaert, R. Nagel, D. Wernaart, S. Ramalho, J. Kenski, O.B. Bleijerveld, E.A. Zaal, C.R. Berkers, Oncogene-dependent sloppiness in mRNA translation, Mol. Cell 81 (22) (2021) 4709–4721, e4709.
[84] C.L. Simms, L.L. Yan, J.K. Qiu, H.S. Zaher, Ribosome collisions result in+ 1 frameshifting in the absence of no-go decay, Cell Rep. 28 (7) (2019) 1679–1689, e1674.
[85] A.M. Smith, M.S. Costello, A.H. Kettring, R.J. Wingo, S.D. Moore, Ribosome collisions alter frameshifting at translational reprogramming motifs in bacterial mRNAs, Proc. Natl. Acad. Sci. 116 (43) (2019) 21769–21779.
[86] A.B. Arpat, A. Liechti, M. De Matos, R. Dreos, P. Janich, D. Gatfield, Transcriptome-wide sites of collided ribosomes reveal principles of translational pausing, Genome Res. 30 (7) (2020) 985–999.
[87] M.R. Lawson, L.N. Lessen, J. Wang, A. Prabhakar, N.C. Corsepius, R. Green, J. D. Puglisi, Mechanisms that ensure speed and fidelity in eukaryotic translation termination, Science 373 (6557) (2021) 876–882.
[88] Y. Zhao, B.G. Lindberg, S.S. Esfahani, X. Tang, S. Piazza, Y. Engström, Stop codon readthrough alters the activity of a POU/Oct transcription factor during Drosophila development, BMC Biol. 19 (1) (2021) 1–25.
[89] J.G. Dunn, C.K. Foo, N.G. Belletier, E.R. Gavis, J.S. Weissman, Ribosome profiling reveals pervasive and regulated stop codon readthrough in Drosophila melanogaster, elife 2 (2013), e01179.
[90] A.M. Hudson, N.L. Szabo, G. Loughran, N.M. Wills, J.F. Atkins, L. Cooley, Tissue-specific dynamic codon redefinition in Drosophila, Proc. Natl. Acad. Sci. 118 (2021) 5.
[91] C. Floquet, I. Hatin, J.-P. Rousset, L. Bidou, Statistical analysis of readthrough levels for nonsense mutations in mammalian cells reveals a major determinant of response to gentamicin, PLoS Genet. 8 (3) (2012) e1002608-e1002608.
[92] G. Loughran, M.-Y. Chou, I.P. Ivanov, I. Jungreis, M. Kellis, A.M. Kiran, P. V. Baranov, J.F. Atkins, Evidence of efficient stop codon readthrough in four mammalian genes, Nucleic Acids Res. 42 (14) (2014) 8928–8938.
[93] G. Loughran, I. Jungreis, I. Tzani, M. Power, R.I. Dmitriev, I.P. Ivanov, M. Kellis, J.F. Atkins, Stop codon readthrough generates a C-terminally extended variant of the human vitamin D receptor with reduced calcitriol response, J. Biol. Chem. 293 (12) (2018) 4434–4444.
[94] J.R. Wangen, R. Green, Stop codon context influences genome-wide stimulation of termination codon readthrough by aminoglycosides, Elife 9 (2020), e52611.
[95] T.P. Miettinen, M. Björklund, Modified ribosome profiling reveals high abundance of ribosome protected mRNA fragments derived from 3′ untranslated regions, Nucleic Acids Res. 43 (2) (2015) 1019–1034.
[96] G. Annibaldis, M. Domanski, R. Dreos, L. Contu, S. Carl, N. Kläy, O. Mühlemann, Readthrough of stop codons under limiting ABCE1 concentration involves frameshifting and inhibits nonsense-mediated mRNA decay, Nucleic Acids Res. 48 (18) (2020) 10259–10279.
[97] A.V. Anzalone, S. Zairis, A.J. Lin, R. Rabadan, V.W. Cornish, Interrogation of Eukaryotic Stop Codon Readthrough Signals by in Vitro RNA Selection, Biochemistry 58 (8) (2019) 1167–1178.
[98] K.M. Vattem, R.C. Wek, Reinitiation involving upstream ORFs regulates ATF4 mRNA translation in mammalian cells, Proc. Natl. Acad. Sci. 101 (31) (2004) 11269–11274.
[99] M. Kozak, Constraints on reinitiation of translation in mammals, Nucleic Acids Res. 29 (24) (2001) 5226–5232.
[100] L. Rajkowitsch, C. Vilela, K. Berthelot, C.V. Ramirez, J.E. McCarthy, Reinitiation and recycling are distinct processes occurring downstream of translation termination in yeast, J. Mol. Biol. 335 (1) (2004) 71–85.
[101] S. Schleich, K. Strassburger, P.C. Janiesch, T. Koledachkina, K.K. Miller, K. Haneke, Y.-S. Cheng, K. Kuechler, G. Stoecklin, K.E. Duncan, DENR–MCT-1 promotes translation re-initiation downstream of uORFs to control tissue growth, Nature 512 (7513) (2014) 208–212.
[102] X.E. Shu, Y. Mao, L. Jia, S.-B. Qian, Dynamic eIF3a O-GlcNAcylation controls translation reinitiation during nutrient stress, Nat. Chem. Biol. (2021) 1–8.
[103] M.P. Mohammad, V. Munzarová Pondělíčková, J. Zeman, S. Gunišová, L. S. Valášek, In vivo evidence that eIF3 stays bound to ribosomes elongating and terminating on short upstream ORFs to promote reinitiation, Nucleic Acids Res. 45 (5) (2017) 2658–2674.
[104] P.S. Gould, N.P. Dyer, W. Croft, S. Ott, A.J. Easton, Cellular mRNAs access second ORFs using a novel amino acid sequence-dependent coupled translation termination–reinitiation mechanism, RNA 20 (3) (2014) 373–381.
[105] T.G. Johnstone, A.A. Bazzini, A.J. Giraldez, Upstream ORF s are prevalent translational repressors in vertebrates, EMBO J. 35 (7) (2016) 706–723.
[106] A. Giess, Y.N. Torres Cleuren, H. Tjeldnes, M. Krause, T.T. Bizuayehu, S. Hiensch, A. Okon, C.R. Wagner, E. Valen, Profiling of small ribosomal subunits reveals modes and regulation of translation initiation, Cell Rep. 31 (3) (2020), 107534.
[107] S.K. Young, L.R. Palam, C. Wu, M.S. Sachs, R.C. Wek, Ribosome elongation stall directs gene-specific translation in the integrated stress response, J. Biol. Chem. 291 (12) (2016) 6546–6558.
[108] A. Gaba, A. Jacobson, M.S. Sachs, Ribosome occupancy of the yeast CPA1 upstream open reading frame termination codon modulates nonsense-mediated mRNA decay, Mol. Cell 20 (3) (2005) 449–460.
[109] H.A. Meijer, A.A. Thomas, Ribosomes stalling on uORF1 in the Xenopus Cx41 5' UTR inhibit downstream translation initiation, Nucleic Acids Res 31 (12) (2003) 3174–3184.
[110] A. Wiese, N. Elzinga, B. Wobbes, S. Smeekens, A conserved upstream open reading frame mediates sucrose-induced repression of translation, Plant Cell 16 (7) (2004) 1717–1729.
[111] J. Bohlen, L. Harbrecht, S. Blanco, K. Clemm von Hohenberg, K. Fenzl, G. Kramer, B. Bukau, A.A. Teleman, DENR promotes translation reinitiation via ribosome recycling to drive expression of oncogenes including ATF4, Nat. Commun. 11 (1) (2020) 1–15.
[112] X.E. Shu, Y. Mao, L. Jia, S.-B. Qian, Dynamic eIF3a O-GlcNAcylation controls translation reinitiation during nutrient stress, Nat. Chem. Biol. 18 (2) (2022) 134–141.
[113] S.K. Young, R.C. Wek, Upstream open reading frames differentially regulate gene-specific translation in the integrated stress response, J. Biol. Chem. 291 (33) (2016) 16927–16935.
[114] M. Costa-Mattioli, P. Walter, The integrated stress response: From mechanism to disease, Science 368 (6489) (2020) eaat5314.
[115] S. Suresh, B. Chen, J. Zhu, R.J. Golden, C. Lu, B.M. Evers, N. Novaresi, B. Smith, X. Zhan, V. Schmid, eIF5B drives integrated stress response-dependent translation of PD-L1 in lung cancer, Nat. Cancer 1 (5) (2020) 533–545.
[116] Y. Xu, M. Poggio, H.Y. Jin, Z. Shi, C.M. Forester, Y. Wang, C.R. Stumpf, L. Xue, E. Devericks, L. So, Translation control of the immune checkpoint in cancer and its therapeutic targeting, Nat. Med. 25 (2) (2019) 301–311.
[117] Q. Wu, M. Wright, M.M. Gogol, W.D. Bradford, N. Zhang, A.A. Bazzini, Translation of small downstream ORFs enhances translation of canonical main open reading frames, EMBO J. 39 (17) (2020), e104763.
[118] J.-P. Couso, P. Patraquim, Classification and function of small open reading frames, Nat. Rev. Mol. Cell Biol. 18 (9) (2017) 575–589.
[119] H. Zhang, S. Dou, F. He, J. Luo, L. Wei, J. Lu, Genome-wide maps of ribosomal occupancy provide insights into adaptive evolution and regulatory roles of uORFs during Drosophila development, PLoS Biol. 16 (7) (2018), e2003903.
[120] N. Huang, F. Li, M. Zhang, H. Zhou, Z. Chen, X. Ma, L. Yang, X. Wu, J. Zhong, F. Xiao, An upstream open reading frame in phosphatase and tensin homolog encodes a circuit breaker of lactate metabolism, Cell Metab. 33 (1) (2021) 128–144, e129.
[121] P.A. Dumesic, D.F. Egan, P. Gut, M.T. Tran, A. Parisi, N. Chatterjee, M. Jedrychowski, M. Paschini, L. Kazak, S.E. Wilensky, An evolutionarily conserved uORF regulates PGC1α and oxidative metabolism in mice, flies, and bluefin tuna, Cell Metab. 30 (1) (2019) 190–200, e196.
[122] T. Ouspenskaia, T. Law, K.R. Clauser, S. Klaeger, S. Sarkizova, F. Aguet, B. Li, E. Christian, B.A. Knisbacher, P.M. Le, Unannotated proteins expand the MHC-I-restricted immunopeptidome in cancer, Nat. Biotechnol. (2021) 1–9.
[123] C. Chong, G. Coukos, M. Bassani-Sternberg, Identification of tumor antigens with immunopeptidomics, Nat. Biotechnol. 40 (2) (2022) 175–188.
[124] S.R. Schwab, K.C. Li, C. Kang, N. Shastri, Constitutive display of cryptic translation products by MHC class I molecules, Science 301 (5638) (2003) 1367–1371.
[125] S.R. Starck, V. Jiang, M. Pavon-Eternod, S. Prasad, B. McCarthy, T. Pan, N. Shastri, Leucine-tRNA initiates at CUG start codons for protein synthesis and presentation by MHC class I, Science 336 (6089) (2012) 1719–1723.
[126] S.R. Starck, J.C. Tsai, K. Chen, M. Shodiya, L. Wang, K. Yahiro, M. Martins-Green, N. Shastri, P. Walter, Translation from the 5' untranslated region shapes the integrated stress response, Science 351 (6272) (2016) aad3867.
[127] S. Lykke-Andersen, T.H. Jensen, Nonsense-mediated mRNA decay: an intricate machinery that shapes transcriptomes, Nat. Rev. Mol. Cell Biol. 16 (11) (2015) 665–677.
[128] T. Kurosaki, M.W. Popp, L.E. Maquat, Quality and quantity control of gene expression by nonsense-mediated mRNA decay, Nat. Rev. Mol. Cell Biol. 20 (7) (2019) 406–420.
[129] A. Celik, R. Baker, F. He, A. Jacobson, High-resolution profiling of NMD targets in yeast reveals translational fidelity as a basis for substrate selection, Rna 23 (5) (2017) 735–748.
[130] V. Presnyak, N. Alhusaini, Y.-H. Chen, S. Martin, N. Morris, N. Kline, S. Olson, D. Weinberg, K.E. Baker, B.R. Graveley, Codon optimality is a major determinant of mRNA stability, Cell 160 (6) (2015) 1111–1124.
[131] J.-C. Shen, A.S. Kamath-Loeb, B.F. Kohrn, K.R. Loeb, B.D. Preston, L.A. Loeb, A high-resolution landscape of mutations in the BCL6 super-enhancer in normal human B cells, Proc. Natl. Acad. Sci. 116 (49) (2019) 24779–24785.
[132] R.A. Katz, A.M. Skalka, Generation of diversity in retroviruses, Annu. Rev. Genet. 24 (1) (1990) 409–443.
[133] N. Netzer, J.M. Goodenbour, A. David, K.A. Dittmar, R.B. Jones, J.R. Schneider, D. Boone, E.M. Eves, M.R. Rosner, J.S. Gibbs, A. Embry, B. Dolan, et al., Innate immune and chemically triggered oxidative stress modifies translational fidelity, Nature 462 (7272) (2009) 522–526.
[134] X.E. Shu, R.V. Swanda, S.-B. Qian, Nutrient control of mRNA translation, Annu. Rev. Nutr. 40 (2020) 51–75.
[135] G.A. Wyant, M. Abu-Remaileh, E.M. Frenkel, N.N. Laqtom, V. Dharamdasani, C. A. Lewis, S.H. Chan, I. Heinze, A. Ori, D.M. Sabatini, NUFIP1 is a ribosome receptor for starvation-induced ribophagy, Science 360 (6390) (2018) 751–758.
[136] H. An, A. Ordureau, M. Korner, J.A. Paulo, J.W. Harper, Systematic quantitative analysis of ribosome inventory during nutrient stress, Nature 583 (7815) (2020) 303–309.