

Mini review

Open Access

Gene expression regulation by upstream open reading frames in rare diseases

Joana Silva^{1,2}, Rafael Fernandes^{1,2}, and Luísa Romão^{1,2*}

¹Department of Human Genetics, Instituto Nacional de Saúde Doutor Ricardo Jorge, Lisboa, Portugal

²Gene Expression and Regulation Group, Biosystems & Integrative Sciences Institute (BioISI), Faculdade de Ciências, Universidade de Lisboa, Lisboa, Portugal

Article Info

Article Notes

Received: June 19, 2017

Accepted: July 21, 2017

*Correspondence:

Dr. Luísa Romão, Department of Human Genetics, Instituto Nacional de Saúde Doutor Ricardo Jorge, Av. Padre Cruz, 1649-016, Lisboa, Portugal; Tel: (+351) 21 750 8155; Fax: (+351) 21 752 6410; E-mail: luisa.romao@insa.min-saude.pt

© 2017 Romão L. This article is distributed under the terms of the Creative Commons Attribution 4.0 International License.

Keywords

Upstream open reading frame (uORF)

Non-AUG initiation codon

Translation initiation

Translational regulation

Human rare disease

Stress

ABSTRACT

Upstream open reading frames (uORFs) constitute a class of *cis*-acting elements that regulate translation initiation. Mutations or polymorphisms that alter, create or disrupt a uORF have been widely associated with several human disorders, including rare diseases. In this mini-review, we intend to highlight the mechanisms associated with the uORF-mediated translational regulation and describe recent examples of their deregulation in the etiology of human rare diseases. Additionally, we discuss new insights arising from ribosome profiling studies and reporter assays regarding uORF features and their intrinsic role in translational regulation. This type of knowledge is of most importance to design and implement new or improved diagnostic and/or treatment strategies for uORF-related human disorders.

Introduction

Over the past few years, many genome-wide studies [RNA deep sequencing, ribosome profiling (RiboSeq), mass spectrometry-based methodologies] pointed out translation as a major regulator of gene expression, being recognized as a key post-transcriptional mechanism by which cells rapidly change their expression patterns in response to a wide variety of stimuli¹⁻⁹. RiboSeq is the most promising genome-wide approach to monitor *in vivo* translation, providing also new information about mechanisms of protein synthesis and its control^{5,10}.

Translation is a tightly controlled process that comprises four different steps: initiation, elongation, termination and ribosome recycling. Translation initiation is the most regulated step of translation¹¹. In eukaryotes, translation initiation starts with the recruitment of the cap-binding protein complex, namely eukaryotic initiation factor 4F (eIF4F), which comprises eIF4E, eIF4A and eIF4G, to the mRNA 5' end. The unwinding of the 5'UTR by the helicase eIF4A, enables binding of the 40S ribosomal subunit. The association of eIF1, eIF1A and eIF3 to the 40S subunit facilitates the binding of the ternary complex eIF2-GTP-Met-tRNA_i. The resulting 43S preinitiation complex can land next to the cap and scans in a 5' to 3' direction until it recognizes an initiation codon base-pairing with methionine initiator-tRNA (Met-tRNA_i). Upon recognition of the start codon, eIF5 stimulates GTP hydrolysis, resulting in the release of eIF2-GDP and probably of other 40S-bound initiation factors. eIF5B catalyzes the joining of 60S subunit to form an 80S ribosome, and elongation can start (reviewed in¹¹).

There are several cis-acting elements involved in the regulation of translation initiation, for instance, internal ribosome entry sites (IRESs) and upstream open reading frames (uORFs)^{1,11}. IRESs are highly structured RNA sequences that allow the recruitment of the 40S ribosomal subunit directly to the initiation codon or to its vicinity, promoting translation initiation via a cap-independent mechanism^{12,13}. On the other hand, a uORF is defined as a sequence beginning at an initiation codon, within the 5' untranslated region (5'UTR) of a transcript, in frame with a termination codon positioned upstream or downstream (overlapped uORF) of the main ORF initiation codon^{1,14,15}. uORFs, the most abundant and the best understood class of small ORFs (sORFs), are sequences that encode for peptides up to 100 amino acids, and play different biological roles in the cell^{16,17}. uORFs are typically described as repressors of translation initiation at the main ORF¹⁸⁻²². These cis-regulatory elements are prevalent genome-wide being estimated that approximately half of the human transcripts contain at least one uORF, and many of them are conserved among species, suggesting an evolutionary selection of functional uORFs^{18,23-27}. Genes that need a highly controlled translational regulation, such as oncogenes and genes involved in cell growth, differentiation, development and stress response, are the typical classes of genes harboring uORFs^{1,14,23}. Thus, it is easy to understand that mutations or polymorphisms that disrupt, create or modify uORFs can potentially be associated with the development of several disorders, including rare diseases^{1,14}. Additionally, uORFs and IRESs in the same transcript can cooperate to regulate protein synthesis, although with an antagonist effect¹². For example, translation of fibroblast growth factor 9 (FGF9) is repressed by a uORF in physiological conditions, and induced by a IRES in hypoxia²⁸.

By including the new contributions from RiboSeq analyses and reporter assays, here, we review the mechanisms about uORF-mediated translational regulation, and show how their deregulation can cause human rare disorders.

uORFs and translational regulation

For a uORF to function as a translational regulator its initiation codon needs to be recognized. This process requires the recruitment of the 43S pre-initiation complex (PIC) to the mRNA 5'-cap that, as mentioned above, allows the scanning of the 5'UTR and the recognition of the upstream AUG (uAUG) codon to start translation (reviewed in¹¹). When the ribosome reaches the uORF stop codon, it can either: (i) dissociate and be recycled, which induces translational repression of the downstream ORF(s), or (ii) the 40S subunit does not dissociate from the mRNA and is able to reinitiate translation at a downstream initiation codon^{1,14,22,25,29}. Translation of a uORF may also trigger nonsense-mediated mRNA decay, if the uORF stop codon is recognized as a premature translation termination codon¹.

In these circumstances, translation reinitiation at the main ORF cannot occur^{1,30}. An efficient translation repression mediated by translatable uORF(s) is positively correlated with: (i) a strong uAUG context, (ii) a large distance from the 5' cap to the uAUG, (iii) a great number of uORFs, (iv) a long uORF, and (v) a short distance between the uORF and the main coding sequence (CDS)^{23,31-33}. Additionally, the uORF-encoded peptide can exert an inhibitory effect in translation through a potential interaction with the translational machinery by stalling the translating ribosomes in a sequence-dependent manner or/and in an indirect way through interactions with other small molecules³⁴⁻³⁶. Moreover, the uORF-encoded peptides can have additional biological functions in the cell, working as trans-regulatory factors³⁷.

The cell microenvironment influences the recognition of the AUG initiation codon by the 43S PIC¹¹. During stress conditions, such as hypoxia, endoplasmic reticulum (ER) stress or nutrient depletion, eIF2 α subunit is phosphorylated at serine 51 (eIF2 α -P) by specific kinases³⁸⁻⁴⁰. This phosphorylation prevents eIF2 recycling by the guanine nucleotide exchange factor eIF2B, thus impairing the formation of the ternary complex and reducing the global rate of translation as part of the cell response to stress^{11,31,41}. However, facing this global translational repression a group of transcripts escape and increase their translational rates via uORF-mediated mechanisms, specifically the ones that are involved in cell stress-response^{3,32,38-41}. In the context of high levels of eIF2 α -P, uORFs are usually bypassed by the scanning ribosome that will then access the main AUG initiation codon³¹. This phenomenon is called leaky scanning and is responsible for allowing expression of, for instance, proteins involved in the ER stress response, like the growth arrest and DNA damage-inducible protein (*GADD34*) and the C/EBP homologous protein (*CHOP*), which are encoded by mRNAs with two and one uORFs, respectively^{3,20,21}. The leaky scanning process in these two transcripts is mainly potentiated by a weak uAUG sequence context^{3,20,21,32,42}.

In other cases, translation reinitiation at the main ORF occurs after translation of at least one uORF. The yeast general control protein (*GCN4*) and the activating transcription factor 4 (*ATF4*), with four and two uORFs in their 5'UTRs, respectively, are good examples of transcripts encoding stress related-proteins whose expression depends on translation reinitiation^{22,43}. In amino acid starvation conditions, the first uORF of *GCN4* is efficiently translated, but low levels of available ternary complex force the ribosome to bypass the other uORFs, granting the time to acquire a newly formed ternary complex and reinitiate translation at the main initiation codon⁴³. In the case of *ATF4*, only the first uORF is translated in stress conditions and, again, the ternary complex will eventually be formed

by the time the ribosome reaches the initiation codon of the main ORF²².

The advent of RiboSeq brought new insights about translational regulation, allowing the large-scale identification of mRNAs harboring uORFs as observed by the higher ribosome occupancy at the 5'UTR of many transcripts^{5,10,44,45}. Interestingly, a great number of non-AUG initiation codon-carrying uORFs has been identified, a feature that was not possible to be computationally predicted before RiboSeq has emerged. Ribosome occupancy patterns revealed that CUG is the most prevalent non-AUG initiation codon in uORFs^{5,45,46}. These analyses also revealed a positive correlation between mRNAs carrying non-AUG uORFs and their main ORF translation⁴⁷. This new information raises the question about the initiation factors that regulate the recognition of a non-AUG initiation codon in conditions of an overall repression of translation by eIF2 α phosphorylation. eIF2A seems to act as an alternative to eIF2 α -P since it allows the recruitment of a leucine-tRNA (Leu-tRNA) to the vicinity of a CUG (and UUG) initiation codon of a uORF to start its translation. eIF2A is functionally different from eIF2 α : its knockdown does not repress global translation but markedly impairs expression of uORFs containing leucine-initiation codons^{37,45,48}. Thus, the depletion of eIF2A compromises uORF translation and consequently the main coding sequence expression of transcripts encoding stress-responsive proteins. Binding immunoglobulin protein (*BiP*) mRNA, which encodes a chaperone involved in the ER stress recovery response, was recently described to be regulated by this mechanism of uORF-mediated translation. *BiP* has two leucine initiation codons at position -190 (UUG) and -61 (CUG) nucleotides upstream the main AUG. Upon ER stress (and hence eIF2 α phosphorylation), eIF2A-dependent translation of the -190 UUG uORF is essential for the *BiP* main ORF expression⁴⁸. The same mechanism operates in mRNAs encoding proteins implicated in the tumorigenic process. In fact, eIF2A was directly associated with tumor formation in mice with squamous cell-carcinoma (SCC) xenografts, a phenomenon that was accompanied by the eIF2A-dependent translation of a cohort of cancer-associated transcripts⁴⁵.

Changes in the expression profiles in response to stress conditions can be a consequence of a transition of the translational machinery to different subsets of mRNAs harboring uORFs or other cis-regulatory elements. This phenomenon was also hypothesized to be due to an intrinsic regulation of the ribosome⁴⁹. In fact, an increasing number of evidence point out the heterogeneity in ribosomal protein composition (riboproteome), ribosomal RNA (rRNA) and ribosomal-associated co-factors depending on localization, cell type and environmental conditions, as a new layer of complexity in gene expression regulation⁵⁰⁻⁵². Furthermore, even little variation in the core of ribosomal proteins seems to regulate translation of specific mRNAs by direct interaction with specific cis-regulatory elements within the 5' and 3'UTRs^{49,53}. An example of this regulatory mechanism was described in *Arabidopsis thaliana*, where a ribosomal protein plays a critical role in translation reinitiation of polycistronic mRNAs and of mRNAs harboring uORFs⁵⁴. This raises the possibility that, similarly to what happens in plants, ribosomal proteins can modulate uORF-mediated translation in human mRNAs and once more, alterations on those mechanisms may be disease-associated, being a promising field of study in the future.

uORFs and human rare diseases

When modified, disrupted or created due to mutations or polymorphisms, uORFs can deregulate the downstream main ORF expression and hence be the cause of several pathologies that can include metabolic, hematologic, endocrine and neurodegenerative disorders, and susceptibility to cancer¹. Among them, numerous rare diseases can be found. According to the European legislation, a disease is considered rare when it affects up to five people per 10000⁵⁵. Several examples of deregulated uORFs associated with the onset or development of rare diseases were well reviewed by Barbosa and co-workers¹. Meanwhile, other cases have been described (Table 1), highlighting and reinforcing the impact of uORFs in mediating translational regulation in human health and disease⁵⁶⁻⁶¹.

Belonging to the set of rare diseases associated with

Rare disease	Gene	Pathogenesis	Reference
Familial DOPA responsive dystonia (DRD)	GCH1	The c.-22C>T polymorphism in the 5'UTR creates an out-of-frame uORF that reduces the main ORF translation efficiency; additionally, the 73-amino acid uORF-encoded peptide has cytotoxic effects.	56,57
Complete androgen insensitivity syndrome (CAIS)	AR	The c.-547C>T mutation in the 5'UTR creates an out-of-frame uORF that reduces the main ORF translation efficiency.	58
Acapomelic campomelic dysplasia (ACD)	SOX9	The c.-185G>A mutation in the 5'UTR creates an out-of-frame and overlapped uORF that reduces the main ORF translation efficiency.	59
Multiple endocrine neoplasia syndrome type 4 (MEN4)	CDKN1B	A 4bp deletion in the uORF shifts its termination codon and impairs translation reinitiation at the main ORF.	60
Shwachman-Diamond syndrome (SDS)	C/EBP α C/EBP β	uORFs in the 5'UTR of C/EBP α and C/EBP β regulate translation reinitiation of C/EBP α -p30 and C/EBP β -LIP, respectively; low levels of SBDS impairs this translation reinitiation.	61

Table 1. Recently described cases of human rare diseases associated with deregulated uORF-mediated translation.

the creation of a uORF is complete androgen insensitivity syndrome (CAIS), part of the group of sex developmental disorders. CAIS is characterized by low levels of androgen receptor (AR) that impairs the response to the androgen dihydrotestosterone, compromising the male phenotype^{62,63}. In 1994, when studying the regulation of AR expression, a group of investigators showed that the 5'UTR is involved in its translational regulation and hypothesized that mutations in this 5'UTR could explain, in part, the etiology of androgen insensitivity syndrome, since the AR mRNA levels are maintained despite the reduced protein levels⁶⁴. Only in 2016, identification by next generation sequencing of a germline mutation in the 5'UTR of AR gene (c.-547C>T) proven to create a translatable uORF responsible for the low AR protein levels in CAIS was possible. In the mutant AR transcript, the 43S preinitiation complex recognizes and initiates translation at the uAUG, leading to the formation of a small peptide. Then, the ribosome dissociates, promoting low rate of translation reinitiation at the main ORF⁵⁸. Other examples of rare diseases that can be related to the creation of uORFs include familial DOPA responsive dystonia (DRD) and acampomelic campomelic dysplasia (ACD)^{56,59}. In the first case the polymorphism c.-22C>T in the 5'UTR of the human guanosine triphosphate cyclohydrolase 1 (*GCH1*) gene creates an out-of-frame uORF that encodes a 73-amino acid peptide, impairing the main ORF translation^{56,57}. The subsequent low expression of GCH1 impairs the dopamine biosynthesis pathway that, ultimately, results in reduced levels of dopamine and dopaminergic dysfunction in the brain, typical of DRD. In addition, the synthesized 73-amino acid peptide is localized in the nucleus where it promotes cytotoxic effects that are accentuated by proteasome impairment⁵⁷. Regarding ACD, a *de novo* mutation, c.-185G>A, in the transcription factor *SRY-box 9 (SOX-9)* gene creates an overlapped uORF that reduces SOX-9 translation that is responsible for the ACD, the milder phenotype of CD⁵⁹.

A new example of a disease which phenotype is related to mutations that disrupt uORFs is multiple endocrine neoplasia syndrome type 4 (MEN4). A 4 base pair (4bp) deletion in the sequence of the highly conserved uORF of the cyclin dependent kinase inhibitor 1B (*CDKN1B*) gene was reported to lengthen the uORF by shifting the uORF stop codon and to reduce the intercistronic space. This event seems to prevent translation reinitiation at the main ORF and therefore the expression of p27^{KIP1}, a tumor suppressor with a crucial role in cell cycle and proliferation regulation that when downregulated increases susceptibility to tumor development⁶⁰.

In addition to alterations that create or disrupt uORFs in a disease context, there are genetic alterations that can indirectly impair the uORF-mediated translational regulation. As an example there is Shwachman-Diamond syndrome (SDS), a rare congenital disease caused by

mutations in the Shwachman-Bodian-Diamond syndrome (SBDS) gene⁶⁵. This disorder is a ribosomopathy since defective SBDS protein impairs large ribosome subunit maturation^{61,66}. Due to its function, SBDS can regulate translation of other transcripts, such as CCAAT/enhancer-binding protein- α (*C/EBP α*) and - β (*C/EBP β*), involved in granulocyte differentiation^{61,67}. Both transcripts have alternative initiation codons that result in three different N-terminal protein isoforms: extended, p42 and p30 for *C/EBP α* , and LAP*, LAP and LIP for *C/EBP β* ^{14,61,67}. Low levels of SBDS expression impair the translation of *C/EBP α* -p30 and *C/EBP β* -LIP truncated proteins, which can explain the hematological phenotype of SDS consisting of bone marrow failure with neutropenia^{61,67,68}. *C/EBP α* and *C/EBP β* mRNA have a uORF within their 5'UTR and Kyungmin and co-workers showed that SBDS is crucial for uORF-mediated translation reinitiation of *C/EBP α* -p30 and *C/EBP β* -LIP. Thus, although the origin of SDS does not depend on alterations in the uORF sequence, the hematological picture is determined by disruption of a uORF-dependent translational mechanism⁶¹. This prompts us to look not only for the uORF sequence context but also to understand the processes, factors and/or networks that drive uORF-mediated translational regulation.

Conclusions

Growing evidence from RiboSeq analyses and reporter assays have brought new insights about the existence of uORFs and their translational regulatory functions, reinforcing the importance of these *cis*-acting regulatory elements in the pathophysiology of several human disorders, including rare diseases. In addition, recent studies have revealed many uORFs harboring non-AUG initiation codons and alternative mechanisms of translation initiation associated with pathological conditions. These new data emphasize the importance of understanding the detailed molecular mechanisms through which a disease relies on in order to develop and implement new strategies for disease diagnosis and treatment.

Acknowledgements

This work was partially supported by Fundação para a Ciência e a Tecnologia (UID/MULTI/04046/2013 to BioISI from FCT/MCTES/PIDDAC). JS and RF are supported by fellowships from Fundação para a Ciência e a Tecnologia (SFRH/BD/106081/2015 and SFRH/BD/114392/2016, respectively).

References

- Barbosa C, Peixeiro I, Romão L. Gene expression regulation by upstream open reading frames and human disease. *PLoS Genetics*. 2013; 9: e1003529.
- Sajjanar B, Deb R, Raina SK, et al. Untranslated regions (UTRs) orchestrate translation reprogramming in cellular stress responses. *J Therm Biol*. 2017; 65: 69–75.

3. Young SK, Wek RC. Upstream open reading frames differentially regulate gene-specific translation in the integrated stress response. *J Biol Chem.* 2016; 291: 16927–16935.
4. Schwanhäusser B, Gossen M, Dittmar G, et al. Global analysis of cellular protein translation by pulsed SILAC. *Proteomics.* 2009; 9: 205–209.
5. Ingolia NT, Ghaemmaghami S, Newman JRS, et al. Genome-wide analysis *in vivo* of translation with nucleotide resolution using ribosome profiling. *Science.* 2009; 324: 218–223.
6. Aviner R, Geiger T, Elroy-Stein O. Novel proteomic approach (PUNCH-P) reveals cell cycle-specific fluctuations in mRNA translation. *Genes Dev.* 2013; 27: 1834–1844.
7. Baird TD, Palam LR, Fusakio ME, et al. Selective mRNA translation during eIF2 phosphorylation induces expression of IBTK α . *Mol Biol Cell.* 2014; 25: 1686–1697.
8. Gao X, Wan J, Liu B, et al. Quantitative profiling of initiating ribosomes *in vivo*. *Nat Methods.* 2014; 12: 147–153.
9. Loayza-Puch F, Drost J, Rooijers K, et al. p53 induces transcriptional and translational programs to suppress cell proliferation and growth. *Genome Biol.* 2013; 14: R32.
10. Andreev DE, O'Connor PBF, Loughran G, et al. Insights into the mechanisms of eukaryotic translation gained with ribosome profiling. *Nucleic Acids Res.* 2017; 45: 513–526.
11. Sonenberg N, Hinnebusch AG. Regulation of translation initiation in eukaryotes: mechanisms and biological targets. *Cell.* 2009; 136: 731–745.
12. Lacerda R, Menezes J, Romão L. More than just scanning the importance of cap-independent mRNA translation initiation for cellular stress response and cancer. *Cell Mol Life Sci.* 2017; 74: 1659–1680.
13. Jackson RJ. The current status of vertebrate cellular mRNA IRESs. *Cold Spring Harb Perspect Biol.* 2013; 5: a011569.
14. Wethmar K, Smink JJ, Leutz A. Upstream open reading frames: molecular switches in (patho)physiology. *Bioessays.* 2010; 32: 885–893.
15. Wethmar K, Schulz J, Muro EM, et al. Comprehensive translational control of tyrosine kinase expression by upstream open reading frames. *Oncogene.* 2016; 35: 1736–1742.
16. Olexiouk V, Crappé J, Verbruggen S, et al. sORFs.org: a repository of small ORFs identified by ribosome profiling. *Nucleic Acids Res.* 2016; 44: D324–D329.
17. Cabrera-Quio LE, Herberg S, Pauli A. Decoding sORF translation from small proteins to gene regulation. *RNA Biol.* 2016; 13: 1051–1059.
18. Barbosa C, Romão L. Translation of the human erythropoietin transcript is regulated by an upstream open reading frame in response to hypoxia. *RNA.* 2014; 20: 594–608.
19. Onofre C, Tomé F, Barbosa C, et al. Expression of human Hemojuvelin (HJV) is tightly regulated by two upstream open reading frames in HJV mRNA that respond to iron overload in hepatic cells. *Mol Cell Biol.* 2015; 35: 1376–1389.
20. Lee YY, Cevallos RC, Jan E. An upstream open reading frame regulates translation of GADD34 during cellular stresses that induce eIF2 α phosphorylation. *J Biol Chem.* 2009; 284: 6661–6673.
21. Palam LR, Baird TD, Wek RC. Phosphorylation of eIF2 facilitates ribosomal bypass of an inhibitory upstream ORF to enhance CHOP translation. *J Biol Chem.* 2011; 286: 10939–10949.
22. Vattem KM, Wek RC. Reinitiation involving upstream ORFs regulates ATF4 mRNA translation in mammalian cells. *Proc Natl Acad Sci USA.* 2004; 101: 11269–11274.
23. Calvo SE, Pagliarini DJ, Mootha VK. Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. *Proc Natl Acad Sci USA.* 2009; 106: 7507–7512.
24. Mignone F, Gissi C, Liuni S, et al. Untranslated regions of mRNAs. *Genome Biol.* 2002; 3: reviews0004.
25. Iacono M, Mignone F, Pesole G. uAUG and uORFs in human and rodent 5' untranslated mRNAs. *Gene.* 2005; 349: 97–105.
26. Sathirapongsasuti JF, Sathira N, Suzuki Y, et al. Ultraconserved cDNA segments in the human transcriptome exhibit resistance to folding and implicate function in translation and alternative splicing. *Nucleic Acids Res.* 2011; 39: 1967–1979.
27. Morris DR, Geballe AP. Upstream open reading frames as regulators of mRNA translation. *Mol Cell Biol.* 2000; 20: 8635–8642.
28. Chen TM, Shih YH, Tseng JT, et al. Overexpression of FGF9 in colon cancer cells is mediated by hypoxia-induced translational activation. *Nucleic Acids Res.* 2014; 42: 2932–2944.
29. Meijer HA, Thomas AAM. Control of eukaryotic protein synthesis by upstream open reading frames in the 5'-untranslated region of an mRNA. *Biochem J.* 2002; 367: 1–11.
30. Mendell JT, Sharifi NA, Meyers JL, et al. Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise. *Nat Genet.* 2004; 36: 1073–1078.
31. Jackson RJ, Hellen CUT, Pestova TV. The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat Rev Mol Cell Biol.* 2010; 11: 113–127.
32. Kozak M. Possible role of flanking nucleotides in recognition of the AUG initiator codon by eukaryotic ribosomes. *Nucleic Acids Res.* 1981; 9: 5233–5252.
33. Kozak M. Constraints on reinitiation of translation in mammals. *Nucleic Acids Res.* 2001; 29: 5226–5232.
34. Wen Y, Liu Y, Xu Y, et al. Loss-of-function mutations of an inhibitory upstream ORF in the human hairless transcript cause Marie Unna hereditary hypotrichosis. *Nat Genet.* 2009; 41: 228–233.
35. Young SK, Willy JA, Wu C, et al. Ribosome reinitiation directs gene-specific translation and regulates the integrated stress response. *J Biol Chem.* 2015; 290: 28257–28271.
36. Andrews SJ, Rothnagel JA. Emerging evidence for functional peptides encoded by short open reading frames. *Nat Rev Genet.* 2014; 15: 193–204.
37. Starck SR, Jiang V, Pavon Eternod M, et al. Leucine-tRNA initiates at CUG start codons for protein synthesis and presentation by MHC Class I. *Science.* 2012; 336: 1719–1723.
38. Novoa I, Zeng H, Harding HP, et al. Feedback inhibition of the unfolded protein response by GADD34-mediated dephosphorylation of eIF2 α . *J Cell Biol.* 2001; 153: 1011–1022.
39. Harding HP, Zhang Y, Bertolotti A, et al. Perk is essential for translational regulation and cell survival during the unfolded protein response. *Mol Cell.* 2000; 5: 897–904.
40. Han AP, Yu C, Lu L, et al. Heme-regulated eIF2 α kinase (HRI) is required for translational regulation and survival of erythroid precursors in iron deficiency. *EMBO J.* 2001; 20: 6909–6918.
41. Blais JD, Filipenko V, Bi M, et al. Activating transcription factor 4 is translationally regulated by hypoxic stress. *Mol Cell Biol.* 2004; 24: 7469–7482.
42. Kozak M. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell.* 1986; 44: 283–292.
43. Abastado JP, Miller PF, Jackson BM, et al. Suppression of ribosomal

- reinitiation at upstream open reading frames in amino acid-starved cells forms the basis for GCN4 translational control. *Mol Cell Biol.* 1991; 11: 486–496.
44. Ivanov IP, Firth AE, Michel AM, et al. Identification of evolutionarily conserved non-AUG-initiated N-terminal extensions in human coding sequences. *Nucleic Acids Res.* 2011; 39: 4220–4234.
45. Sendoel A, Dunn JG, Rodriguez EH, et al. Translation from unconventional 5' start sites drives tumour initiation. *Nature.* 2017; 541: 494–499.
46. Ingolia NT, Lareau LF, Weissman JS. Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell.* 2011; 147: 789–802.
47. Brar GA, Yassour M, Friedman N, et al. High-Resolution View of the Yeast Meiotic Program Revealed by Ribosome Profiling. *Science.* 2012; 335: 552–557.
48. Starck SR, Tsai JC, Chen K, et al. Translation from the 5' untranslated region shapes the integrated stress response. *Science.* 2016; 351: aad3867.
49. Xue S, Barna M. Specialized ribosomes a new frontier in gene regulation and organismal biology. *Nat Rev Mol Cell Biol.* 2012; 13: 355–369.
50. Xue S, Tian S, Fuji K, et al. RNA regulons in Hox 5' UTRs confer ribosome specificity to gene regulation. *Nature.* 2015; 517: 33–38.
51. Slavov N, Semrau S, Airoidi E, et al. Differential Stoichiometry among Core Ribosomal Proteins. *Cell Rep.* 2015; 13: 865–873.
52. Brombin A, Joly JS, Jamen F. New tricks for an old dog: ribosome biogenesis contributes to stem cell homeostasis. *Curr Opin Genet Dev.* 2015; 34: 61–70.
53. Kraushar ML, Popovitchenko T, Volk NL, et al. The frontier of RNA metamorphosis and ribosome signature in neocortical development. *Int J Dev Neurosci.* 2016; 55: 131–139.
54. Nishimura T, Wada T, Yamamoto KT, et al. The Arabidopsis STV1 protein, responsible for translation reinitiation, is required for auxin-mediated gynoecium patterning. *Plant Cell.* 2005; 17: 2940–2953.
55. List of rare diseases and synonyms listed in alphabetical order, 2016. http://www.orpha.net/orphacom/cahiers/docs/GB/List_of_rare_diseases_in_alphabetical_order.pdf. Accessed 22 April, 2017.
56. Sharma N, Armata IA, Mulhaupt Buell TJ, et al. Mutation in 5' upstream region of GCHI gene causes familial dopa-responsive dystonia. *Mov Disord.* 2011; 26: 2140–2141.
57. Jones L, Goode L, Davila E, et al. Translational effects and coding potential of an upstream open reading frame associated with DOPA Responsive Dystonia. *Biochim Biophys Acta - Mol Basis Dis.* 2017; 1863: 1171–1182.
58. Hornig NC, De Beaufort C, Denzer F, et al. A recurrent germline mutation in the 5'UTR of the androgen receptor causes complete androgen insensitivity by activating aberrant uORF translation. *PLoS One.* 2016; 11: e0154158.
59. von Bohlen AE, Böhm J, Pop R, et al. A mutation creating an upstream initiation codon in the SOX9 5' UTR causes acampomelic campomelic dysplasia. *Mol Genet Genomic Med.* 2017; 5: 261–268.
60. Occhi G, Regazzo D, Trivellin G, et al. A novel mutation in the upstream open reading frame of the CDKN1B Gene Causes a MEN4 Phenotype. *PLoS Genet.* 2013; 9: e1003350.
61. In K, Zaini M, Müller C, et al. Shwachman-Bodian-Diamond syndrome (SBDS) protein deficiency impairs translation re-initiation from C/EBP α and C/EBP β mRNAs. *Nucleic Acids Res.* 2016; 44: 4134–4146.
62. Lonergan PE, Tindall DJ. Androgen receptor signaling in prostate cancer development and progression. *J Carcinog.* 2011; 10: 20.
63. Gottlieb B, Beitel LK, Nadarajah A, et al. The androgen receptor gene mutations database: 2012 update. *Hum Mutat.* 2012; 33: 887–894.
64. Mizokami A, Chang C. Induction of translation by the 5'-untranslated region of human androgen receptor mRNA. *J Biol Chem.* 1994; 269: 25655–25659.
65. Boocock GRB, Morrison JA, Popovic M, et al. Mutations in SBDS are associated with Shwachman–Diamond syndrome. *Nat Genet.* 2002; 33: 97–101.
66. Wong CC, Traynor D, Basse N, et al. Defective ribosome assembly in Shwachman-Diamond syndrome. *Blood.* 2011; 118: 4305–4312.
67. Calkhoven CF, Müller C, Leutz A. Translational control of C/EBP α and C/EBP β isoform expression. *Genes Dev.* 2000; 14: 1920–1932.
68. Dror Y, Donadieu J, Kogelmeier J, et al. Draft consensus guidelines for diagnosis and treatment of Shwachman-Diamond syndrome. *Ann NY Acad Sci.* 2011; 1242: 40–55.