Data in brief 28 (2020) 104943

Contents lists available at ScienceDirect

Data in brief

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

Data Article

Experimental supporting data on DIS3L2 over nonsense-mediated mRNA decay targets in human cells

Paulo J. da Costa [a, b], Juliane Menezes [a, b], Margarida Saramago [c], Juan F. García-Moreno [a, b], Hugo A. Santos [b], Margarida Gama-Carvalho [b], Cecília M. Arraiano [c], Sandra C. Viegas [c, **], Luísa Romão [a, b, *]

[a] Department of Human Genetics, Instituto Nacional de Saúde Doutor Ricardo Jorge, Lisboa, Portugal
[b] University of Lisbon, Faculty of Sciences, BioISI - Biosystems and Integrative Sciences Institute, Lisboa, Portugal
[c] Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Oeiras, Portugal

ARTICLE INFO

Article history:
Received 7 November 2019
Received in revised form 25 November 2019
Accepted 28 November 2019
Available online 6 December 2019

Keywords:
DIS3L2
UPF1
NMD
NMD-targets
mRNA surveillance
mRNA degradation

ABSTRACT

In this article, we present supportive data related to the research article "A role for DIS3L2 over natural nonsense-mediated mRNA decay targets in human cells" [1], where interpretation of the data presented here is available. Indeed, here we analyze the impact of the DIS3L2 exoribonuclease over nonsense-mediated mRNA decay (NMD)-targets. Specifically, we present data on: a) the expression of various reporter human β-globin mRNAs, monitored by Northern blot and RT-qPCR, before and after altering DIS3L2 levels in HeLa cells, and b) the gene expression levels of deregulated transcripts generated by re-analyzing publicly available data from UPF1-depleted HeLa cells that were further cross-referenced with a dataset of transcripts upregulated in DIS3L2-depleted cells. These analyses revealed that DIS3L2 regulates the levels of a subset of NMD-targets. These data can be valuable for researchers interested in the NMD mechanism.

* Corresponding author. Department of Human Genetics, Instituto Nacional de Saúde Doutor Ricardo Jorge, Lisboa, Portugal.
** Corresponding author.
E-mail addresses: sviegas@itqb.unl.pt (S.C. Viegas), luisa.romao@insa.min-saude.pt (L. Romão).

https://doi.org/10.1016/j.dib.2019.104943

Specifications Table

Subject	Molecular Biology
Specific subject area	Gene expression, nonsense-mediated mRNA decay
Type of data	Table, Chart, Graph, Figure
How data were acquired	Western blotting, reverse-transcription quantitative PCR (RT-qPCR) and Northern blot Raw data for Illumina paired-end mRNA-seq of Hela cells treated with UPF1 siRNA and control were retrieved from SRA (identifiers DRS001615 - DRS001618) [2]. Deregulated transcripts in DIS3L2 depleted HeLa cells were obtained directly from Lubas et al. (2013) [3].
Data format	Raw, analyzed, filtered and graph
Parameters for data collection	HeLa cells under DIS3L2 or UPF1 knockdowns
Description of data collection	After siRNA treatment, HeLa cells were transfected with the plasmids expressing wild type (βWT), NMD-resistant (β15), or NMD-sensitive (β26, β39) human β-globin mRNAs. Knockdown efficiency was analyzed by Western blot and mRNA levels by RT-qPCR and Northern blot. Raw mRNA-seq data was re-analyzed using the pipelined described by Lubas et al. (2013) [3], using the DESeq algorithm for estimation of differentially expressed genes after siRNA-mediated silencing of UPF1 in HeLa cells.
Data source location	Instituto Nacional de Saúde Doutor Ricardo Jorge, Lisboa, Portugal
Data accessibility	The raw data are available with this article
Related research article	Paulo J. da Costa, Juliane Menezes, Margarida Saramago, Juan F. García-Moreno, Hugo Santos, Margarida Gama-Carvalho, Cecília Arraiano, Sandra Viegas and Luísa Romão A role for DIS3L2 over natural nonsense-mediated mRNA decay targets in human cells. *Biochemical and Biophysical Research Communications*, 2019 Oct 22; 518 (4):664-671 https://doi.org/10.1016/j.bbrc.2019.08.105

Value of the Data

- The data presented here will be useful for further studies aiming to characterize the specificity of DIS3L2-targets.
- The involvement of DIS3L2 in the degradation of specific natural nonsense-mediated mRNA decay (NMD)-targets is relevant for researchers in related fields, particularly researchers investigating the mechanism of NMD.
- These data can be compared to other scientific data addressing the mechanism of mRNA degradation inherent to NMD.
- The data shown here reveals that NMD-sensitive human β-globin mRNA reporters are resistant to DIS3L2; nevertheless, a wider analysis shows that DIS3L2 acts over specific NMD-targets.
- The integrative analysis performed here gives us biological information about the role of DIS3L2 in NMD allowing a wider understanding of its regulatory network.

1. Data

Here we report experimental data on the mRNA expression levels of nonsense-mediated mRNA decay (NMD) reporters after DIS3L2 knockdown (KD). We also present a transcriptome profile of a publicly available mRNA-seq data of UPF1-depleted and control HeLa cells, including gene differential expression analysis.

The mechanism of NMD has been widely studied using the human β-globin mRNA as reporter [4–9]. Indeed, it has been shown that a premature termination codon (PTC) at position 26 or 39 of the human β-globin mRNA induces NMD [4–6,9]; however, mRNAs carrying a PTC in close proximity to the translation initiation codon are NMD-resistant [6–8]. Based on this information, we used nonsense-mutated human β-globin mRNAs as well-documented NMD-reporters. HeLa cells depleted of DIS3L2 by siRNAs were transiently transfected with constructs containing different human β-globin variants: βWT (wild type), β15 (PTC at codon 15; NMD-resistant), β26 (PTC at codon 26; NMD-sensitive), β39 (PTC at codon 39; NMD-sensitive) [6]. After monitoring the KD efficiency by Western blot (Fig. 1A), the impact of altering DIS3L2 levels on the expression of the various reporter mRNAs was monitored by Northern blot (Fig. 1B) and quantitative RT-qPCR (raw data in Table S1; Fig. 1C). As a control, cells were treated with Luciferase (LUC) siRNA. In control conditions, without DIS3L2 KD, the mRNA levels of β15 were comparable (1.22-fold) to those of the βWT mRNA levels (arbitrarily set to 1), and both NMD-sensitive variants were expressed at significant low levels. In addition, in conditions of DIS3L2 KD,

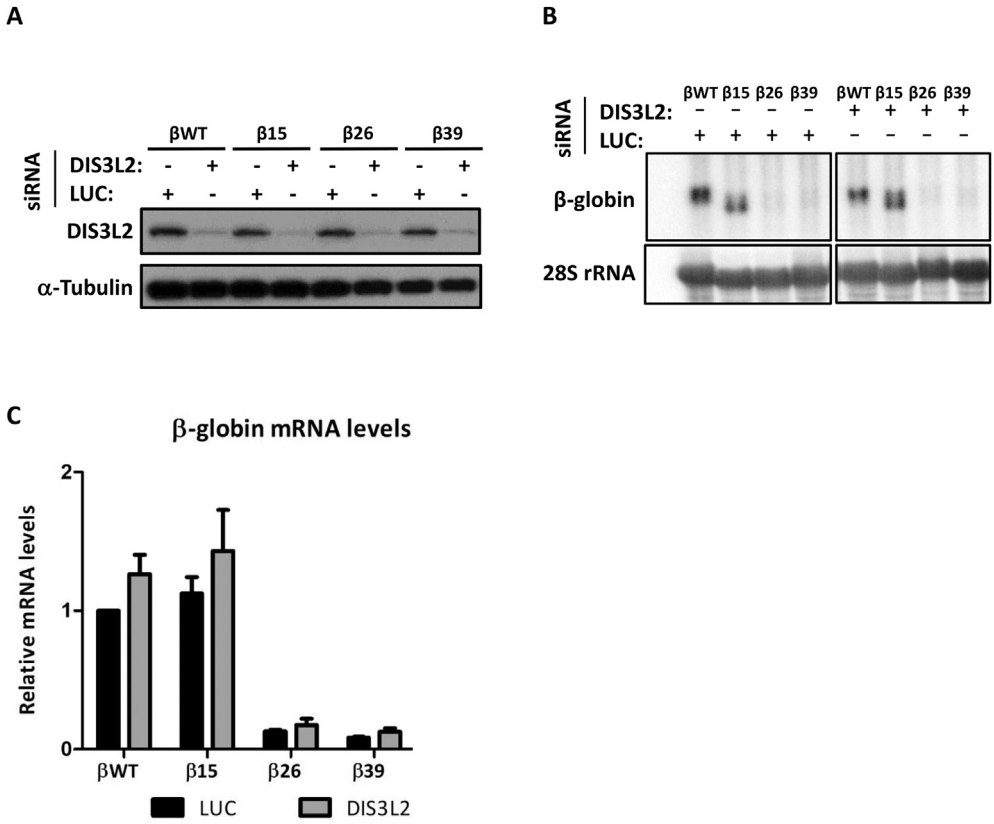

Fig. 1. Levels of normal or nonsense-mutated human β-globin transcripts are not affected by depletion of DIS3L2 in HeLa cells. (A) Western blot analysis of HeLa cells lysates transfected (+) with control Luciferase siRNA (LUC siRNA), or with siRNA targeting the human DIS3L2. Following to the siRNA-mediated silencing, cells were transfected with the plasmids expressing wild type (βWT), NMD-resistant (β15), or NMD-sensitive (β26, β39) human β-globin mRNAs. Primary antibodies against DIS3L2 and α-tubulin (loading control) were used in this blotting to monitor DIS3L2 knockdown. Identification of each band is indicated on the left side of the image. (B) Representative Northern blot analysis of total RNA obtained from HeLa cells with (+) or without (−) depletion of DIS3L2. Fifteen micrograms of total RNA were separated under denaturing conditions on an MOPS/formaldehyde agarose gel (1%) and human β-globin variants βWT, β15, β26 and β39 were detected with a specific probe for exon 1. Loading was controlled by detection of 28S rRNA. (C) Bar plot shows fold-change of each sample mRNA level relative to the control [βWT at Luciferase (LUC) siRNA-treated cells], arbitrarily set to 1. RT-qPCR was used to determine mRNA levels by using primers specific for human β-globin gene, and for glyceraldehyde-3-phosphate dehydrogenase gene. Mean and standard deviation (SD) of four independent biological replicates are shown in the bar plot.

the β-globin variants mRNA levels did not change significantly: the βWT mRNA levels had a 1.34-fold increase, β15 was 1.43-fold higher, β26 increased 1.19-fold, and β39 increased 1.13-fold, when compared to the normal reporter in control conditions (Fig. 1C). Our data using the β-globin reporter variants do not suggest involvement of DIS3L2 in NMD.

Having demonstrated that DIS3L2 regulates some NMD-targets [1], we decided to address if this finding could represent a wider regulatory network by taking advantage of publicly available data from transcriptome profiling of cells depleted of DIS3L2 or UPF1. In order to have a bona fide comparison between datasets, we restricted our analysis to datasets generated in the same cell line, HeLa cells, using the same siRNA silencing-method [2,3]. The analysis pipeline used for the identification of DIS3L2-dependent mRNA targets by Lubas and colleagues, accessible as supplementary data in Ref. [3], was applied to the UPF1 KD raw mRNA-seq dataset from the work by Tani and colleagues [2] available through Sequence Read Archive (SRA). The resulting transcriptome profile is provided here as

supplemental Dataset 1, including estimates of average expression level between conditions, fold change and statistical testing performed with DESeq. We found 3693 genes to be differentially expressed upon UPF1 depletion (adj pval <0.05; Table S2), of which 1828 genes were upregulated. After cross-referencing these gene sets, we found that 76 out of 680 (11%) upregulated genes after DIS3L2 KD are within the pool of the upregulated transcripts detected in UPF1-depleted cells, representing a statistically significant overlap (p-value < 0.001; hypergeometric test) (Fig. 2A, Table S2). A set of 12 natural-NMD targets (BAG1, SLC7A11, ANTXR1, PLXNA1, ARFRP1, GADD45A, GADD45B, SLC1A3, SMG5, GABARAPL1, SMG1 and ATF3) were evaluated by da Costa et al., upon depletion of DIS3L2 to address its role in NMD [1]. Interestingly, 11 out of these 12 NMD-targets are included in the pool of the upregulated transcripts in UPF1-depleted cells, supporting the fidelity of the UPF1 silencing in the analyzed raw mRNA-seq dataset, according to the key role of this protein in the NMD pathway. Remarkably, three of them (ANTXR1, PLXNA1, SLC1A3) are common UPF1/DIS3L2-targets, according to both RNA-seq experiments re-evaluated here (Table S2).

We further examined 785 transcripts that were stabilized in UPF1-depleted cells according to one of the transcriptomic datasets published by Tani and colleagues [2]. In this analysis, the decay rates of the whole transcriptome were measured by 5'bromo-uridine immunoprecipitation chase-deep sequencing (BRIC-seq) analysis [10]. Interestingly, 69 out of the 680 upregulated transcripts after DIS3L2 KD (10%) are included in this set of presumably direct UPF1 targets (Fig. 2B), with a statistically significant overlap between both gene sets (p-value < 0.001; hypergeometric test).

The overlapped portion, as shown in Fig. 2A and B, represents a small percentage of total UPF1 and DIS3L2 affected targets, suggesting that the role of DIS3L2 in NMD is limited to a small subset of targets.

2. Experimental design, materials, and methods

2.1. Plasmid constructs

The pTRE2pur vector (BD Biosciences) containing each of the human β-globin gene variants βWT (wild type), β15 (PTC at codon 15; TGG→TGA), β26 (PTC at codon 26; GAG→TAG), β39 (PTC at codon 39; CAG→TAG) have been used before in our laboratory and were obtained as previously described [6–8].

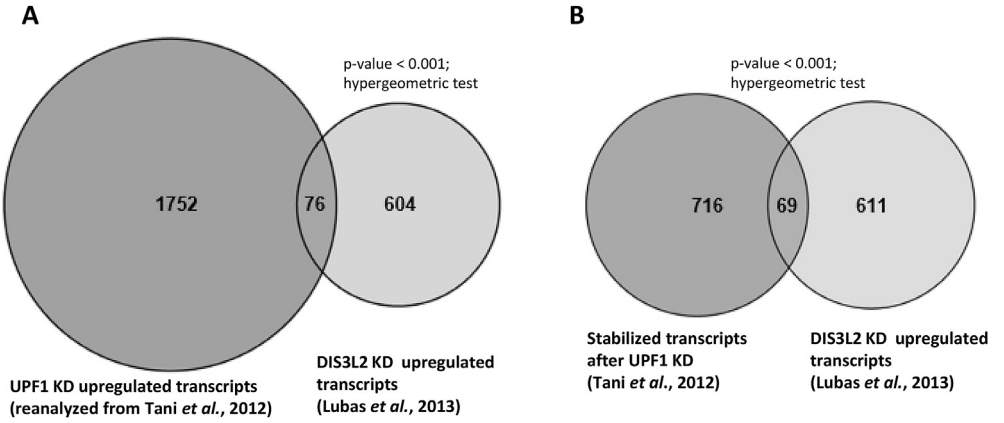

Fig. 2. DIS3L2 regulates a set of transcripts that are also under the regulation of the NMD key factor, UPF1. (A) Analysis of DIS3L2-and UPF1-dependent mRNA targets obtained from mRNA-seq [2,3]. Overlap between transcripts that are upregulated upon UPF1 (left circle) and DIS3L2 (right circle) knockdown (KD). The number of transcripts that are upregulated in both gene sets (76) is represented by the overlap between circles; a hypergeometric test was applied to estimate the statistical significance of this overlap in comparison with the genetic background (p-value < 0.001). (B) Venn diagram of overlap between transcripts from BRIC-seq analysis [3] that are stabilized after UPF1 KD (left circle) and upregulated transcripts in DIS3L2 KD [2] (right circle). The overlap between the circles (69 transcripts) indicates upregulated transcripts after DIS3L2 KD that are also stabilized when HeLa cells are UPF1-depleted (p-value < 0.001; hypergeometric test).

2.2. Cell culture, plasmid and siRNA transfections

HeLa cells were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum. Cells in 35-mm plates were transiently transfected with siRNAs using 200 pmol of siRNA oligonucleotides and 4 μl of Lipofectamine 2000 reagent (Invitrogen) according to the manufacturer's instructions. Twenty-four hours after first siRNA transfection, cells were transfected with additional 50 pmol of siRNAs and 500 ng of the tested construct DNA. The siRNA oligonucleotides were designed with 3′-dTdT overhangs, and purchased as annealed, ready-to-use duplexes from Thermo Fisher Scientific. All sequences are available in da Costa et al. [1]. Cells were harvested for RNA and protein expression analysis, 24 hours after the second transfection, as previously described [11].

2.3. Isolation of total RNA and protein lysates

To obtain total RNA and protein extracts, transfected cells were lysed with 100 μl of NP40 buffer [50 mM Tris-HCl pH = 7.5, 10 mM $MgCl_2$, 100 mM NaCl, 10% (v/v) glycerol and 1% (v/v) Nonidet P-40 (Roche)], supplemented with Proteinase and RNase inhibitors. Then, cells were spun down at maximum speed for 1 minute. For protein analysis, 20 μl of lysate were used and RNA was extracted from remaining supernatant with the Nucleospin RNA extraction II Kit (Macherey-Nagel), as previously described [11].

2.4. Western blot analysis

Protein lysates were collected in 1x SDS-PAGE Sample Loading Buffer (NZYtech), denatured at 95 °C for 10 minutes, and centrifuged at 500g for 5 minutes. Then, lysates were resolved in a 10% SDS-PAGE gel and transferred to polyvinylidene difluoride (PVDF) membranes (Bio-Rad), according to standard protocols. After transfer, membranes were treated with blocking buffer [TBS-0.5% TW20: 50 mM Tris-HCl pH7.5, 150 mM NaCl, 5% non-fat dried milk, and 0.5% (v/v) Tween 20] for 1 hour at room temperature (RT). Membranes were incubated overnight (O/N) at 4 °C with mouse anti-α-tubulin antibody (Roche, loading control) at 1:4000 and rabbit anti-DIS3L2 antibody (Novus Biologicals) at 1:200 dilution in blocking buffer and, then, rinsed 4 × 5 minutes with washing buffer [TBS-0.05% TW20: 50 mM Tris-HCl pH7.5, 150 mM NaCl and 0.05% (v/v) Tween 20]. For detection, the membranes were incubated with the secondary antibody (goat anti-rabbit horseradish peroxidase conjugate; Sigma) diluted 1:3000 (or 1:4000 goat anti-mouse (Bio-Rad) for α-tubulin) in 12 ml of blocking buffer for 1 hour at RT, followed by enhanced chemiluminescence detection, as previously described [11].

2.5. Northern blot analysis

Fifteen micrograms of RNA were separated under denaturing conditions in an agarose MOPS/formaldehyde gel (1%) following the procedure described in Ref. [12]. Riboprobe templates were generated by standard PCR with β-globin specific primers (antisense: GTTTTTTTTAATACGACTCACTATAGGCAGTAACGGCAGACTTCTCCTCAG and sense: ACATTTGCTTCTGACACAAC) and carried out on human cDNA. The *in vitro* transcription reaction was performed in the presence of an excess of [^{32}P]-α-UTP (PerkinElmer) over unlabeled UTP, using T7 polymerase from Promega (a T7 RNA polymerase promoter sequence was introduced in the template DNA by the antisense primer). 28S rRNA was detected using an oligoprobe (AACGATCAGAGTAGTGGTATTTCACC) labelled on its 5'end with Υ-^{32}P from [^{32}P]-Υ-ATP (PerkinElmer) in a PNK phosphorylation reaction (Promega). The riboprobe and oligoprobe were subjected to gel filtration on G-50 and G-25 columns (GE Healthcare) respectively, to remove unincorporated nucleotides, as previously described [11].

2.6. Reverse-transcription coupled to quantitative PCR (RT-qPCR)

One microgram of total RNA was used to synthesize first-strand cDNA using reverse transcriptase (NZYTech) according to the manufacturer's instructions. Real-time PCR was performed with the ABI7000 Sequence Detection System (Applied Biosystems) using SYBR Green PCR Master Mix (Applied

Biosystems). The relative expression levels of β-globin mRNA were normalized to the endogenous GAPDH mRNA and calculated using the comparative Ct method ($2^{-\Delta\Delta Ct}$) [13]. The Ct values of variant β-globin mRNA amplicons were compared to the respective wild-type counterpart or to wild-type at LUC siRNA conditions, as indicated in figures and normalized with the reference amplicon Ct value. Dilution series were performed to determine the amplification efficiencies of each primer pair. The forward and reverse primer sequences used are HBB-F: GTGGATCCTGAGAACTTCAGGC; HBB-R: CAGCACACA-GACCAGCACGT; GAPDH-F: CCATGAGAAGTATGACAACAGCC and GAPDH-R: GGGTGCTAAGCAGTTGGTG. Technical triplicates from four independent experiments were evaluated in all cases. Quantitative PCR without reverse transcription was also performed to discard DNA contamination, as previously described [11].

2.7. Bioinformatic analysis

Transcriptome data for DIS3L2 KD in HeLa cells was retrieved from the supplementary material provided by Lubas et al. [3] and refers to a list of 873 genes identified by differential expression analysis using the DESeq algorithm [3] (Table S2). A dataset of UPF1 KD in HeLa cells was obtained by Tani et al. [2]. To generate a list of genes that can be directly compared to the DIS3L2-dependent transcriptome, raw data for Illumina paired-end mRNA-seq were retrieved from SRA (identifiers DRS001615 - DRS001618) and analyzed as described by Lubas et al. [3]. Data obtained are presented in Table S2. In Supplemental Dataset 1 and Table S2B, the calculated baseMean indicates normalized counts, averaged over all samples from considered conditions, and baseMeanA and B are calculated from normalized counts from condition A and B, respectively. Condition A represents the control condition (LUC siRNA-transfected cells) and condition B corresponds to the KD of DIS3L2 for Table S2B and the KD of UPF1 in the Supplemental Dataset 1. Hypergeometric test was used for assessment of statistical significance of overlaps between gene sets.

2.8. Statistical analysis

When applicable, Student's unpaired, two-tailed *t*-test was used for estimation of statistical significance. Significance for statistical analysis was defined as follows: $p < 0.05$ (*), $p < 0.01$ (**) and $p < 0.001$ (***). Data are expressed as mean ± standard deviation from at least three independent experiments.

Acknowledgments

This work was partially supported by Fundação para a Ciência e a Tecnologia (FCT) (PTFC/BIM-MEC/3749/2014 to LR and UID/MULTI/04046/2013 to BioISI). PJdC, HAS and JFG-M are recipients of a fellowship from BioSys PhD programme (SFRH/BD/52495/2014, SFRH/BD/52492/2014, and PD/BD/142898/2018, respectively) and JM is a postdoctoral fellow (SFRH/BPD/98360/2013) from FCT.

Work at ITQB-NOVA was financially supported by: Project LISBOA-01-0145-FEDER-007660 funded by the European Regional Development Fund (FEDER) through COMPETE2020 - Programa Operacional Competitividade e Internacionalização (POCI) and by FCT funds: PTDC/BIA-MIC/1399/2014 to CMA and PTFC/BIM-MEC/3749/2014 to SCV. SCV was financed by program IF of FCT (IF/00217/2015). MS was financed by an FCT contract according to DL57/2016 [SFRH/BPD/109464/2015].

Conflict of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.dib.2019.104943.

References

[1] P.J. da Costa, J. Menezes, M. Saramago, J.F. García-Moreno, H.A. Santos, M. Gama-Carvalho, C.M. Arraiano, S.C. Viegas, L. Romão, A role for DIS3L2 over natural nonsense-mediated mRNA decay targets in human cells, Biochem. Biophys. Res. Commun. 518 (2019) 664–671.

[2] H. Tani, N. Imamachi, K.A. Salam, R. Mizutani, K. Ijiri, T. Irie, T. Yada, Y. Suzuki, N. Akimitsu, Identification of hundreds of novel UPF1 target transcripts by direct determination of whole transcriptome stability, RNA Biol. 9 (2012) 1370–1379.

[3] M. Lubas, C.K. Damgaard, R. Tomecki, D. Cysewski, T.H. Jensen, A. Dziembowski, Exonuclease hDIS3L2 specifies an exosome-independent 3′-5′ degradation pathway of human cytoplasmic mRNA, EMBO J. 32 (2013) 1855–1868.

[4] R. Thermann, G. Neu-Yilik, A. Deters, U. Frede, K. Wehr, C. Hagemeier, M.W. Hentze, A.E. Kulozik, Binary specification of nonsense codons by splicing and cytoplasmic translation, EMBO J. 17 (1998) 3484–3494.

[5] J. Zhang, X. Sun, Y. Qian, L.E. Maquat, Intron function in the nonsense-mediated decay of b-globin mRNA: indications that pre-mRNA splicing in the nucleus can influence mRNA translation in the cytoplasm, RNA 4 (1998) 801–815.

[6] L. Romão, Â. Inácio, S. Santos, M. Ávila, P. Faustino, P. Pacheco, J. Lavinha, Nonsense mutations in the human β-globin gene lead to unexpected levels of cytoplasmic mRNA accumulation, Blood 96 (2000) 2895–2902.

[7] Â. Inácio, A.L. Silva, J. Pinto, X. Ji, A. Morgado, F. Almeida, P. Faustino, J. Lavinha, S.A. Liebhaber, L. Romão, Nonsense mutations in close proximity to the initiation codon fail to trigger full nonsense-mediated mRNA decay, J. Biol. Chem. 279 (2004) 32170–32180.

[8] A.L. Silva, P. Ribeiro, Â. Inácio, S.A. Liebhaber, L. Romão, Proximity of the poly(A)-binding protein to a premature termination codon inhibits mammalian nonsense-mediated mRNA decay, RNA 14 (2008) 563–576.

[9] W.K. Kim, S. Yun, Y. Kwon, K.T. You, N. Shin, J. Kim, H. Kim, mRNAs containing NMD-competent premature termination codons are stabilized and translated under UPF1 depletion, Sci. Rep. 7 (2017) 15833.

[10] H. Tani, R. Mizutani, K.A. Salam, K. Tano, K. Ijiri, A. Wakamatsu, T. Isogai, Y. Suzuki, N. Akimitsu, Genome-wide determination of RNA stability reveals hundreds of short-lived noncoding transcripts in mammals, Genome Res. 22 (2012) 947–956.

[11] P.J. da Costa, J. Menezes, M. Saramago, J.F. García-Moreno, H.A. Santos, M. Gama-Carvalho, C.M. Arraiano, S.C. Viegas, L. Romão, A Role for DIS3L2 over Human Nonsense-Mediated mRNA Decay Targets, bioRxiv, 2019, https://doi.org/10.1101/722702.

[12] M. Malecki, S.C. Viegas, T. Carneiro, P. Golik, C. Dressaire, M.G. Ferreira, C.M. Arraiano, The exoribonuclease Dis3L2 defines a novel eukaryotic RNA degradation pathway, EMBO J. 32 (2013) 1842–1854.

[13] M.W. Pfaffl, A new mathematical model for relative quantification in real-time RT-PCR, Nucleic Acids Res. 29 (2001) 45e.