Transcriptional profile of ribosome-associated quality control components and their associated phenotypes in mammalian cells

Otávio Augusto Leitão Dos Santos[1#], Rodolfo L. Carneiro[1#], Rodrigo D. Requião[2], Marcelo Ribeiro-Alves[3], Tatiana Domitrovic[4] and Fernando L. Palhano[1]*

[1] Programa de Biologia Estrutural, Instituto de Bioquímica Médica Leopoldo de Meis, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, 21941-902, Brazil.

[2] Departamento de Genética, Evolução, Microbiologia e Imunologia, Instituto de Biologia, Universidade Estadual de Campinas, Campinas, SP, Brazil.

[3] Fundação Oswaldo Cruz, Instituto Nacional de Infectologia Evandro Chagas, Rio de Janeiro, 21040-900, Brazil.

[4] Departamento de Virologia, Instituto de Microbiologia Paulo de Góes, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 21941-902, Brazil.

# These authors contributed equally.

* To whom correspondence should be addressed: Tel.: +55 21 3938-6761; Email: palhano@bioqmed.ufrj.br

Keywords: proteostasis, ribosome quality control, protein aggregation,

Abstract

During protein synthesis, organisms constantly survey defective translation events that may lead to ribosome stalling with consequent ribosome collisions. Translation problems

can result in misfolded protein products that sequester ribosomes into intracellular protein aggregates. Cells have co-translational quality control mechanisms to eliminate these aberrant protein products before they are released from the ribosome. One of these mechanisms is the Ribosome-associated Quality Control (RQC) complex. To better understand the role of the RQC complex in mammals, we used public datasets of transcriptome and proteomics analysis to study the tissue distribution and profile of the three RQC components, namely, LTN1, TCF25, and NEMF, together with the ubiquitin ligase ZNF598, responsible for RQC recruitment. Our data shows that the RQC components are widely expressed in human tissues with no preference and their silencing lead to mild effects. Interestingly, TCF25 shows a particular profile with high mRNA levels that is not reflected in its protein content, suggesting a post-translational regulation. Finally, cellular aging caused a mild impact on their expression, causing a decrease in mRNA levels in some human tissues. Our data suggest that the RQC is essential in maintaining proteostasis in all mammalian tissues but is not vital for cell survival.

Introduction

Protein synthesis is a highly regulated cellular process necessary for maintaining cellular homeostasis (Yip & Shao, 2021). However, this process is prone to different errors that can cause ribosome stalling followed by ribosome collisions and premature termination. Stalled ribosomes can be a source of aberrant misfolded and dysfunctional polypeptides, which can aggregate and disrupt cellular homeostasis. In addition, stalled ribosomes are prevented from initiating new translation cycles, affecting the entire dynamics of protein synthesis (Howard & Frost, 2021). Ribosome stalling may occur when the ribosome encounters structural defects in the mRNA molecule, such as the presence of premature termination codons, or the absence of termination codons, which lead to the mRNA nonsense-mediated decay and non-stop decay, respectively (Losson & Lacroute, 1979; Maquat *et al.*, 1981; Frischmeyer *et al.*, 2002; van Hoof *et al.*, 2002; Lykke-Andersen & Bennett, 2014). Additionally, the presence of strong structural elements, rare codons or polylysine coding tracts in the mRNA are also able to cause ribosome stalling, which, in this case, may lead to mRNA no-go decay (Ito-Harashima *et al.*, 2007; Lu & Deutsch, 2008; Tsuboi *et al.*, 2012; Lykke-Andersen & Bennett, 2014; Joazeiro, 2017; Garzia *et al.*, 2017; Guydosh & Green, 2017; Barros *et al.*, 2021). These mRNA degradation pathways lead to the recycling of the 40S ribosomal subunit, but leave the 60S subunit

with an incomplete nascent peptide occupying the exiting tunnel and a tRNA linked to the incomplete polypeptide (Lykke-Andersen & Bennett, 2014; Joazeiro, 2017).

Initially identified in yeast (Brandman et al., 2012), the Ribosome-associated Quality Control (RQC) complex acts during the elongation step of translation to correct the ribosome stall in eukaryotes. It is responsible for the ubiquitination and extraction of the nascent peptide, which also leads to recycling the 60S subunit. The RQC complex comprises the proteins LTN1, NEMF, and TCF25, which show molecular interaction with each other and the 60S subunit bound to uncompleted polypeptides (Brandman et al., 2012; Defenouillère et al., 2013). In the RQC pathway, ribosomes stalled and collided during translation are first sensed by factors that eventually promote subunit dissociation, a process mediated mainly by ZNF598 (Matsuo & Inada, 2021). ZNF598 is an E3 ligase that recognizes the unique structural architecture at the collision interface and ubiquitinates the 40S subunit, which leads to subunit dissociation and permits the association of the RQC with the 60S (Ikeuchi et al., 2019; Juszkiewicz et al., 2018; Matsuo et al., 2020; Matsuo & Inada, 2021; Winz et al., 2019). After ribosome dissociation, NEMF binds to the aberrant 60S subunit contacting the polypeptide-bound tRNA, a conformation that confers specificity to aberrant 60S and prevents 40S subunit reassociation (Filbeck et al., 2022). Moreover, NEMF promotes the mRNA-independent synthesis of the C-terminal Alanine Threonine (CAT) tail (Shen *et al.*, 2015). This CAT tail has two distinct functions: it signals stress to the cell due to its hydrophobic nature, and it aids nascent peptide ubiquitination through the exposure of lysine residues that may be enclosed inside the ribosomal exit tunnel (Brandman *et al.*, 2012; Kostova *et al.*, 2017). Following NEMF biding to the 60S, the ubiquitination of the nascent peptide is performed by the E3 ligase LTN1, a process that apparently depends on TCF25 (Kuroha et al., 2018; Osuna et al., 2017). The molecular function of TCF25 is still not fully understood, but as previously shown by our group, the expression of its yeast homolog, Rqc1, is regulated post-translationally by LTN1, perhaps as a negative feedback mechanism to ensure that the RQC complex is not over activated (Barros *et al.*, 2021). It is worth noting that TCF25 was originally described as a transcriptional factor, and it was recently implicated in the expression regulation of NFAT (Nuclear factor of activated T-cells) in cardiomyocytes and the activation of T-cells during papillary renal cell carcinoma (Olsson *et al.*, 2002; Cai *et al.*, 2006; Zhang *et al.*, 2020; Wang *et al.*, 2021).

Defects in the RQC complex have been associated with neurodegeneration and protein aggregation. Chu *et al.*, 2009, demonstrated that homozygous LTN1 knockout mice exhibit age-dependent loss of locomotor ability. Besides, pathological analyses evidenced dystrophic neurites, gliosis, vacuolated mitochondria, and accumulation of soluble hyperphosphorylated tau in these mice. Although LTN1 is widely expressed in all tissues, motor and sensory neurons and neuronal processes were the most affected (Chu et al., 2009). Choe *et al.*, 2016, showed that deletion of the LTN1 in yeast causes stalled nascent peptides to form detergent-resistant aggregates and inclusions. This aggregation process depended on CAT tail formation by Rqc2, the yeast homolog of NEMF, while the double knockout ΔLTN1ΔRqc2 prevented the formation of these aggregates in yeast (Choe et al., 2016).

Similarly, CAT tailing by NEMF in HeLa cells can form aggregates when LTN1 is depleted (Udagawa et al., 2021). The dysfunction of NEMF is also associated with neurological disorders. Martin *et al.*, 2020, reported that mice with NEMF mutations R86S and R487G present many neuromuscular phenotypes. Moreover, R86S mice died prematurely, with a median lifespan of 20 days. Furthermore, it was already observed that people with mutations in NEMF display intellectual disability, speech delay, neuropathy, and motor dysfunction with varying severities and progression (Martin et al., 2020).

Recently, it was shown that aged *Saccharomyces cerevisiae* and *Caenorhabditis elegans* under chronological aging downregulate the mRNA of some RQC components and have increased ribosome stalling and collisions at polybasic tracts (Stein et al., 2022). It was proposed that the RQC complex is saturated due to an age-dependent increase in ribosome stalling, which could lead to impaired proteostasis observed in older individuals. However, we do not know if aging affects the activity and expression of RQC components in mammalian tissues.

The above evidence indicates that neuronal cells are affected when the RQC complex is perturbed. It is unclear, however, if the expression of the components of the RQC complex is tissue-specific, if their mRNA levels are co-regulated, and if aging plays a role in their expression. To address this issue, we used publicly available transcriptome and proteome profiles (**Table 1**) to evaluate the expression of the LTN1, NEMF, and TCF25. In addition, even though ZNF598 is not regarded as an integrant part of the RQC complex (Juszkiewicz et al., 2018), its function is responsible for many of RQC subtracts.

Therefore, ZNF598 was also included in our analysis and, for the sake of simplicity, will be referred as a RQC component.

We found that all RQC components are ubiquitously expressed in the analyzed tissues. Interestingly, TCF25 showed a high number of mRNA transcripts, which does not translate into high protein levels in most tissues analyzed. In addition, when compared to the rest of the genome, TCF25 had increased mRNA half-lives coupled with shorter protein half-lives, suggesting a post-translational regulation, as previously proposed for the yeast ortholog Rqc1 (Barros et al., 2021). We also looked for the consequences of silencing RQC components in different cellular lineages and observed mild effects. Finally, we dissected the effect of aging on the mRNA levels of the RQC components in humans and mice. In general, aging caused a mild effect on their expression, causing a decrease in mRNA levels in some human tissues. Moreover, we found that LTN1 and NEMF expression and phenotypes tend to be similar, while TCF25 and ZNF598 had specific, uncorrelated patterns. In conclusion, our data points out that the RQC components are ubiquitously expressed in mammalian tissues, and we found support for aging-dependent downregulation that might regulate RQC function. Additionally, although TCF25 interacts with LTN1 and NEMF, it is subjected to different expression regulations.

Data	Source	Identifier
Proteome and transcriptome abundance	Wang et al., 2019	
Proteome and transcriptome abundance	Jiang et al., 2020	
mRNA half-life	Agarwal & Kelley, 2022	
Translation efficiency	Ingolia et al., 2011	
Protein half-life	Mathieson et al., 2018	
Dependency Map	Tsherniak et al., 2017	https://depmap.org/portal/
Haploinsufficiency score	Collins et al, 2022	
Perturb-Seq CRISPRi gene silencing	Replogle et al. 2022	https://gwps.wi.mit.edu/

CRISPR-Cas9-based screening for GFP fluorescence level	Hickey et al., 2020	
Transcriptome of human organs and tissues in age clusters	The GTEx Consortium, 2020.	https://www.gtexportal.org
Transcriptome of mouse organs and tissues in age clusters	The Tabula Muris Consortium, 2018	https://tabula-muris.ds.czbiohub.org/

Materials and Methods

Transcriptome and translatome profile

To assess the transcription and translation profile of the components of the RQC complex, we obtained data from Wang et al., 2019 obtained from 29 healthy human tissue samples. For each RQC complex component (LTN1, TCF25, NEMF, and ZNF598), the transcripts per million (TPM) values were used to estimate the Z-score for each gene across the tissue, and it is shown as a heat map. To study the proteome in each tissue, we used the dataset containing relative protein abundances (log_2) and estimated the Z-score for each gene. A similar approach was performed with Jiang et al. 2020 dataset in the GTEx Consortium. From this data, we obtained the transcriptome and proteome of 114 samples from 32 major organs from 14 individuals, with replicates ranging from 1 to 22.

RNA and protein half-life

To assess the half-life of transcripts for each gene in the RQC complex, we used the dataset produced by Vikram Agarwal and David Kelley, 2022. They generated a compendium of mammalian (mouse and human) mRNA half-life datasets from 33 publications obtained from cell culture and normalized the values using the Z-score. For protein half-life, we assessed the data from five different non-dividing cell types (human B-cells, monocytes, NK cells, hepatocytes, and mouse embryonic neurons) by Mathieson et al. 2018. We calculated mean protein half-live values considering all cells and

compared them with the values for each component of the RQC complex in each cell line analyzed.

Translational efficiency

To analyze the mRNA translation efficiency (TE), we downloaded the dataset obtained from ribosome profiling experiments of mouse embryonic stem cells published by Ingolia et al., 2011. Then, we compared the TE of the RQC components with the global profile and divided the data into quartiles.

Impact on growth analyses

The impact of the deletion of RQC complex components genes was obtained using the large-scale gene essentiality data from the DepMap project (Tsherniak et al., 2017). They performed genome-scale library of 100,000 short hairpin RNAs (shRNAs) to analyze a genome-scale loss-of-function screens performed in 1086 human cell lines. A scoring system was created to quantify the effect of gene inactivation on cellular growth in a specific cell lineage. If the gene inactivation causes no effect in a specific cell lineage, a score equal to zero is attributed to this pair (cell line/gene inactivated). On the other side, a score of -1 means that this pair leads to cell death, and it corresponds to the median of all common essential genes. Data were obtained directly from the DepMap portal (https://depmap.org). We searched in "perturbation effects" and downloaded the dataset (DepMap 22Q2 Public+Score, Chronos) for each RQC complex gene. Furthermore, we plotted the correlation between growth score values obtained for NEMF and LTN1 silencing on different cell types.

Haploinsufficiency score

The haploinsufficiency score was created through the meta-analysis of approximately one million individuals in both disease and control cohorts to identify disease-associated dosage-sensitive regions, allowing for dosage sensitivity for each gene and creating a haploinsufficiency score (Collins et al, 2022). We used the score obtained for each gene

of interest for comparison purposes. Genes with values equal to or greater than 0.86 are considered haploinsufficient.

Perturb-Seq CRISPRi gene silencing

To evaluate the effects of RQC impairment on a transcriptional level, we analyzed data from Replogle et al., 2022, which performed a genome-scale disturb-seq in leukemia K562 cell line and investigated the mRNA expression profile of individually silenced genes. The data was obtained from the online platform https://gwps.wi.mit.edu/, which separates the top 30 up and downregulated genes for each selected gene. We used the K562 Genome-Wide Perturb-Seq database, selected the top 30 up and downregulated component genes of the RQC complex, and separated them into clusters to understand the transcriptional profile. Also, we used the 10 genes with the most similar transcriptional profiles provided by the tool for each RQC complex component.

Gene ontology

We selected gene clusters (up and downregulated) that present similar transcriptional profile between NEMF and LTN1 and performed a gene ontology analysis. The gene ontology analyses were performed in the Gene Ontology Consortium (http://geneontology.org/) in the GO Enrichment Analysis tool, searching for "Biological process". We selected the results with the highest Fold enrichment.

FACS-based CRISPRi screen

To screen if the genes participate in the RQC complex, we used data from Hickey et al., 2020, which measures the expression of a fluorescence reporter lacking a stop codon (leading to non-stop decay) in a genome-wide CRISPRi screening. Positive values indicate that the knockdown of that gene stabilizes the GFP signal. We used the RQC complex component as a control.

Age-dependent transcriptome

To assess if the transcriptional profile of RQC components changes with age, we used data from the GTEx consortium for human tissues (The GTEx Consortium, 2020) and The Tabula Muris Consortium for mouse tissues (The Tabula Muris Consortium, 2018). We downloaded the data from the GTEX consortium data on the website https://www.gtexportal.org for each human tissue and separated them into age clusters (20-29, 30-39, 40-49, 50-59, 60-69, and 70-79). We used the transcript per million (TPM) to determine the mRNA abundance across the human tissues at different ages. The data was normalized by the reference age value (20-29). For mouse data, we used The Tabula Muris Consortium, which offers transcriptomes of 20 organs from four male and three female mice from 1 to 27 weeks old. Data handling and data presentation were similar to that of human samples.

Results and Discussion

RQC mRNA and protein expression across human tissues

To describe and compare the expression levels of the RQC components in different human tissues, we curated previously published work that generated paired transcriptome and proteome datasets of human samples. Wang et al., 2019, investigated the transcriptome and proteome from 29 different healthy human tissues from the Human Protein Atlas project (Wang et al., 2019). Using this dataset, we investigated the mRNA and protein levels of LTN1, TCF25, NEMF, and ZNF598 (**Figure 1**). The Z-score for each gene was calculated using the transcripts per million (TPM) values when analyzing the transcriptome and the relative protein abundance ($\log_2$) when analyzing the proteome. (**Figure 1**). All RQC components were expressed in most tissues analyzed at roughly similar levels. So, even though RQC malfunction impacts the nervous system in vertebrates (Chu et al., 2009; Martin et al., 2020), this specificity could not be explained by particular expression levels of RQC components in the nervous tissue when compared to others. In general, LTN1 and ZNF598 showed lower mRNA levels than the average gene expression of each tissue. The mRNA and protein levels showed similar expression levels for LTN1, NEMF, and ZNF598. The exception was TCF25. The mRNA levels for TCF25 were higher than the average gene expression in every tissue analyzed, and it was the most abundant transcript among the RQC components (**Figure 1**). However, this high

amount of mRNA was not reflected in protein levels, which was generally similar to the other RQC components.

In this study, Wang et al., 2019 used only one sample per tissue to generate the protein and mRNA expression data. Thus, the small sample size undermines conclusions about subtle expression variations between tissues. However, despite this limitation, this dataset allowed us to conclude that: 1) both mRNA and proteins of the RQC components are expressed in most tissues, 2) TCF25 mRNA is expressed in high levels while its protein has expression levels closer to the proteome average in most tissues (**Figure 1**).

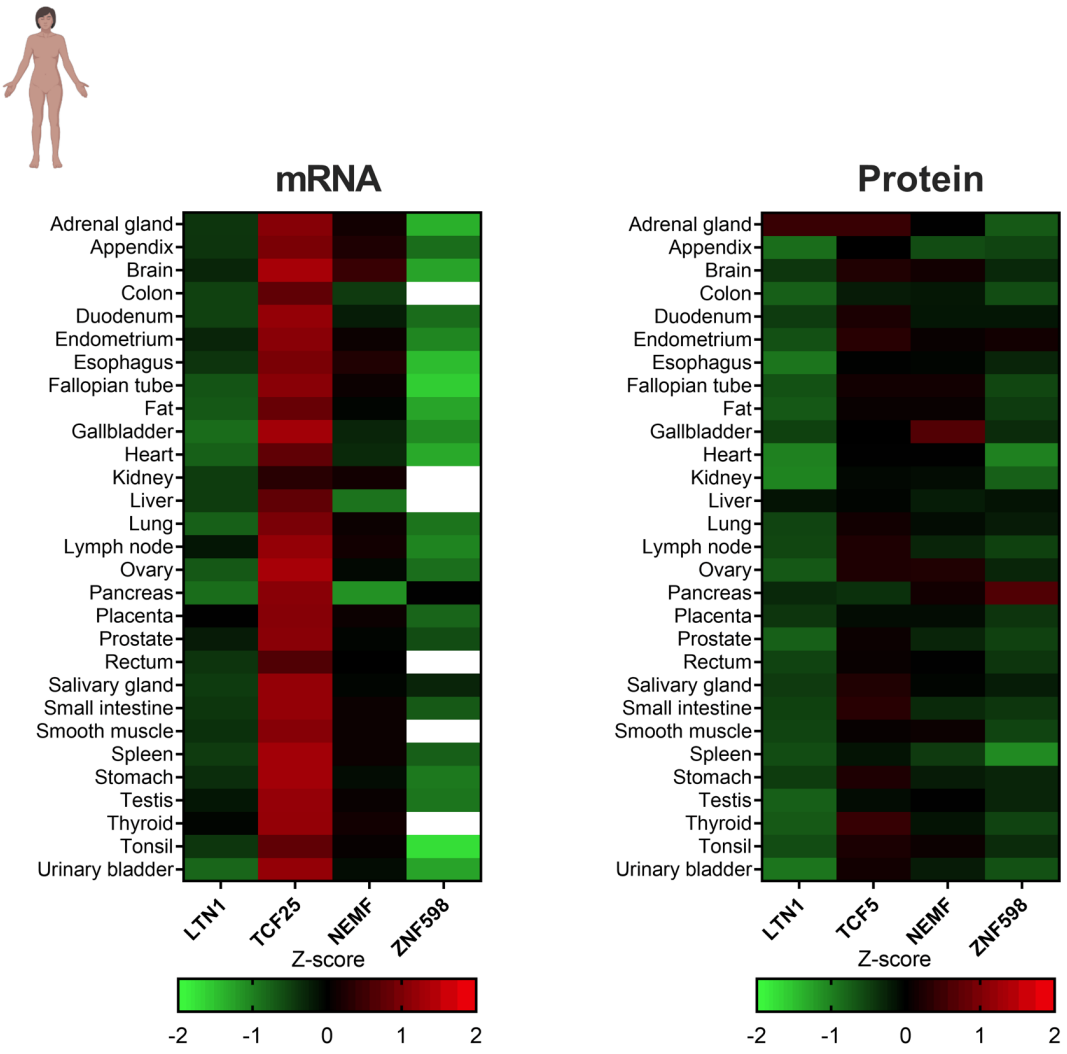


Figure 1. Heat map of the Z-score distribution of the RQC components calculated from the transcriptome and proteome of human tissue samples. Data obtained from Wang et al., 2019. White rectangles denote absence of data.

We repeated the same analysis using another data set with a bigger sample size. Jiang et al., 2020 analyzed the transcriptome and proteome of 114 samples from 32 major organs from 14 different individuals, with biological replicates ranging from 1-22 (Jiang et al., 2020). As in figure 1, we used the Z-score to compare the mRNA and protein levels of the RQC components (**Figure 2**). We observed that, overall, the expression levels were similar to or higher than the average of the genome (**Figure 2**). For NEMF and ZNF598, protein levels were slightly higher than the global average, showing a positive correlation with their mRNA levels. TCF25 mRNA levels were above the 75th percentile in all 32 tissues studied. On the other hand, TCF25 protein expression levels were above the 75th percentile just in 10 out of 32 tissues, namely, the adrenal gland, brain cortex, minor salivary, ovary, pancreas, pituitary, prostate, thyroid, uterus, and vagina. For the other 22 tissues, the protein Z-scores were relatively lower than the mRNA Z-score (**Figure 2**).

However, the two datasets pointed out that TCF25 protein levels are relatively low for the mRNA level detected in many tissues compared to other RQC components (**Figures 1 and 2**). This observation suggests that TCF25 may be subjected to a post-transcriptional expression control.

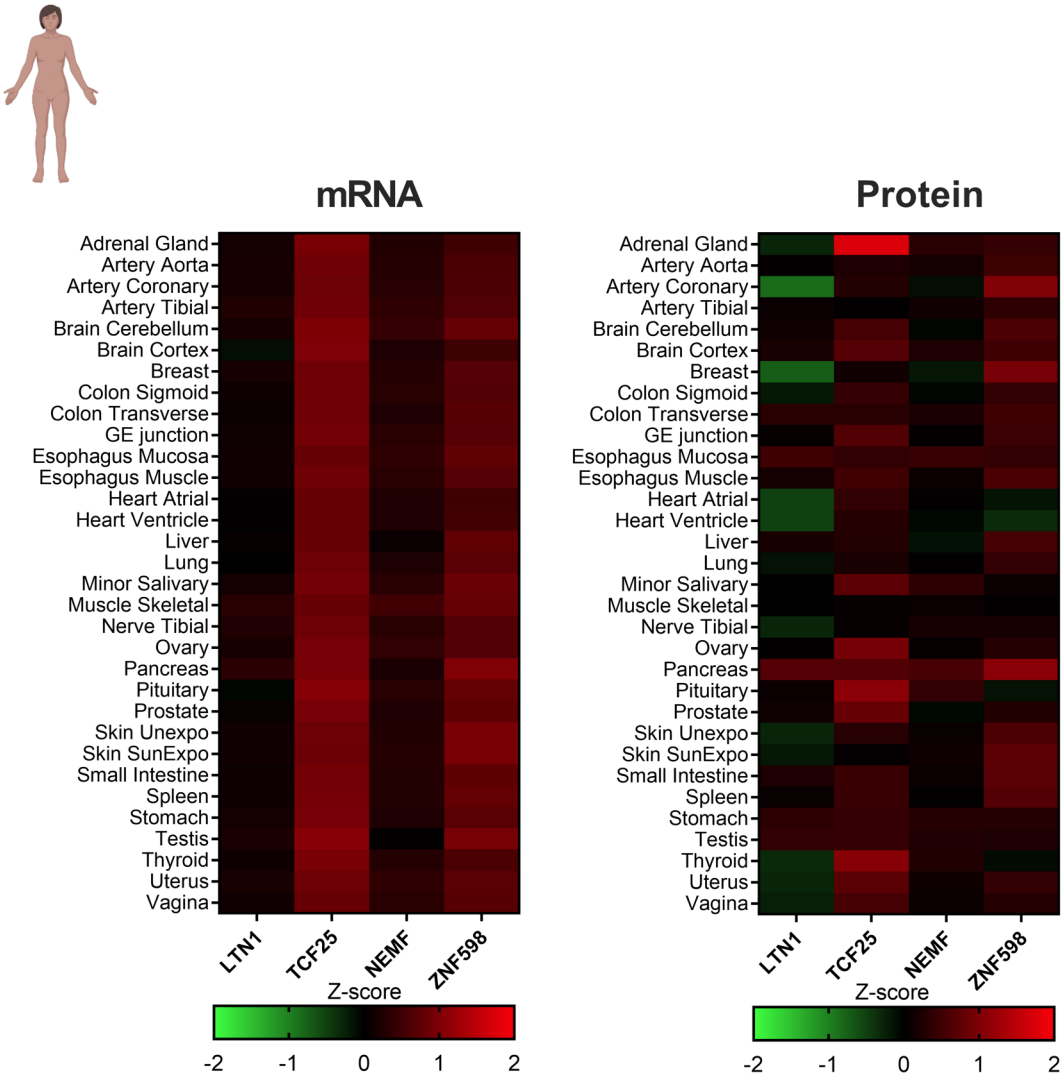


Figure 2. Heat map of the Z-score distribution of the RQC components calculated from the transcriptome and proteome of 114 samples of 32 major organs from 14 different individuals, with replicates ranging from 1-22. Data obtained from Jiang et al. 2020.

The RNA transcripts are determinants for protein abundance, and, in most cases, higher mRNA levels lead to higher protein abundance (Vogel & Marcotte, 2012). However, many factors can affect the mRNA/protein ratio in a dynamic and complex process (Greenbaum et al., 2003). The amount of protein generated by mRNA results from different, non-mutually exclusive factors or regulatory processes such as codon usage, mRNA half-live, and protein turnover (Maier et al., 2009; Vogel & Marcotte, 2012).

To assess if TCF25's relatively low protein levels observed in Figures 1 and 2 could be related to processes involving mRNA stability, we evaluated the half-live of the mRNA of each RQC component (**Figure 3A**). We used data produced by Vikram Agarwal and David Kelley, 2022, which generated a compendium of mammalian (mouse and human) mRNA half-life datasets from 33 publications obtained from cell culture (Agarwal & Kelley, 2022). For both human and mouse cells, the mRNA half-life for LTN1, NEMF, and ZNF598 tended to be close to most transcripts in these cells (**Figure 3 dotted line**). In contrast, TCF25 showed a slightly higher mRNA half-life with mean Z-score values around 0.9 and 0.6 for human and mouse cells, respectively. These findings indicate that the decrease in TCF25 protein levels is not due to a short half-life of its mRNA. It is important to note that while mRNA and protein quantification presented herein were obtained from studies using human tissues (**Figure 1 and 2**), the mRNA half-life data were obtained using cultured cells (**Figure 3A**).

Another factor that can modulate the protein amount produced by mRNA is the translation efficiency (TE). To investigate the effect of TE on TCF25 expression, we used a dataset from mouse embryonic stem cells published by Ingolia *et al.*, 2011. They determined the translation elongation rates for the entire transcriptome through ribosome profiling. The TE values distribution showed that TCF25 accompanied NEMF and ZNF598 in the first quartile of the distribution (**Figure 3B)**. Therefore, the translation efficiency is probably not the factor behind the relatively low TCF25 protein levels. Lastly, we investigated the protein half-life and turnover using data from five different non-dividing cell types (human B-cells, monocytes, NK cells, hepatocytes, and mouse embryonic neurons) by Mathieson *et al.*, 2018. All RQC components and ZNF598 showed a shorter half-life than the average, but TCF25 showed the lowest half-life with less than 20 hours for all cell types tested (**Figure 3C)**.

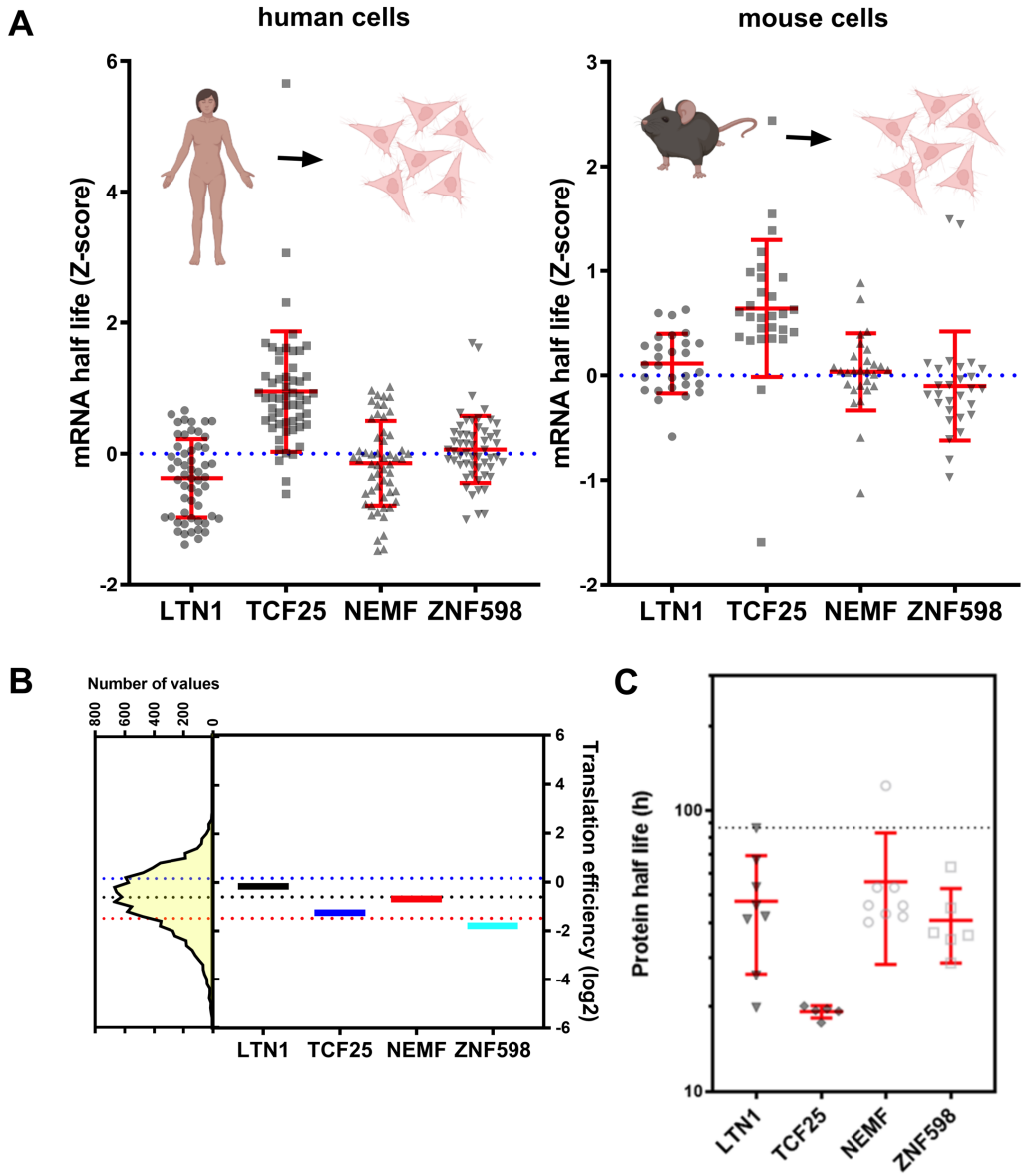


Figure 3. Factors that affect the abundance of RQC complex components. (**A**) Z-score calculated for the half-life of RQC components transcripts in human and mouse cells. Data obtained from Agarwal & Kelley, 2022. (**B**) Log_2 of the mRNA translation efficiency of each transcript in mouse embryonic stem cells. Data obtained from Ingolia et al. 2011. (**C**) Protein half-life for each RQC component. The dotted line refers to the average half-life for all proteins considering all cell lines tested (86.95). Data obtained from Mathieson et al. 2018.

The effect of RQC components disruption in tissue cells survival and gene expression profiling

We saw that the components of the RQC complex are ubiquitously expressed in mammals, which indicates a housekeeping role. Besides the well-established co-translational quality control function described above, RQC components can be implicated in other processes, as in the case of TCF25. Nevertheless, gene deletion of RQC components showed mixed results depending on the organism. Deletion of LTN1 has no detectable effect in yeast cells growing in optimal condition, and double knockout of LTN1 and Rqc1 are also viable (Brandman *et al.*, 2012; Barros *et al.*, 2021). These observations suggest that activation of the RQC complex may be only necessary for specific stress situations. On the other hand, knockout mice for LTN1 exhibit strong neuromotor phenotype, and NEMF mutations in humans and mice also lead to neuromotor impairment (Chu *et al.*, 2009, Martin *et al.*, 2020), which points that the RQC is necessary during neuronal development.

To further examine to what extent the RQC genes are fundamental components of living cells, we utilized the large-scale gene essentiality data from the DepMap project (Tsherniak et al., 2017). A scoring system was created to quantify the effect of gene inactivation on cellular growth in a specific cell lineage (Tsherniak et al., 2017). If the gene inactivation causes no effect on cellular growth in a specific cell lineage, a score equal to zero **(dotted black line, Figure 4A)** is attributed to this pair (cell line/gene inactivated). On the other side, a score of -1 **(blue dotted line, Figure 4A)** means that this pair leads to cell death and corresponds to the median of all common essential genes. The lowest score for the RCQ components was -0.7, meaning that they are not essential for the *in vitro* growth of mammalian cells (**Figure 4A**). However, the median score for all RQC components analyzed was below zero, suggesting that the perturbation of RQC causes some negative effect on cellular growth (**Figure 4A**).

Another feature extracted from this dataset is the correlation between growth score values obtained with different silenced genes on different cell types. We found a positive correlation between NEMF and LTN1, meaning that the silencing of these genes causes a similar effect on growth depending on the cellular lineage analyzed (**Figure 4B**). To see which cellular lineages are more affected by NEMF and LTN1 silencing, we arbitrarily defined a cutoff of -0.25 (**blue circles of Figure 4B**) for both genes. Soft tissue and nervous system cellular lineages were the most affected by NEMF and LTN1 silencing (**Figure 4C**), which agrees with the neurological impairments observed *in vivo* (Chu *et al.*, 2009, Martin *et al.*, 2020).

Another important consideration is to assess the haploinsufficiency of the genes of the RQC. *Haploinsufficiency* is defined as a dominant gene phenotype in diploid organisms, in which a single copy of the wild-type allele at a locus in a heterozygous combination with a variant allele is insufficient to produce the wild-type phenotype. In this context, we evaluated whether defects in one of the alleles of the genes of the RQC complex would be sufficient to alter the wild-type phenotype and thus correlate with the growth impairment observed in the knockout cells **(Figure 4C)**. In addition, recent work performed a meta-analysis of approximately one million human individuals in disease and control cohorts to identify disease-associated dosage-sensitive regions (Collins et al, 2022). The authors predicted dosage sensitivity for each gene, creating a haploinsufficiency score. In other words, a score value equal to or higher than 0.86 (**blue dotted line, Figure 4D**) means that a single copy of the wild-type allele at a locus in a heterozygous combination with a variant allele is insufficient to produce the wild-type phenotype. When we analyzed the haploinsufficiency score of the RQC components, we observed that all genes had a score lower than 0.86, meaning that these genes are not haploinsufficient because just a single copy of the wild-type allele is enough to produce the wild-type phenotype (**Figure 4D**). This data agrees with the description of families with pathogenic biallelic NEMF variants with a spectrum of central neurological disorders (Martin et al., 2020). In conclusion, in vertebrate cellular lineages, knockdown of the RQC components causes a mild phenotype that can be exacerbated in specific tissues (**Figure 4A**). This result suggests that even though the RQC components are ubiquitously expressed genes (**Figures 1 and 2**), the cell can tolerate the translational perturbations caused by RQC impairment and keep growing, perhaps by adapting the transcriptional program.

The effect of RQC components disruption in tissue cells transcriptome

To understand how cells respond to RQC impairment on a transcriptional level, we analyzed data from Replogle *et al.*, 2022, which performed a genome-scale CRISPRi disturb-seq in leukemia K562 cell line to determine the mRNA expression profile of individually knockdown genes. The transcriptional responses were compared with the control cell. (Replogle et al. 2022). Using these data, we separated the top 30 up and downregulated genes of the RQC components in clusters **(Figure 5A)**. As expected, LTN1 and NEMF showed similarities in both up and downregulated genes, but the

silencing of TCF25 showed different patterns (**Figure 5A**). These results further support that TCF25 has additional functions other than co-translational control, as the transcriptional regulation described above, and additional studies are necessary to evaluate which genes are being regulated directly by this protein. We performed a gene ontology analysis of the gene clusters that showed the greatest similarity in expression profile when we silenced LTN1 and NEMF. The gene ontology of the upregulated genes was from genes involved in negative regulation of the ubiquitination process and ribosome assembly (**Figure 5B**), while the gene ontology for the downregulated genes was from genes involved in the electron transport chain and free ubiquitin chain polymerization genes (**Figure 5B**).

Another interesting feature extracted from the Perturb-seq data is a list of silenced genes presenting a similar transcriptional response. Figure 6 shows a Venn diagram with the top 10 genes that generated the most similar transcriptional phenotype compared to the knockdown of the RQC components. None of the 4 RQC genes produced a similar transcriptional response (top 10). Moreover, ZNF598 and TCF25 lists showed no overlap whatsoever, but LTN1 and NEMF had 7 hits in common (TMA16, SPATA5L1, HERC1, TEX10, LAS1L, DDX55, AND SURF6). Interestingly, none of these 7 genes were previously described in the RQC pathway and have diversified functions with no obvious connections with ribosome quality control or proteostasis **(Supplementary table 1)**. To check if these genes could participate in the co-translational quality control, we used data from Hickey *et al.*, 2020, which measures the expression of a fluorescence reporter lacking a stop codon (leading to non-stop decay) in a genome-wide CRISPRi screening. As expected, the knockdown of LTN1 and NEMF greatly enhanced GFP fluorescence, while TCF25 and ZNF598 showed smaller but statistically significant results (**Figure 6**). In contrast, none of the seven genes identified in Perturb-seq analyses (**Figure 6**) affected the GFP fluorescence level, indicating that they are not directly involved in co-translational quality control pathways (**Figure 6**).

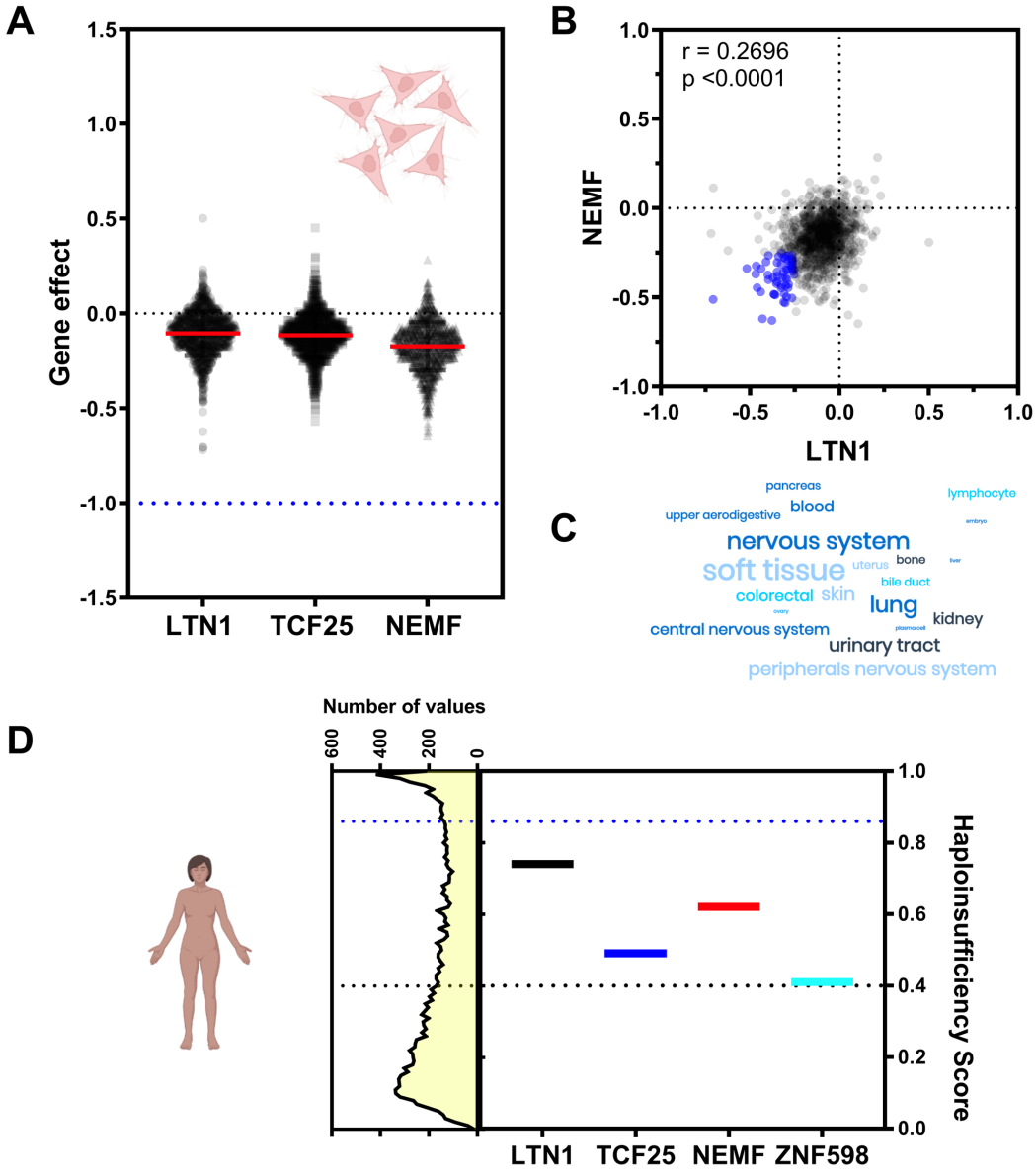


Figure 4. The genes of the components of the RQC complex are not essential. (**A**) Score attributed to the impact of gene inactivation on growth rate in 1086 cancer cell lines for LTN1, NEMF, and TCF25. Data obtained from Tsherniak et al., 2017. (**B**) Correlation graph of silencing effect between NEMF and LTN1. (**C**) Word cloud representation of the cellular lineages is more affected by NEMF and LTN1 silencing. (**D**) Haploinsufficiency score for each component of the RQC complex and ZNF598. Genes with values equal to or greater than 0.86 are considered haploinsufficient. Data obtained from Collins et al, 2022.

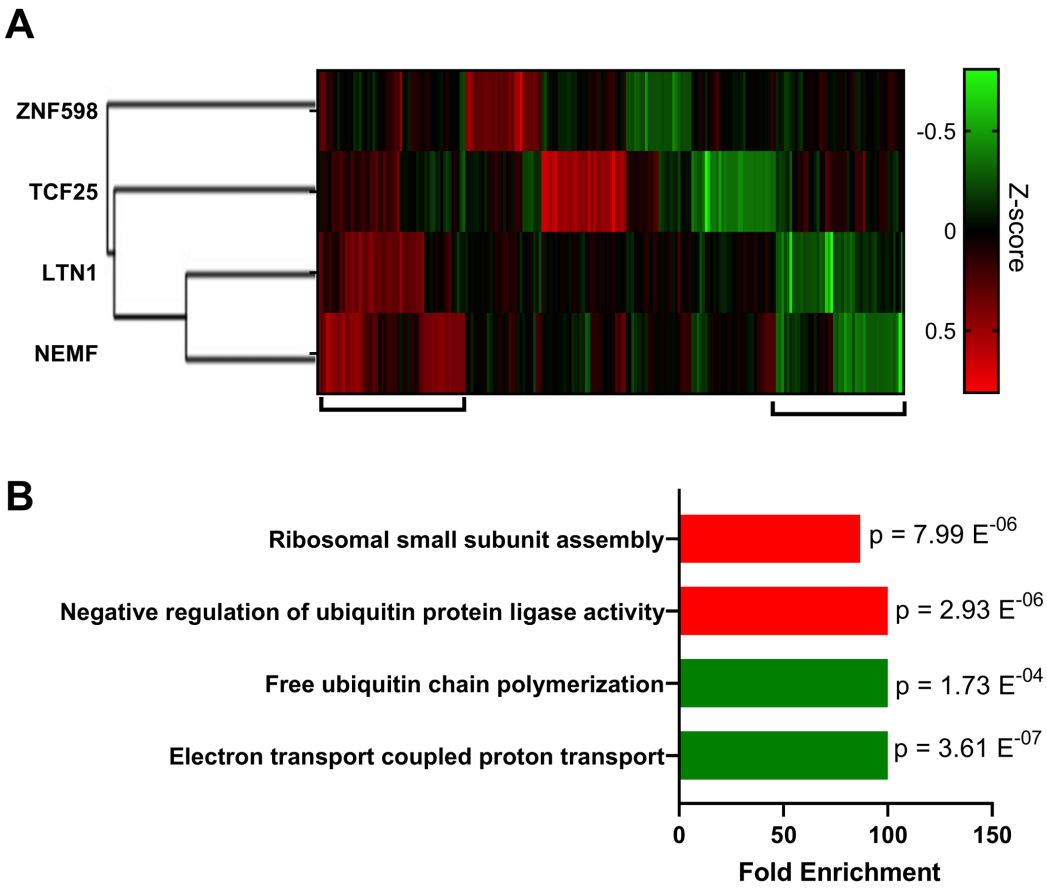


Figure 5. Transcriptional regulation elicited by knockdown of RQC in human cells (**A**) Heat map of the Z-score distribution of the transcriptional profile when RQC genes were downregulated by CRISPR interference (CRISPRi) in leukemia K562 cell line. Data obtained from Replogle et al. 2022. (**B**) Gene ontology of the gene clusters that showed the greatest similarity in expression profile when we silenced LTN1 and NEMF, shown in panel A. The results for up and downregulated gene clusters are shown in **red** and **green**, respectively.

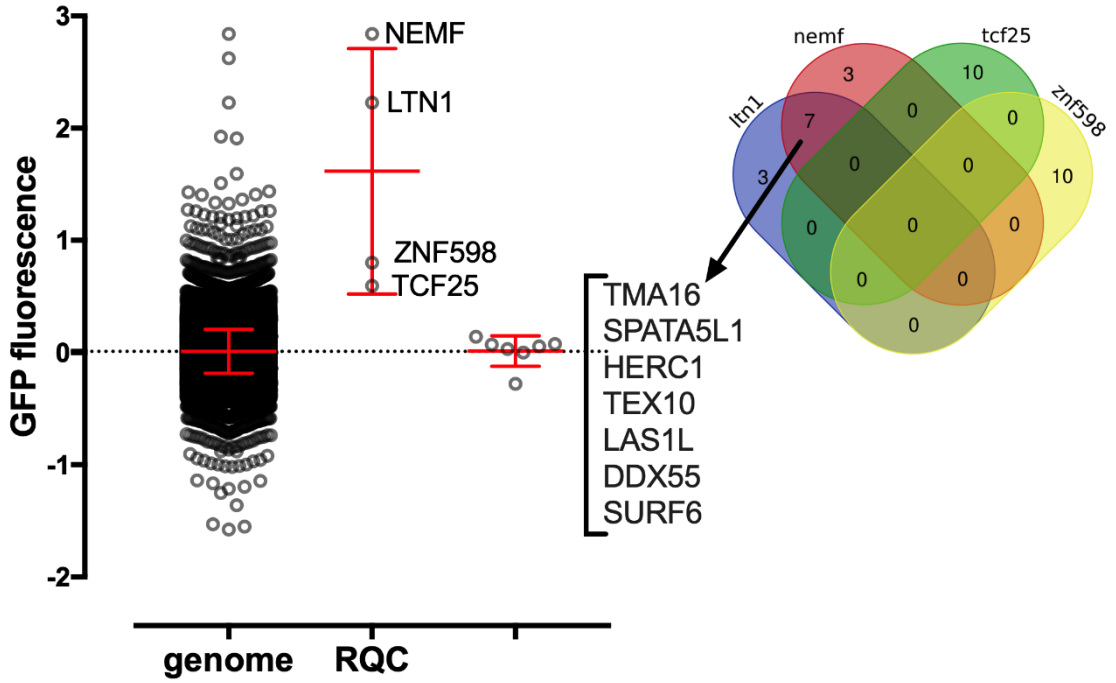


Figure 6. CRISPRi knockdown genes that produced a similar transcriptional phenotype to LTN1 and NEMF are not directly involved in RQC function (Right). Venn diagram shows the top 10 knockdown genes with a similar transcriptional response for each RQC component and ZNF598, of which 7 genes have a similar profile to NEMF and LTN1 knockdown. (**Left**) The graph of GFP fluorescence shows an increase in fluorescence levels when the components of the RQCs complex are deleted, which was not observed for any of the 7 genes with a similar transcriptional profile. Data obtained from Hickey et al., 2020.

The effect of aging on the mRNA levels of the RQC components

Proteostasis failure with the accumulation of insoluble aggregates is a hallmark of several neurodegenerative diseases (Chiti & Dobson, 2006; López-Otín *et al.*, 2013). Failure of the RQC, more specifically of LTN1, can lead to the accumulation of CATylated peptides that promote the formation of protein aggregates due to the CAT tail hydrophobic nature. These aggregates can disrupt neuronal morphogenesis (Defenouillère *et al.*, 2016; Sitron *et al.*, 2020; Udagawa *et al.*, 2021). Based on this information, we asked if the dysfunction of the RQC contributes to the formation of the insoluble protein aggregates related to old age. To assess if the mRNA levels of the RQC components are affected by aging, we

evaluated the transcriptome from several tissues, from donors with different ages, obtained from the GTEx consortium. We used the transcript per million (TPM) to determine the mRNA abundance across the human tissues at different ages (**Figure 7 A-D**). Compared with the youngest individuals, mRNA levels of the RQC components in aged tissues showed a slight downward trend, which was more evident for NEMF, followed by LTN1 (**Figure 7A-D**). TCF25 also showed a statistically significant decrease (**Figures 7B and 7C**), while no difference was observed for ZNF598 (**Figure 7D**). We did a cluster analysis to analyze tissue-specific age-dependent regulations. NEMF and LTN1 showed a similar pattern of transcriptional regulation. Both genes showed repression in 5 tissues with significant p-values (**Figures 7E and S2)**. TCF25 and ZNf598 showed a more complex and diverse pattern of the regulation **(Figures S3 and S4)**. While there was no statistically significant difference in the brain, we could see a reduction trend of the TCF25, LTN1, and ZNF598 mRNA.

We sought another model to test the age-dependent regulation observed with GTEx data. We repeated the analyses described in Figure 7 now using data from the *tabula muris* project (The Tabula Muris Consortium, 2018), which offers transcriptomes of 20 organs from four male and three female mice ranging from 1 to 27 weeks old. However, no age-dependent regulation was observed for the RQC components **(Figure S5-8).**

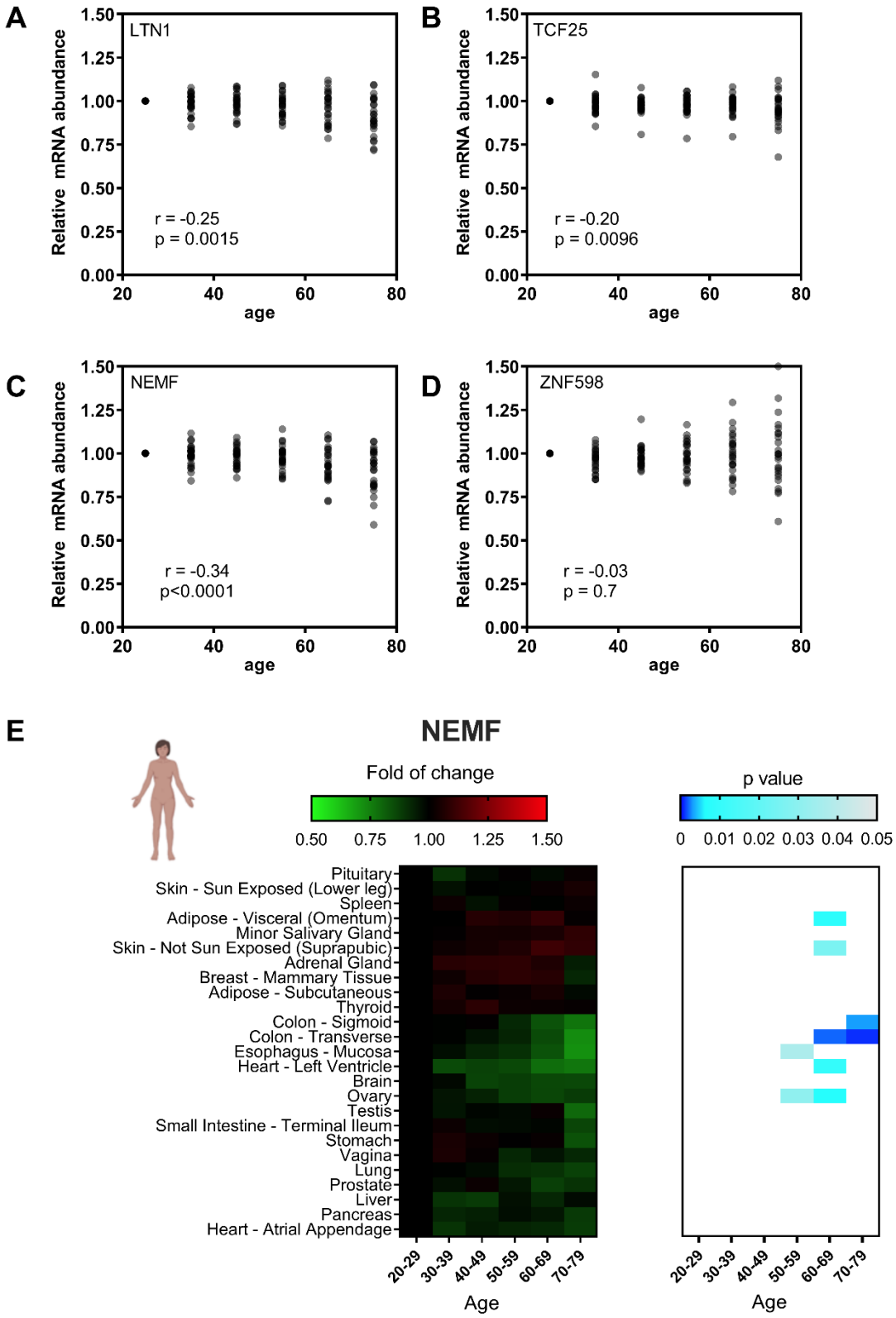


Figure 7. Effects of cellular aging on the transcriptional profile of NEMF, LTN1, TCF25, and ZNF598 in different human tissues. (A-D) Comparison of the relative abundance of mRNA levels in human tissues in different age groups (20-80 years). Data

obtained from GTEx consortium. (E) Distribution of the Z-score value of the relative abundance of NEMF mRNA in main human organs in different age clusters. Statistical significance was calculated using p-value and presented in a blue gradient for those relevant samples. All data were normalized by the reference age value (20-29).

Conclusion

Our data suggest that the RQC complex components are widely expressed in human tissues, suggesting that the RQC is part of the housekeeping processes of the cell **(Figures 1 and 2)**. LTN1 and NEMF share similarities in their expression levels **(Figures 1 and 2)** and their functional and transcriptional phenotypes. However, TCF25, the third essential component of the RQC complex, showed a distinct regulation profile that produced different phenotypes compared with LTN1 and NEMF. This observation can be explained by TCF25 functions outside of the RQC. (Olsson *et al.*, 2002; Cai *et al.*, 2006; Zhang *et al.*, 2020; Wang *et al.*, 2021).

Our data showed that TCF25 has a high level of mRNA, low levels of proteins, a high half-life of mRNA, and low half-life of proteins **(Figure 3).** These results support a post-transcriptional regulation of TCF25 expression, which has already been experimentally examined in the literature in yeast cells (Barros *et al.*, 2021). It has been proposed that the RQC complex regulates Rqc1/TCF25 expression co-translationally through ubiquitination of its polybasic region promoted by RQC activity (Brandman *et al.*, 2012). However, our group has recently shown that this regulation occurs post-translationally and through the activity of LTN1 alone (Barros *et al.*, 2021). TCF25 regulation may impact the RQC complex activity or the transcriptional control exerted by TCF25 in the nucleus. Further studies are necessary to understand how these functions are coordinated and if there are tissue specificities for TCF25 molecular roles.

Further, we observed significant differences in mRNA expression levels of the RQC components when we compared young with aged tissues in humans. A similar effect was observed with *S. cerevisiae* and *C. elegans* (Stein *et al.*, 2022). NEMF and LTN1 showed the most consistent age-dependent down-regulation pattern in humans. In *C. elegans*, NEMF and TCF25 homologs showed repression in aged specimens, while in yeast, LTN1, NEMF, and ZNF598 homologs were affected (Stein *et al.*, 2022).

We observed a trend of decreased expression of the RQC components in the human brain (although not statistically significant), which could be relevant to age-related neuronal impairment **(Figures 7E, S2 and S3)**; ZNF598, contrarily, shows a trend toward increased expression, perhaps as a response to the failure of the RQC **(Figure S4)**. These findings bring insights into the regulation of the RQC complex's components and its tissue-specific pattern, which could help clarify some phenotypes observed when the activity of the complex is affected by mutations or during aging.

ACKNOWLEDGMENT

We thank Caio Oliveira and Elis Moll for helpful discussions.

FUNDING

This work was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Fundação de Amparo a Pesquisa do Estado do Rio de Janeiro (FAPERJ) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

CONFLICT OF INTEREST

The authors declare that they have no conflicts of interest with the contents of this article.

Supplemental data for this article can be accessed on the publisher's website.

Reference

Agarwal, V., & Kelley, D. (2022). The genetic and biochemical determinants of mRNA degradation rates in mammals. *Genome Biology, 23,* https://doi: 10.1186/s13059-022-02811-x.

Barros, G. C., Requião, R. D., Carneiro, R. L., Masuda, C. A., Moreira, M. H., Rossetto, S., Domitrovic, T., & Palhano, F. L. (2021). Rqc1 and other yeast proteins containing highly positively charged sequences are not targets of the RQC complex. *Journal of Biological Chemistry*, *296*. https://doi.org/10.1016/j.jbc.2021.100586

Brandman, O., Stewart-Ornstein, J., Wong, D., Larson, A., Williams, C. C., Li, G. W., Zhou, S., King, D., Shen, P. S., Weibezahn, J., Dunn, J. G., Rouskin, S., Inada, T., Frost, A., & Weissman, J. S. (2012). A Ribosome-Bound Quality Control Complex Triggers Degradation of Nascent Peptides and Signals Translation Stress. *Cell*, *151*(5). https://doi.org/10.1016/j.cell.2012.10.044

Choe, Y. J., Park, S. H., Hassemer, T., Körner, R., Vincenz-Donnelly, L., Hayer-Hartl, M., & Hartl, F. U. (2016). Failure of RQC machinery causes protein aggregation and proteotoxic stress. *Nature*, *531*(7593). https://doi.org/10.1038/nature16973

Chu, J., Hong, N. A., Masuda, C. A., Jenkins, B. v., Nelms, K. A., Goodnow, C. C., Glynne, R. J., Wu, H., Masliah, E., Joazeiro, C. A. P., & Kay, S. A. (2009). A mouse forward genetics screen identifies LISTERIN as an E3 ubiquitin ligase involved in neurodegeneration. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(7). https://doi.org/10.1073/pnas.0812819106

Collins, R. L., Glessner, J. T., Porcu, E., Lepamets, M., Brandon, R., Lauricella, C., Han, L., Morley, T., Niestroj, L. M., Ulirsch, J., Everett, S., Howrigan, D. P., Boone, P. M., Fu, J., Karczewski, K. J., Kellaris, G., Lowther, C., Lucente, D., Mohajeri, K., … Talkowski, M. E. (2022). A cross-disorder dosage sensitivity map of the human genome. *Cell*, *185*(16). https://doi.org/10.1016/j.cell.2022.06.036

Filbeck, S., Cerullo, F., Pfeffer, S., & Joazeiro, C. A. P. (2022). Ribosome-associated quality-control mechanisms from bacteria to humans. *Molecular Cell*, *82*(8), 1451–1466. https://doi.org/10.1016/J.MOLCEL.2022.03.038

Greenbaum, D., Colangelo, C., Williams, K., & Gerstein, M. (2003). Comparing protein abundance and mRNA expression levels on a genomic scale. In *Genome Biology* (Vol. 4, Issue 9). https://doi.org/10.1186/gb-2003-4-9-117

Hickey, K. L., Dickson, K., Cogan, J. Z., Replogle, J. M., Schoof, M., D'Orazio, K. N., Sinha, N. K., Hussmann, J. A., Jost, M., Frost, A., Green, R., Weissman, J. S., & Kostova, K. K. (2020). GIGYF2 and 4EHP Inhibit Translation Initiation of Defective Messenger RNAs to Assist Ribosome-Associated Quality Control. *Molecular Cell*, *79*(6). https://doi.org/10.1016/j.molcel.2020.07.007

Howard, C. J., & Frost, A. (2021). Ribosome-associated quality control and CAT tailing. In *Critical Reviews in Biochemistry and Molecular Biology* (Vol. 56, Issue 6). https://doi.org/10.1080/10409238.2021.1938507

Ikeuchi, K., Tesina, P., Matsuo, Y., Sugiyama, T., Cheng, J., Saeki, Y., Tanaka, K., Becker, T., Beckmann, R., & Inada, T. (2019). Collided ribosomes form a unique structural interface to induce Hel2-driven quality control pathways. *The EMBO Journal*, *38*(5). https://doi.org/10.15252/embj.2018100276

Ingolia, N. T., Lareau, L. F., & Weissman, J. S. (2011). Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell*, *147*(4). https://doi.org/10.1016/j.cell.2011.10.002

Jiang, L., Wang, M., Lin, S., Jian, R., Li, X., Chan, J., Dong, G., Fang, H., Robinson, A. E., Aguet, F., Anand, S., Ardlie, K. G., Gabriel, S., Getz, G., Graubert, A., Hadley, K., Handsaker, R. E., Huang, K. H., Kashin, S., … Snyder, M. P. (2020). A Quantitative Proteome Map of the Human Body. *Cell*, *183*(1). https://doi.org/10.1016/j.cell.2020.08.036

Juszkiewicz, S., Chandrasekaran, V., Lin, Z., Kraatz, S., Ramakrishnan, V., & Hegde, R. S. (2018). ZNF598 Is a Quality Control Sensor of Collided Ribosomes. *Molecular Cell*, *72*(3). https://doi.org/10.1016/j.molcel.2018.08.037

Kuroha, K., Zinoviev, A., Hellen, C. U. T., & Pestova, T. v. (2018). Release of Ubiquitinated and Non-ubiquitinated Nascent Chains from Stalled Mammalian Ribosomal Complexes by ANKZF1 and Ptrh1. *Molecular Cell*, *72*(2). https://doi.org/10.1016/j.molcel.2018.08.022

Maier, T., Güell, M., & Serrano, L. (2009). Correlation of mRNA and protein in complex biological samples. In *FEBS Letters* (Vol. 583, Issue 24). https://doi.org/10.1016/j.febslet.2009.10.036

Martin, P. B., Kigoshi-Tansho, Y., Sher, R. B., Ravenscroft, G., Stauffer, J. E., Kumar, R., Yonashiro, R., Müller, T., Griffith, C., Allen, W., Pehlivan, D., Haral, T., Zenker, M., Howting, D., Schanze, D., Faqeih, E. A., Almontashiri, N. A. M., Maroofian, R., Houlden, H., … Cox, G. A. (2020). NEMF mutations that impair ribosome-associated quality control are associated with neuromuscular disease. *Nature Communications*, *11*(1). https://doi.org/10.1038/s41467-020-18327-6

Mathieson, T., Franken, H., Kosinski, J., Kurzawa, N., Zinn, N., Sweetman, G., Poeckel, D., Ratnu, V. S., Schramm, M., Becher, I., Steidel, M., Noh, K. M., Bergamini, G., Beck, M., Bantscheff, M., & Savitski, M. M. (2018). Systematic analysis of protein turnover in primary cells. *Nature Communications*, *9*(1). https://doi.org/10.1038/s41467-018-03106-1

Matsuo, Y., & Inada, T. (2021). The ribosome collision sensor Hel2 functions as preventive quality control in the secretory pathway. *Cell Reports*, *34*(12). https://doi.org/10.1016/j.celrep.2021.108877

Matsuo, Y., Tesina, P., Nakajima, S., Mizuno, M., Endo, A., Buschauer, R., Cheng, J., Shounai, O., Ikeuchi, K., Saeki, Y., Becker, T., Beckmann, R., & Inada, T. (2020). RQT complex dissociates ribosomes collided on endogenous RQC substrate SDD1. *Nature Structural and Molecular Biology*, *27*(4). https://doi.org/10.1038/s41594-020-0393-9

Osuna, B. A., Howard, C. J., KC, S., Frost, A., & Weinberg, D. E. (2017). In vitro analysis of RQC activities provides insights into the mechanism and function of CAT tailing. *ELife*, *6*. https://doi.org/10.7554/elife.27949

Replogle, J. M., Saunders, R. A., Pogson, A. N., Hussmann, J. A., Lenail, A., Guna, A., Mascibroda, L., Wagner, E. J., Adelman, K., Lithwick-Yanai, G., Iremadze, N., Oberstrass, F., Lipson, D., Bonnar, J. L., Jost, M., Norman, T. M., & Weissman, J. S. (2022). Mapping information-rich genotype-phenotype landscapes with genome-scale Perturb-seq. *Cell*, *185*(14). https://doi.org/10.1016/j.cell.2022.05.013

Stein, K. C., Morales-Polanco, F., van der Lienden, J., Rainbolt, T. K., & Frydman, J. (2022). Ageing exacerbates ribosome pausing to disrupt cotranslational proteostasis. *Nature*, *601*(7894). https://doi.org/10.1038/s41586-021-04295-4

The GTEx Consortium (2020). The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science*, *369*(6509). https://doi.org/10.1126/SCIENCE.AAZ1776

The Tabula Muris Consortium (2018). Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris. *Nature*, *562*(7727).

Tsherniak, A., Vazquez, F., Montgomery, P. G., Weir, B. A., Kryukov, G., Cowley, G. S., Gill, S., Harrington, W. F., Pantel, S., Krill-Burger, J. M., Meyers, R. M., Ali, L., Goodale, A., Lee, Y., Jiang, G., Hsiao, J., Gerath, W. F. J., Howell, S., Merkel, E., … Hahn, W. C. (2017).

Defining a Cancer Dependency Map. *Cell*, *170*(3). https://doi.org/10.1016/j.cell.2017.06.010

Udagawa, T., Seki, M., Okuyama, T., Adachi, S., Natsume, T., Noguchi, T., Matsuzawa, A., & Inada, T. (2021). Failure to Degrade CAT-Tailed Proteins Disrupts Neuronal Morphogenesis and Cell Survival. *Cell Reports*, *34*(1). https://doi.org/10.1016/j.celrep.2020.108599

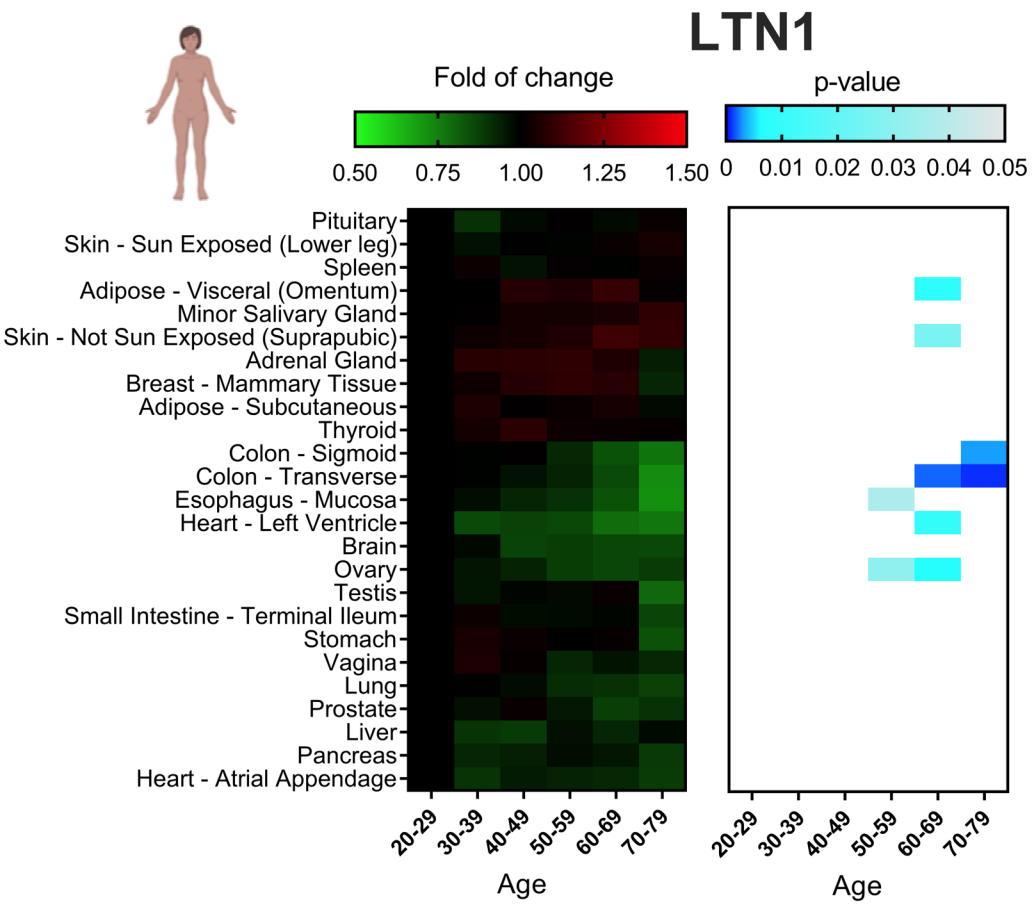
Vogel, C., & Marcotte, E. M. (2012). Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nature Reviews Genetics*, *13*(4). https://doi.org/10.1038/nrg3185

Wang, D., Eraslan, B., Wieland, T., Hallström, B., Hopf, T., Zolg, D. P., Zecha, J., Asplund, A., Li, L., Meng, C., Frejno, M., Schmidt, T., Schnatbaum, K., Wilhelm, M., Ponten, F., Uhlen, M., Gagneur, J., Hahne, H., & Kuster, B. (2019). A deep proteome and transcriptome abundance atlas of 29 healthy human tissues. *Molecular Systems Biology*, *15*(2). https://doi.org/10.15252/msb.20188503

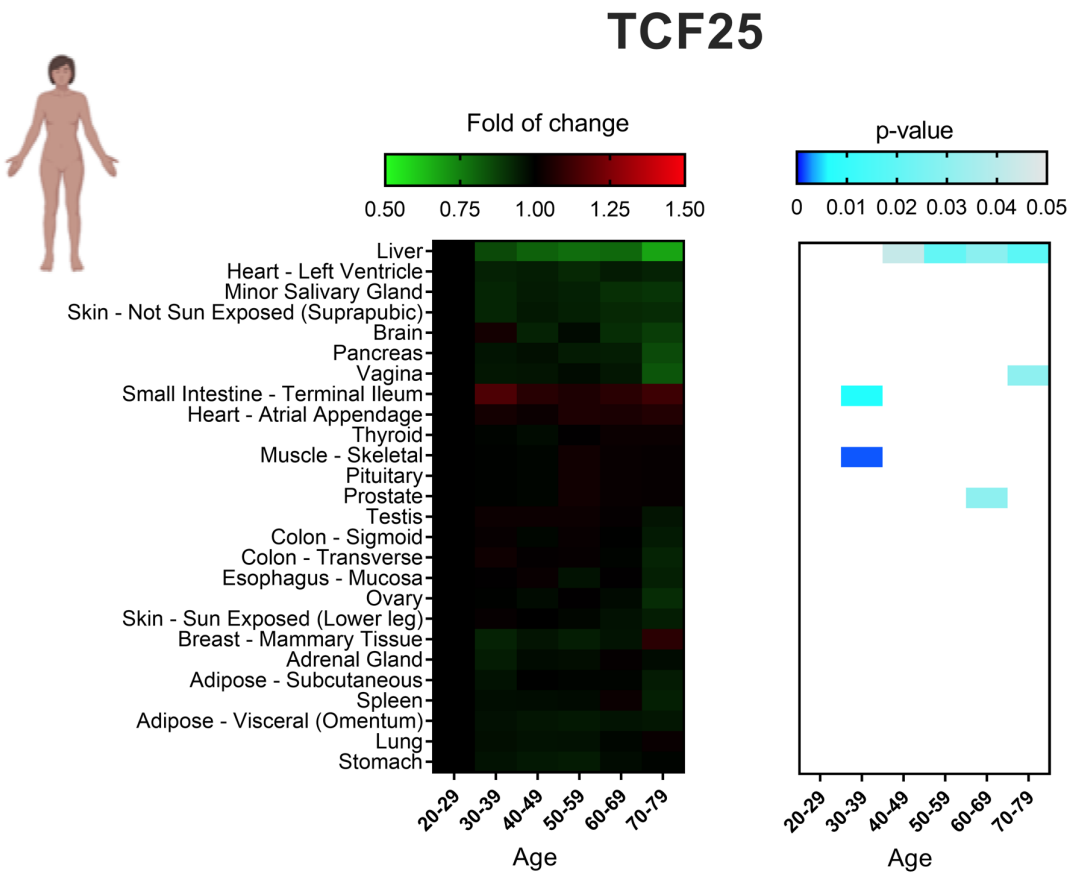
Winz, M. L., Peil, L., Turowski, T. W., Rappsilber, J., & Tollervey, D. (2019). Molecular interactions between Hel2 and RNA supporting ribosome-associated quality control. *Nature Communications*, *10*(1). https://doi.org/10.1038/s41467-019-08382-z

Yip, M. C. J., & Shao, S. (2021). Detecting and Rescuing Stalled Ribosomes. In *Trends in Biochemical Sciences* (Vol. 46, Issue 9). https://doi.org/10.1016/j.tibs.2021.03.008

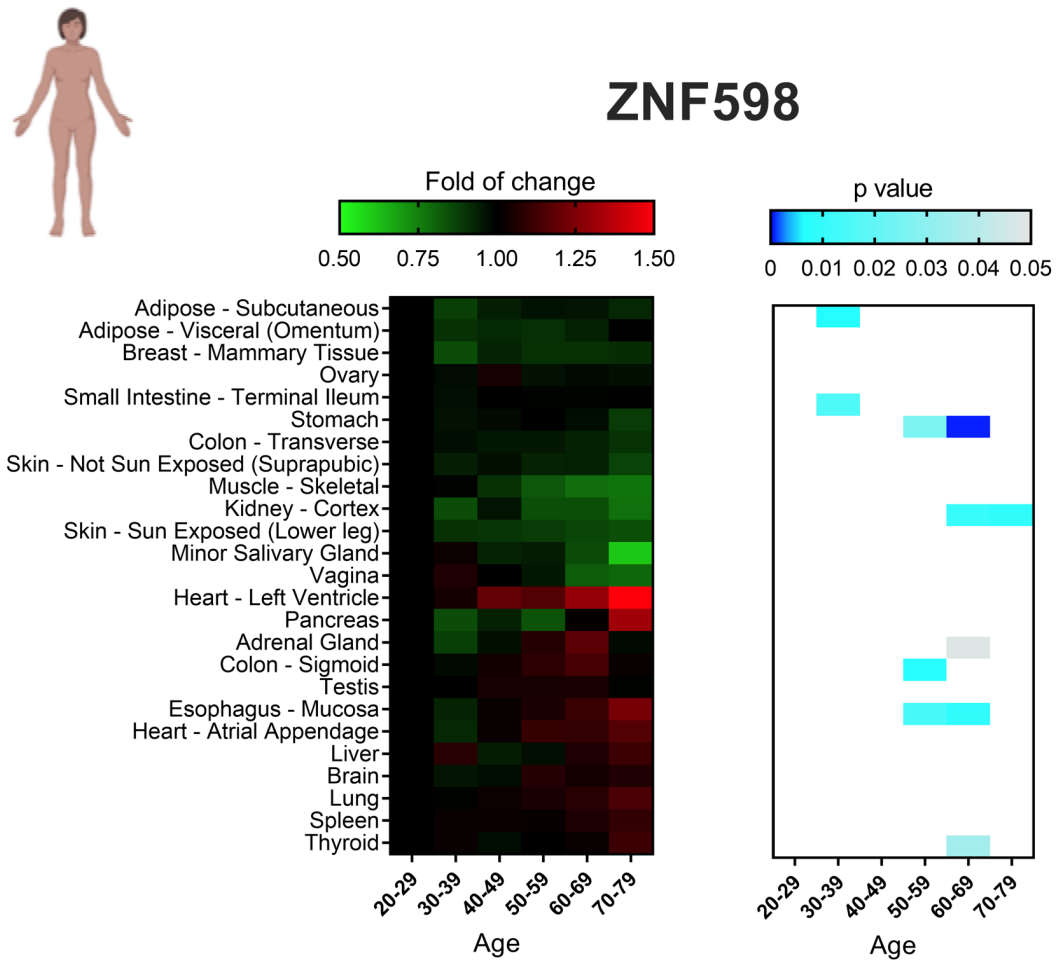
Supplementary Figure 1. Word cloud representation of the cellular lineages used to determine the impact of CRISPR-individual gene inactivation for LTN1, NEMF, and TCF25. Data obtained from Tsherniak et al., 2017.



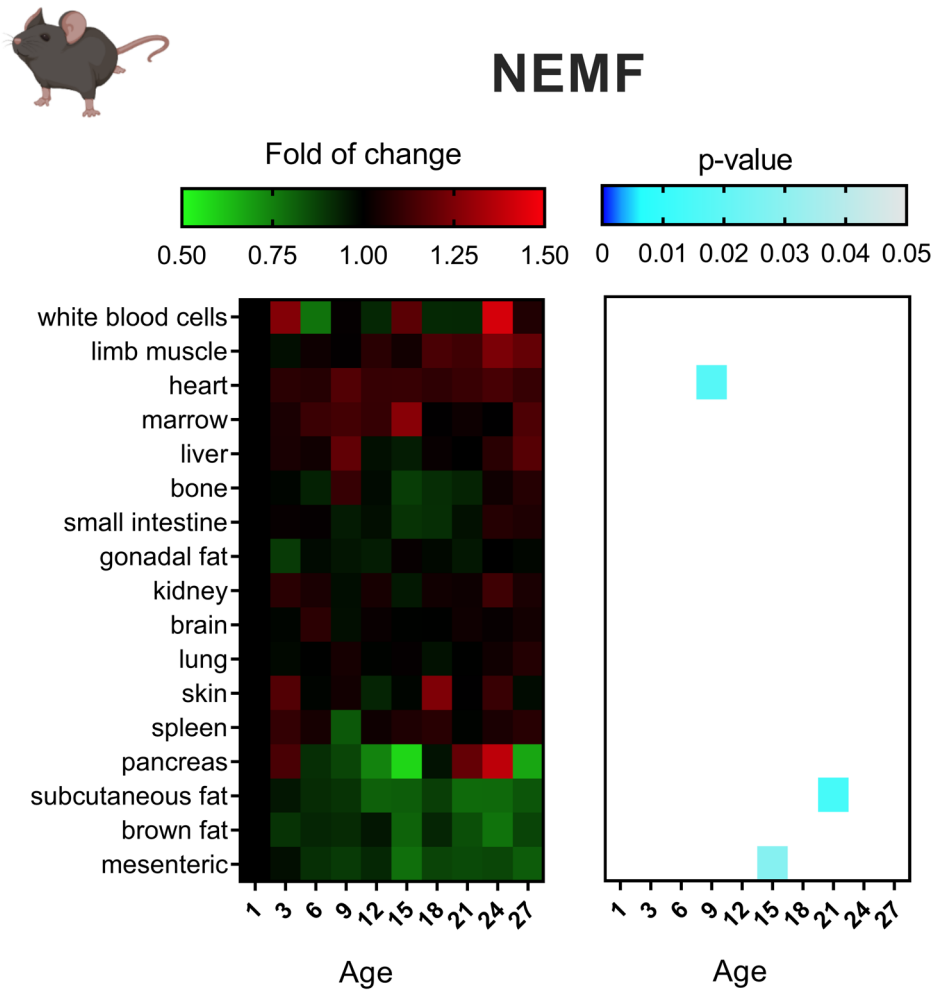
Supplementary Figure 2. Distribution of the Z-score value of the relative abundance of LTN1 mRNA in main human organs in different age clusters. Statistical significance was calculated using p-value and presented in a blue gradient for those samples that showed relevance. All data were normalized by the reference age value (20-29).



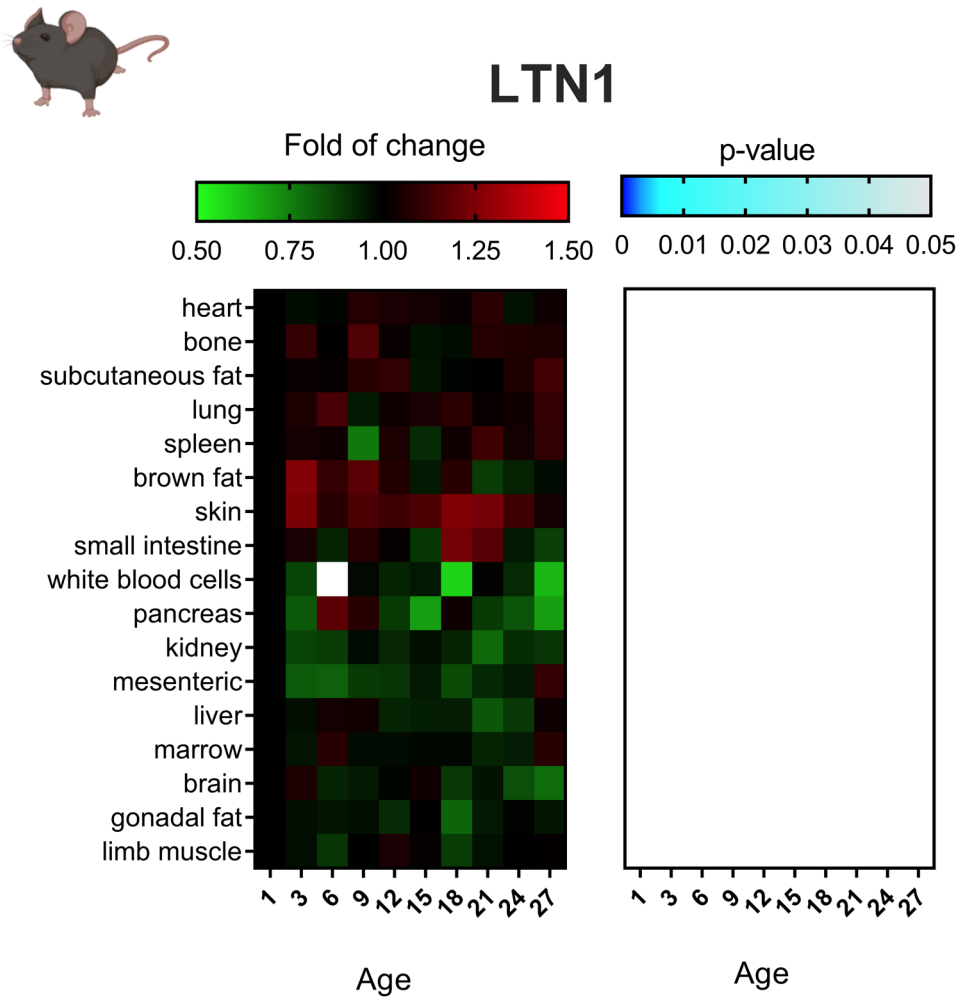
Supplementary Figure 3. Distribution of the Z-score value of the relative abundance of TCF25 mRNA in main human organs in different age clusters. Statistical significance was calculated using p-value and presented in a blue gradient for those samples that showed relevance. All data were normalized by the reference age value (20-29).



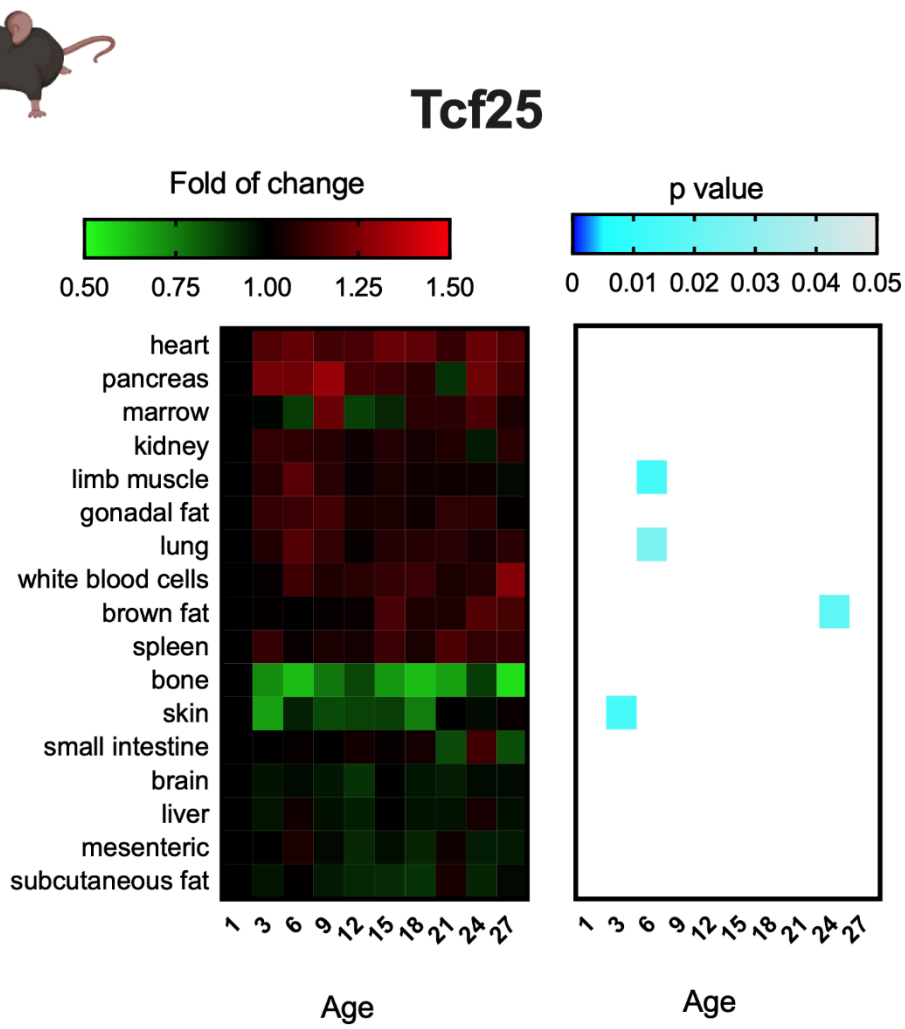
Supplementary Figure 4. Distribution of the Z-score value of the relative abundance of ZNF598 mRNA in main human organs in different age clusters. Statistical significance was calculated using p-value and presented in a blue gradient for those samples that showed relevance.



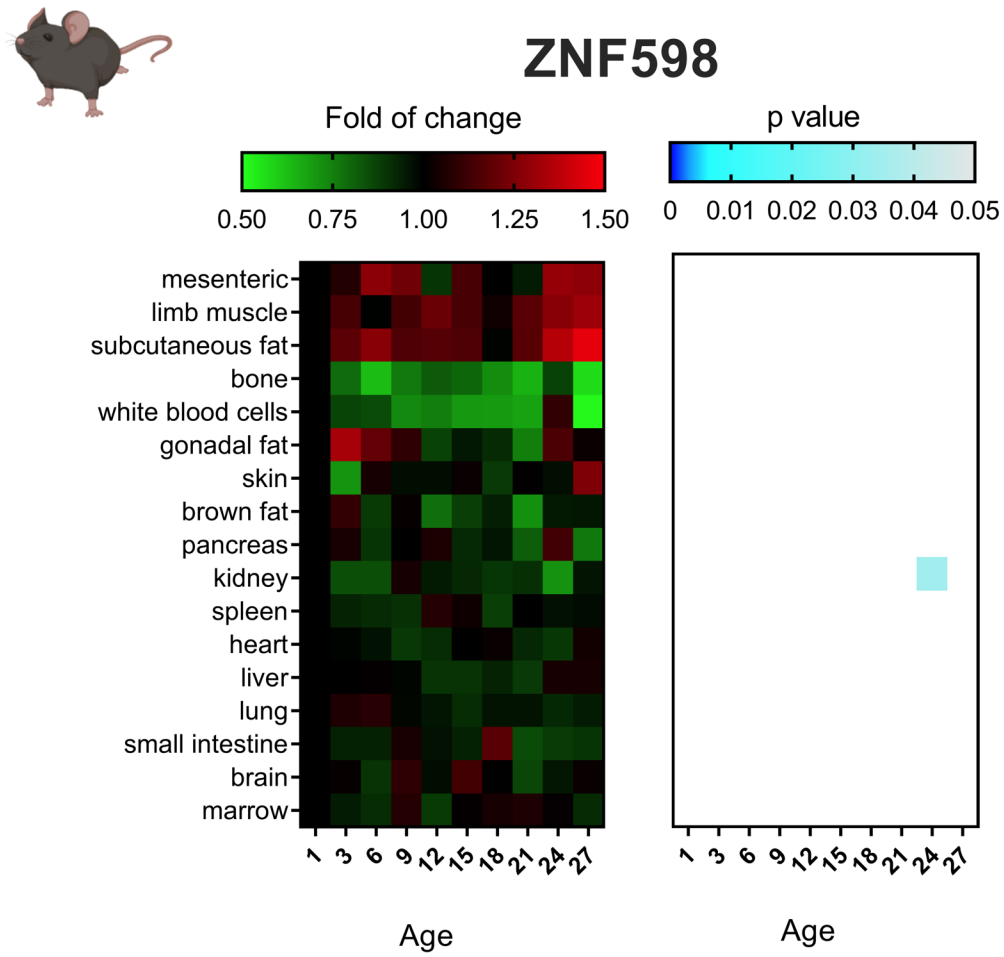
Supplementary Figure 5. Distribution of the Z-score value of the relative abundance of NEMF mRNA in main mouse organs in different age clusters. Statistical significance was calculated using p-value and presented in a blue gradient for those samples that showed relevance. All data were normalized by the reference age value (1 week).



Supplementary Figure 6. Distribution of the Z-score value of the relative abundance of LTN1 mRNA in main mouse organs in different age clusters. Statistical significance was calculated using p-value and presented in a blue gradient for those samples that showed relevance. All data were normalized by the reference age value (1 week).



Supplementary Figure 7. Distribution of the Z-score value of the relative abundance of TCF25 mRNA in main mouse organs in different age clusters. Statistical significance was calculated using p-value and presented in a blue gradient for those samples that showed relevance. All data were normalized by the reference age value (1 week).



Supplementary Figure 8. Distribution of the Z-score value of the relative abundance of ZNF598 mRNA in main mouse organs in different age clusters. Statistical significance was calculated using p-value and presented in a blue gradient for those samples that showed relevance. All data were normalized by the reference age value (1 week).

Supplementary Table 1: Gene function of the 7 genes with similar transcriptional profile with Nemf and Ltn1 obtained from Ensembl.

Gene name	Ensembl ID	GO: Biological process
TMA16	ENSG00000198498	Ribosome biogenesis
SAPTA5L1	ENSG00000171763	Ribosome biogenesis
HERC1	ENSG00000103657	Autophagy Brain development Protein ubiquitination Cerebellar Purkinje cell differentiation Corpus callosum development Neuron projection development Regulation of catalytic activity Neuromuscular process controlling balance
TEX10	ENSG00000136891	No available
LAS1L	ENSG00000001497	rRNA processing and maturation Nucleic acid phosphodiester bond hydrolysis
DDX55	ENSG00000111364	No available
SURF6	ENSG00000148296	Ribosome biogenesis