Multi-faceted deregulation of gene expression and protein synthesis with age

Aleksandra S. Anisimova, Mark B. Meerson, Maxim V. Gerashchenko,

Ivan V. Kulakovskiy, Sergey E. Dmitriev, Vadim N. Gladyshev

SUPPLEMENTARY DATA

Content

www.pnas.org/cgi/doi/10.1073/pnas.2001788117

SUPPLEMENTAL METHODS

Section 1.

Ribo-Seq and RNA-Seq sequencing, data processing and bioinformatic analysis

Ribo-Seq and RNA-Seq sequencing and data processing

Quality control of sequencing reads was performed with FastQC v0.11.5 (1). For Ribo-Seq data, the adapter sequences were cut with Cutadapt 1.14 (2), and quality trimming was performed with Sickle 1.33 (3). The Ribo-Seq and RNA-Seq reads were aligned to the mouse genome assemblies (mm10, downloaded from the UCSC Genome Browser) with using GENCODE M13 transcripts annotation (4). Alignment and basic read counting were performed with STAR 2.5.3 (5) with default parameters except for the number of allowed mismatches: up to 5% and 3% for Ribo-Seq and RNA-Seq, respectively, i.e. allowing approximately 1 mismatch per ~30nt (a Ribo-Seq footprint) and 2 per 75nt (an RNA-Seq read).

Ribo-Seq and RNA-Seq differential expression analysis

Raw gene counts processing and statistical analysis was performed in R Environment using edgeR Bioconductor package (6). RNA-Seq and Ribo-Seq data were RLE-normalized, separately for kidney and liver. Genes not reaching 1 read count per million (CPM) in at least one library were excluded from the analysis resulting in 8,992 and 11,461 genes in liver and kidney samples, respectively.

To analyze age-dependent gene expression dynamics, we used Ribo-Seq data, as ribosome footprint counts combines transcript abundance and translation efficiency. First, we used the generalized linear model (GLM, glmQLFit, and glmQLFTest of the edgeR package) with the age as a categorical variable (see Supplementary Methods, Section 2). To identify differentially expressed genes, we compared gene expression in the samples of various groups (from 11- to 2-month-old) against 3-month-old samples and considered genes to be differentially expressed if they passed the 0.05 Benjamini-Hochberg (FDR) adjusted P-values threshold. The number of genes differentially expressed at least in one age was 689 in liver (excluding 1-month group) and 2,001 in kidney. To identify and characterize the groups of genes with similar patterns of age-dependent changes, the coexpression clusters, we clustered the genes using the expression fold change vectors (11, 20 and 26 versus 3 months) with ward.D clustering algorithm and Euclidean

distance (Fig. 2*A, C*). For up- and down-regulated genes we calculated Gene Ontology biological process (BP) and cellular compartment (CC) enrichment with the clusterProfiler package (7) (q-value threshold of 0.05, minimum gene set size of 2). The obtained GO enrichment data were visualized with REVIGO (8) (Fig. 2*B*, *D*).

Second, we assessed a linear trend of age-dependent changes in gene expression (Figs. 3*C*, *D*; 4*C*, *D*; S6*C*, *D*.). To this end, we used the edgeR GLM with an alternative design matrix where age as a predictor was considered a continuous variable (see Supplementary Methods, Section 3). In all cases, genes were considered differentially expressed if the respective Benjamini-Hochberg (FDR) adjusted P-values passed the 0.05 threshold.

Principal component analysis of Ribo-Seq and RNA-Seq expression profiles

To generally characterize RNA- and Ribo-Seq data, we applied principal component analysis (PCA) to 8,562 genes reaching at least one CPM in all considered samples of mouse liver and kidney Ribo- and RNA-Seq datasets (Figs. 1*D*, S1*C*). For this purpose, the gene counts were RLE-normalized together for liver and kidney.

Distribution of the Ribo-Seq and RNA-Seq coverage to the genome regions

The Ribo-Seq and RNA-Seq read coverage was estimated in 5' and 3' untranslated regions (UTRs), coding sequences (CDS), and introns (Fig. 1*C*) by using `read_distribution.py` from RSeQC-2.3.7 (9).

Identifying putative transcriptional regulators of coexpressed genes

For sets of up- and downregulated genes (Fig. 2, Table S2), we searched for potential transcriptional regulators and pairs of regulators with binding sites enriched at promoters and thus potentially involved in transcription regulation of genes exhibiting age-related expression changes (Table S4). To this end, we used the ChIP-Seq-based mouse cistrome data (10) for 315 transcription factors. As putative target genes of a particular transcription factor (TF), we considered those for which a cistrome segment with binding motif occurrences was within 500 nt from the transcription start site annotated in GENCODE (4). For each TF, we considered a 2x2 contingency table classifying genes by two binary features, being a putative TF target and being included in a gene list of interest. Then, we applied right-sided Fisher's exact test (alternative

'greater') and adjusted the resulting P-values using Benjamini-Hochberg correction for multiple tested TFs, separately for up- and down-regulated genes. We then selected putative transcriptional regulators from the set of TFs passing adjusted P-values < 0.05 and odds ratios of 1.5 or more. The same analysis was performed for pairs of TFs. To select the meaningful pairs, we used those of physically interacting TFs, based on the BioGRID dataset ALL-3.5.166 (11). The sets of targets were intersected between the TFs forming particular TF pair and used for the analysis with Fisher's exact test as for single TFs.

Analysis of transcript isoform composition changes

To analyze changes in transcript isoform abundance, we compared shares of transcript isoform RNA-Seq coverage estimated with RSEM program (12) for young (3 months for kidney, and 3 and 11 months for liver) and old animals (32 months for kidney, and 26 and 32 months for liver) using Student's t-test with Benjamini-Hochberg correction for multiple tested transcripts. Genes reaching one CPM in each sample were used in the analysis (8,992 and 11,461 genes for liver and kidney, respectively). No genes passed 0.05 adjusted P-value threshold (Table S5). To visualize possible relation between RO changes and changes in isoform abundance, we selected genes with a mean share of a single isoform changing for more than 10% from total and with the standard deviation of a single isoform share within young and old groups being no more than 5% of the mean for kidney (34 genes) and 15% of the mean for liver (56 genes). The genes were highlighted among the list of genes sorted by their linear RO Fold Change with age (Fig. S4, Tables S2 and S5).

Ribosome occupancy analysis

To assess the contribution of translation control to age-dependent gene expression changes, we estimated the transcript ribosome occupancy (RO, often named after and used as an estimate of the translation efficiency, TE). RO is computed as the ratio or log-ratio of Ribo-Seq and RNA-Seq read counts. We used edgeR GLM to estimate the contribution of Ribo-Seq relative to RNA-Seq. The primary analysis was performed with the following model: `~exp.type+age:exp.type`, where `exp.type` is a categorical variable (Ribo-Seq or RNA-Seq) (see Supplementary Methods, Section 4). The age was considered a continuous predictor variable. Such design not only allowed to subtract contribution of the RNA-Seq changes to Ribo-Seq read counts and thus obtain RO

estimates, but also to detect age-dependent ribosome occupancy changes and assess their statistical significance by using edgeR contrasts (10). Samples from 1-month-old mice were excluded from the analysis.

In individual cases which are specially indicated, a simple paired comparison of RO between 32- or 26-month-old mice and 3-month-old mice was used, with the age treated as categorical variable.

Functional enrichment of the groups of transcripts whose translation changed with age was determined for ranked gene lists sorted according to the "signed P-values" ($-\log_{10}$(P-value)*sign($\log_2$(Fold change))) of their RO linear fold change with age (from 3- to 32-month-old mice, 14 samples in total). To perform Gene Set Enrichment Analysis (GSEA) we used the fgsea R package (13) (Fig. 3*C, D*) and original Java desktop application (14) (Fig. 4*C,* S6*C*). The GO terms with q-value less than 0.05 were clustered with REVIGO (8) and manually curated to obtain a non-redundant list.

Construction of metagene profiles

Metagene profiles of ribosome footprints at the start and stop codons were constructed from the Ribo-Seq read alignment to reveal systematic age-induced changes in ribosomal coverage (Fig. 1*B* (footprint 5' end coverage), Fig. 5*A* (full-length footprint coverage)). Transcript isoforms expression was estimated with RSEM (12). The major (the most expressed) isoforms of each gene reaching no less than 2.5 TPM (transcripts per kilobase million) in all libraries were selected for the subsequent analysis. From those, we selected only transcripts with both start and stop codons explicitly annotated in GENCODE M13. The Ribo-Seq footprint coverage was calculated in 200nt windows centered at the start and stop codons using bedtools 2.27.0-1 (15). Liver (2,920 genes) and kidney (4,566 genes) data were analyzed separately. To construct metagene profiles, for each transcript, the coverage profile was normalized to the respective (5'end or full-length footprint) average whole transcript Ribo-Seq coverage. The normalized values were summed for all transcripts at each position relative to the start or stop codon thus forming the metagene profile.

Analysis of ribosomal positional changes

To analyze the dynamics of age-related changes in ribosomal coverage of protein-coding regions we, first, performed segmentation of transcript profiles to decrease the level of noise. We pooled

ribosomal footprint coverage profiles for each age and replicate with an exception for 1 month-old mice samples and separately for organs and processed the pooled profiles with Poisson segmentation (for a detailed algorithm see Supplementary Methods, Section 5) to split each transcript onto windows of different lengths, with stable ribosome footprints coverage within each window. Then, for each window, we calculated the log-ratio of the average ribosome footprint coverage at particular age to the mean coverage at 3 months using the list of transcripts selected previously for metagene profiles. This allowed to de-trend the profiles and construct a linear model for the ribosome footprint coverage with the relative transcript coordinate as a predictor variable. The distribution of linear regression slopes for each age was then visualized (Fig. 5*B*, Fig. S9) and the significance of shift between the distribution of slopes of different ages was assessed with the SIGN.test function from BSDA R package.

Section 2.

Ribo-Seq and RNA-Seq differential expression analysis

Paired comparisons of Ribo-Seq read counts in tissue samples of each age group with samples of 3 months old mice. Age is taken as a categorical variable. Below we show minimal extract of the R code for liver samples to serve as a simplified formal description of the performed analysis.

```
> groups
                              Name Organ Age Replicate Seqtype
ribo_liver_1m_1   ribo_liver_1m_1 liver   1         1    ribo
ribo_liver_1m_3   ribo_liver_1m_3 liver   1         3    ribo
ribo_liver_1m_2   ribo_liver_1m_2 liver   1         2    ribo
ribo_liver_3m_2   ribo_liver_3m_2 liver   3         2    ribo
ribo_liver_3m_1   ribo_liver_3m_1 liver   3         1    ribo
ribo_liver_3m_3   ribo_liver_3m_3 liver   3         3    ribo
ribo_liver_10m_1 ribo_liver_10m_1 liver  10         1    ribo
ribo_liver_10m_2 ribo_liver_10m_2 liver  10         2    ribo
ribo_liver_10m_3 ribo_liver_10m_3 liver  10         3    ribo
ribo_liver_20m_1 ribo_liver_20m_1 liver  20         1    ribo
ribo_liver_20m_2 ribo_liver_20m_2 liver  20         2    ribo
ribo_liver_20m_3 ribo_liver_20m_3 liver  20         3    ribo
ribo_liver_26m_1 ribo_liver_26m_1 liver  26         1    ribo
ribo_liver_26m_3 ribo_liver_26m_3 liver  26         3    ribo
ribo_liver_26m_2 ribo_liver_26m_2 liver  26         2    ribo
ribo_liver_32m_2 ribo_liver_32m_2 liver  32         2    ribo
```

```
ribo_liver_32m_3 ribo_liver_32m_3 liver  32         3    ribo
```

```
> design <- model.matrix(~Age, groups)
> design
                 (Intercept) Age1 Age10 Age20 Age26 Age32
ribo_liver_1m_1            1    1     0     0     0     0
ribo_liver_1m_3            1    1     0     0     0     0
ribo_liver_1m_2            1    1     0     0     0     0
ribo_liver_3m_2            1    0     0     0     0     0
ribo_liver_3m_1            1    0     0     0     0     0
ribo_liver_3m_3            1    0     0     0     0     0
ribo_liver_10m_1           1    0     1     0     0     0
ribo_liver_10m_2           1    0     1     0     0     0
ribo_liver_10m_3           1    0     1     0     0     0
ribo_liver_20m_1           1    0     0     1     0     0
ribo_liver_20m_2           1    0     0     1     0     0
ribo_liver_20m_3           1    0     0     1     0     0
ribo_liver_26m_1           1    0     0     0     1     0
ribo_liver_26m_3           1    0     0     0     1     0
ribo_liver_26m_2           1    0     0     0     1     0
ribo_liver_32m_2           1    0     0     0     0     1
ribo_liver_32m_3           1    0     0     0     0     1
```

```
> dgList_filtered_norm
An object of class "DGEList"
```

```
> dgList_filtered_norm <- estimateDisp(dgList_filtered_norm, design)
> fit <- glmQLFit(dgList_filtered_norm, design)
```

```
> qlf.1vs3 <- glmQLFTest(fit, coef=2)
> qlf.11vs3 <- glmQLFTest(fit, coef=3)
> qlf.20vs3 <- glmQLFTest(fit, coef=4)
> qlf.26vs3 <- glmQLFTest(fit, coef=5)
> qlf.32vs3 <- glmQLFTest(fit, coef=6)
```

Section 3.

Ribo-Seq and RNA-Seq differential expression analysis

Analysis of linear age-dependent changes in gene expression using Ribo-Seq data with the age predictor as a continuous variable. Below we show minimal extract of the R code for liver samples to serve as a simplified formal description of the performed analysis.

```
> groups
                              Name Organ Age Replicate Seqtype
ribo_liver_3m_2   ribo_liver_3m_2 liver   3         2    ribo
ribo_liver_3m_1   ribo_liver_3m_1 liver   3         1    ribo
ribo_liver_3m_3   ribo_liver_3m_3 liver   3         3    ribo
ribo_liver_10m_1 ribo_liver_10m_1 liver  10         1    ribo
ribo_liver_10m_2 ribo_liver_10m_2 liver  10         2    ribo
ribo_liver_10m_3 ribo_liver_10m_3 liver  10         3    ribo
ribo_liver_20m_1 ribo_liver_20m_1 liver  20         1    ribo
ribo_liver_20m_2 ribo_liver_20m_2 liver  20         2    ribo
ribo_liver_20m_3 ribo_liver_20m_3 liver  20         3    ribo
ribo_liver_26m_1 ribo_liver_26m_1 liver  26         1    ribo
ribo_liver_26m_3 ribo_liver_26m_3 liver  26         3    ribo
ribo_liver_26m_2 ribo_liver_26m_2 liver  26         2    ribo
ribo_liver_32m_2 ribo_liver_32m_2 liver  32         2    ribo
ribo_liver_32m_3 ribo_liver_32m_3 liver  32         3    ribo

> design <- model.matrix(~Age, groups)
> design
                 (Intercept) Age
ribo_liver_3m_2            1   3
ribo_liver_3m_1            1   3
ribo_liver_3m_3            1   3
ribo_liver_10m_1           1  10
ribo_liver_10m_2           1  10
ribo_liver_10m_3           1  10
ribo_liver_20m_1           1  20
ribo_liver_20m_2           1  20
ribo_liver_20m_3           1  20
ribo_liver_26m_1           1  26
ribo_liver_26m_3           1  26
ribo_liver_26m_2           1  26
ribo_liver_32m_2           1  32
ribo_liver_32m_3           1  32

> dgList_filtered_norm
An object of class "DGEList"

> dgList_filtered_norm <- estimateDisp(dgList_filtered_norm, design)
> fit <- glmQLFit(dgList_filtered_norm, design)
> qlf.cont.age <- glmQLFTest(fit)
```

Section 4.

Analysis of mRNA ribosome occupancy (RO) linear changes with age

Contribution of ribosome occupancy to gene expression is assessed by estimating Ribo-Seq changes relative to the respective RNA-Seq changes. Here age is considered as a continuous variable and the type of the experiment (Ribo-Seq or RNA-Seq) is taken as a categorical variable. Below we show minimal extract of the R code for liver samples to serve as a simplified formal description of the performed analysis.

```
> groups
```

	Name	Organ	Age	Replicate	Seqtype
ribo_liver_3m_2	ribo_liver_3m_2	liver	3	2	ribo
ribo_liver_3m_1	ribo_liver_3m_1	liver	3	1	ribo
ribo_liver_3m_3	ribo_liver_3m_3	liver	3	3	ribo
ribo_liver_10m_1	ribo_liver_10m_1	liver	10	1	ribo
ribo_liver_10m_2	ribo_liver_10m_2	liver	10	2	ribo
ribo_liver_10m_3	ribo_liver_10m_3	liver	10	3	ribo
ribo_liver_20m_1	ribo_liver_20m_1	liver	20	1	ribo
ribo_liver_20m_2	ribo_liver_20m_2	liver	20	2	ribo
ribo_liver_20m_3	ribo_liver_20m_3	liver	20	3	ribo
ribo_liver_26m_1	ribo_liver_26m_1	liver	26	1	ribo
ribo_liver_26m_3	ribo_liver_26m_3	liver	26	3	ribo
ribo_liver_26m_2	ribo_liver_26m_2	liver	26	2	ribo
ribo_liver_32m_2	ribo_liver_32m_2	liver	32	2	ribo
ribo_liver_32m_3	ribo_liver_32m_3	liver	32	3	ribo
rna_liver_3m_2	rna_liver_3m_2	liver	3	2	rna
rna_liver_3m_1	rna_liver_3m_1	liver	3	1	rna
rna_liver_3m_3	rna_liver_3m_3	liver	3	3	rna
rna_liver_10m_1	rna_liver_10m_1	liver	10	1	rna
rna_liver_10m_2	rna_liver_10m_2	liver	10	2	rna
rna_liver_10m_3	rna_liver_10m_3	liver	10	3	rna
rna_liver_20m_1	rna_liver_20m_1	liver	20	1	rna
rna_liver_20m_2	rna_liver_20m_2	liver	20	2	rna
rna_liver_20m_3	rna_liver_20m_3	liver	20	3	rna
rna_liver_26m_1	rna_liver_26m_1	liver	26	1	rna
rna_liver_26m_3	rna_liver_26m_3	liver	26	3	rna
rna_liver_26m_2	rna_liver_26m_2	liver	26	2	rna
rna_liver_32m_2	rna_liver_32m_2	liver	32	2	rna
rna_liver_32m_3	rna_liver_32m_3	liver	32	3	rna

```
> design <- model.matrix(~Seqtype+Age:Seqtype, groups)
> design
                 (Intercept) Seqtyperibo Seqtyperna:Age Seqtyperibo:Age
ribo_liver_3m_2            1           1              0               3
ribo_liver_3m_1            1           1              0               3
ribo_liver_3m_3            1           1              0               3
ribo_liver_10m_1           1           1              0              10
ribo_liver_10m_2           1           1              0              10
ribo_liver_10m_3           1           1              0              10
ribo_liver_20m_1           1           1              0              20
ribo_liver_20m_2           1           1              0              20
ribo_liver_20m_3           1           1              0              20
ribo_liver_26m_1           1           1              0              26
ribo_liver_26m_3           1           1              0              26
ribo_liver_26m_2           1           1              0              26
ribo_liver_32m_2           1           1              0              32
ribo_liver_32m_3           1           1              0              32
rna_liver_3m_2             1           0              3               0
rna_liver_3m_1             1           0              3               0
rna_liver_3m_3             1           0              3               0
rna_liver_10m_1            1           0             10               0
rna_liver_10m_2            1           0             10               0
rna_liver_10m_3            1           0             10               0
rna_liver_20m_1            1           0             20               0
rna_liver_20m_2            1           0             20               0
rna_liver_20m_3            1           0             20               0
rna_liver_26m_1            1           0             26               0
rna_liver_26m_3            1           0             26               0
rna_liver_26m_2            1           0             26               0
rna_liver_32m_2            1           0             32               0
rna_liver_32m_3            1           0             32               0

> dgList_filtered_norm
An object of class "DGEList"

> dgList_filtered_norm <- estimateDisp(dgList_filtered_norm, design)
> fit <- glmQLFit(dgList_filtered_norm, design)
> qlf.cont.age <- glmQLFTest(fit, contrast=c(0,0,-1,1))
```

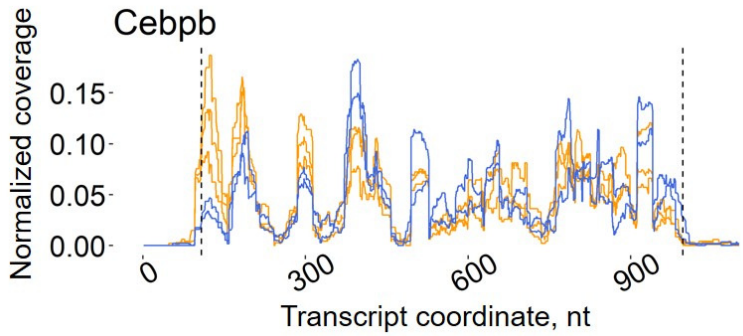
Section 5.

Analysis of age-related changes in ribosomal footprint coverage of CDS (coding segments)

To analyze ribosome occupancy along the transcripts, we used the following approach illustrated by ENSMUST00000070642.3 Cebpb transcript using the liver Ribo-Seq data (data for liver and kidney were considered separately). Of note, an extended and automated Python+BASH workflow for this type of analysis can be found at https://github.com/autosome-ru/papolarity.

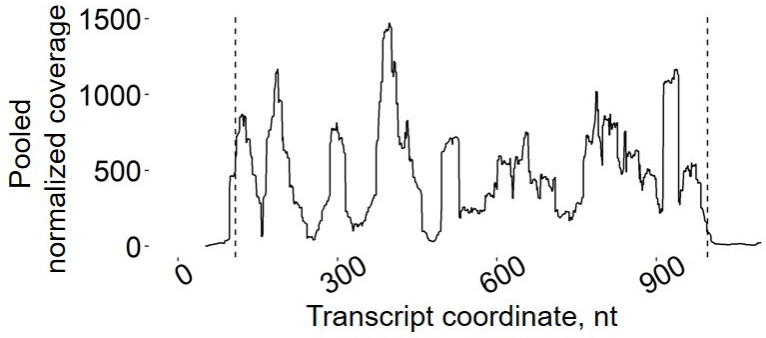
Panel A: Ribo-Seq full-length footprint coverage normalized to the total transcript coverage (3 months - orange, 32 months - blue, dashed lines represent the start and stop codons).

A.



Panel B: Ribosome footprint coverage profiles were pooled across all age and replicates excluding 1-month-old mice samples.

B.



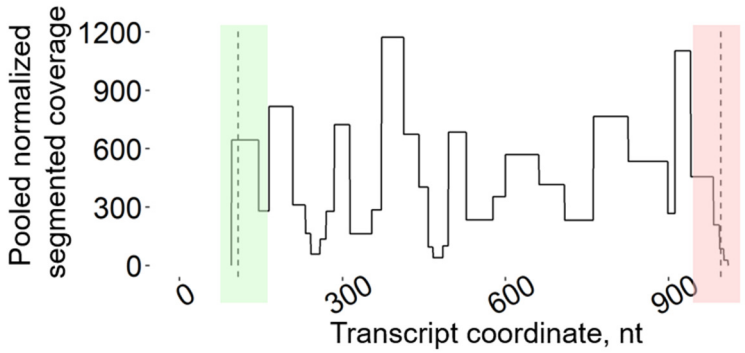
The pooled profiles were split into optimal segments of different lengths with *pasio* (https://github.com/autosome-ru/pasio, inspired by (16)) using default parameters (alpha = 1 and

beta = 1). This allowed obtaining a set of non-overlapping windows resembling the non-uniform distribution of ribosome footprints along each transcript.

For the further analyses we used the transcript CDS in two variants: (1) considering all segments including (-/+ 14nt) or (2) excluding (+/- 42nt) windows surrounding start and stop codons.

Panel C: The pooled profile processed with *pasio* with segments in the vicinity of the start and stop codons highlighted in green and red respectively.

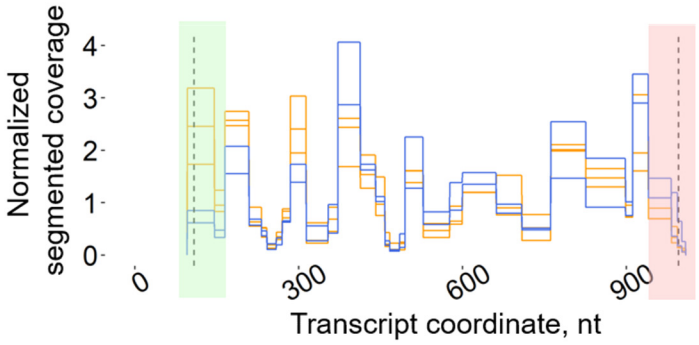
C.



The pooled segmented profile provided a set of non-overlapping windows along the transcript characterized by particular ribosome occupancy. Next, for each library, we computed the average coverage of each segment and normalized the resulting profiles to the total coverage of a particular transcript (**Panel D**, 3- (orange) and 32-month-old (blue) mice samples). A pseudocount of 1 was added to the coverage value of each segment.

Panel D: The segmented normalized coverage profile for individual replicates.

D.

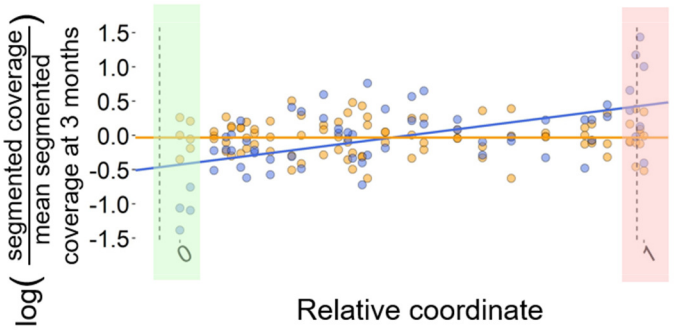


Next, we calculated the log-ratio of segmented ribosomal footprint coverage at each age versus the mean coverage profile at 3 months (**Panel E**, 3- (orange) and 32-month-old (blue) samples). We considered all transcripts that were previously used for metagene profile construction. The procedure allowed to obtain de-trended profiles (**Panel E**, each point shows an average coverage of a particular segment in a particular sample). The transcript coordinates (in nucleotides) of the segmented profile were projected into [0:1].

Next, for each transcript, we aggregated values across replicates and fitted a linear model of coverage depending on the relative transcript coordinate. The resulting linear trends for 3- (orange) and 32-month-old (blue) mice are shown in the **panel E** as the solid lines. The distribution of regression slopes across genes for each age is shown in Fig. 5*B* and Fig. S9.

Panel E: Fitting a linear model of the de-trended segmented ribosome footprint coverage depending on the relative transcript coordinate.

E.



Results of linear regression analysis for ENSMUST00000070642.3 Cebpb transcript including start and stop codons are shown in **panel F**.

Panel F:

Age	Term	Slope	Std. error	Statistic	P-value
3	Intercept	-0.028	0.046	-0.618	5.38E-01
3	x	-0.004	0.080	-0.054	9.57E-01
32	Intercept	-0.427	0.096	-4.426	4.12E-05
32	x	0.866	0.170	5.093	3.75E-06

SUPPLEMENTARY FIGURES

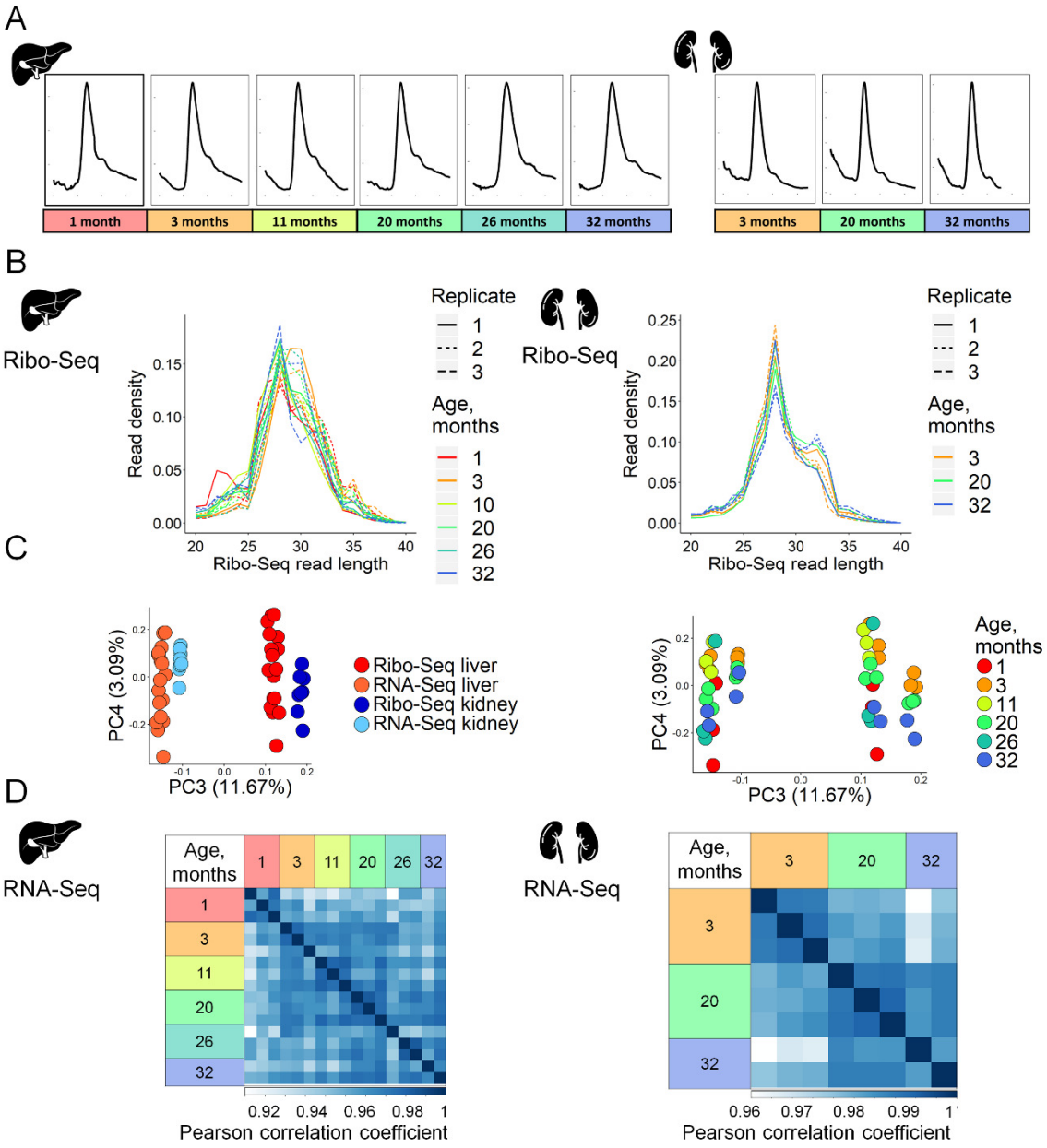


Figure S1. Quality and reproducibility of Ribo-Seq and RNA-Seq. (*A*) Examples of sucrose gradient profiles illustrating monosome appearance after RNA digestion with RNase T1 and S7. (*B*) Ribo-Seq footprint length distribution. (*C*) Principal component analysis (PCA) of 8,562 genes in Ribo-Seq and RNA-Seq datasets of mouse liver and kidney showing third (PC3) and fourth (PC4) components, percentage of explained variance shown in axes labels. PC4 separates the samples according to the age groups. (*D*) Pearson correlation coefficient matrices of mouse liver (8,992 genes) and kidney (11,461 genes) RNA-Seq data.

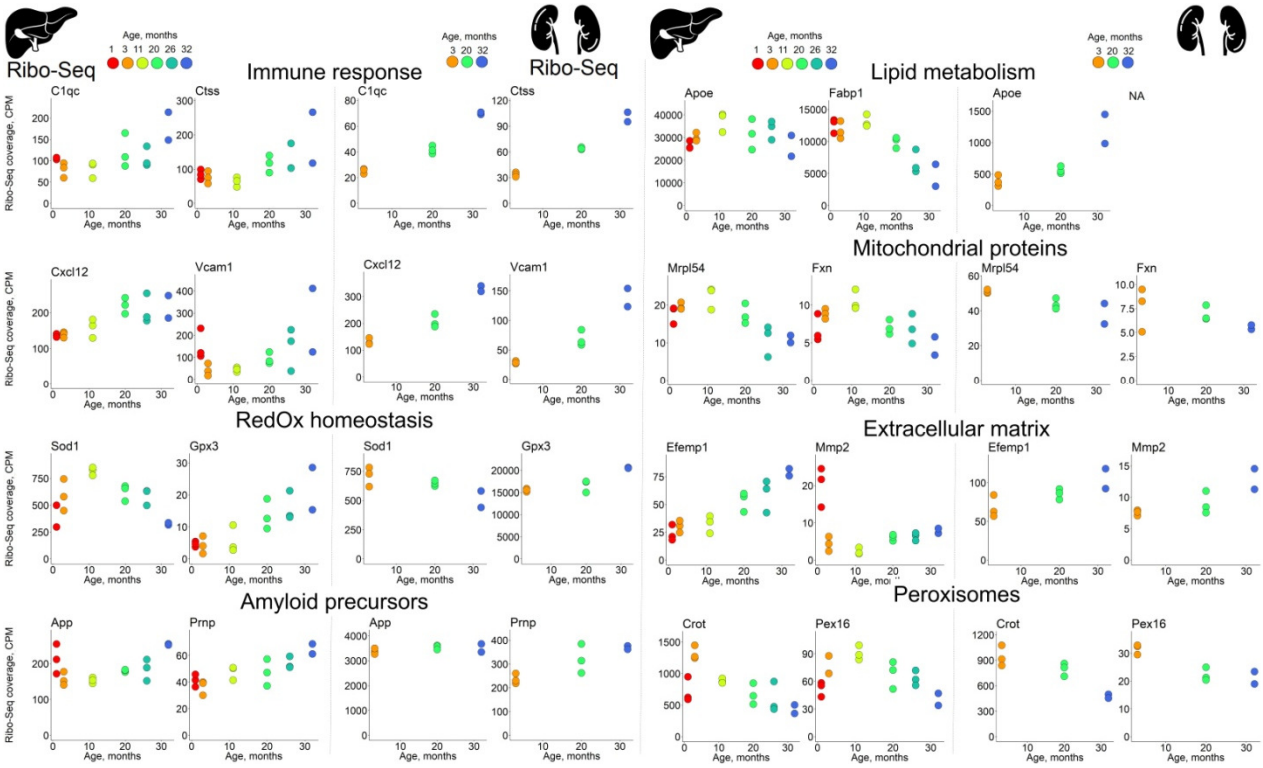
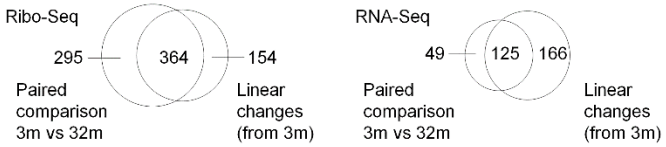


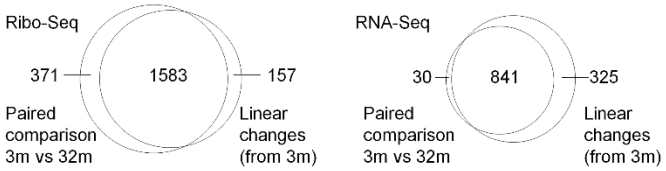
Figure S2. Examples of age-dependent expression changes in mouse liver and kidney. The plots illustrate age-related patterns of ribosome profiling counts of genes belonging to selected functional groups found to be associated with aging in liver and kidney.

A

From 8992 genes analyzed	Months / Direction	Differentially expressed for each age vs 3m					Linearly changed from 9110 genes analyzed
		1	11	20	26	32	
Ribo-Seq	Up	347	26	13	27	442	369
	Down	296	7	2	10	217	149
RNA-Seq	Up	259	56	21	28	126	211
	Down	295	15	4	9	48	80



From 11461 genes analyzed	Months / Direction	Differentially expressed for each age vs 3m		Linearly changed from 11461 genes analyzed
		20	32	
Ribo-Seq	Up	199	1006	920
	Down	95	948	820
RNA-Seq	Up	134	477	610
	Down	90	394	556



B

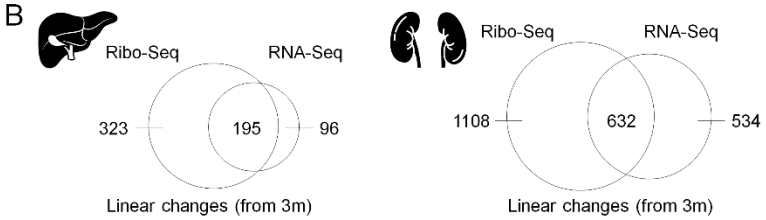


Figure S3. Summary of differential age-associated gene expression. (*A*) The number of differentially expressed genes in Ribo-Seq and RNA-Seq data found in paired comparisons of each age versus 3 months samples and linear age-dependent changes estimated from 3 to 32 months. (*B*) Venn diagrams illustrating the intersections between gene sets identified with linear age-dependent and basic paired differential expression analysis.

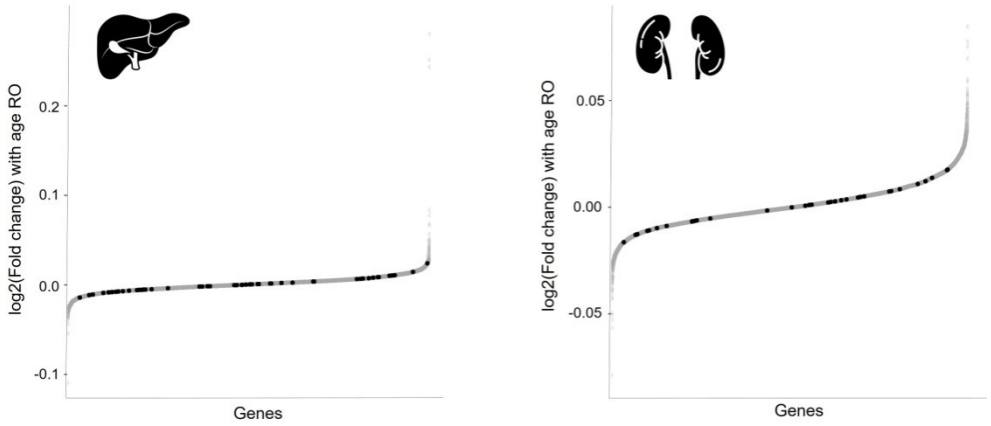


Figure S4. Association of RO changes with changes in transcript isoform abundance. Genes with isoform abundance changes with age shown in black among the list of all tested genes sorted according to their linear RO Fold Change with age (Tables S2 and S5). Genes were considered to exhibit age-related isoform abundance changes if the mean percentage of at least one of their isoforms changed more than 10% and standard deviation of a single isoform share within young (3 months for kidney, 3 and 11 months for liver) and old (32 months for kidney, 26 and 32 months for liver) animals being no more than 5% of the mean for kidney (34 genes) and 15% of the mean for liver (56 genes).

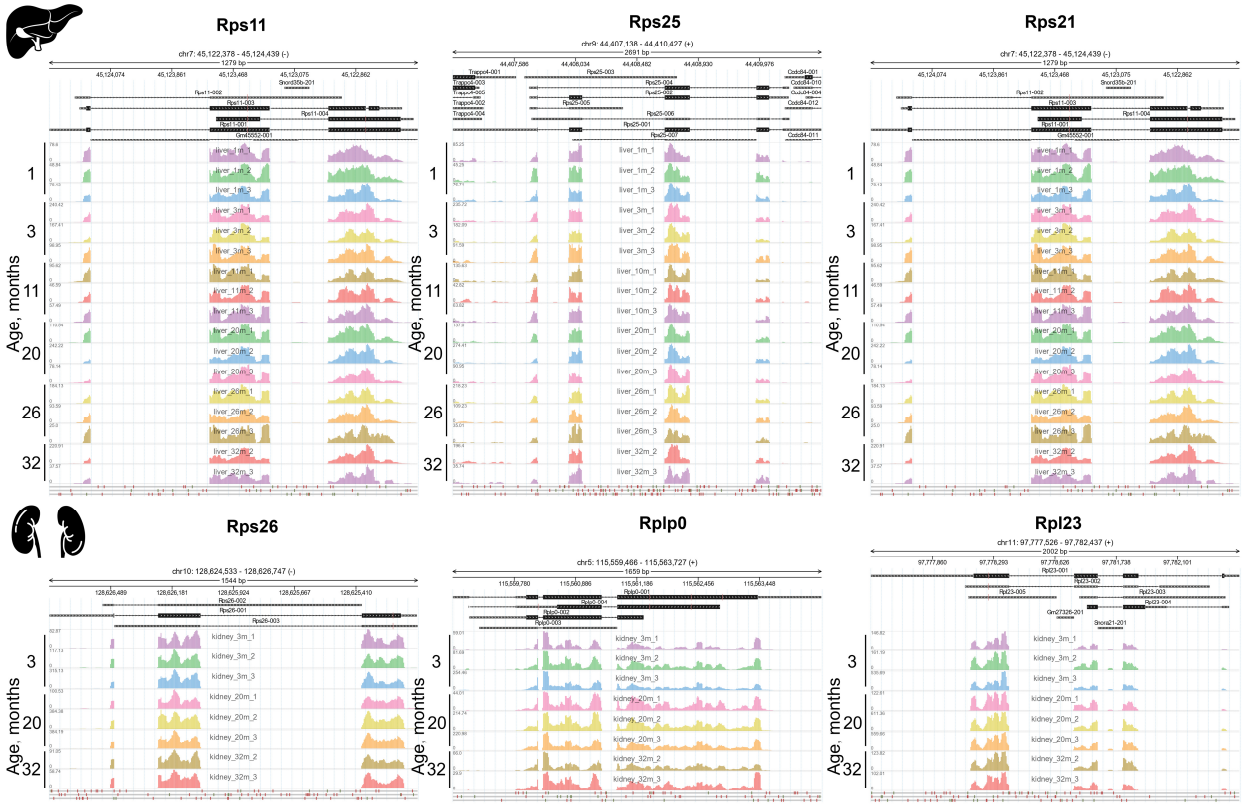


Figure S5. Genes encoding ribosomal proteins with the highest age-related ribosome occupancy changes do not show age-related isoform switching. Ribo-Seq read coverage visualization is plotted with svist4get (17).

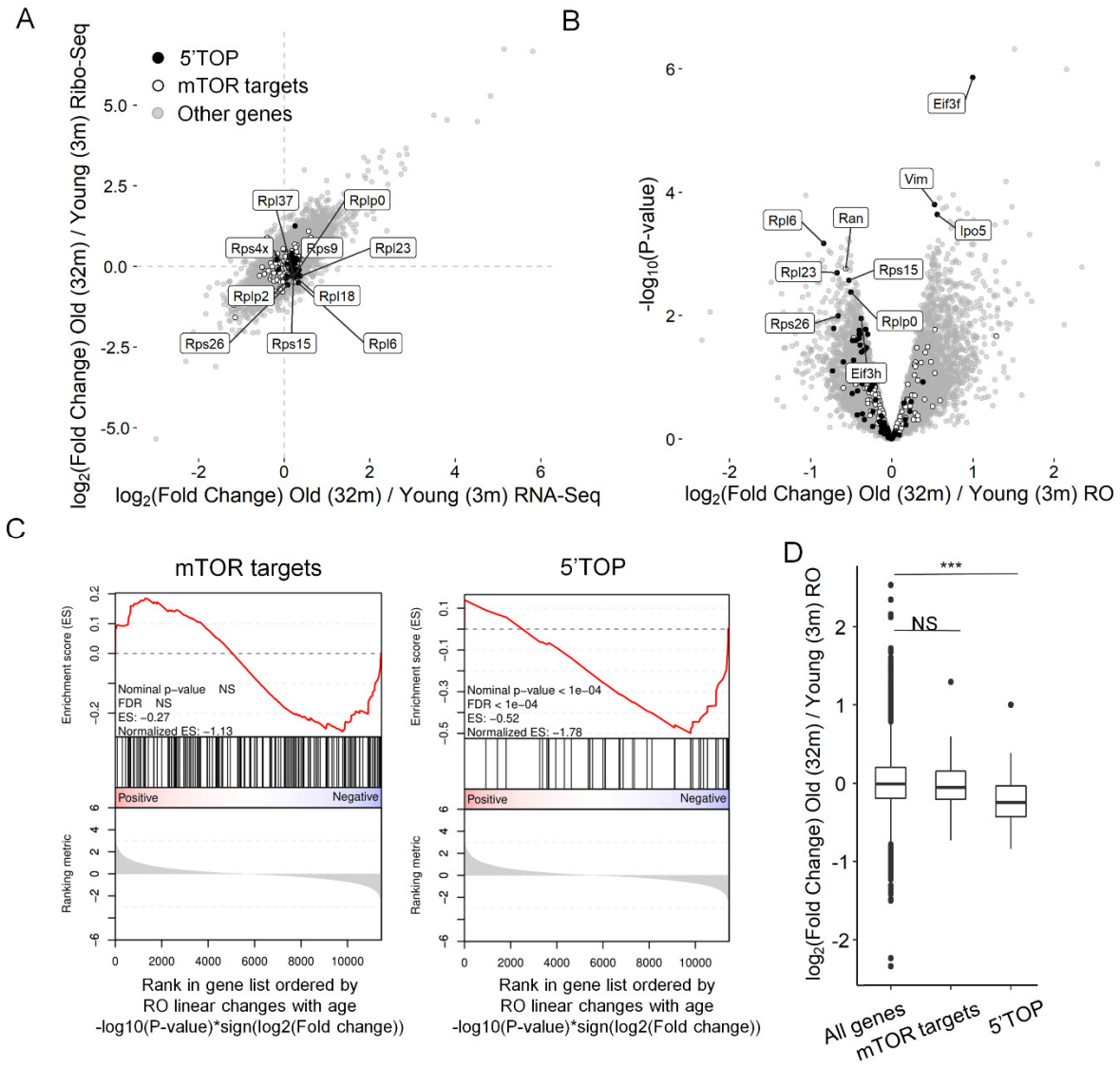


Figure S6. Decreased ribosome occupancy of transcripts encoding ribosomal and other translation-related proteins with age in kidney. (*A*) Comparison of transcriptome (RNA-Seq) and translatome (Ribo-Seq) $\log_2$(Fold change) between 32-month-old mice (2 mice) and 3-month-old-mice (3 mice). (*B*) The volcano plot shows the $\log_2$(Fold change) and P-values of the RO changes between 32-month-old mice (2 mice) and 3-month-old-mice (3 mice). (*C*) GSEA of age-related changes in RO of 44 5' TOP and 176 mTOR-sensitive genes (18) in kidney. RO linear changes with age (from 3- to 32-month-old mice) were estimated with edgeR (14 mice in total). Genes were sorted according to the signed P-values ($-\log_{10}$(P-value)*sign($\log_2$(Fold change)). (*D*) Box plot showing distribution of ROs for mTOR targets at the translation level and 5'TOP genes. Statistical significance was estimated with Mann–Whitney U test.

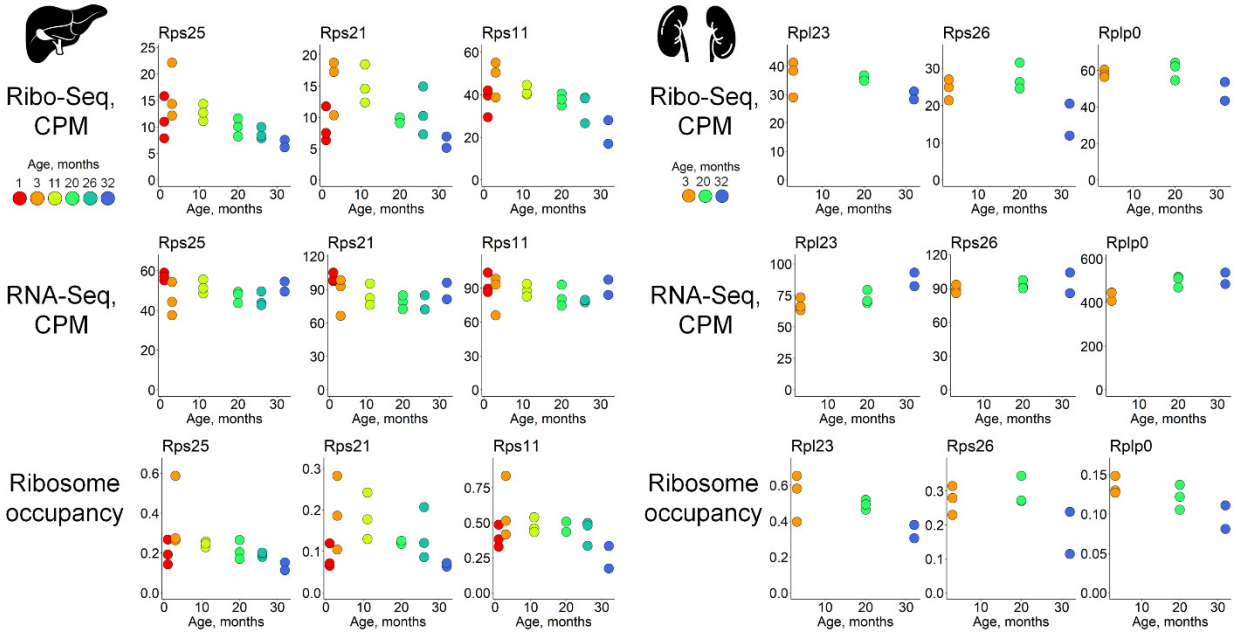


Figure S7. Age-dependent changes in Ribo- and RNA-Seq counts and ribosome occupancy (RO) of selected genes encoding proteins with functions associated with translation.

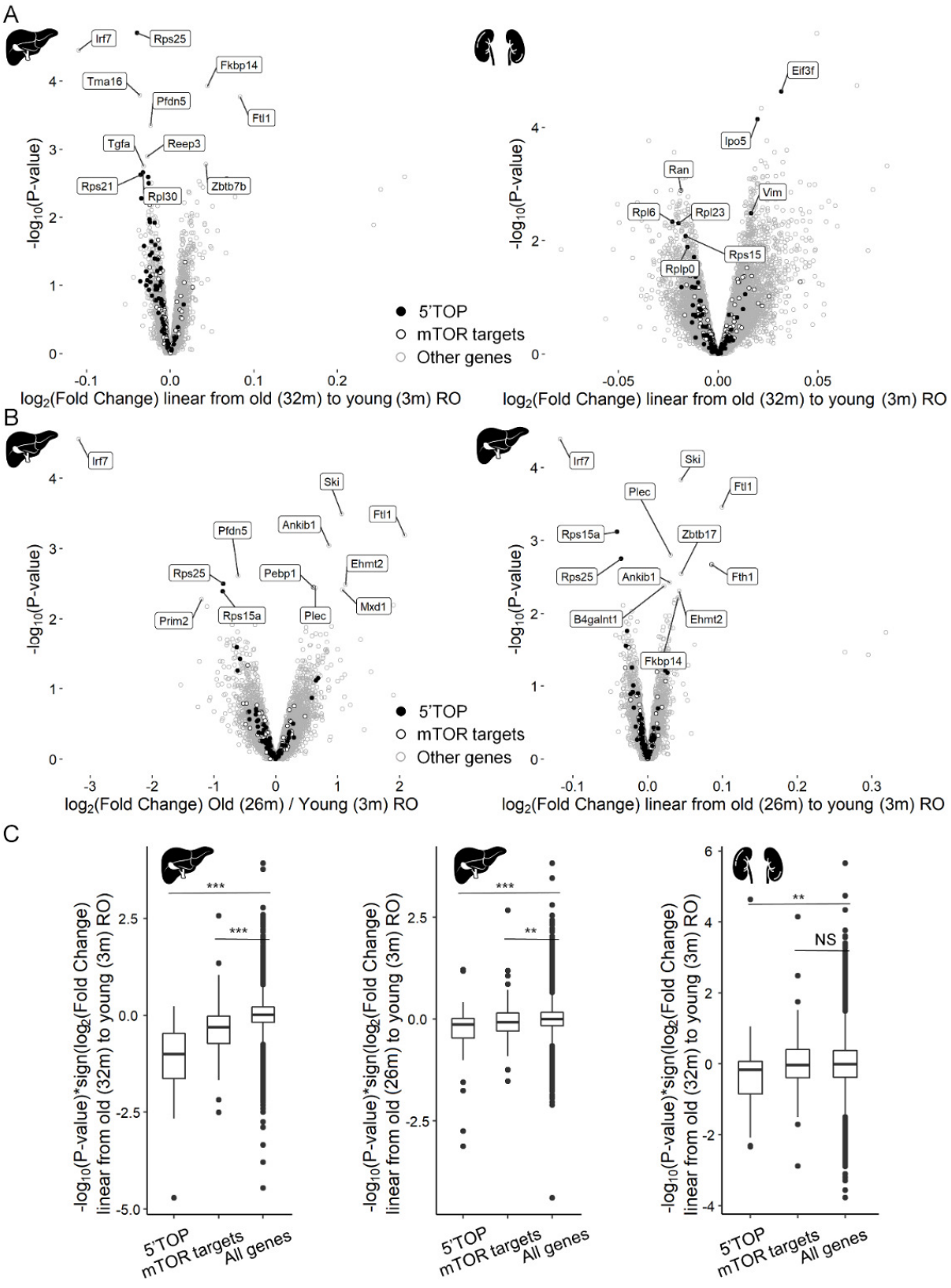
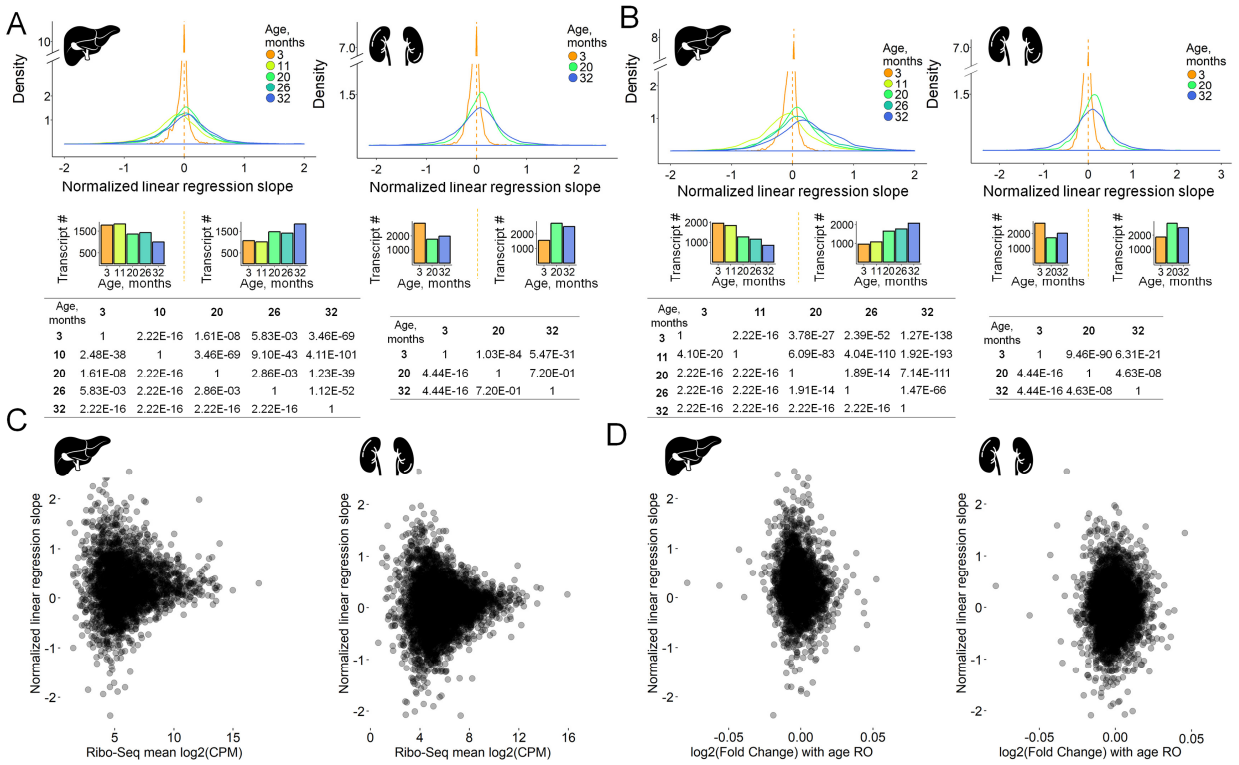

A
Irf7
Rps25
Fkbp14
Tma16
Pfdn5
Ftl1
Tgfa
Reep3
Rps21
Rpl30
Zbtb7b
-log10(P-value)
5'TOP
mTOR targets
Other genes
log2(Fold Change) linear from old (32m) to young (3m) RO
Eif3f
Ipo5
Ran
Vim
Rpl6
Rpl23
Rps15
Rplp0
B
Irf7
Ski
Ftl1
Pfdn5
Ankib1
Ehmt2
Rps25
Pebp1
Mxd1
Prim2
Rps15a
Plec
log2(Fold Change) Old (26m) / Young (3m) RO
Irf7
Ski
Plec
Ftl1
Rps15a
Zbtb17
Rps25
Ankib1
Fth1
B4galnt1
Ehmt2
Fkbp14
log2(Fold Change) linear from old (26m) to young (3m) RO
C
-log10(P-value)*sign(log2(Fold Change) linear from old (32m) to young (3m) RO)
***
***
5'TOP
mTOR targets
All genes
-log10(P-value)*sign(log2(Fold Change) linear from old (26m) to young (3m) RO)
***
**
-log10(P-value)*sign(log2(Fold Change) linear from old (32m) to young (3m) RO)
**
NS

Figure S8. Analysis of age-related linear changes in ribosome occupancy of transcripts encoding ribosomal and other translation-related proteins. RO changes between 3- and 26-month-old mice are shown for the liver. (*A*) Volcano plots showing the $\log_2$(Fold change) and P-values of the RO changes with age (from 3- to 32-month-old mice, 14 mice in total). RO linear changes with age were estimated using edgeR, and log2(Fold change) is the slope of linear regression. (*B*) Volcano plots showing the $\log_2$(Fold change) and P-values of the RO changes with age excluding 32-month-old mice. Paired (26- vs 3-month-old mice, 3 samples for each age group) and linear (from 3- to 26-month-old mice, 11 samples in total) comparisons are shown. (*C*) Box plots showing distribution of signed P-values for RO linear changes with age. Statistical significance was calculated with Mann–Whitney U test. The gene set sizes are: 44 for the 5' TOP genes, 176 for mTOR-sensitive genes, and 11,461 for all genes analyzed in kidney; 41 for 5' TOP, 160 for mTOR-sensitive genes, and 8,992 for all genes analyzed in liver. A list of the 5' TOP and mTOR-sensitive genes was taken from (18).



Age, months	3	10	20	26	32
3	1	2.22E-16	1.61E-08	5.83E-03	3.46E-69
10	2.48E-38	1	3.46E-69	9.10E-43	4.11E-101
20	1.61E-08	2.22E-16	1	2.86E-03	1.23E-39
26	5.83E-03	2.22E-16	2.86E-03	1	1.12E-52
32	2.22E-16	2.22E-16	2.22E-16	2.22E-16	1

Age, months	3	20	32
3	1	1.03E-84	5.47E-31
20	4.44E-16	1	7.20E-01
32	4.44E-16	7.20E-01	1

Age, months	3	11	20	26	32
3	1	2.22E-16	3.78E-27	2.39E-52	1.27E-138
11	4.10E-20	1	6.09E-83	4.04E-110	1.92E-193
20	2.22E-16	2.22E-16	1	1.89E-14	7.14E-111
26	2.22E-16	2.22E-16	1.91E-14	1	1.47E-66
32	2.22E-16	2.22E-16	2.22E-16	2.22E-16	1

Age, months	3	20	32
3	1	9.46E-90	6.31E-21
20	4.44E-16	1	4.63E-08
32	4.44E-16	4.63E-08	1

Figure S9. Age-related gradual rearrangement of ribosome footprints towards the 3' end of the coding sequence. (*A, B*) Distribution of linear regression slopes for smoothed ribosome footprint profiles normalized to mean coverage at 3 months with the relative transcript coordinate as the predictor variable, excluding (*A*) or including (*B*) the start and stop codons (see Supplementary Materials, Section 5). The data were pooled in each age group and the linear regression was performed. Bar plots depict the number of transcripts with negative (left) and positive (right) linear regression slope. Tables below the plots show P-values of the sign test comparing differences of transcript linear regression slopes (slopes for the ages in rows were subtracted from the slopes of the ages in columns). (*C, D*) Comparison between the ribosome footprint rearrangement along the transcript with age (3-months normalized linear regression slopes) with (*C*) the mean Ribo-Seq coverage and (*D*) ribosome occupancy (RO) linear changes with age (from 3 to 32 months).

SUPPLEMENTAL REFERENCES

1. Andrews S. FastQC: a quality control tool for high throughput sequence data. [Software] 2010; Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc
2. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 2011; 17:10.
3. Joshi N.A. and Fass J.N. Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33). [Software] 2011; Available online at: https://github.com/najoshi/sickle
4. GENCODE Release Files M13 (GRCm38.p5).
5. Dobin A., Davis C.A., Schlesinger F., Drenkow J., Zaleski C., Jha S., Batut P., Chaisson M. and Gingeras T.R. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 2013; 29:15–21.
6. Robinson M.D., McCarthy D.J. and Smyth G.K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 2010; 26:139–40.
7. Yu G., Wang L.-G., Han Y. and He Q.-Y. clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *Omi. A J. Integr. Biol.* 2012; 16:284–287.
8. Supek F., Bošnjak M., Škunca N. and Šmuc T. REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms. *PLoS One* 2011; 6:e21800.
9. Wang L., Wang S. and Li W. RSeQC: quality control of RNA-seq experiments. *Bioinformatics* 2012; 28:2184–5.
10. Vorontsov I.E., Fedorova A.D., Yevshin I.S., Sharipov R.N., Kolpakov F.A., Makeev V.J. and Kulakovskiy I. V. Genome-wide map of human and mouse transcription factor binding sites aggregated from ChIP-Seq data. *BMC Res. Notes* 2018; 11:756.
11. Stark C., Breitkreutz B.-J., Reguly T., Boucher L., Breitkreutz A. and Tyers M. BioGRID: a general repository for interaction datasets. *Nucleic Acids Res.* 2006; 34:D535–D539.
12. Li B. and Dewey C.N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 2011; 12:323.
13. Sergushichev A. An algorithm for fast preranked gene set enrichment analysis using cumulative statistic calculation. *bioRxiv* 2016; 10.1101/060012.
14. Subramanian A., Tamayo P., Mootha V.K., Mukherjee S., Ebert B.L., Gillette M.A.,

Paulovich A., Pomeroy S.L., Golub T.R., Lander E.S., *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* 2005; 102:15545–50.

15. Quinlan A.R. and Hall I.M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 2010; 26:841–2.
16. Ramensky V.E., Makeev V.J., Roytberg M.A. and Tumanyan V.G. Segmentation of long genomic sequences into domains with homogeneous composition with BASIO software. *Bioinformatics* 2001; 17:1065–1066.
17. Egorov A.A., Sakharova E.A., Anisimova A.S., Dmitriev S.E., Gladyshev V.N. and Kulakovskiy I. V. svist4get: a simple visualization tool for genomic tracks from sequencing experiments. *BMC Bioinformatics* 2019; 20:113.
18. Thoreen C.C., Chantranupong L., Keys H.R., Wang T., Gray N.S. and Sabatini D.M. A unifying model for mTORC1-mediated regulation of mRNA translation. *Nature* 2012; 485:109–13.