Molecular Cell, Volume *83*

Supplemental information

mRNA stability and m^{6}A are major determinants of subcellular mRNA localization in neurons

Inga Loedige, Artem Baranovskii, Samantha Mendonsa, Sayaka Dantsuji, Niko Popitsch, Laura Breimann, Nadja Zerna, Vsevolod Cherepanov, Miha Milek, Stefan Ameres, and Marina Chekulaeva

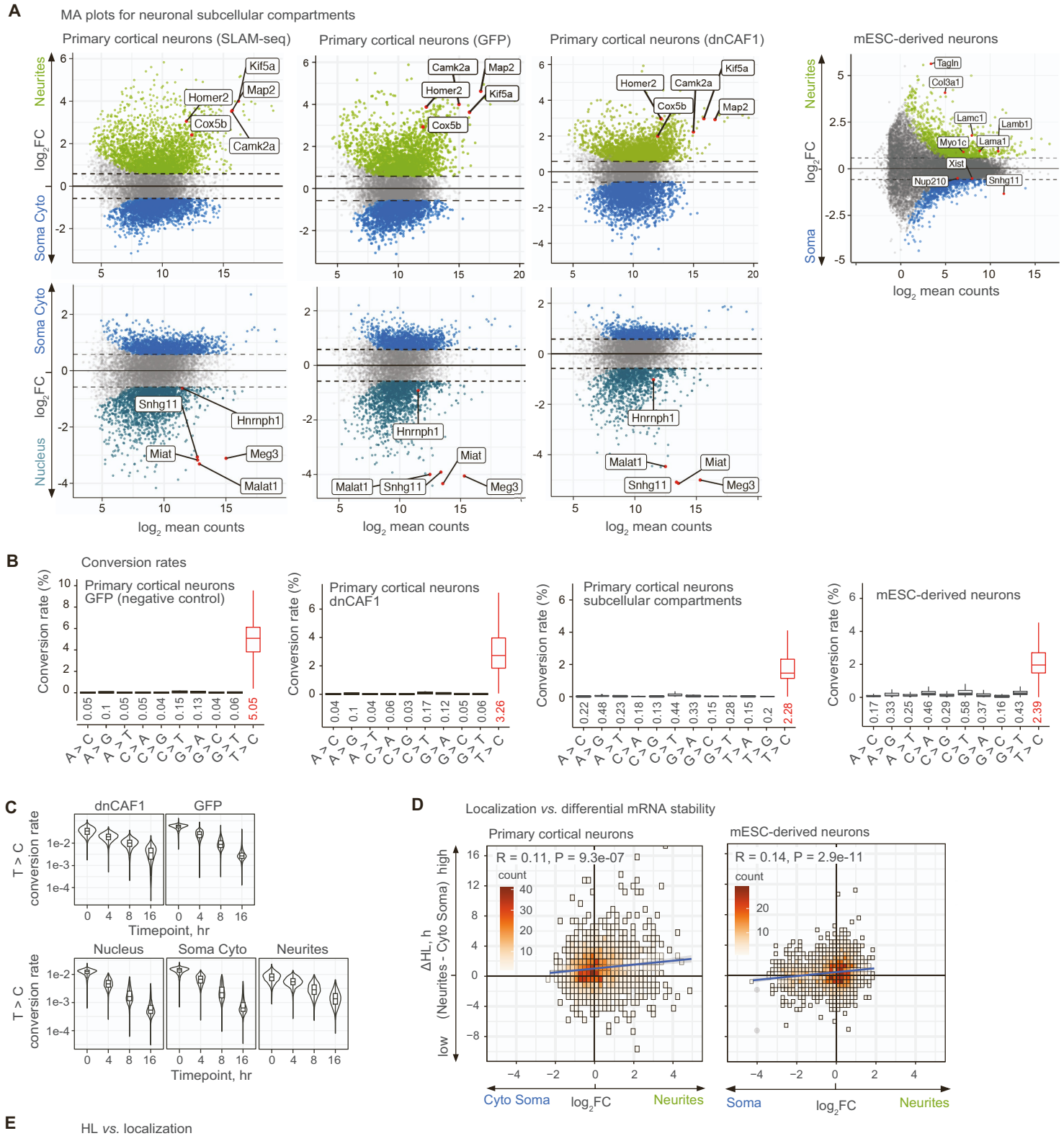


Figure S1, related to Figure 1. Quality controls for SLAM-seq and separation of subcellular neuronal compartments. (**A**) Neurons are efficiently separated on subcellular compartments. RNA-seq data are presented as log-transformed RNA enrichment in indicated subcellular compartments (Y) plotted against log-transformed mean counts (X). Selected labeled neurite-localized (green) and nucleus-localized transcripts (dark blue) serve as markers of efficient separation. (**B**) Boxplots showing the distribution of conversion rates in percentages (Y) for all possible conversions (X) among transcripts for which half-lives were estimated in primary cortical and mESC-derived neurons. (**C**) SLAM-seq measures mRNA degradation in subcellular neuronal compartments. Violin plots showing T to C conversion rates (Y) at different time points (X) for neurites, soma cytoplasm and nuclei of primary cortical neurons, and for primary cortical neurons expressing dnCAF1 or GFP (negative control). The data represent means of biological triplicates. (**D**) Differences in mRNA half-lives between neurites and soma positively correlate with localization to neurites. Scatterplots showing differences in mRNA half-lives between neurites and soma or soma cytoplasm (Y) are plotted against log-transformed RNA enrichment in indicated subcellular compartments (X). Pearson correlation coefficient is shown on the plot. To avoid overplotting, the color scale shows the number of transcripts per bin. (**E**) Higher overall mRNA stability favors localization to neurites. Transcripts stratified by the percentiles of overall half-lives (in the whole neuron, Y) are plotted as boxes against neuronal localization in PCNs (X). The difference with Figure 1D is that here nuclei were not excluded from soma preparations. P-value was computed with Pearson correlation test.

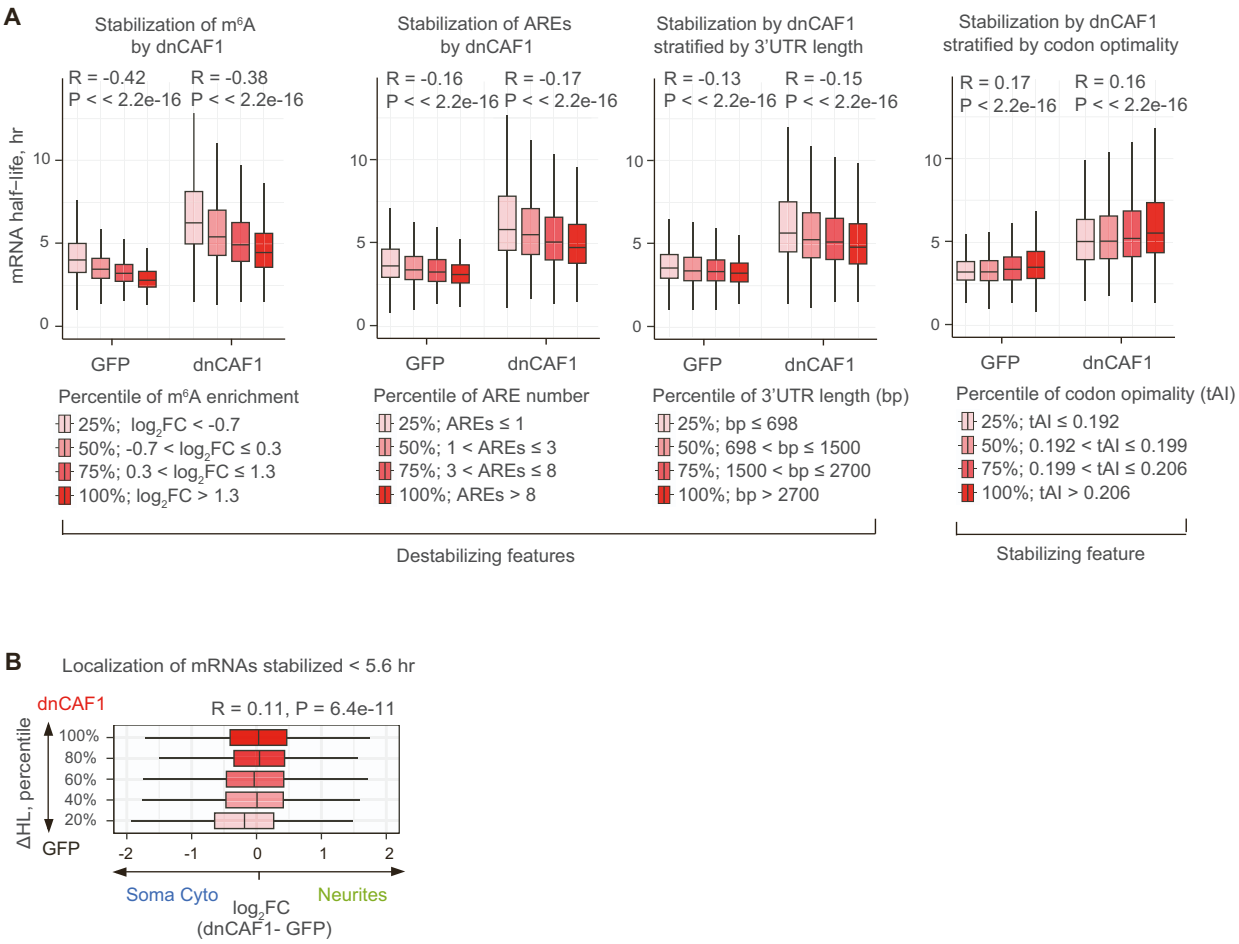


Figure S2, related to Figure 3. mRNA stabilization by dnCAF1 expression. (**A**) dnCAF1 globally stabilizes mRNAs, without preference for transcripts enriched in destabilizing elements. Boxplots showing mRNA half-lives for neurons expressing GFP or dnCAF1 (Y) categorized according to indicated features that influence mRNA stability (left to right: m^{6}A, AREs, 3'UTR length, codon optimality). The transcripts are divided into quartiles from the least to the most frequent occurrence of these features, represented by shades of red. Left to right: For m^{6}A analysis, mRNAs are grouped according to their enrichment in m^{6}A-RNA-IP. For ARE analysis, mRNAs are grouped according to the number of AREs in their 3'UTRs. For 3'UTR length analysis, the transcripts are grouped based on the length of their 3'UTR. For analysis of codon optimality, mRNAs are grouped according to their gene-wise tRNA Adaptation Index (tAI). Refer to the figure for more details. P-values were computed with Pearson correlation test. (**B**) Changes in mRNA localization upon dnCAF1 expression correlate with the extent of their stabilization. Transcripts with the final half-life below 5.6 hours in dnCAF1 neurons were stratified by the percentiles of their stabilization in dnCAF1-expressing neurons (Y) and plotted as boxes against changes in localization between dnCAF1- and GFP-expressing PCNs (X). Boxes are colored by the degree of stabilization. P-value was computed with Pearson correlation test.

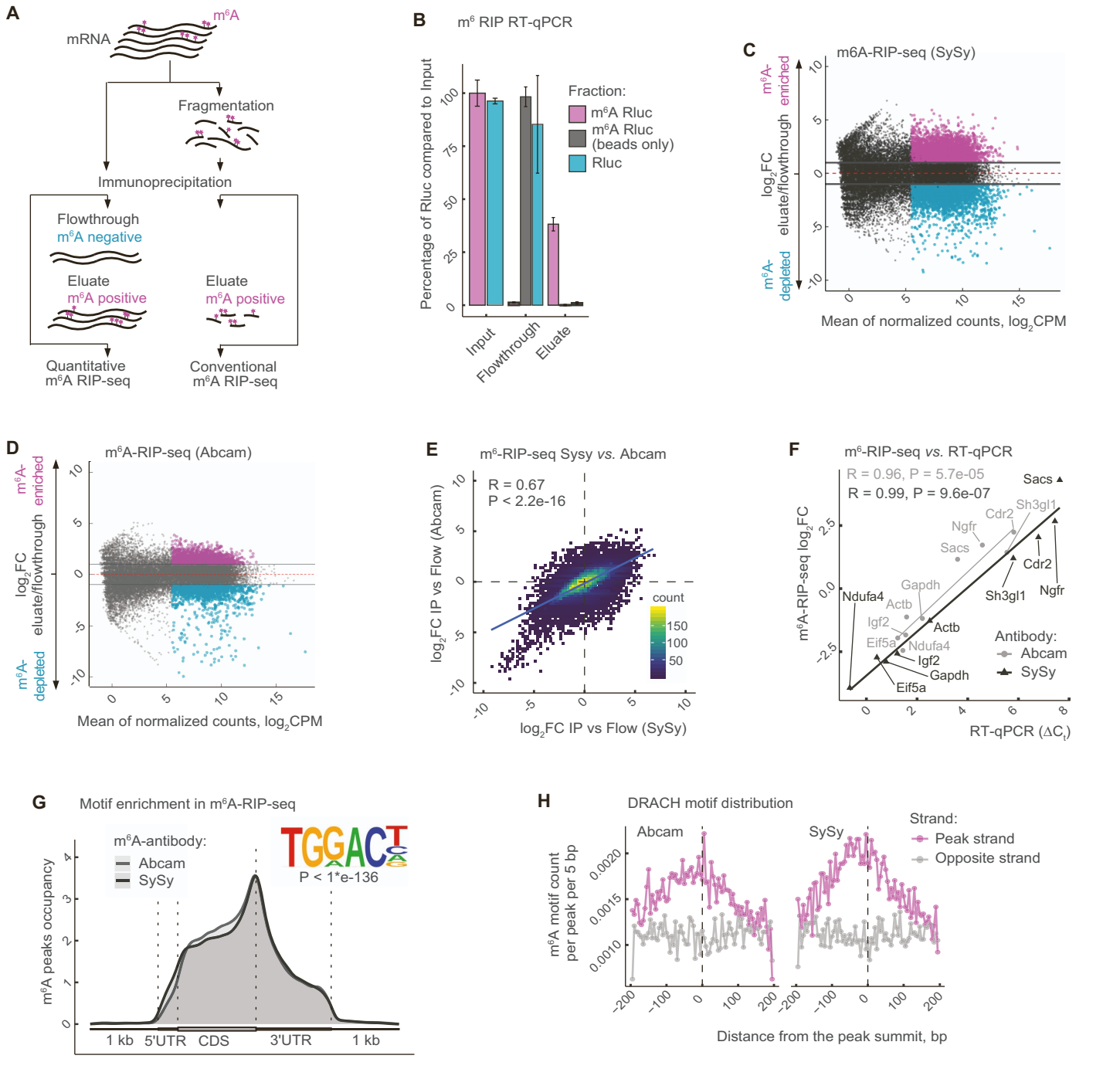
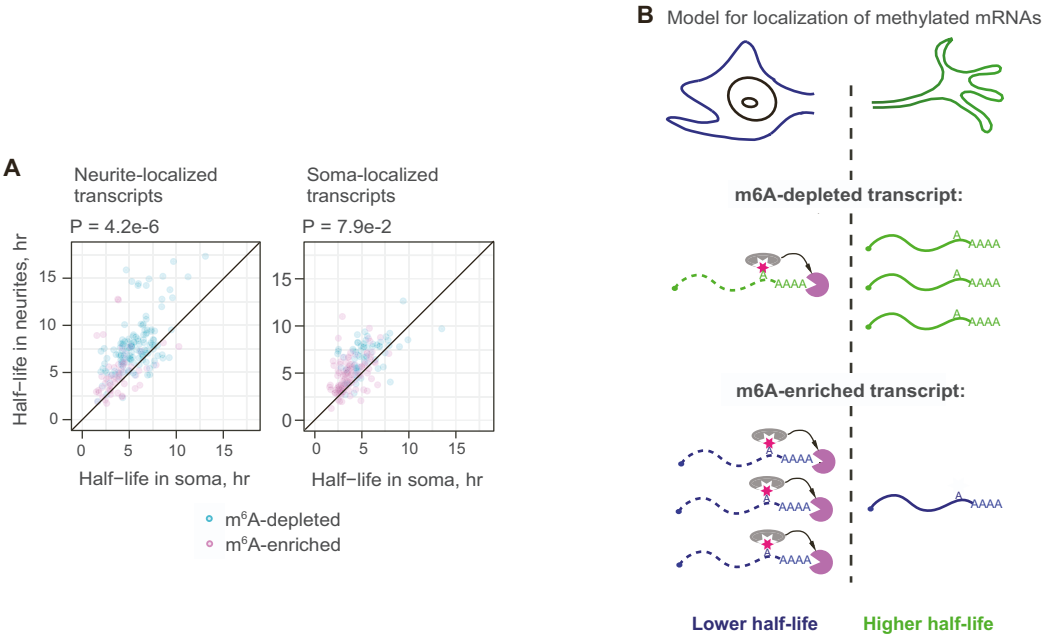
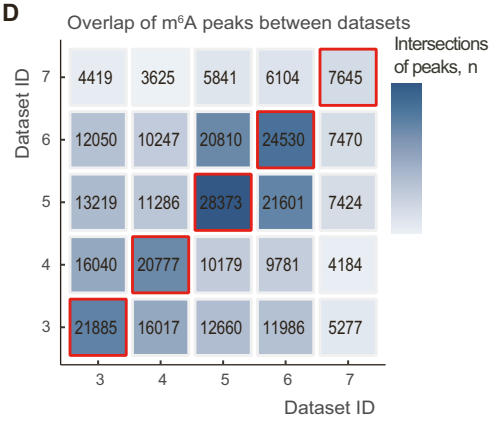


Figure S3, related to Figure 4. Transcriptome-wide analysis of m6A levels and distribution by m⁶A-RIP-seq. (**A**) Scheme of quantitative (left) and conventional (right) m⁶A-RIP-seq. For quantitative m⁶A-RIP-seq, m⁶A-containing transcripts are pulled down with antibodies against m⁶A, eluted with m⁶A nucleoside and mRNA levels in eluate and flowthrough are quantified by RNA-seq. Methylation levels are estimated as a log2FC between eluate and flowthrough fractions. For conventional m⁶A-RIP-seq, RNA is fragmented prior to pulldown and the eluate is analyzed by RNA-seq, allowing identification of a consensus motif and motif distribution along the transcript. Quantitative m⁶A-RIP-seq was performed in experiments shown in **Figure S3C-F** and conventional m⁶A-RIP-seq in **Figure S3G-H**. (**B**) Validation of m6A-RIP protocol by RT-qPCR on methylated (m⁶A-Rluc, magenta) and unmethylated RNA spike-ins (Rluc, cyan) in the input, flowthrough and eluate fractions. m⁶A-Rluc level in the input is taken for 100% and the rest is normalized relative to the m⁶A-Rluc input. Error bars show standard deviation of technical duplicates. (**C-D**) m⁶A-methylome of mESC-derived neurons, generated using SySy (**C**) or Abcam (**D**) m⁶A antibodies. The data are presented as RNA enrichment in eluate versus flowthrough (Y) plotted against average RNA abundance (X), expressed in log2 CPMs (counts per million mapped reads). Magenta: m⁶A-enriched transcripts, log2FC > 1, p-value < 0.05. Cyan: m⁶A-depleted transcripts, log2FC < 1, p-value < 0.05. (**E**) m⁶A-methylomes evaluated with two different antibodies show strong correlation. (**F**) Validation of m⁶A-methylation levels from m⁶A-RIP-seq by RT-qPCR. mRNA log2FC between eluate and flowthrough measured by m6A-RIP-seq (Y) is plotted against the same log2FC measured by RT-qPCR (X). The linear least squares fit between m⁶A-RIP-seq and RT-qPCR data is shown. R is the Pearson correlation coefficient. (**G**) Motif discovery and distribution of m⁶A-peaks along transcript. UTR: untranslated region; CDS: coding regions. m⁶A peaks from fragmented m⁶A-RIP-seq were called with R package exomePeak and plotted over gene features using R package Guitar. (Inset) *De novo* motif discovery identified an m⁶A consensus DRACH motif in m⁶A peaks. (**H**) Distributions of the discovered m6A motif DRACH (D=A, G or U; H=A, C or U) around the peak summit. Motif count (Y) is plotted over the 200 bp flanks around the peaks summit (X).



C m⁶A-RIP-seq datasets

ID:	Study:	m⁶A AB:	RNA treatment:
1	This study	Sysy	Unfragmented
2	This study	Abcam	Unfragmented
3	This study	Sysy	Fragmented
4	This study	Abcam	Fragmented
5	Merkurjev et al. 2019	Sysy	Fragmented
6	Merkurjev et al. 2019	Sysy	Fragmented
7	Engel et al. 2018	Sysy	Fragmented



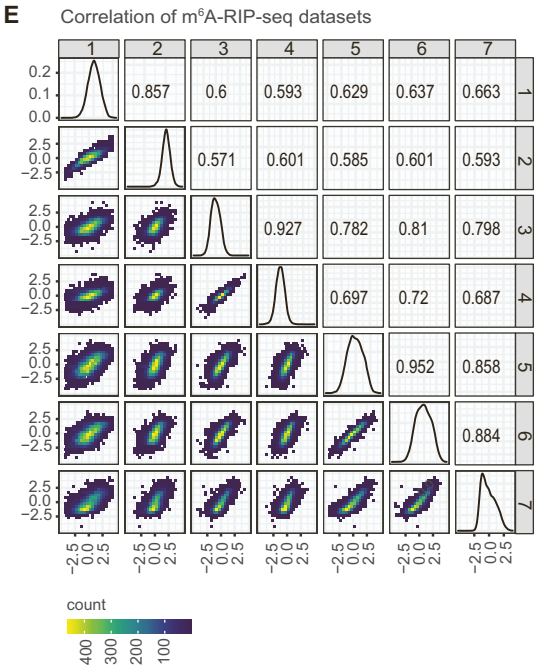
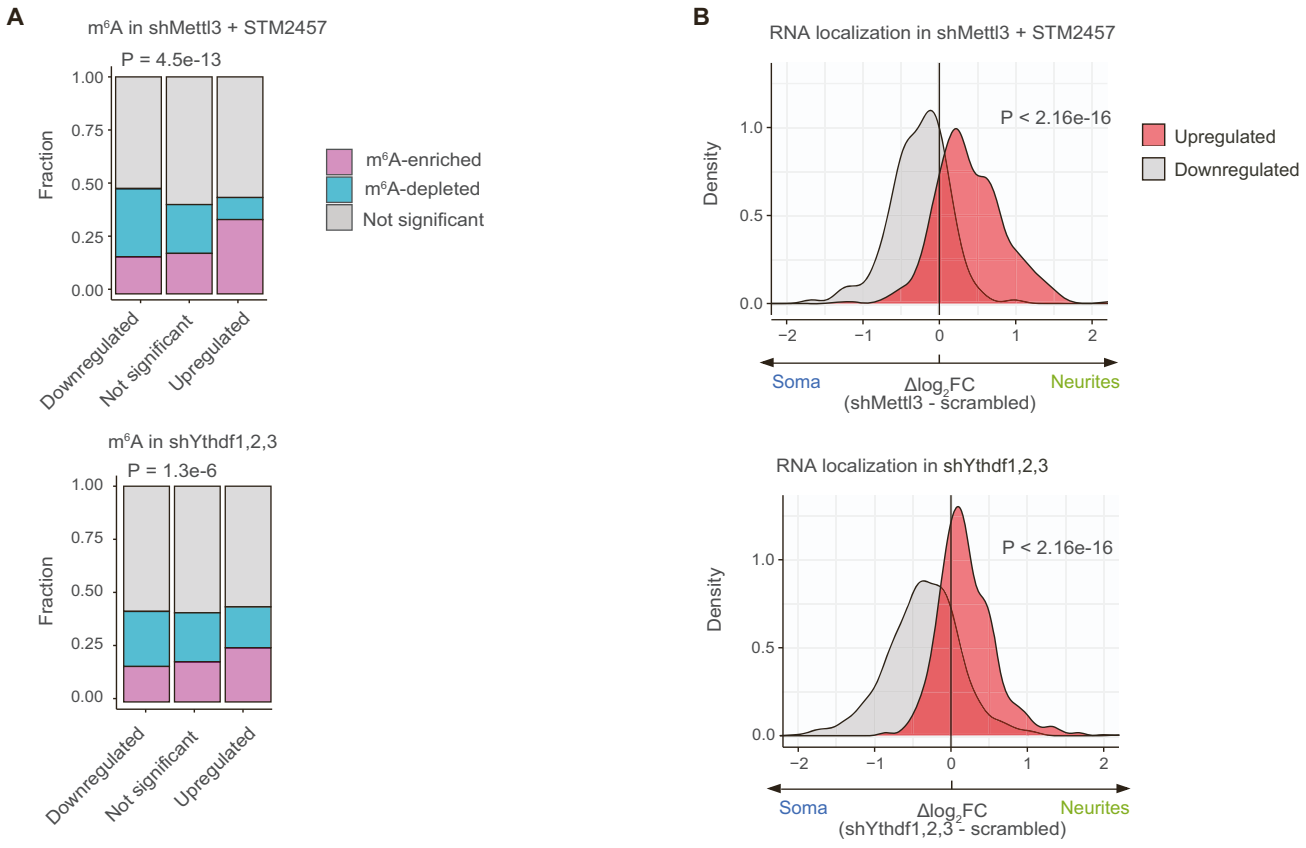


Figure S4, related to Figure 4. m⁶A levels correlate between multiple neuronal m⁶A-RIP-seq datasets and anti-correlate with mRNA stability. (**A**) m⁶A contributes to the differences in mRNA half-lives between neurites and soma. Scatterplots showing half-lives for neurite- (left) and soma-localized mRNA (right) in neurites (X) and soma (Y). Cyan: m⁶A-depleted, magenta: m⁶A-enriched transcripts. P-values for half-live differences between neurites and soma for m⁶A-enriched and m⁶A-depleted transcripts were calculated using Wilcoxon test. (**B**) Model for distribution of mRNAs between neurites and soma depending on their methylation status. (**C**) Table with annotation of m⁶A-RIP-seq datasets used in this study, including the source of the data, antibody used for the m⁶A-RIP and RNA treatment (either fragmented before RIP or not). (**D**) Heatmap showing the number of intersections between m⁶A peaks identified in fragmented m⁶A-RIP-seq datasets. Intersections below the diagonal were computed using dataset on X as a reference, while intersections above the diagonal use dataset on Y as a reference. Labels on both X and Y correspond to the annotation of the datasets shown in (**C**). (**E**) Pairwise correlations between log2FC identified in each of the m⁶A-RIP-seq datasets. Below the diagonal, 2-dimensional density plots are presented, showing the distribution of transcripts relative to the log2FC (Y - left, X - bottom) from the corresponding pair of datasets (Y - right, X - top). Pearson correlation coefficients between the log2FC (Y - left, X - bottom) of the corresponding datasets (Y - right, X - top) are reported above the diagonal. Diagonal shows the log2FC distributions in each individual dataset.



Upregulated: RNA-seq log2FC shRNA vs. scrambled > 0.58, p-value < 0.05
Downregulated: RNA-seq log2FC shRNA vs. scrambled < -0.58, p-value < 0.05
Not significant: the rest

Figure S5, related to Figure 4. Transcripts upregulated after perturbation of m^{6}A machinery shift their localization to neurites. (**A**) m^{6}A-enriched transcripts are overrepresented among transcripts that are upregulated after the perturbation of m^{6}A machinery. Percentages of m^{6}A-enriched (magenta), m^{6}A-depleted (cyan) and the rest (grey) of mRNAs (Y) among groups of transcripts which are upregulated (RNA-seq log2FC shRNA vs. scrambled > 0.58, p-value < 0.05 in both soma and neurites), downregulated (RNA-seq log2FC shRNA vs. scrambled < -0.58, p-value < 0.05 in both soma and neurites transcripts) or not significantly changing their levels upon depletion of *Mettl3* (top) or *Ythdf* (bottom). Statistical significance was estimated using χ2-test (Chi-squared test). (**B**) Transcripts that are upregulated upon *Mettl3* and *Ythdf* depletion shift their localization to neurites. Changes in localization (log2FC neurites vs. soma) upon knockdown of specified transcripts (X) are plotted as densities for two groups of transcripts: grey, upregulated transcripts; red, downregulated transcripts (as defined in **A**). Statistical significance of difference between means of two distributions was estimated using Kolmogorov-Smirnov test.

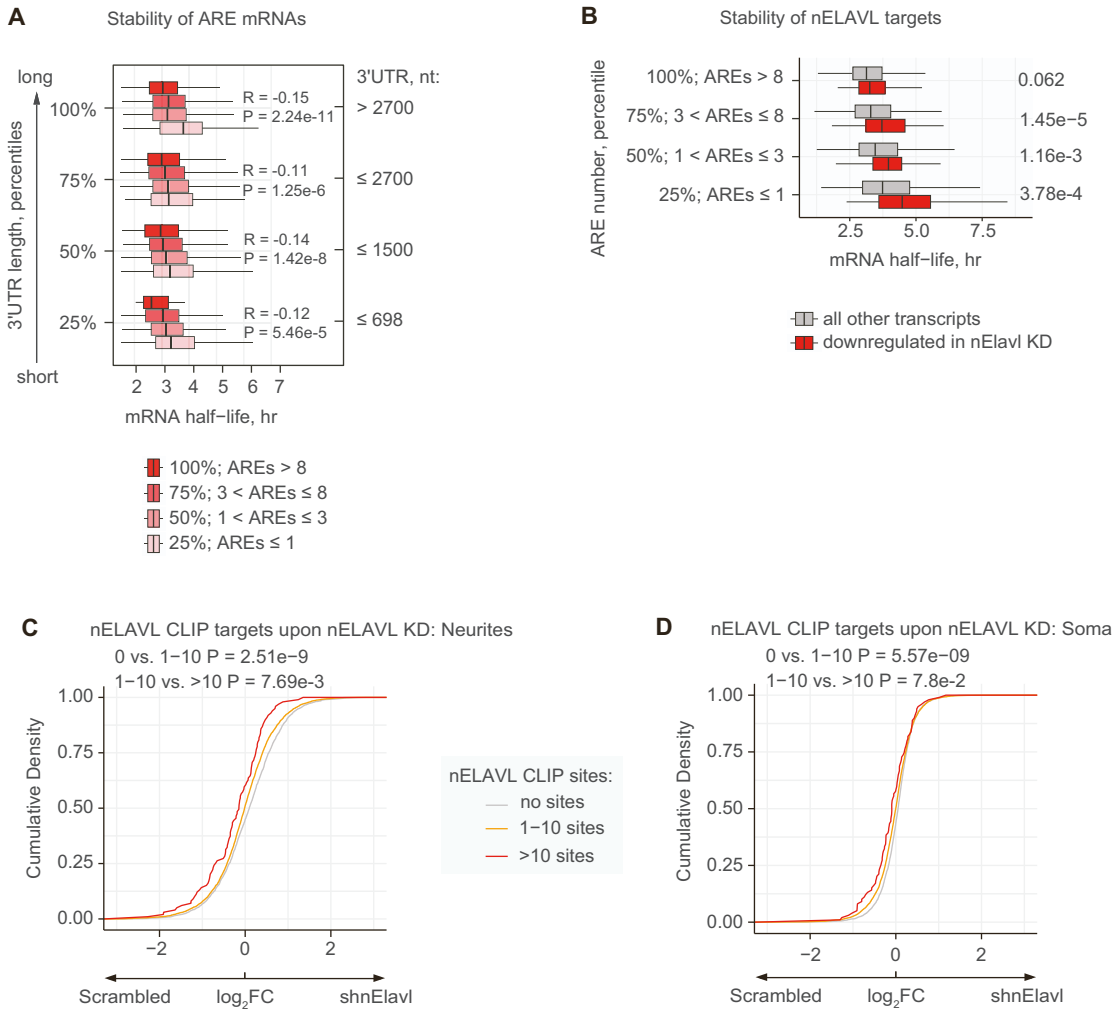


Figure S6, related to Figure 5. AREs are enriched in unstable mRNA and are upregulated by neuronal ELAVL proteins. (**A**) AREs are enriched in unstable mRNAs independently of 3'UTR length. Boxplots showing the distribution of mRNA half-lives (X) for mRNAs with different number of AREs (shades of red) in their 3'UTR. Transcripts are stratified by the percentiles of 3'UTR length. P-values were computed with Pearson correlation test. (**B**) nELAVL targets are more stable than other transcripts. Boxplots showing the distribution of mRNA half-lives (X) for nELAVL targets (red, RNA-seq log2FC shnElavl vs. scrambled < -0.58, p-value < 0.05 in soma and neurites) and all other transcripts (grey). Transcripts are grouped into quartiles according to the number of AREs (Y). P-values were computed with Wilcoxon test. (**C-D**) Transcripts carrying a higher number of nELAVL CLIP sites exhibit stronger downregulation in neurites (**C**) and soma (**D**) following nELAVL knockdown than transcripts with fewer or no CLIP sites. Cumulative distribution functions (CDF) showing fractions of transcripts (Y) with no nELAVL CLIP sites (grey), 1 to 10 nELAVL CLIP sites (yellow) or > 10 nELAVL CLIP sites (red), plotted against changes in transcript levels upon *nElavl* knockdown (X). nELAVL CLIP data are from Ince-Dunn et al.[1]. P-values computed with Wilcoxon test.

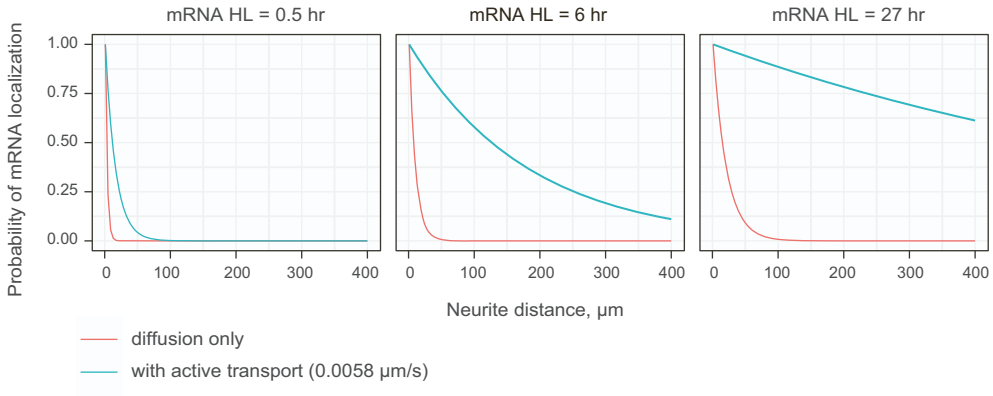


Figure S7, related to Figure 7. Diffusion alone cannot provide for localization at the distance of 100 µm and further. Modelled distribution of mRNA transcripts along the neurite distance (X) with diffusion only (red line) and active transport (cyan line). The distribution is calculated for a range of mRNA half-lives measured in our experiments, from the minimal to the maximal. The computational framework of Fonkeu et al.[2] was used.

Supplementary References

1 Ince-Dunn, G. *et al.* (2012). Neuronal Elav-like (Hu) proteins regulate RNA splicing and abundance to control glutamate levels and neuronal excitability. *Neuron* **75**, 1067-1080, doi:10.1016/j.neuron.2012.07.009.

2 Fonkeu, Y. *et al.* (2019). How mRNA Localization and Protein Synthesis Sites Influence Dendritic Protein Distribution and Dynamics. *Neuron* **103**, 1109-1122 e1107, doi:10.1016/j.neuron.2019.06.022.