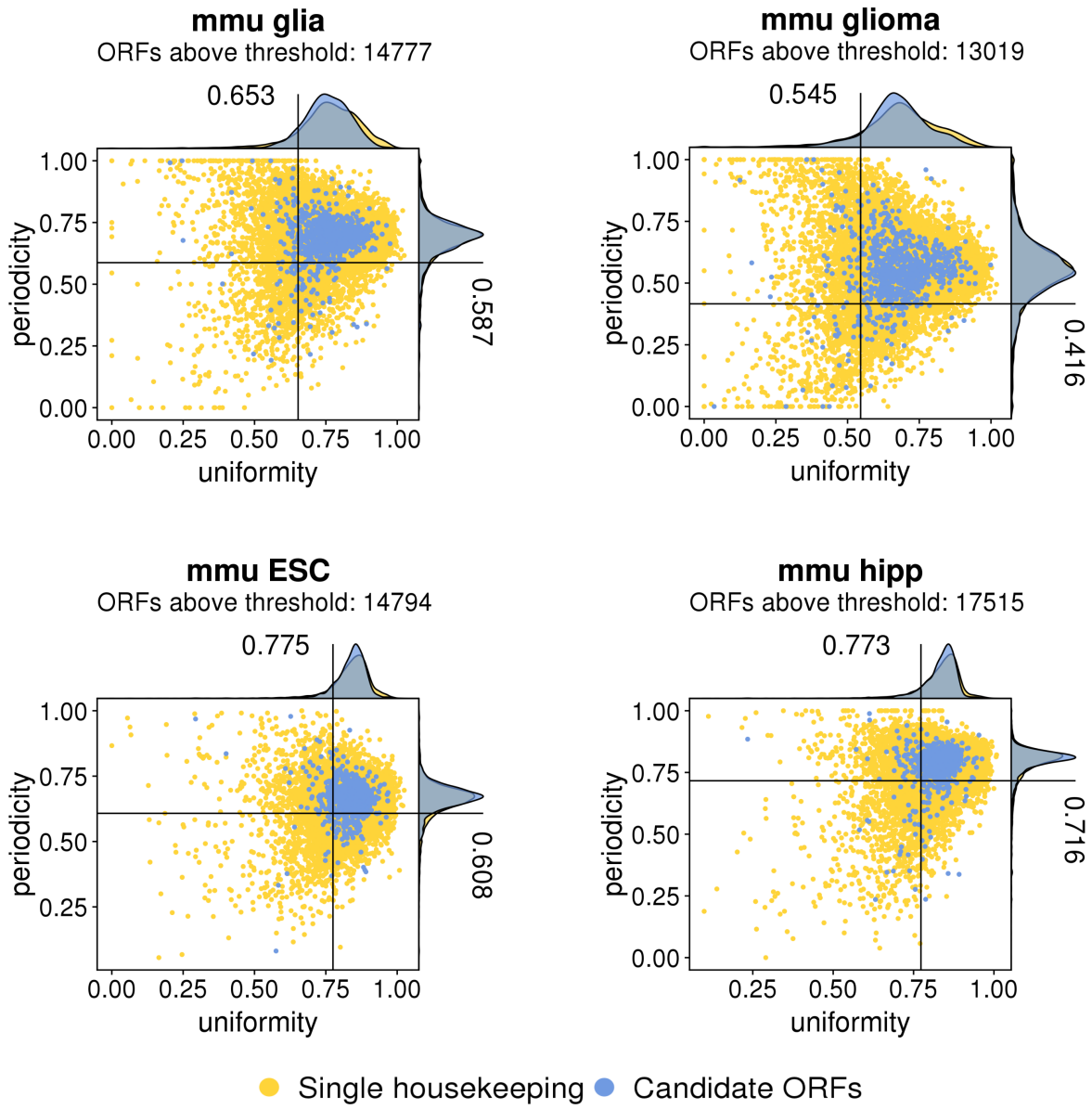
Supplementary Figures

Ribosome profiling at isoform level reveals an evolutionary conserved impact of differential splicing on the proteome

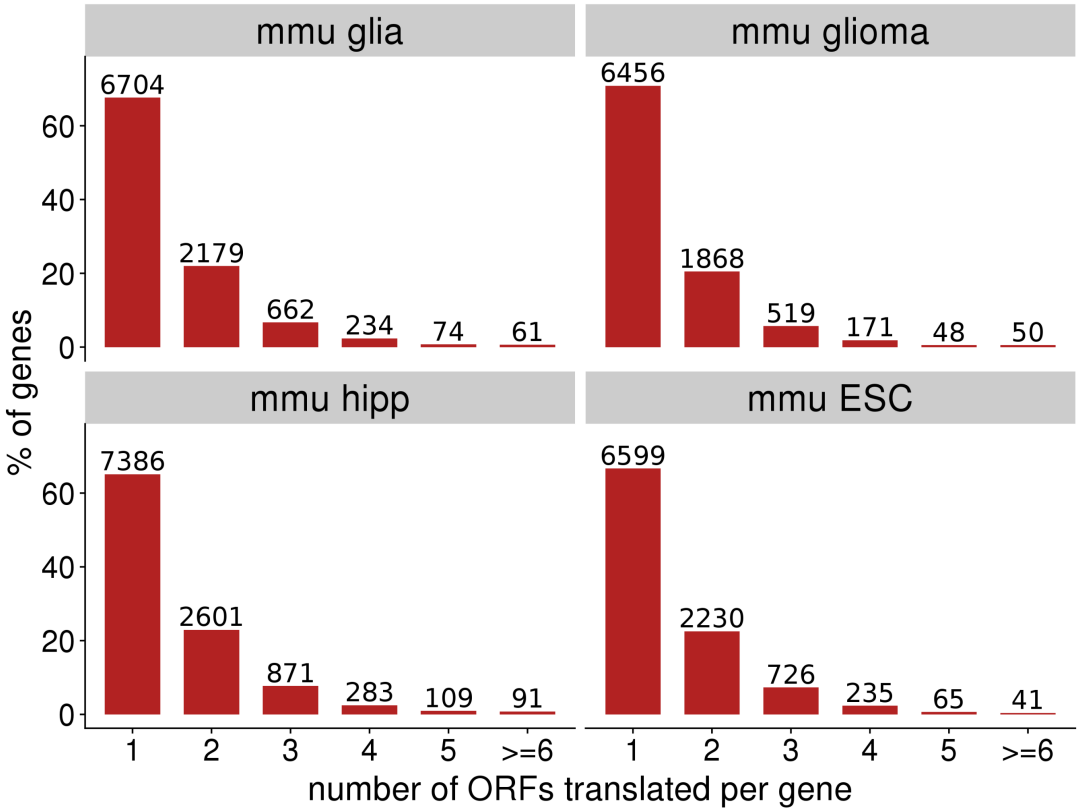
Marina Reixachs-Solé[1], Jorge Ruiz-Orera[2], M Mar Alba[2,3], Eduardo Eyras[2,4,*],

[1]Pompeu Fabra University, E08003 Barcelona, Spain.
[2]IMIM - Hospital del Mar Medical Research Institute. E08003 Barcelona, Spain.
[3]Catalan Institution for Research and Advanced Studies. E08010 Barcelona, Spain
[4]The John Curtin School of Medical Research, Australian National University, Acton ACT 2601, Canberra, Australia

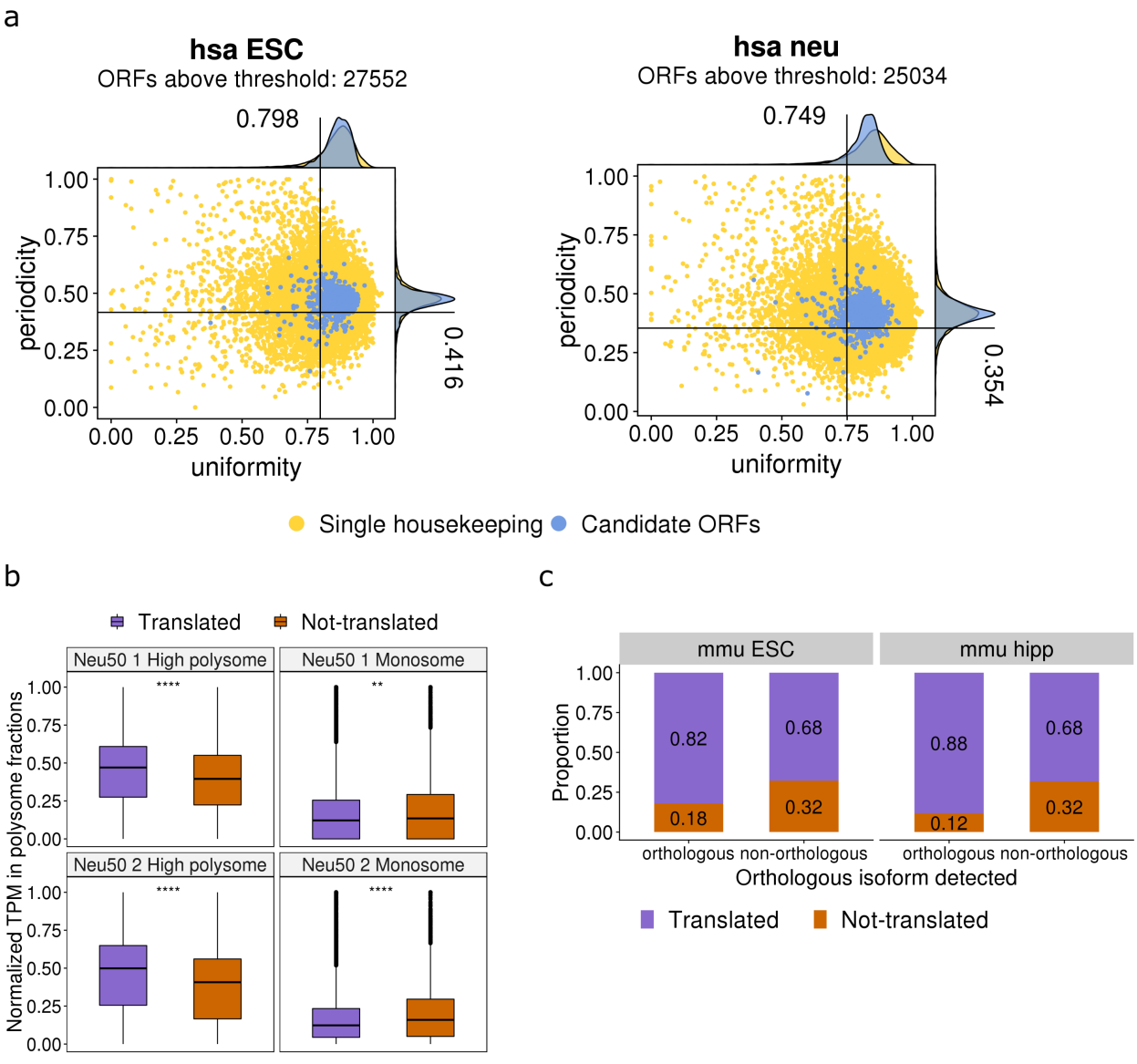
*correspondence to: eduardo.eyras@anu.edu.au



Supplementary Figure 1. Uniformity (x axis) versus periodicity (y axis) for the ORFs with with RNA expression TPM > 0.1 and at least 10 Ribo-seq reads assigned. In blue we indicate ORFs from genes defined as housekeeping singletons, and in yellow the rest of ORFs considered. Uniformity is measured as the percentage of maximum entropy and periodicity is measured in the first annotated frame. We show the data for mouse samples of glia (mmu glia), glioma (mmu glioma), Embrionic Stem Cells (mmu ESC) and hippocampus (mmu hipp).

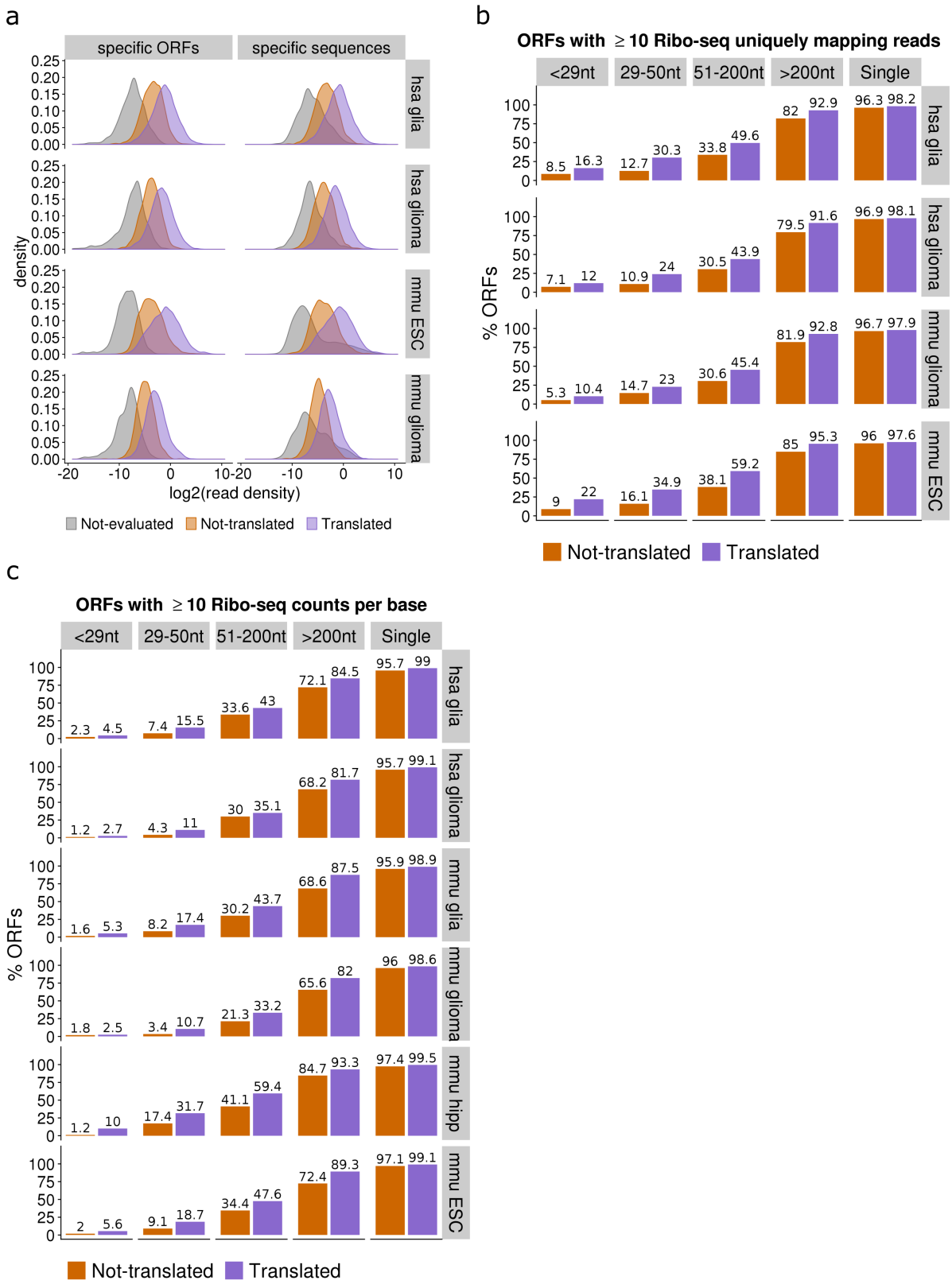


Supplementary Figure 2. Distribution of the number of different ORFs translated per gene in the mouse samples of glia (mmu glia), glioma (mmu glioma), hippocampus (mmu hipp) and Embrionic Stem Cells (mmu ESC).



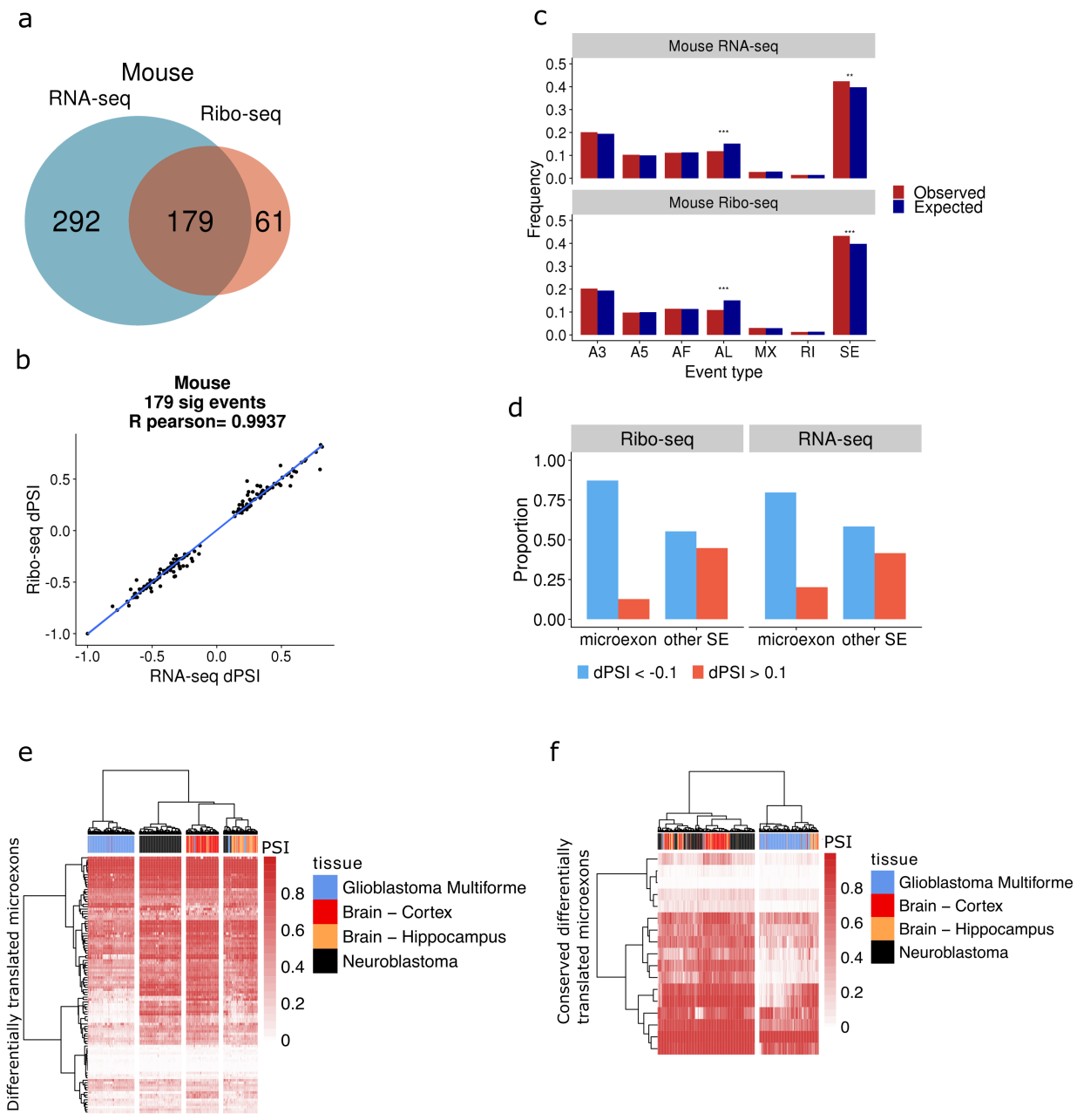
Supplementary Figure 3. (a) Uniformity (x axis) versus periodicity (y axis) for the ORFs with with RNA expression TPM > 0.1 and at least 10 Ribo-seq reads assigned. In blue we indicate ORFs from genes defined as housekeeping singletons, and in yellow the rest of ORFs considered. Uniformity is measured as the percentage of maximum entropy and periodicity is measured in the first annotated frame. We show the data for the samples of human ESCs (left panel) and human differentiated neurons (right panel). **(b)** For isoforms predicted to be translated, and for isoforms with enough RNA expression (TPM>0.1) but predicted as not translated, we show the distributions of the relative abundance in monosomes and high polysome fractions (Methods). The plot shows the results for the two replicates for monosome (replicate 1 p-value = 1.21e-03 and replicate 2 p-value = 2.50e-32) and high polysome (replicate 1 p-value = 6.10e-61 and replicate 2 p-value = 6.15e-87) fractions for neuronal cells. **(c)** For the set of ORFs encoding a human-mouse orthologous protein pair (orthologous) and for those encoding proteins without an orthologous pair in mouse (non-orthologous) we plot the percentage that are predicted to be translated (translated)

and the ones that did not pass the uniformity and periodicity thresholds (not-translated). We show here the results for mouse ESCs (p-value = 5.228e-124 Fisher test) and for mouse hippocampus (p-value = 9.821e-311).



Supplementary Figure 4. (a) For the samples of human glia (hsa glia), human glioma (hsa glioma), mouse Embrionic Stem Cells (mmu ESC) and mouse glioma (mmu glioma), we show the density of Ribo-seq reads per nucleotide over the isoform-specific regions, measured as the uniquely mapping read-count over region length in $\log_2$ scale for isoform-specific sequences (right

panels) and the counts per nucleotide based on the estimated P-site positions over region length in $\log_2$ scale for isoform-specific ORFs (left panels). Distributions are given for predicted translated isoforms, for isoforms that did not pass the threshold of uniformity and periodicity (not translated), and for the isoforms with low expression (TPM<0.1) (not evaluated). **(b)** For the same samples we show the percentage of regions with at least 10 uniquely mapping Ribo-seq reads in isoform-specific sequences over the total number of isoforms with an isoform-specific sequence defined according to the length of the region. **(c)** We show the percentage of regions with at least 10 Ribo-seq counts in the isoform-specific ORFs considering the estimated P-site position according to the length of the region.



Supplementary Figure 5. (a) Overlap of events changing significantly (dPSI > 0.1 and p-value < 0.05) with RNA-seq and with Ribo-seq for mouse. **(b)** Correlation of changes in splicing and translation in events in mouse. **(c) P**roportions of events calculated in RNA space (upper panel) or ribosome space (lower panel). In red we show the proportion of alternative splicing events calculated with SUPPA that overlap coding regions in mouse, whereas in blue we show the events that show a significant change using RNA-seq comparing mouse glia and glioma. Even types are: alternative 3'ss (A3) and 5'ss (A5), alternative first (AF) and last (AL) exon, mutually exclusive (MX) exon, retained introns (RI) and skipping exon (SE). There is significant enrichment of SE events for RNA (Fisher's test p-value = 3.22e-03) and Ribo-seq (p-value = 3.04e-04); and significant depletion of AL events for RNA (2.17e-13) and Ribo-seq (2.86e-18). **(d)** Proportions of

regulated microexons in RNA and Ribosome space for the comparison of glia and glioma samples in mouse. **(e)** Patterns of inclusion of differentially translated microexons in normal brain and hippocampus samples from GTEX, gliblastoma multiforme from TCGA and neuroblastoma from TARGET. The heatmap shows the Percent Spliced In (PSI) values for each sample. **(f)** Patterns of inclusion of conserved differentially translated microexons in normal brain and hippocampus samples from GTEX, gliblastoma multiforme (GBM) from TCGA and neuroblastoma (NB) from TARGET. The heatmap shows the Percent Spliced In (PSI) values for each sample.