

Description of Additional supplementary files

Supplementary Data 1: Output file of the RiboTISH pipeline run on the Hippocampal data. The columns correspond to the details of the translon identified.

Supplementary Data 2: Output file of the transcriptomic pipeline run on the Hippocampal data. Columns correspond to: transcript ID, gene name, strat and stop positions of the translon on the transcript, translon start codon, type of translon and whether it passed the FT filter.

Supplementary Data 3: Output file of the RiboTISH pipeline run on the Embryo data. The columns correspond to the details of the translon identified.

Supplementary Data 4: Output file of the transcriptomic pipeline run on the Embryo data. Columns correspond to: transcript ID, gene name, strat and stop positions of the translon on the transcript, translon start codon, type of translon and whether it passed the FT filter.

Supplementary Data 5: Table with the aggregated expression values of the identified translons across cell lines in Ribobase. The columns correspond to: transcript ID, gene name, strat and stop positions of the translon on the transcript, translon start codon, dataset from which the translon was identified followed by columns for each cell type found in Ribobase.

Supplementary Data 6: Table showing the translons with peptidomic evidence in Wang et al (NAR), and Budamgunta et al (Proteomics). The columns correspond to the translon containing gene name, Publication and translon type.

Supplementary Data 7: Output file from the MAGeCK analysis for the pooled KO screen.

Supplementary Data 8: Lists of the names and nucleotide sequences of guideRNA oligos cloned in plentiCrisprV2 plasmid, sequences of primers used for library amplification, sequences of primers used to check mutations in the Phgdh genomic locus.

Supplementary Data 9: Raw RiboNN predictions