

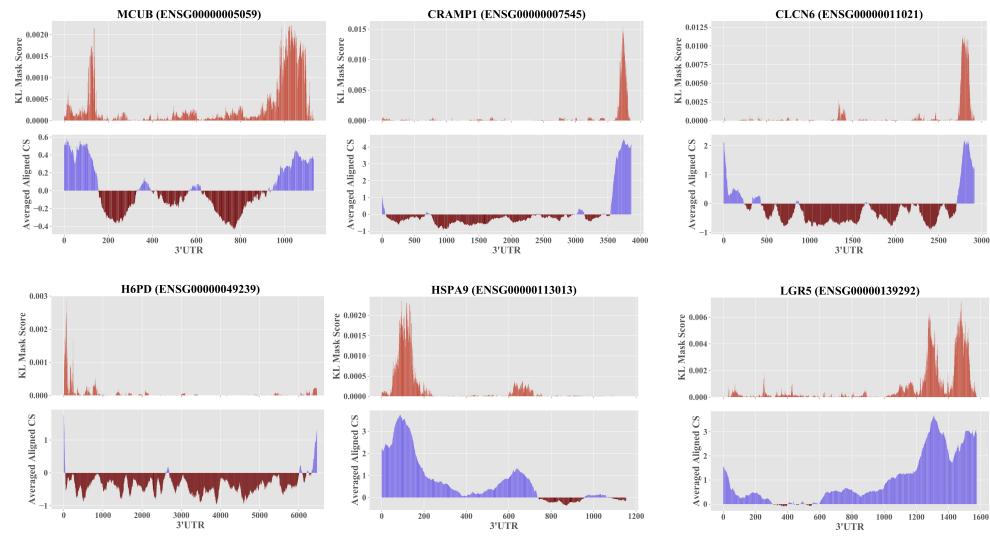
Supplementary Figures

Fig. S1. Alignments of KL scores and average conservation scores for 6 selected genes.

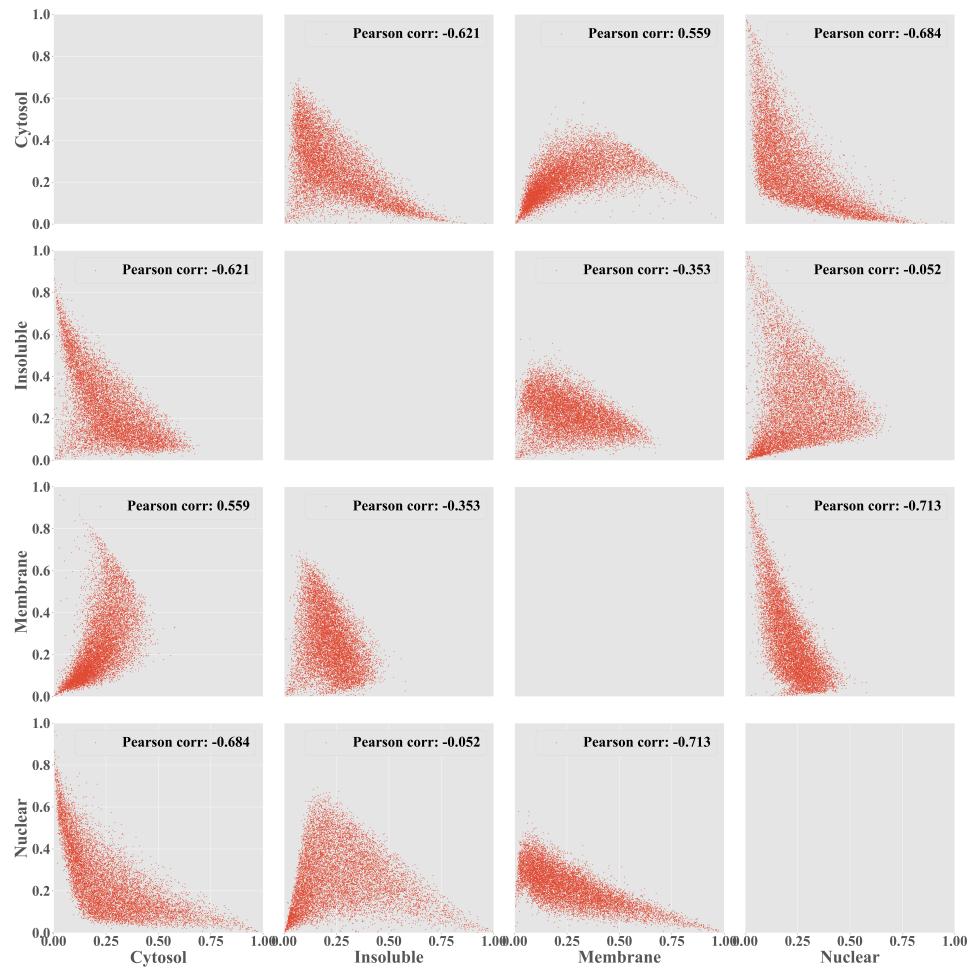


Fig. S2. Pairwise joint distribution of the normalized expression values between different subcellular fractions over all genes in the CeFra-Seq dataset.

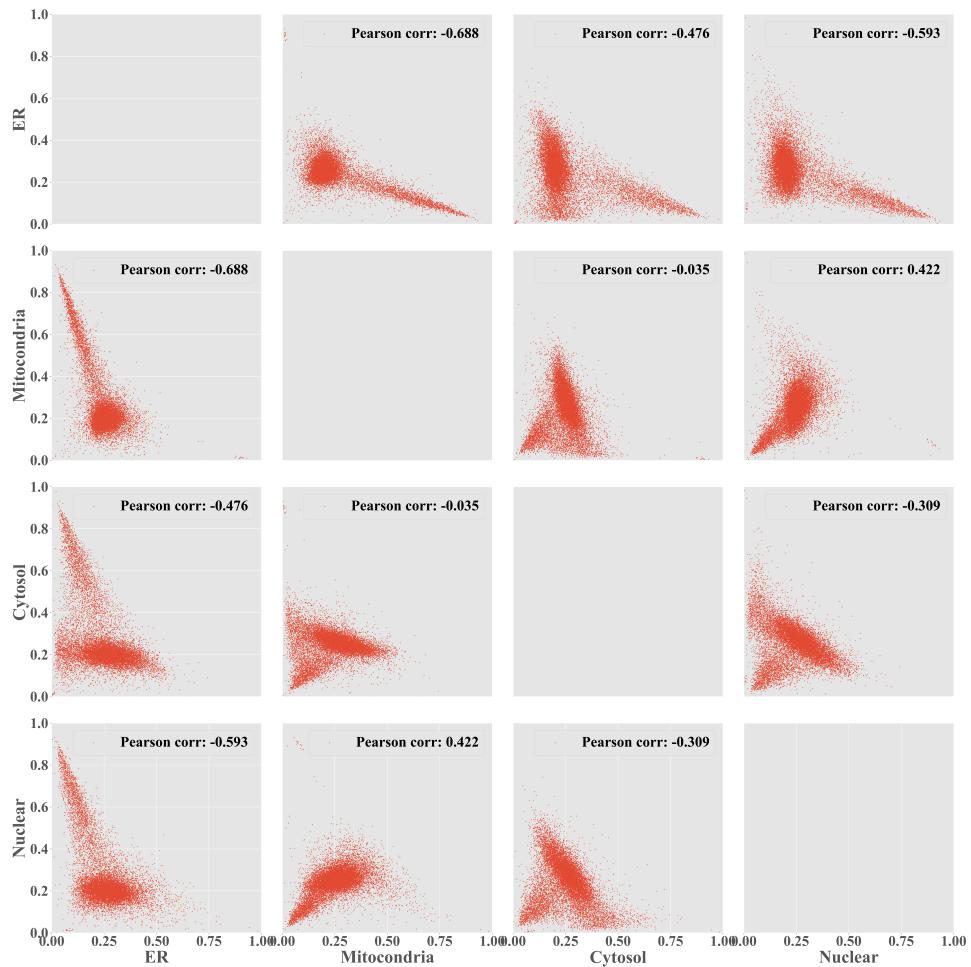


Fig. S3. Pairwise joint distribution of the normalized expression values between different subcellular fractions over all genes in the APEX-RIP dataset.

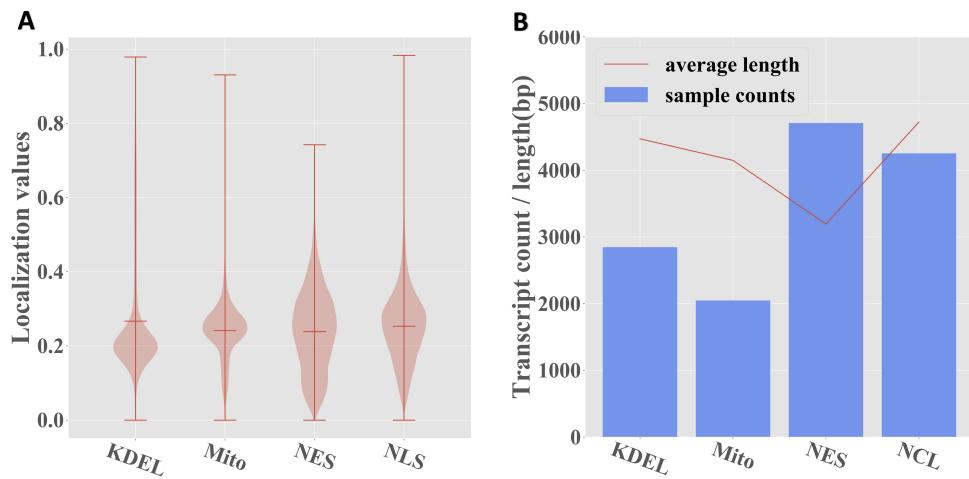


Fig. S4. Distribution of the localization values for the APEX-RIP dataset. *KDEL*: endoplasmic reticulum, *Mito*: mitochondria, *NES*: cytoplasm, *NLS*: nucleus.

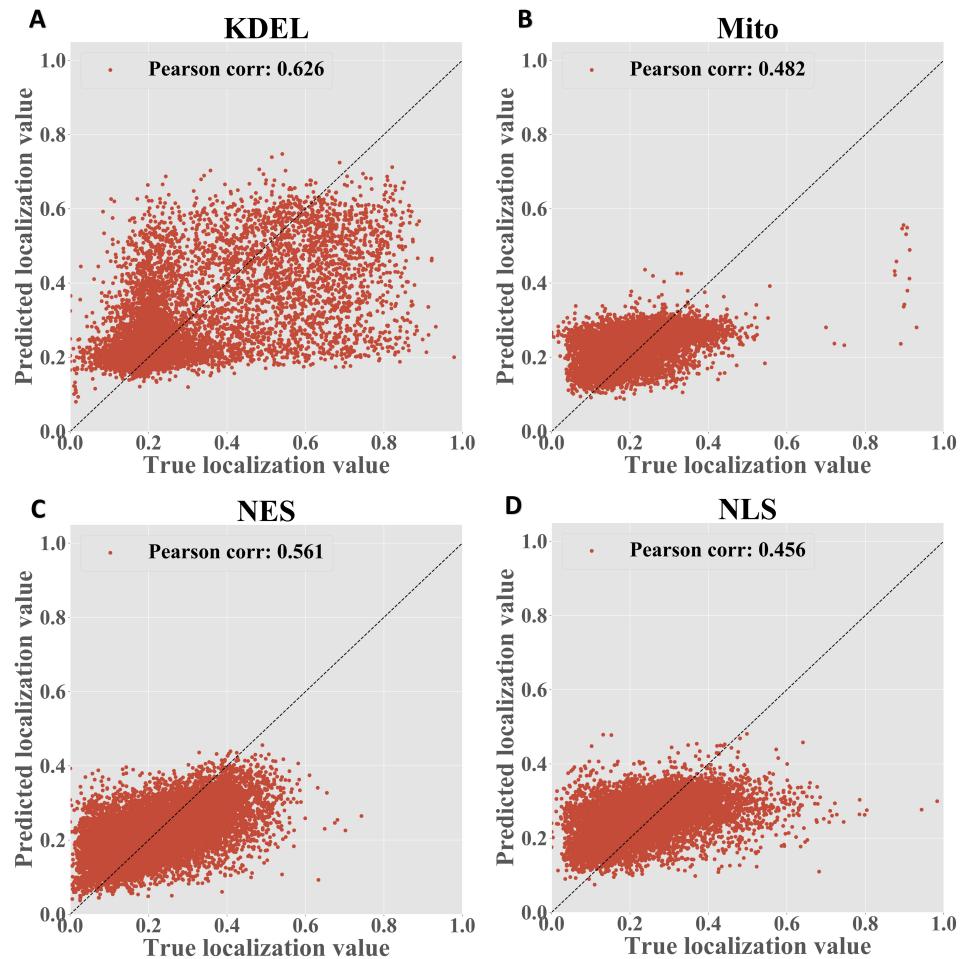


Fig. S5. Scatter plots for RNATracker applied to full-length sequences in the APEX-RIP dataset.

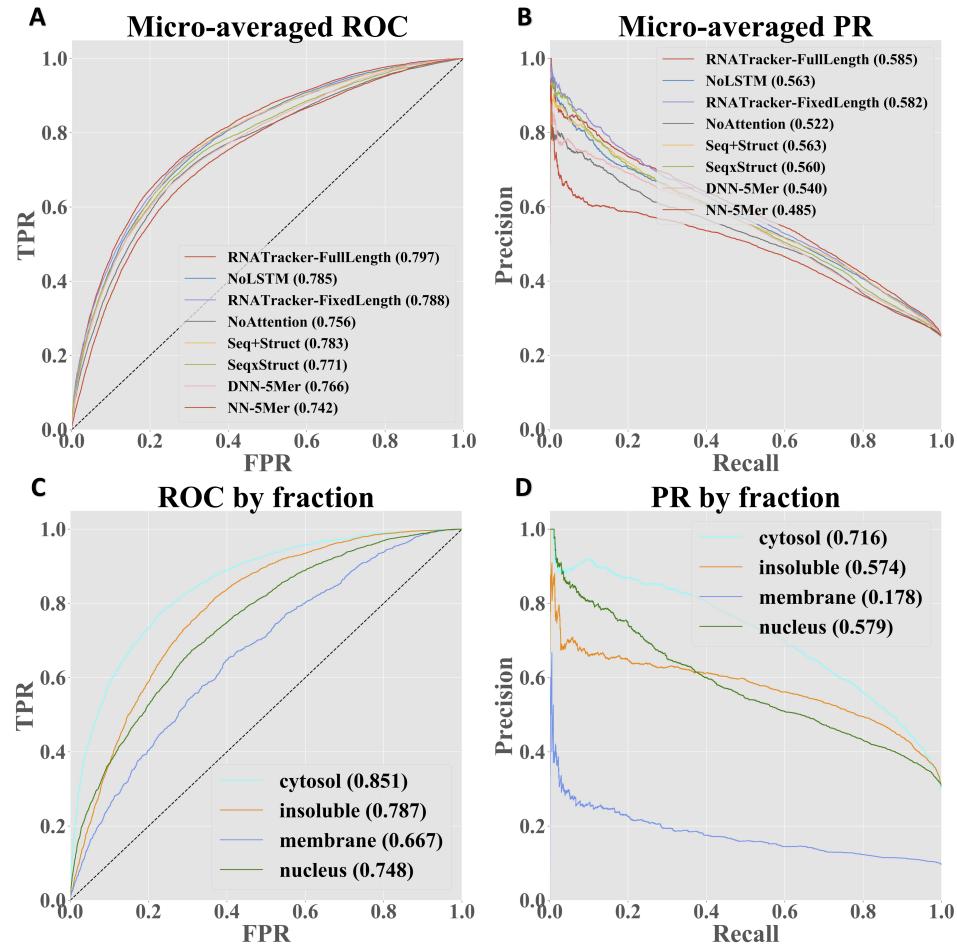


Fig. S6. Analysis of RNATracker variants evaluated on their ability to predict the predominant localization of a transcript. (A) and (B) present the micro-averaged ROC and PR curves for the four fractions of the different modalities of RNATracker and baselines. (C) and (D) ROC curves and PR curves for RNATracker_{Seq} for each fraction.

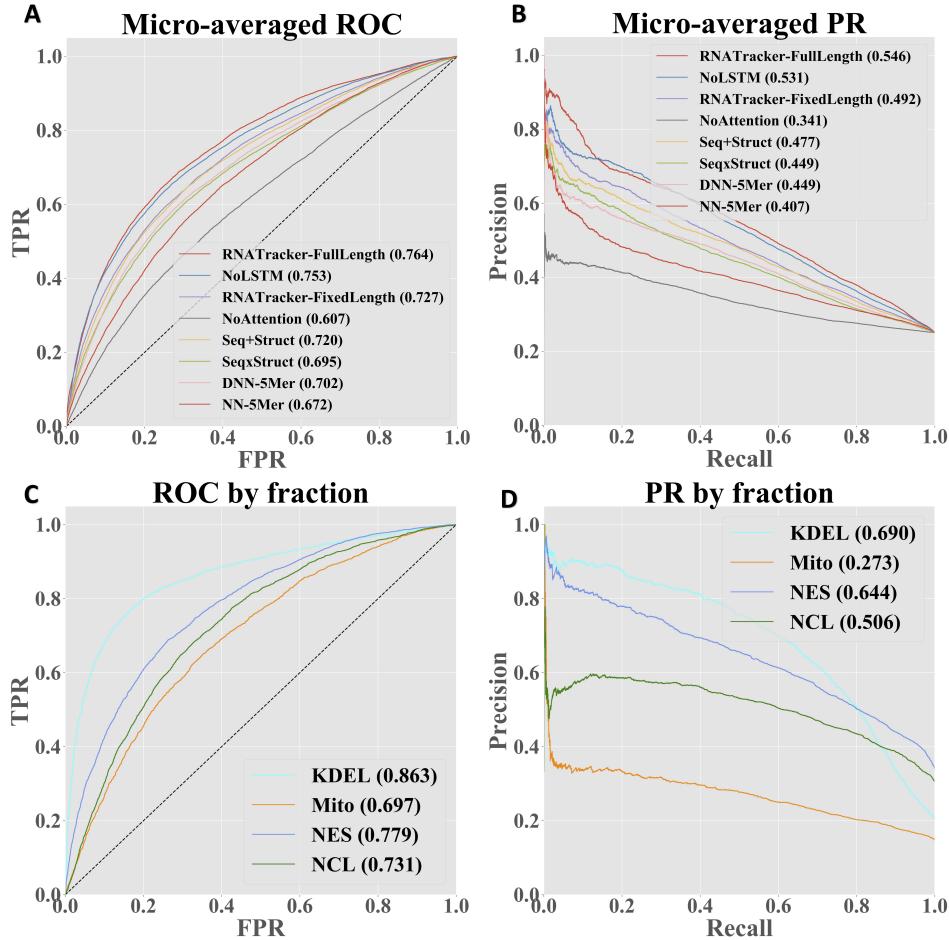


Fig. S7. (A) and (B) compared the micro-averaged ROC and PR curves of different RNATracker modalities and baselines. (C) and (D) presents ROC curve and PR curve on a fraction basis, for the best performing RNATracker model trained with full-length sequences.