

## **SUPPLEMENTARY TABLE LEGENDS**

### **Table S1**

This table shows the model parameters used when training DeepCLIP models on the curated dataset originally used by Maticzka et al, 2014.

### **Table S2**

This table shows performance metrics for the DeepCLIP models trained on the curated dataset originally used by Maticzka et al, 2014.

### **Table S3**

This table shows the raw data used in Figure 1 and Figure S2. The highest AUROC score per dataset is indicated in bold.

### **Table S4**

This table shows DeepCLIP analysis results of exonic point mutations taken from the dataset previously published by Raponi et al, 2011. During preparation of this manuscript we uncovered some errors in the original dataset, which we corrected. This table contains the raw data for Figure 2c-h.

### **Table S5**

This table contains information about the oligos used for SPRi and the model estimates produced by the software applications CLAMP and scrubber, as well as the DeepCLIP prediction scores. This table contains the raw data for Figure 5c-d and Figure S40.

## **SUPPLEMENTARY FIGURE LEGENDS**

### **Figure S1 | Conversion of the BLSTM output to binding profile and DeepCLIP training methodology.**

(a) The output of the forward LSTM layer is shown as a red arrow with a color intensity that increases gradually with the time-step number. The higher the color intensity and the higher the time-step number, the more contextual information has been available to the LSTM layer. The same is shown for the backward LSTM layer (blue arrow). The hidden state produced at the last time-step, , could contain information about the entire sequence. The outputs from the forward and backward LSTM layers are concatenated so hidden-states that are based on the base are combined. This figure is equal to from the BLSTM layer. Here, the time-step order is equal to the time-step order of the forward LSTM layer. (b) The contextual information that may present in BLSTM hidden-state that represents the highlighted cytosine (C in bold) is shown. The cytosine contains information about all the bases that surrounds it. (c) The binding profile is found via a summation of the vector elements in the BLSTM hidden-states. This results in a new vector that has a length equal to the input sequence. Every element of the vector indicates a class; positive value = class 1 (protein) and negative value = class 0 (genomic background). Every sequence has been divided by the vector element that has the highest absolute value. This results in a sequence that ranges from -1 to 1 and where it is easy to locate areas that are important for protein binding and areas that look like random genomic background. At the bottom, the prediction of the sequences is shown. (d) DeepCLIP is trained on input sequences belonging to either the bound class (assigned a score of 1) or the unbound (background) class (assigned a score of 0). The combined dataset is then divided into a training set, a validation set, and a test set. With default options they constitute 80%, 10%, and 10% of the dataset respectively. (e) When running in 10-fold cross-validation (CV) mode, the training set is divided into 10 different segmentations with a non-overlapping distribution of sequences in the validation and test set. These 10 different segmentations are then used to train 10 different models.

### **Figure S2 | Comparison of DeepCLIP AUROC performance on a benchmark dataset with existing tools.**

(a) Peak-sequences from the same input sequences used in Maticzka et al., 2014 were used to train DeepCLIP and iDeepS models using 10-fold cross-validation. Area under receiver operating characteristics (AUROC) for each protein were calculated using the combined predictions of all 10 models. AUROC for other methods were based on the reported scores from these studies. mDBN- indicates Deepnet with secondary structures, mDBN+ indicates Deepnet with secondary and tertiary structures. (b) AUROC metrics for the 12 proteins with DeepBind available were obtained by running DeepBind on the full dataset with all available models for the protein and using the model with the highest performance. All other values are identical to those in (a). Because deepnet with tertiary structures performed better than with just secondary structures, we show only this variant in this plot. (c) Boxplot of AUROC scores from each method for the 12 datasets. DeepBind produced an AUROC score of 0.319 for FUS, this datapoint is outside the plotting areas of both (b) and (c) plots.

### **Figure S3 | DeepCLIP AUROC performance by CLIP method and dataset size.**

(a) AUROC metrics for DeepCLIP grouped into CLIP methods. Significance of pairwise differences was calculated using Wilcoxon rank sum tests. (b) Correlation plot of DeepCLIP AUROC measures and log<sub>10</sub> of the number of input binding sites. Correlation was estimated using Spearman's rank correlation rho. 95% confidence interval of the linear model (red line) is shown in grey.

**Figure S4 | Pseudo-PFMs captured by DeepCLIP.** The column **Protein** contains the names of the analyzed proteins, which have binding preferences as described in the column **Known binding preference**. References to literature can be seen in column **Source**. In the outermost right column, **Pseudo-PFM**, motifs derived from DeepCLIPs convolutional nodes that correlates with known binding preferences are shown. H: A, C or U; N: A, C, G or U.

**Figure S5 | DeepCLIP model characteristics for AGO1-4 GraphProt dataset.** (a) Area under curve analysis of DeepCLIP models trained using 10-fold cross-validation. (b) Visualization of the CNN filters learned by the best performing model based on AUC. Score is equal to the mean information per base. (c-h) Visualizations of the combined model predictions of the 10-fold cross-validation. Scores are scaled to 0-100. (c) Density of background and bound prediction scores. (d) Combined density of prediction scores. (e) Cumulative predictive score of background and bound input sequences. (f) Barplot of cumulative scores of all input sequences. (g) Split prediction scores of background and bound input sequences. (h) Combined AUROC analysis using the pROC R package with DeLong estimation of 95% confidence interval.

**Figure S6 | DeepCLIP model characteristics for AGO2 GraphProt dataset.** (a) Area under curve analysis of DeepCLIP models trained using 10-fold cross-validation. (b) Visualization of the CNN filters learned by the best performing model based on AUC. Score is equal to the mean information per base. (c-h) Visualizations of the combined model predictions of the 10-fold cross-validation. Scores are scaled to 0-100. (c) Density of background and bound prediction scores. (d) Combined density of prediction scores. (e) Cumulative predictive score of background and bound input sequences. (f) Barplot of cumulative scores of all input sequences. (g) Split prediction scores of background and bound input sequences. (h) Combined AUROC analysis using the pROC R package with DeLong estimation of 95% confidence interval.

**Figure S7 | DeepCLIP model characteristics for ALKBH5 GraphProt dataset.** (a) Area under curve analysis of DeepCLIP models trained using 10-fold cross-validation. (b) Visualization of the CNN filters learned by the best performing model based on AUC. Score is equal to the mean information per base. (c-h) Visualizations of the combined model predictions of the 10-fold cross-validation. Scores are scaled to 0-100. (c) Density of background and bound prediction scores. (d) Combined density of prediction scores. (e) Cumulative predictive score of background and bound input sequences. (f) Barplot of cumulative scores of all input sequences. (g) Split prediction scores of background and bound input sequences. (h) Combined AUROC analysis using the pROC R package with DeLong estimation of 95% confidence interval.

**Figure S8 | DeepCLIP model characteristics for C17ORF85 GraphProt dataset.** (a) Area under curve analysis of DeepCLIP models trained using 10-fold cross-validation. (b) Visualization of the CNN filters learned by the best performing model based on AUC. Score is equal to the mean information per base. (c-h) Visualizations of the combined model predictions of the 10-fold cross-validation. Scores are scaled to 0-100. (c) Density of background and bound prediction scores. (d) Combined density of prediction scores. (e) Cumulative predictive score of background and bound input sequences. (f) Barplot of cumulative scores of all input sequences. (g) Split prediction scores of background and bound input sequences. (h) Combined AUROC analysis using the pROC R package with DeLong estimation of 95% confidence interval.

**Figure S9 | DeepCLIP model characteristics for C22ORF28 GraphProt dataset.** (a) Area under curve analysis of DeepCLIP models trained using 10-fold cross-validation. (b) Visualization of the CNN filters learned by the best performing model based on AUC. Score is equal to the mean information per base. (c-h) Visualizations of the combined model predictions of the 10-fold cross-validation. Scores are scaled to 0-100.

(c) Density of background and bound prediction scores. (d) Combined density of prediction scores. (e) Cumulative predictive score of background and bound input sequences. (f) Barplot of cumulative scores of all input sequences. (g) Split prediction scores of background and bound input sequences. (h) Combined AUROC analysis using the pROC R package with DeLong estimation of 95% confidence interval.

**Figure S10 | DeepCLIP model characteristics for CAPRIN1 GraphProt dataset.** (a) Area under curve analysis of DeepCLIP models trained using 10-fold cross-validation. (b) Visualization of the CNN filters learned by the best performing model based on AUC. Score is equal to the mean information per base. (c-h) Visualizations of the combined model predictions of the 10-fold cross-validation. Scores are scaled to 0-100. (c) Density of background and bound prediction scores. (d) Combined density of prediction scores. (e) Cumulative predictive score of background and bound input sequences. (f) Barplot of cumulative scores of all input sequences. (g) Split prediction scores of background and bound input sequences. (h) Combined AUROC analysis using the pROC R package with DeLong estimation of 95% confidence interval.

**Figure S11 | DeepCLIP model characteristics for ELAVL1 GraphProt dataset.** (a) Area under curve analysis of DeepCLIP models trained using 10-fold cross-validation. (b) Visualization of the CNN filters learned by the best performing model based on AUC. Score is equal to the mean information per base. (c-h) Visualizations of the combined model predictions of the 10-fold cross-validation. Scores are scaled to 0-100. (c) Density of background and bound prediction scores. (d) Combined density of prediction scores. (e) Cumulative predictive score of background and bound input sequences. (f) Barplot of cumulative scores of all input sequences. (g) Split prediction scores of background and bound input sequences. (h) Combined AUROC analysis using the pROC R package with DeLong estimation of 95% confidence interval.

**Figure S12 | DeepCLIP model characteristics for ELAVL1A GraphProt dataset.** (a) Area under curve analysis of DeepCLIP models trained using 10-fold cross-validation. (b) Visualization of the CNN filters learned by the best performing model based on AUC. Score is equal to the mean information per base. (c-h) Visualizations of the combined model predictions of the 10-fold cross-validation. Scores are scaled to 0-100. (c) Density of background and bound prediction scores. (d) Combined density of prediction scores. (e) Cumulative predictive score of background and bound input sequences. (f) Barplot of cumulative scores of all input sequences. (g) Split prediction scores of background and bound input sequences. (h) Combined AUROC analysis using the pROC R package with DeLong estimation of 95% confidence interval.

**Figure S13 | DeepCLIP model characteristics for ELAVL1B GraphProt dataset.** (a) Area under curve analysis of DeepCLIP models trained using 10-fold cross-validation. (b) Visualization of the CNN filters learned by the best performing model based on AUC. Score is equal to the mean information per base. (c-h) Visualizations of the combined model predictions of the 10-fold cross-validation. Scores are scaled to 0-100. (c) Density of background and bound prediction scores. (d) Combined density of prediction scores. (e) Cumulative predictive score of background and bound input sequences. (f) Barplot of cumulative scores of all input sequences. (g) Split prediction scores of background and bound input sequences. (h) Combined AUROC analysis using the pROC R package with DeLong estimation of 95% confidence interval.

**Figure S14 | DeepCLIP model characteristics for EWSR1 GraphProt dataset.** (a) Area under curve analysis of DeepCLIP models trained using 10-fold cross-validation. (b) Visualization of the CNN filters learned by the best performing model based on AUC. Score is equal to the mean information per base. (c-h) Visualizations of the combined model predictions of the 10-fold cross-validation. Scores are scaled to 0-100. (c) Density of background and bound prediction scores. (d) Combined density of prediction scores. (e)

Cumulative predictive score of background and bound input sequences. (f) Barplot of cumulative scores of all input sequences. (g) Split prediction scores of background and bound input sequences. (h) Combined AUROC analysis using the pROC R package with DeLong estimation of 95% confidence interval.

**Figure S15 | DeepCLIP model characteristics for FUS GraphProt dataset.** (a) Area under curve analysis of DeepCLIP models trained using 10-fold cross-validation. (b) Visualization of the CNN filters learned by the best performing model based on AUC. Score is equal to the mean information per base. (c-h) Visualizations of the combined model predictions of the 10-fold cross-validation. Scores are scaled to 0-100. (c) Density of background and bound prediction scores. (d) Combined density of prediction scores. (e) Cumulative predictive score of background and bound input sequences. (f) Barplot of cumulative scores of all input sequences. (g) Split prediction scores of background and bound input sequences. (h) Combined AUROC analysis using the pROC R package with DeLong estimation of 95% confidence interval.

**Figure S16 | DeepCLIP model characteristics for hnRNP C GraphProt dataset.** (a) Area under curve analysis of DeepCLIP models trained using 10-fold cross-validation. (b) Visualization of the CNN filters learned by the best performing model based on AUC. Score is equal to the mean information per base. (c-h) Visualizations of the combined model predictions of the 10-fold cross-validation. Scores are scaled to 0-100. (c) Density of background and bound prediction scores. (d) Combined density of prediction scores. (e) Cumulative predictive score of background and bound input sequences. (f) Barplot of cumulative scores of all input sequences. (g) Split prediction scores of background and bound input sequences. (h) Combined AUROC analysis using the pROC R package with DeLong estimation of 95% confidence interval.

**Figure S17 | DeepCLIP model characteristics for HuR GraphProt dataset.** (a) Area under curve analysis of DeepCLIP models trained using 10-fold cross-validation. (b) Visualization of the CNN filters learned by the best performing model based on AUC. Score is equal to the mean information per base. (c-h) Visualizations of the combined model predictions of the 10-fold cross-validation. Scores are scaled to 0-100. (c) Density of background and bound prediction scores. (d) Combined density of prediction scores. (e) Cumulative predictive score of background and bound input sequences. (f) Barplot of cumulative scores of all input sequences. (g) Split prediction scores of background and bound input sequences. (h) Combined AUROC analysis using the pROC R package with DeLong estimation of 95% confidence interval.

**Figure S18 | DeepCLIP model characteristics for IGF2BP1-3 GraphProt dataset.** (a) Area under curve analysis of DeepCLIP models trained using 10-fold cross-validation. (b) Visualization of the CNN filters learned by the best performing model based on AUC. Score is equal to the mean information per base. (c-h) Visualizations of the combined model predictions of the 10-fold cross-validation. Scores are scaled to 0-100. (c) Density of background and bound prediction scores. (d) Combined density of prediction scores. (e) Cumulative predictive score of background and bound input sequences. (f) Barplot of cumulative scores of all input sequences. (g) Split prediction scores of background and bound input sequences. (h) Combined AUROC analysis using the pROC R package with DeLong estimation of 95% confidence interval.

**Figure S19 | DeepCLIP model characteristics for MOV10 GraphProt dataset.** (a) Area under curve analysis of DeepCLIP models trained using 10-fold cross-validation. (b) Visualization of the CNN filters learned by the best performing model based on AUC. Score is equal to the mean information per base. (c-h) Visualizations of the combined model predictions of the 10-fold cross-validation. Scores are scaled to 0-100. (c) Density of background and bound prediction scores. (d) Combined density of prediction scores. (e) Cumulative predictive score of background and bound input sequences. (f) Barplot of cumulative scores of

all input sequences. (g) Split prediction scores of background and bound input sequences. (h) Combined AUROC analysis using the pROC R package with DeLong estimation of 95% confidence interval.

**Figure S20 | DeepCLIP model characteristics for PTBP1 GraphProt dataset.** (a) Area under curve analysis of DeepCLIP models trained using 10-fold cross-validation. (b) Visualization of the CNN filters learned by the best performing model based on AUC. Score is equal to the mean information per base. (c-h) Visualizations of the combined model predictions of the 10-fold cross-validation. Scores are scaled to 0-100. (c) Density of background and bound prediction scores. (d) Combined density of prediction scores. (e) Cumulative predictive score of background and bound input sequences. (f) Barplot of cumulative scores of all input sequences. (g) Split prediction scores of background and bound input sequences. (h) Combined AUROC analysis using the pROC R package with DeLong estimation of 95% confidence interval.

**Figure S21 | DeepCLIP model characteristics for PUM2 GraphProt dataset.** (a) Area under curve analysis of DeepCLIP models trained using 10-fold cross-validation. (b) Visualization of the CNN filters learned by the best performing model based on AUC. Score is equal to the mean information per base. (c-h) Visualizations of the combined model predictions of the 10-fold cross-validation. Scores are scaled to 0-100. (c) Density of background and bound prediction scores. (d) Combined density of prediction scores. (e) Cumulative predictive score of background and bound input sequences. (f) Barplot of cumulative scores of all input sequences. (g) Split prediction scores of background and bound input sequences. (h) Combined AUROC analysis using the pROC R package with DeLong estimation of 95% confidence interval.

**Figure S22 | DeepCLIP model characteristics for QKI GraphProt dataset.** (a) Area under curve analysis of DeepCLIP models trained using 10-fold cross-validation. (b) Visualization of the CNN filters learned by the best performing model based on AUC. Score is equal to the mean information per base. (c-h) Visualizations of the combined model predictions of the 10-fold cross-validation. Scores are scaled to 0-100. (c) Density of background and bound prediction scores. (d) Combined density of prediction scores. (e) Cumulative predictive score of background and bound input sequences. (f) Barplot of cumulative scores of all input sequences. (g) Split prediction scores of background and bound input sequences. (h) Combined AUROC analysis using the pROC R package with DeLong estimation of 95% confidence interval.

**Figure S23 | DeepCLIP model characteristics for SRSF1 GraphProt dataset.** (a) Area under curve analysis of DeepCLIP models trained using 10-fold cross-validation. (b) Visualization of the CNN filters learned by the best performing model based on AUC. Score is equal to the mean information per base. (c-h) Visualizations of the combined model predictions of the 10-fold cross-validation. Scores are scaled to 0-100. (c) Density of background and bound prediction scores. (d) Combined density of prediction scores. (e) Cumulative predictive score of background and bound input sequences. (f) Barplot of cumulative scores of all input sequences. (g) Split prediction scores of background and bound input sequences. (h) Combined AUROC analysis using the pROC R package with DeLong estimation of 95% confidence interval.

**Figure S24 | DeepCLIP model characteristics for TAF15 GraphProt dataset.** (a) Area under curve analysis of DeepCLIP models trained using 10-fold cross-validation. (b) Visualization of the CNN filters learned by the best performing model based on AUC. Score is equal to the mean information per base. (c-h) Visualizations of the combined model predictions of the 10-fold cross-validation. Scores are scaled to 0-100. (c) Density of background and bound prediction scores. (d) Combined density of prediction scores. (e) Cumulative predictive score of background and bound input sequences. (f) Barplot of cumulative scores of

all input sequences. (g) Split prediction scores of background and bound input sequences. (h) Combined AUROC analysis using the pROC R package with DeLong estimation of 95% confidence interval.

**Figure S25 | DeepCLIP model characteristics for TDP-43 GraphProt dataset.** (a) Area under curve analysis of DeepCLIP models trained using 10-fold cross-validation. (b) Visualization of the CNN filters learned by the best performing model based on AUC. Score is equal to the mean information per base. (c-h) Visualizations of the combined model predictions of the 10-fold cross-validation. Scores are scaled to 0-100. (c) Density of background and bound prediction scores. (d) Combined density of prediction scores. (e) Cumulative predictive score of background and bound input sequences. (f) Barplot of cumulative scores of all input sequences. (g) Split prediction scores of background and bound input sequences. (h) Combined AUROC analysis using the pROC R package with DeLong estimation of 95% confidence interval.

**Figure S26 | DeepCLIP model characteristics for TIA1 GraphProt dataset.** (a) Area under curve analysis of DeepCLIP models trained using 10-fold cross-validation. (b) Visualization of the CNN filters learned by the best performing model based on AUC. Score is equal to the mean information per base. (c-h) Visualizations of the combined model predictions of the 10-fold cross-validation. Scores are scaled to 0-100. (c) Density of background and bound prediction scores. (d) Combined density of prediction scores. (e) Cumulative predictive score of background and bound input sequences. (f) Barplot of cumulative scores of all input sequences. (g) Split prediction scores of background and bound input sequences. (h) Combined AUROC analysis using the pROC R package with DeLong estimation of 95% confidence interval.

**Figure S27 | DeepCLIP model characteristics for TIAL1 GraphProt dataset.** (a) Area under curve analysis of DeepCLIP models trained using 10-fold cross-validation. (b) Visualization of the CNN filters learned by the best performing model based on AUC. Score is equal to the mean information per base. (c-h) Visualizations of the combined model predictions of the 10-fold cross-validation. Scores are scaled to 0-100. (c) Density of background and bound prediction scores. (d) Combined density of prediction scores. (e) Cumulative predictive score of background and bound input sequences. (f) Barplot of cumulative scores of all input sequences. (g) Split prediction scores of background and bound input sequences. (h) Combined AUROC analysis using the pROC R package with DeLong estimation of 95% confidence interval.

**Figure S28 | DeepCLIP model characteristics for ZC3H7B GraphProt dataset.** (a) Area under curve analysis of DeepCLIP models trained using 10-fold cross-validation. (b) Visualization of the CNN filters learned by the best performing model based on AUC. Score is equal to the mean information per base. (c-h) Visualizations of the combined model predictions of the 10-fold cross-validation. Scores are scaled to 0-100. (c) Density of background and bound prediction scores. (d) Combined density of prediction scores. (e) Cumulative predictive score of background and bound input sequences. (f) Barplot of cumulative scores of all input sequences. (g) Split prediction scores of background and bound input sequences. (h) Combined AUROC analysis using the pROC R package with DeLong estimation of 95% confidence interval.

**Figure S29 | DeepCLIP model characteristics for hnRNP A1.** (a) Area under curve analysis of DeepCLIP models trained on hnRNP A1 iCLIP data from Bruun et al. using 10-fold cross-validation. (b) Visualization of the CNN filters learned by the best performing model based on AUC. Score is equal to the mean information per base. (c-h) Visualizations of the combined model predictions of the 10-fold cross-validation. Scores are scaled to 0-100. (c) Density of background and bound prediction scores. (d) Combined density of prediction scores. (e) Cumulative predictive score of background and bound input sequences. (f) Barplot of cumulative scores of all input sequences. (g) Split prediction scores of background and bound input

sequences. (h) Combined AUROC analysis using the pROC R package with DeLong estimation of 95% confidence interval.

**Figure S30 | DeepCLIP model characteristics for SRSF1 (eCLIP).** (a) Area under curve analysis of DeepCLIP models trained on SRSF1 eCLIP data from Van Nostrand et al using 10-fold cross-validation. (b) Visualization of the CNN filters learned by the best performing model based on AUC. Score is equal to the mean information per base. (c-h) Visualizations of the combined model predictions of the 10-fold cross-validation. Scores are scaled to 0-100. (c) Density of background and bound prediction scores. (d) Combined density of prediction scores. (e) Cumulative predictive score of background and bound input sequences. (f) Barplot of cumulative scores of all input sequences. (g) Split prediction scores of background and bound input sequences. (h) Combined AUROC analysis using the pROC R package with DeLong estimation of 95% confidence interval.

**Figure S31 | DeepCLIP predictions of *ACADM* exon 5 mutations using SRSF1 (HITS-CLIP) model.** (a) Scatter plot of *ACADM* exon 5 minigene percent spliced in (PSI) values and DeepCLIP SRSF1 score with linear regression (red line) and 95% confidence interval (shaded area). (b) Same as (a) but with the DeepCLIP hnRNP A1 score subtracted from the SRSF1 score. Spearman's rho is indicated in red for both plots in (a-b). (c) Barplot of the change relative to the wt of exon inclusion levels (blue), hnRNP A1 DeepCLIP scores (brown) and SRSF1 HITS-CLIP DeepCLIP model scores (green). The hnRNP A1 results are the same as presented in figure 5 but are included here for reference.

**Figure S32 | Correlation between observed exon skipping and exon skipping model predictions.** (a) Scatter plot of exon inclusion levels vs EX-SKIP predictions. (b) Scatter plot of the exon inclusions levels vs the SPANR prediction exon inclusion levels. Spearman's rho is indicated in red for both plots in (a-b).

**Figure S33 | DeepCLIP model characteristics for TDP-43.** (a) Area under curve analysis of DeepCLIP models trained on TDP43 binding sites from the POSTAR2 database using 10-fold cross-validation. (b) Visualization of the CNN filters learned by the best performing model based on AUC. Score is equal to the mean information per base. (c-h) Visualizations of the combined model predictions of the 10-fold cross-validation. Scores are scaled to 0-100. (c) Density of background and bound prediction scores. (d) Combined density of prediction scores. (e) Cumulative predictive score of background and bound input sequences. (f) Barplot of cumulative scores of all input sequences. (g) Split prediction scores of background and bound input sequences. (h) Combined AUROC analysis using the pROC R package with DeLong estimation of 95% confidence interval.

**Figure S34-36 | Results of SPRi measurements of hnRNP A1 binding to a set of RNA wt and mutant oligos.** Plots showing measured response (RU) versus time of the wt oligo (left) and mutant (right). The combined model fit across concentrations is indicated in black and different concentrations are indicated green, brown, purple, pink, and olive in decreasing order. The fitted model's maximum simulated value is indicated above each plot.

**Figure S37-39 | Results of SPRi measurements of SRSF1 binding to a set of RNA wt and mutant oligos.** Plots showing measured response (RU) versus time of the wt oligo (left) and mutant (right). The combined model fit across concentrations is indicated in black and different concentrations are indicated



green, brown, purple, pink, and olive in decreasing order. The fitted model's maximum simulated value is indicated above each plot.

**Figure S40 | Correlation between observed *in vitro* binding and DeepCLIP hnRNP A1 and SRSF1 (GraphProt dataset) model predictions.** (a) Scatter plot of Rmax binding values from SPRi measurements of hnRNP A1 binding to oligos vs the predicted DeepCLIP hnRNP A1 scores. (b) Scatter plot of Rmax binding values from SPRi measurements of SRSF1 binding to oligos vs the predicted DeepCLIP SRSF1 scores. Spearman's rho is indicated in red in lower right corner and p-value in black in upper left corner for both plots in (a-b). The data for hnRNP A1 in this plot is the same as in Figure 3 and is included here for comparison.

**Figure S41 | DeepCLIP predicted binding at TDP-43 repressed pseudoexons.** (a-c) Heatmap of TDP-43 DeepCLIP binding profiles at the 3'ss and 5'ss regions of TDP-43 repressed pseudoexons in mice. In all heatmaps the plots show 25 nt into the exon indicated by blue bars at the top, and the 50 first and last nt of the introns. The heatmap in (a) shows pseudoexons up-regulated upon TDP-43 depletion in at least two of the examined tissues (stem cells, neurons, and muscle cells), while the heatmap in (b) show muscle specific pseudoexons and the heatmap in (c) shows neuron specific pseudoexons.

**Fig. S1**

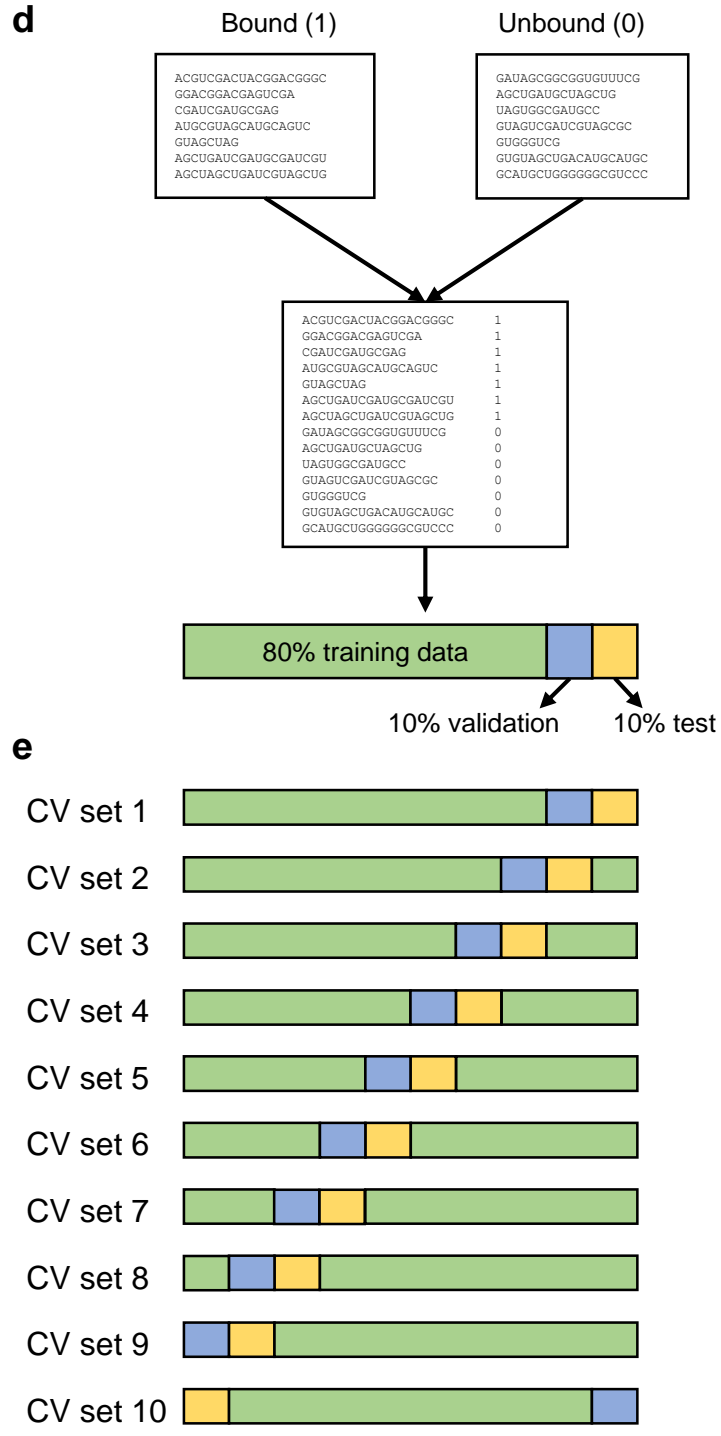
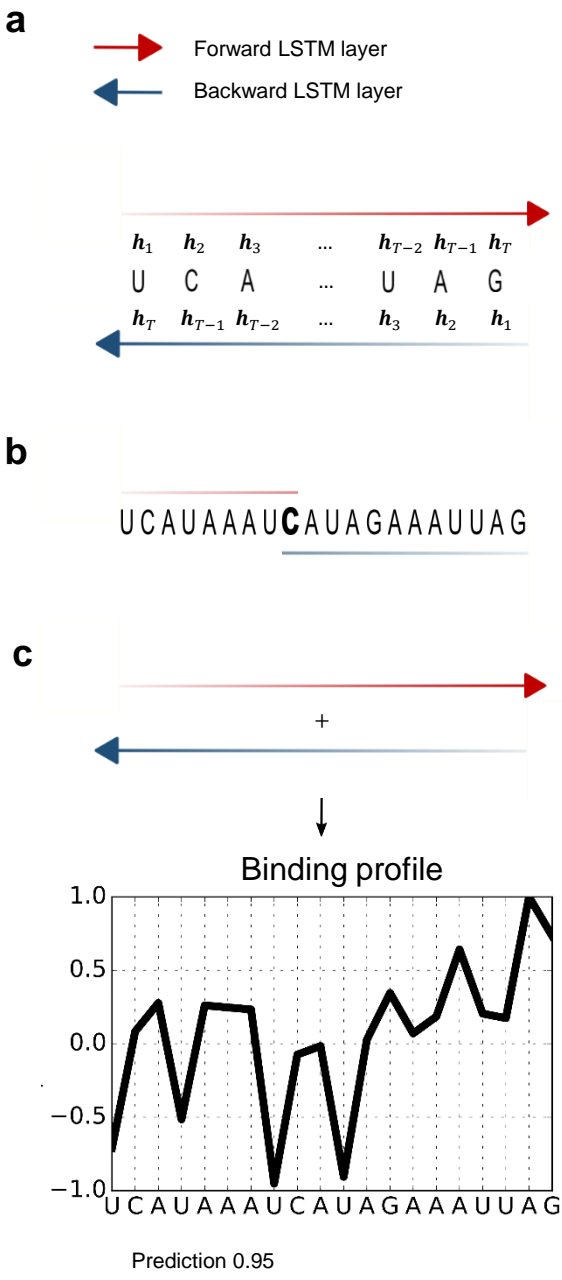
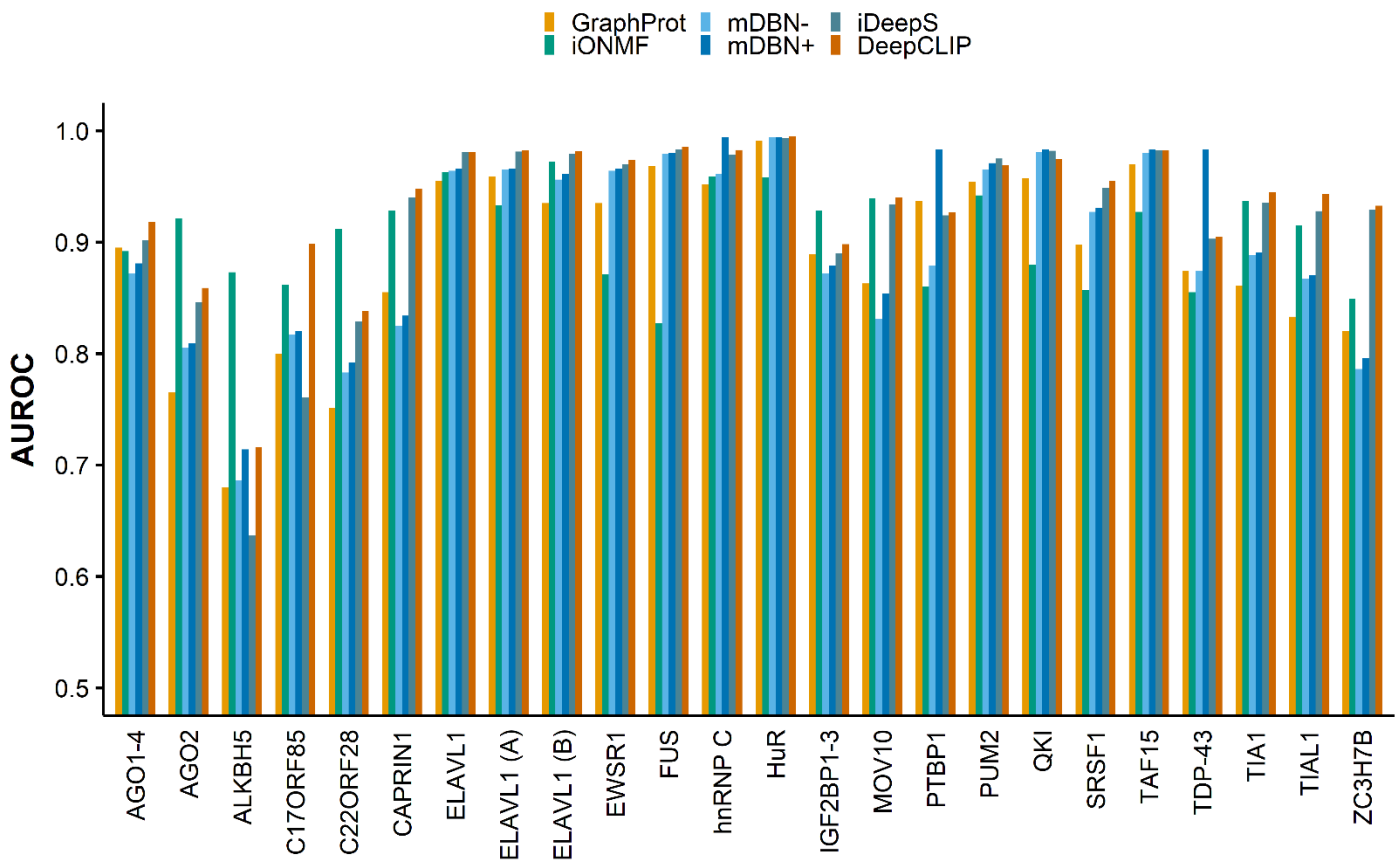
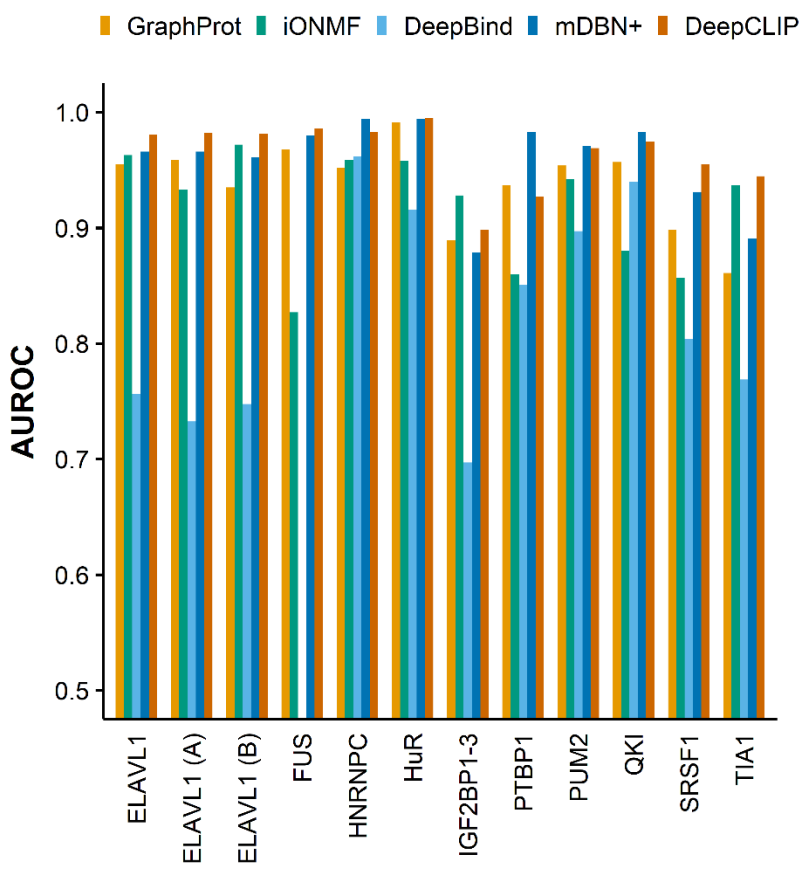


Fig. S2

**a**



**b**



**c**

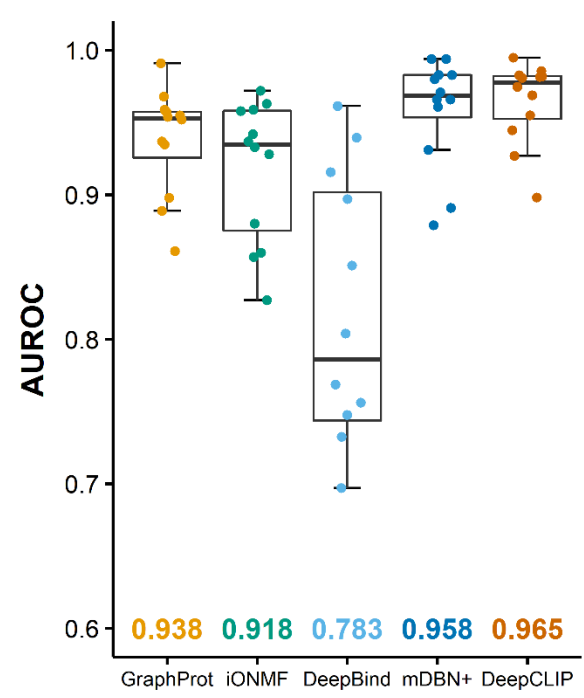


Fig. S3

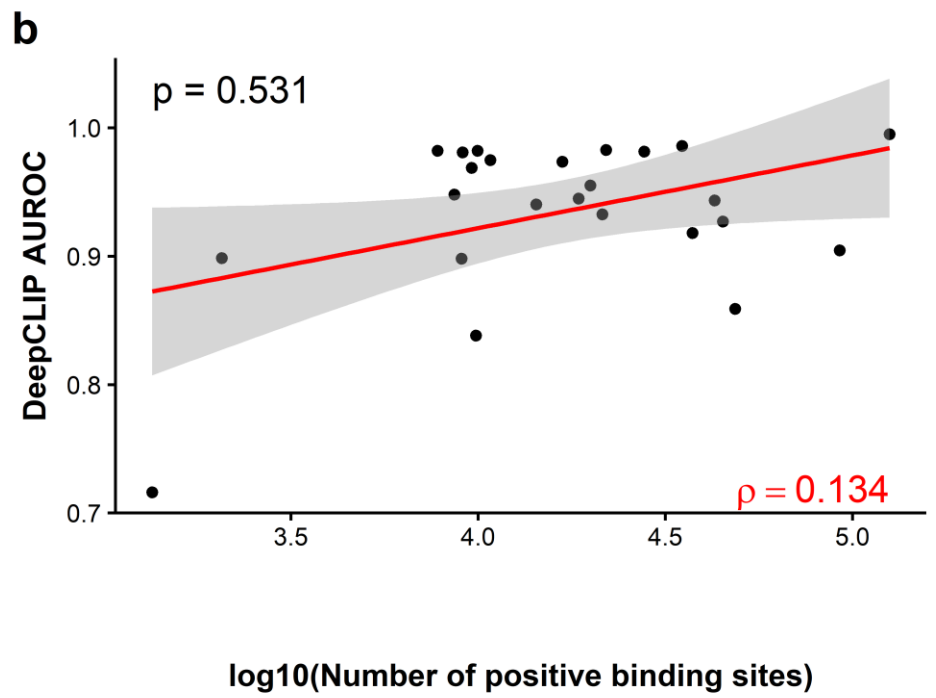
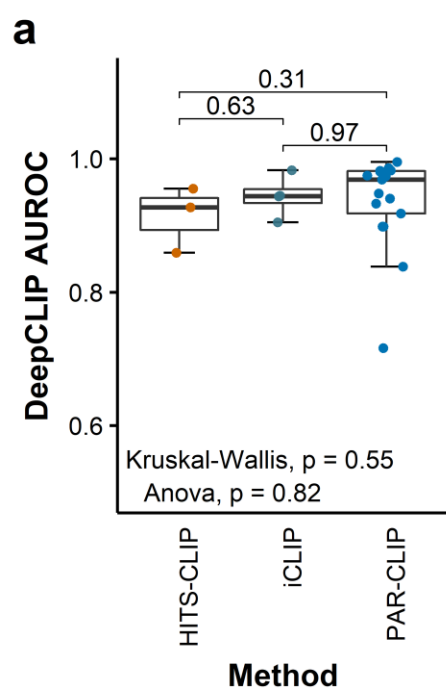
















































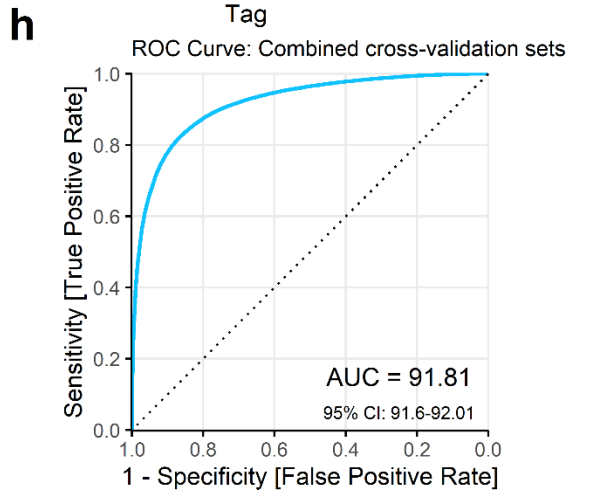
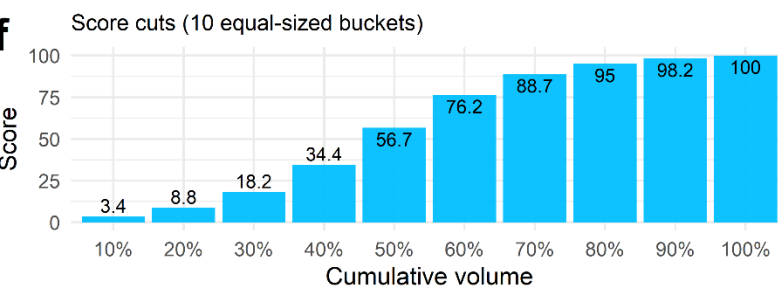
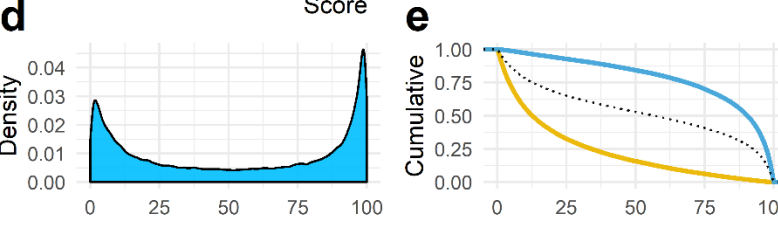
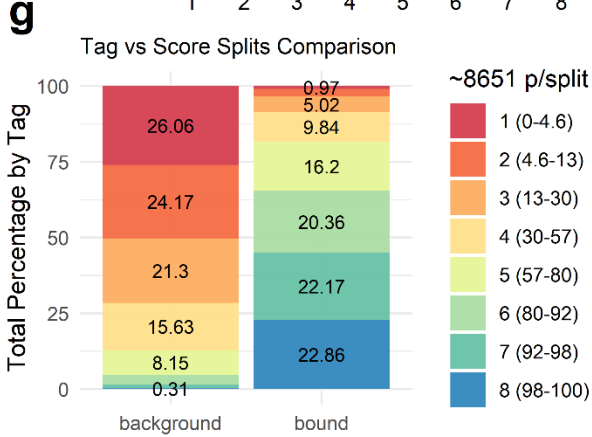
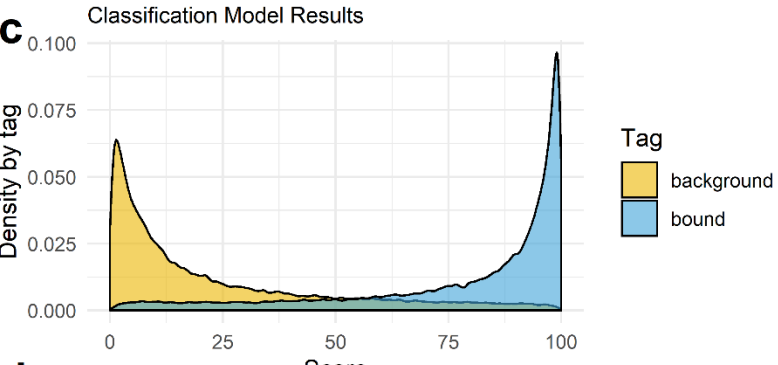
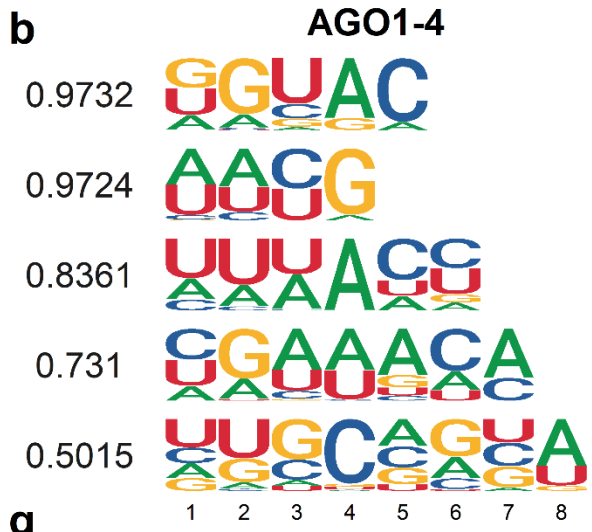
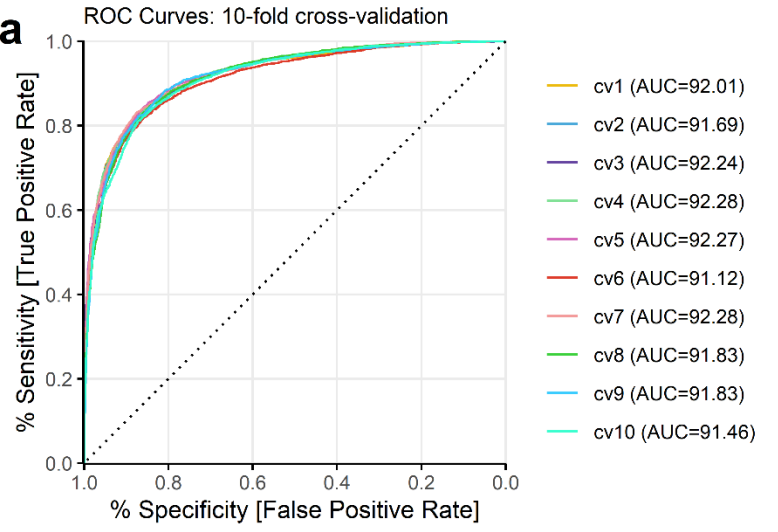


Fig. S4

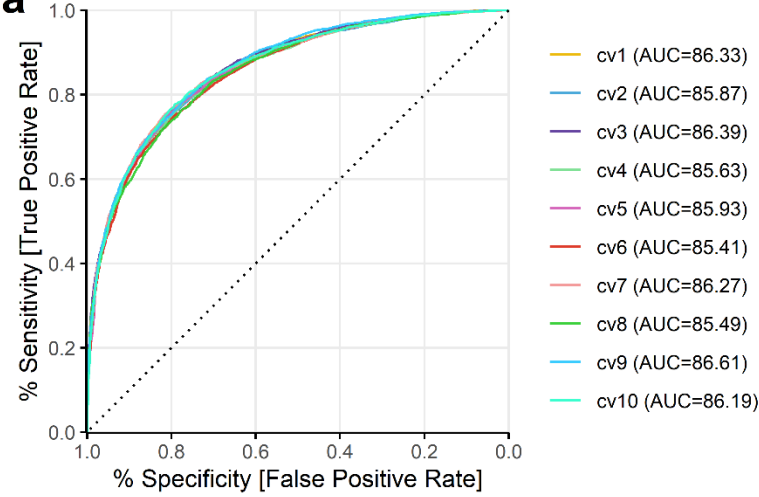
| Protein (dataset) | Known binding preference                                | Add. binding preference source | CLIP data source       | CLIP method | Top-2 CNN filters   |   |
|-------------------|---|--------------------------------|------------------------|-------------|---|---|
| AGO1-4            | Mainly binds coding regions and 3'UTR                   | -                              | Hafner et al., 2010    | PAR-CLIP    |    |    |
| AGO2              | Mainly binds coding regions and 3'UTR                   | -                              | Kishore et al., 2011   | HITS-CLIP   |    |    |
| ALKBH5            | Mainly binds coding regions and 3'UTR                   | -                              | Baltz et al., 2012     | PAR-CLIP    |    |    |
| C17ORF85          | Mainly binds coding regions                             | -                              | Baltz et al., 2012     | PAR-CLIP    |    |    |
| C22ORF28          | Mainly binds coding regions                             | -                              | Baltz et al., 2012     | PAR-CLIP    |    |    |
| CAPRIN1           | Mainly binds coding regions and 3'UTR                   | -                              | Baltz et al., 2012     | PAR-CLIP    |    |    |
| ELAVL1            | Binds single-stranded AU or U-rich regions              | Barker et al., 2012            | Kishore et al., 2011   | HITS-CLIP   |    |    |
| ELAVL1 (A)        | Binds single-stranded AU or U-rich regions              | Barker et al., 2012            | Kishore et al., 2011   | PAR-CLIP    |    |    |
| ELAVL1 (B)        | Binds single-stranded AU or U-rich regions              | Barker et al., 2012            | Lebedeva et al., 2011  | PAR-CLIP    |    |    |
| EWSR1             | Binds AU-rich loop structures                           | -                              | Hoell et al., 2011     | PAR-CLIP    |    |    |
| FUS               | Binds AU-rich loop structures                           | -                              | Hoell et al., 2011     | PAR-CLIP    |   |   |
| hnRNP C           | Binds U-rich regions                                    | -                              | König et al., 2010     | iCLIP       |  |  |
| HuR               | Binds single-stranded AU or U-rich regions              | Barker et al., 2012            | Mukherjee et al., 2011 | PAR-CLIP    |  |  |
| IGF2BP1-3         | Binds strongly to CAUH-patterns                         | -                              | Hafner et al., 2010    | PAR-CLIP    |  |  |
| MOV10             | Binds AU-rich 3' UTR regions                            | Gregersen et al., 2014         | Sievers et al., 2012   | PAR-CLIP    |  |  |
| PTBP1             | Binds CU-rich regions                                   | Perez et al., 1997             | Xue et al., 2009       | HITS-CLIP   |  |  |
| PUM2              | High-affinity binding sites contain UGUANAU             | -                              | Hafner et al., 2010    | PAR-CLIP    |  |  |
| QKI               | High-affinity binding sites contain YUAAAY              | -                              | Hafner et al., 2010    | PAR-CLIP    |  |  |
| SRSF1             | Single-stranded AGAAGAAG motif                          | Wang et al., 2011              | Sanford et al., 2009   | HITS-CLIP   |  |  |
| TAF15             | Binds AU-rich loop structures                           | -                              | Hoell et al., 2011     | PAR-CLIP    |  |  |
| TDP43             | Binds GU-rich regions                                   | Buratti et al., 2004           | Tollervey et al., 2011 | iCLIP       |  |  |
| TIA1              | Binds U-rich regions near 5'ss                          | Dember et al., 1996            | Wang et al., 2010      | iCLIP       |  |  |
| TIAL1             | Binds U-rich regions near 5'ss                          | Dember et al., 1996            | Wang et al., 2010      | iCLIP       |  |  |
| ZC3H7B            | Mainly binds coding regions, 3'UTR and intronic regions | -                              | Baltz et al., 2012     | PAR-CLIP    |  |  |

**Fig. S5**

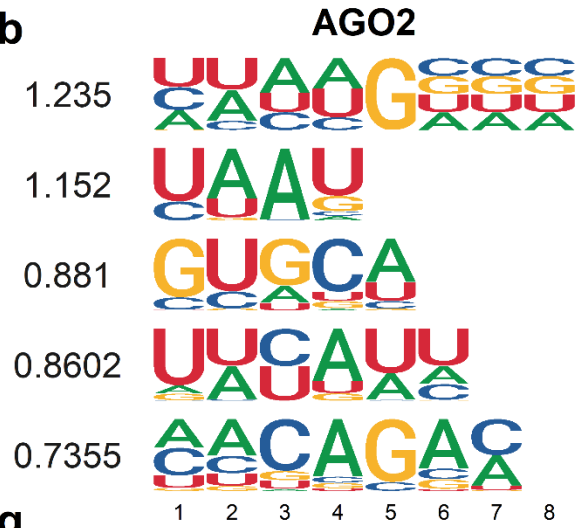


**Fig. S6**

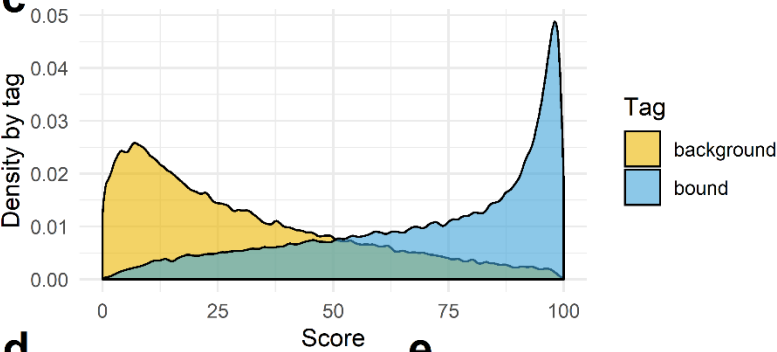
**a** ROC Curves: 10-fold cross-validation



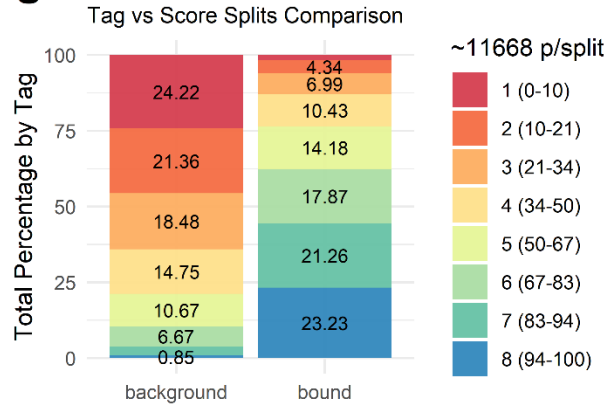
**b**



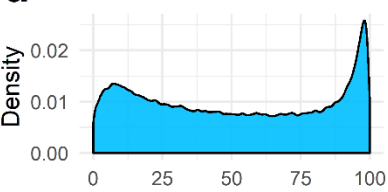
**c** Classification Model Results



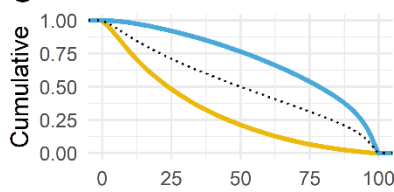
**g**



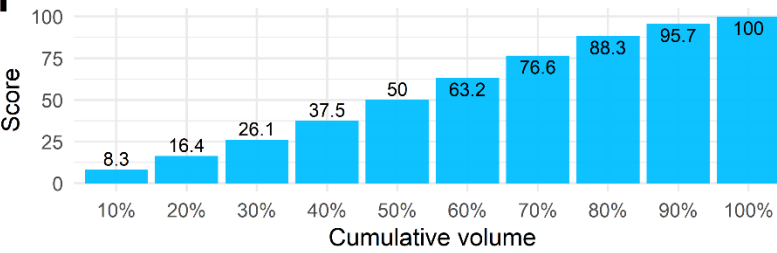
**d** Score



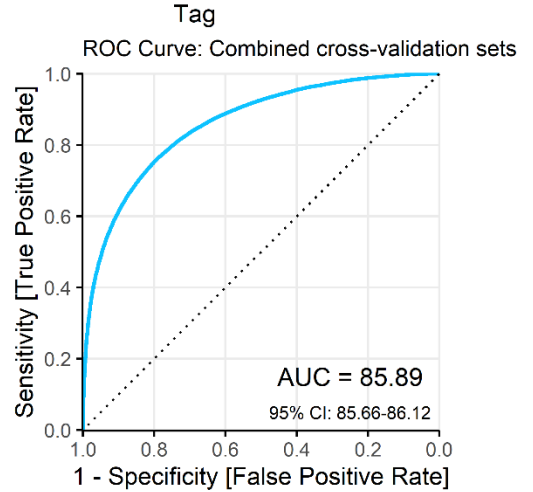
**e**



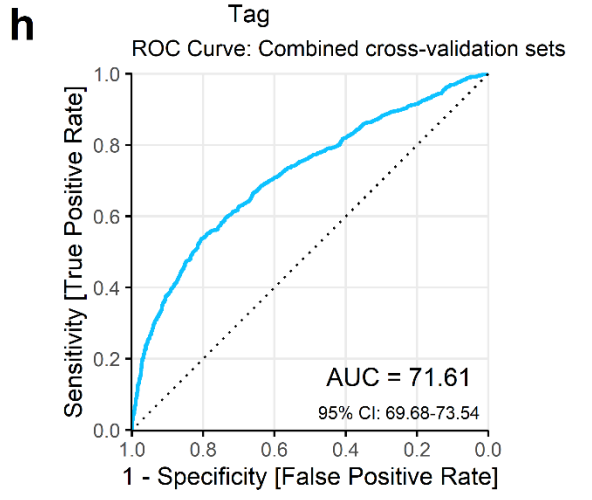
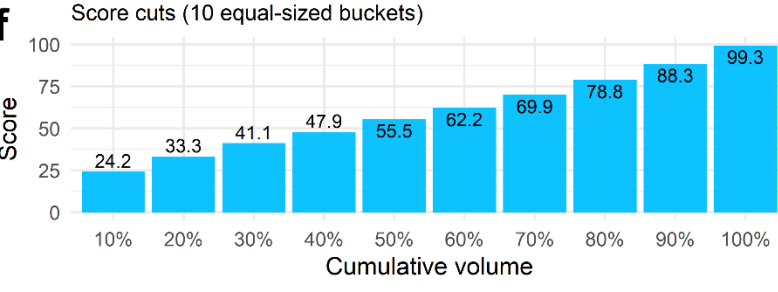
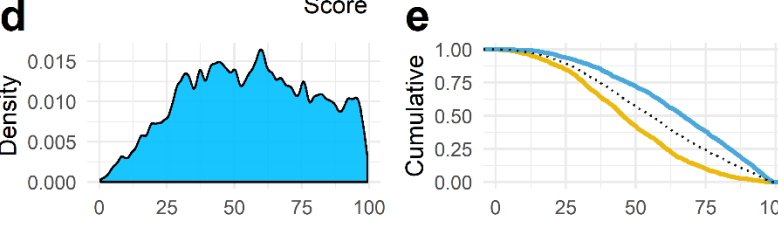
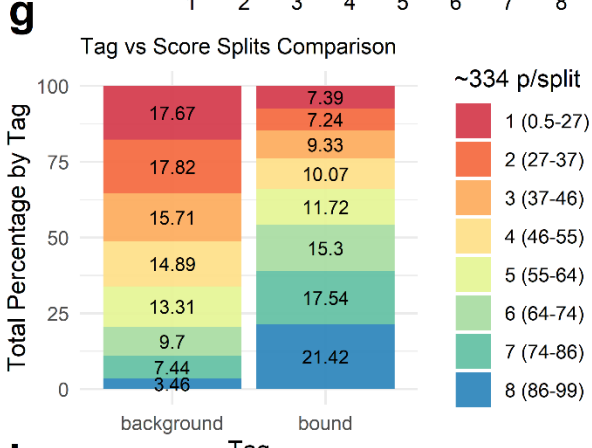
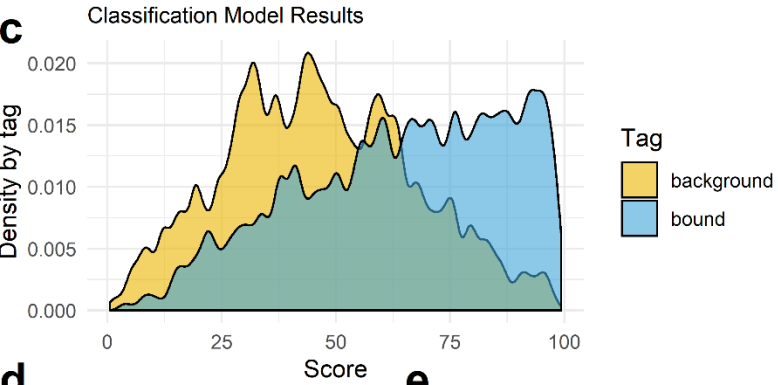
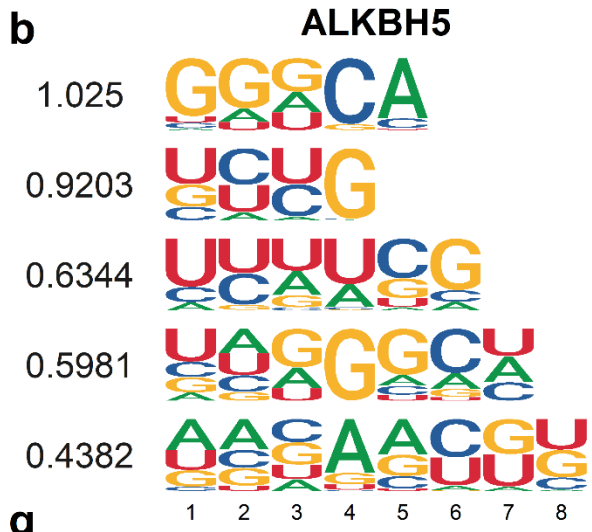
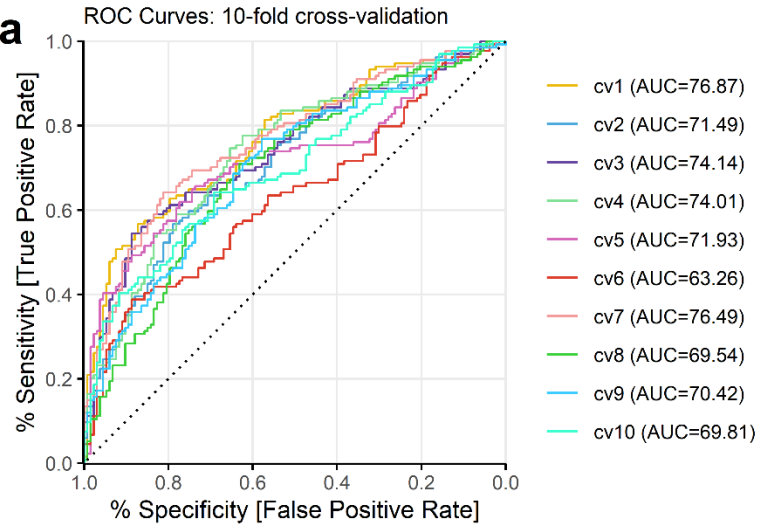
**f** Score cuts (10 equal-sized buckets)



**h**

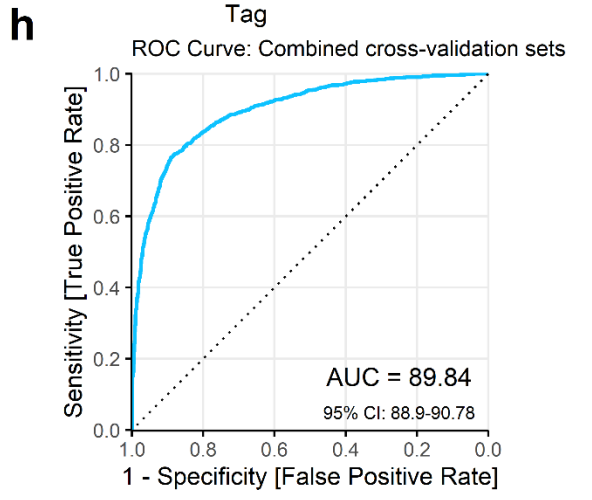
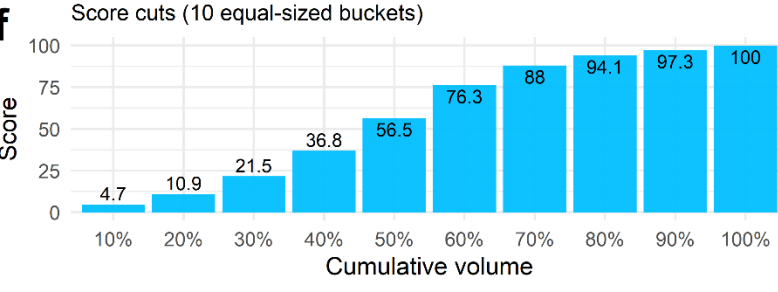
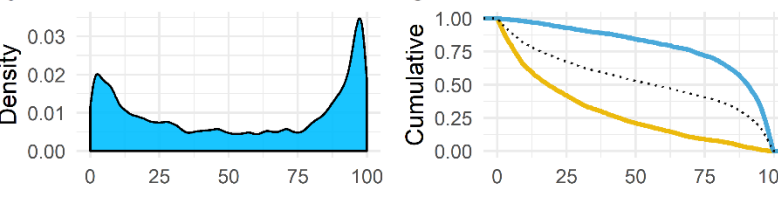
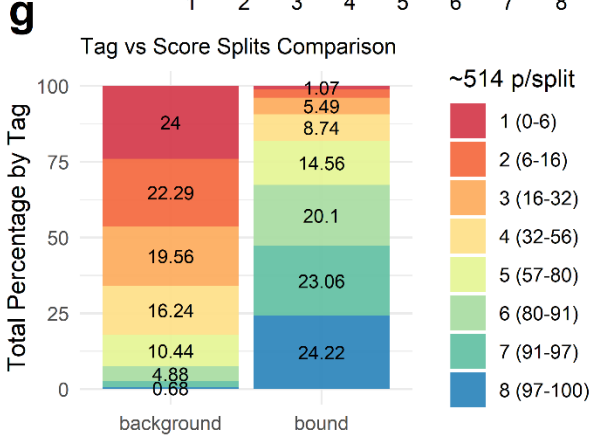
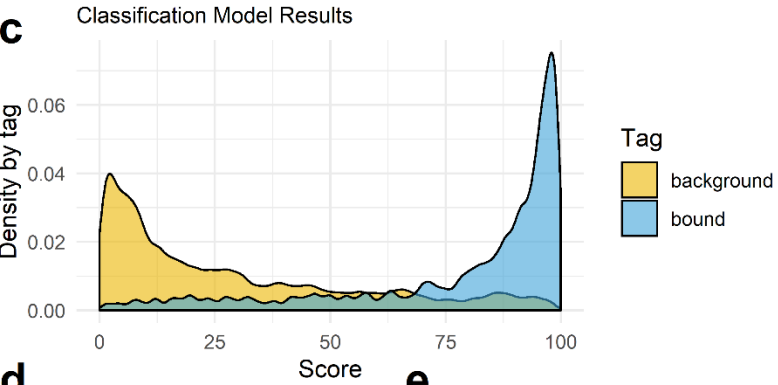
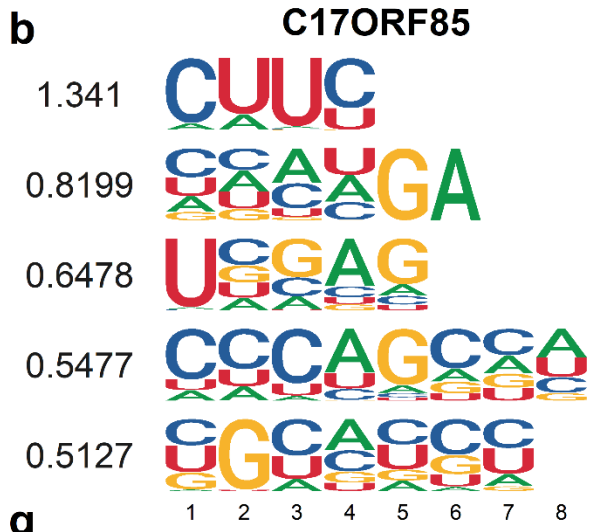
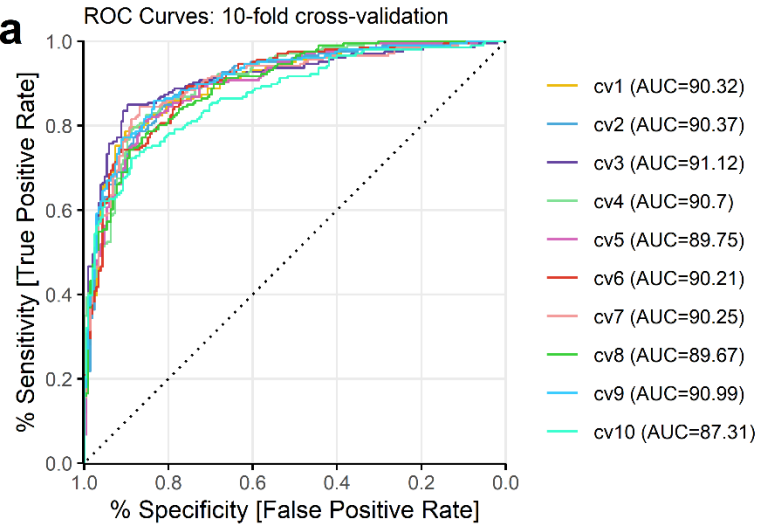


**Fig. S7**

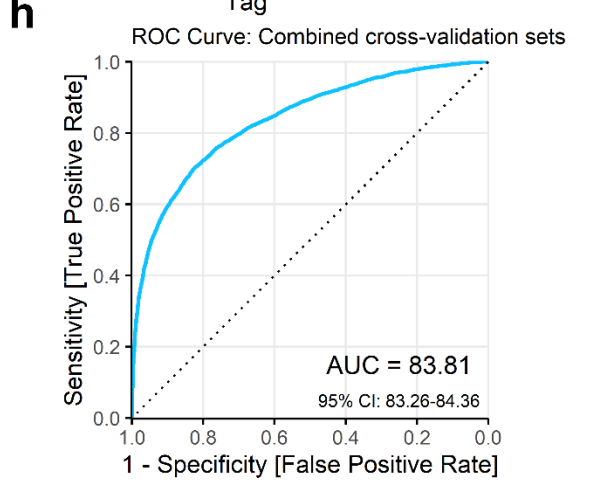
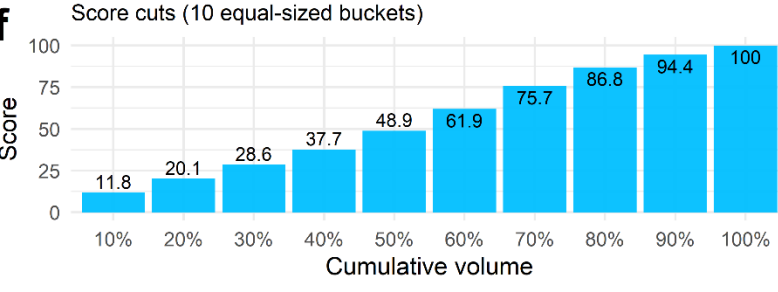
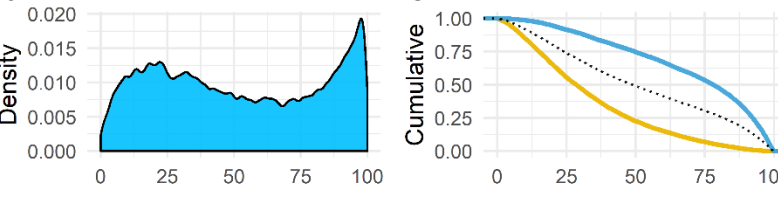
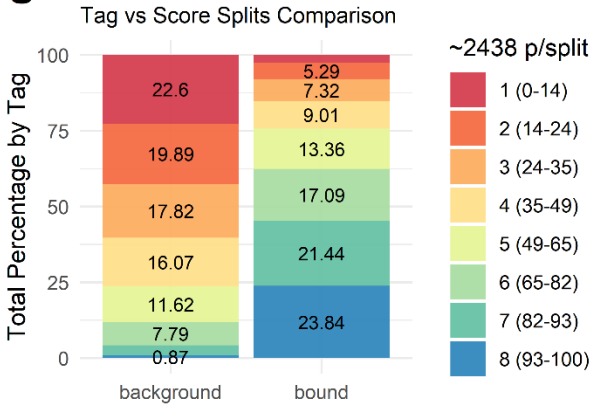
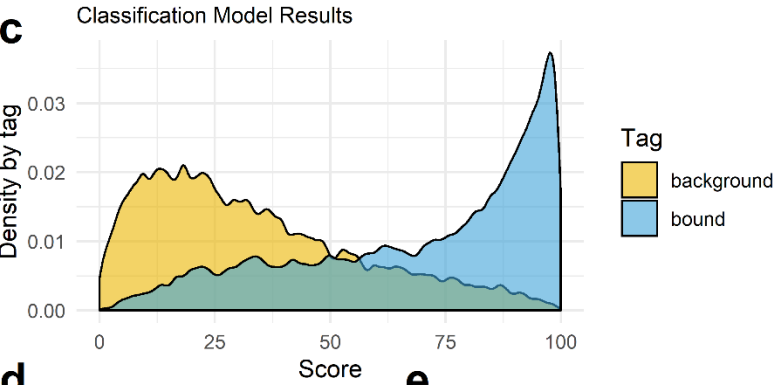
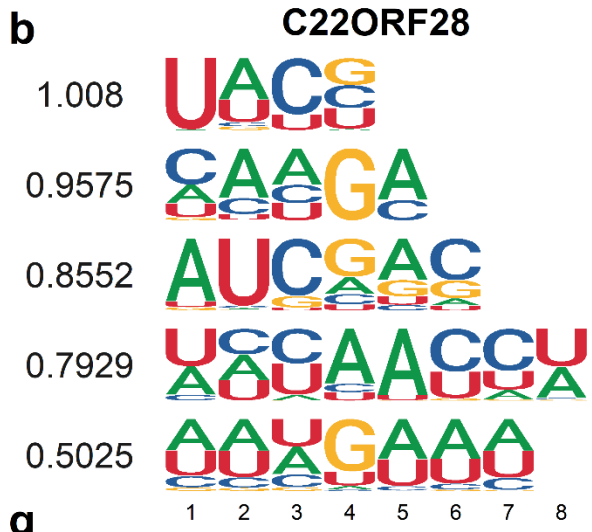
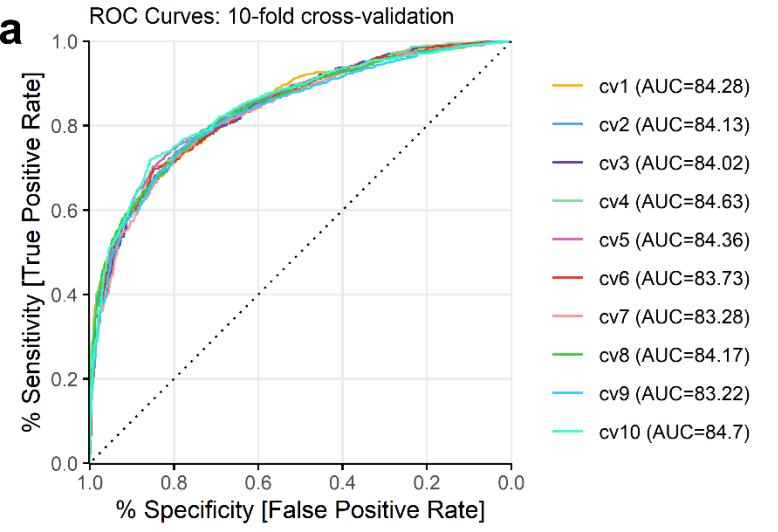




**Fig. S8**

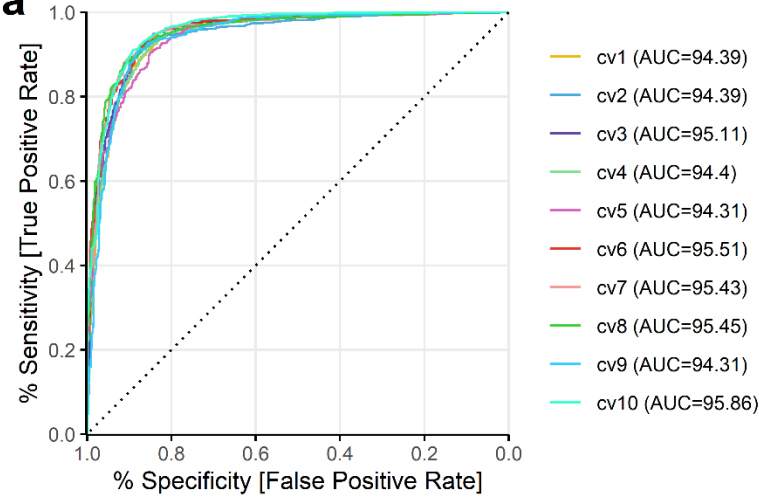


**Fig. S9**

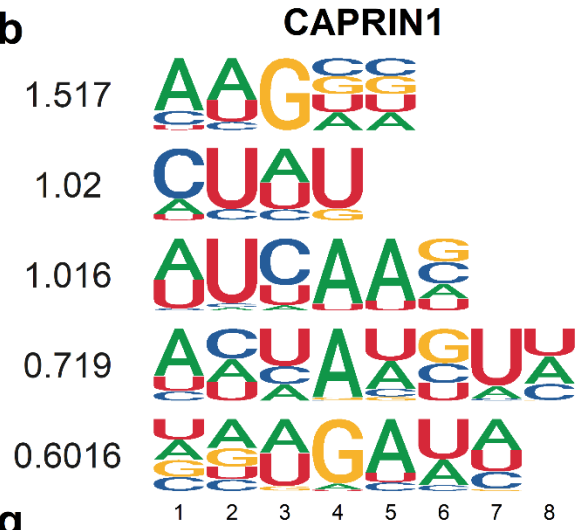


**Fig. S10**

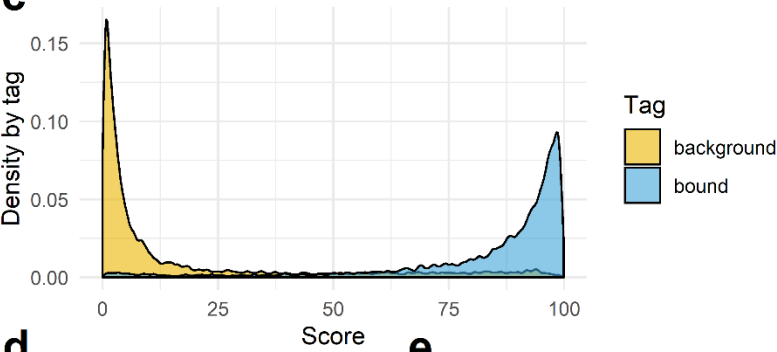
**a** ROC Curves: 10-fold cross-validation



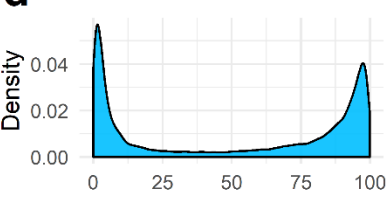
**b**



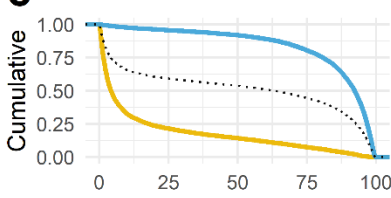
**c** Classification Model Results



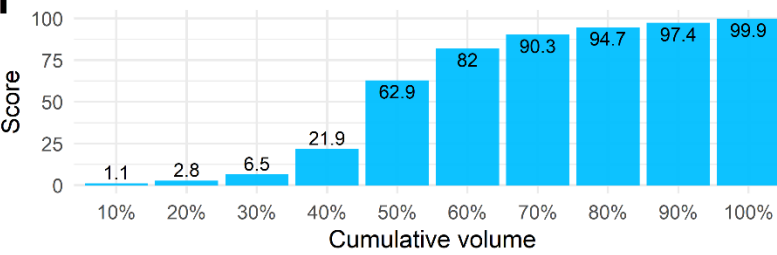
**d**



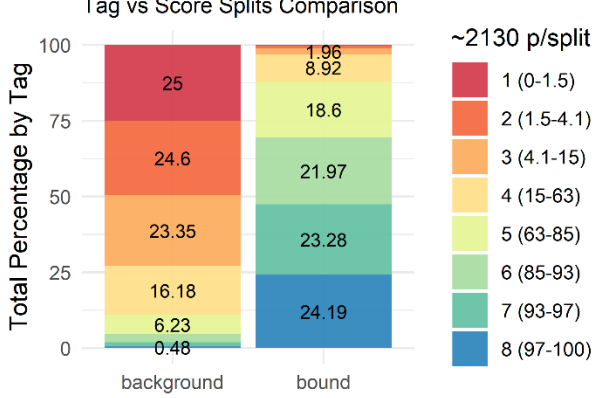
**e**



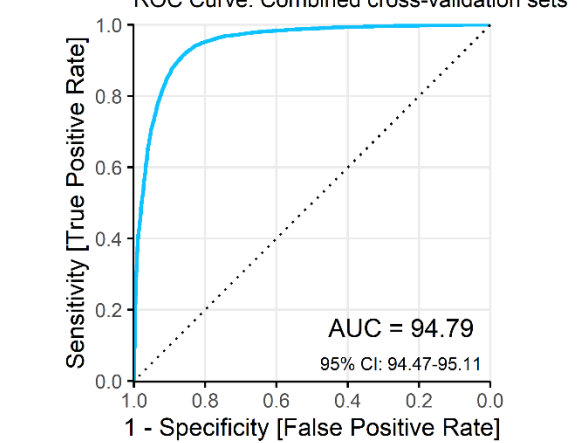
**f** Score cuts (10 equal-sized buckets)



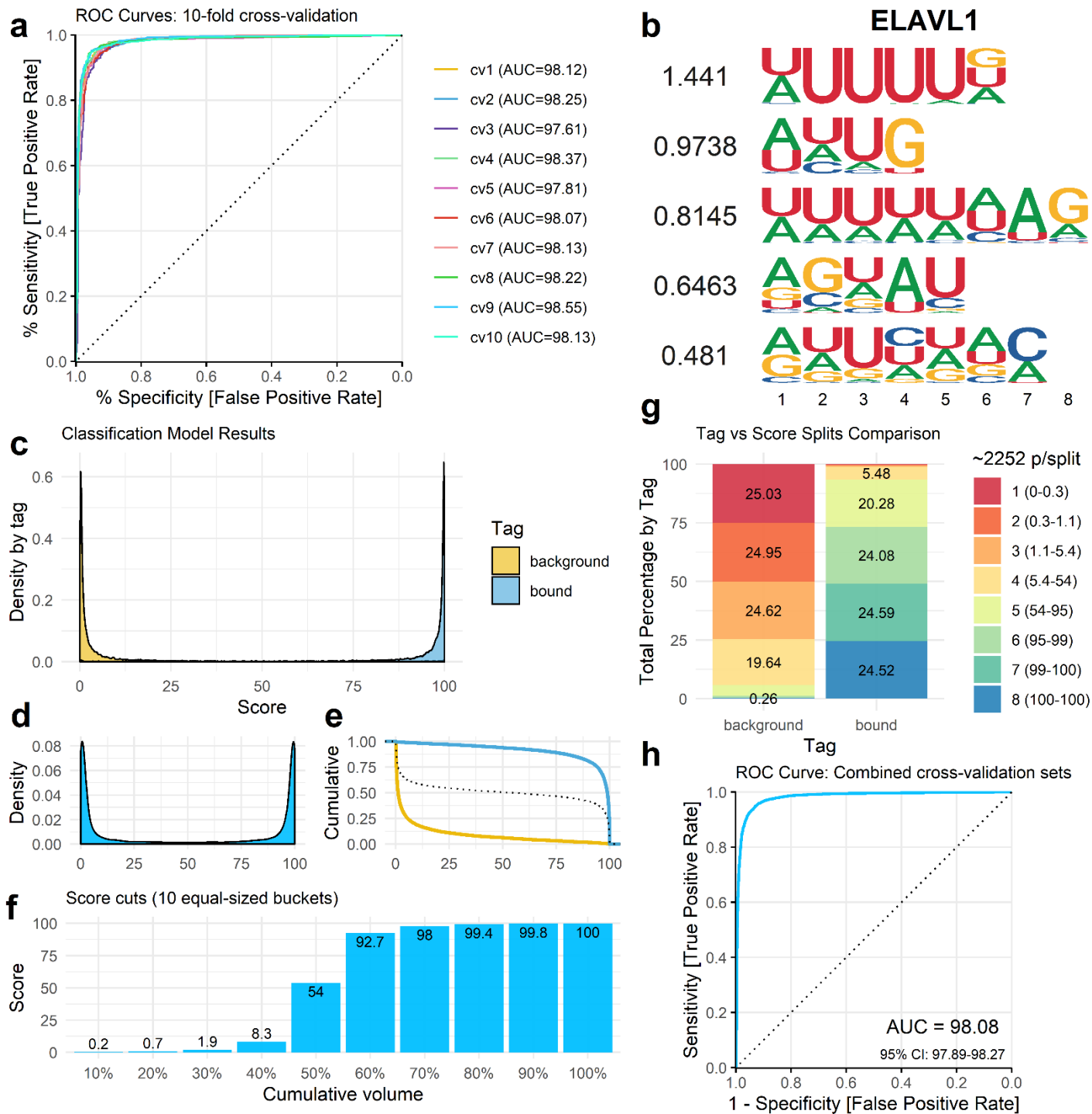
**g**



**h**

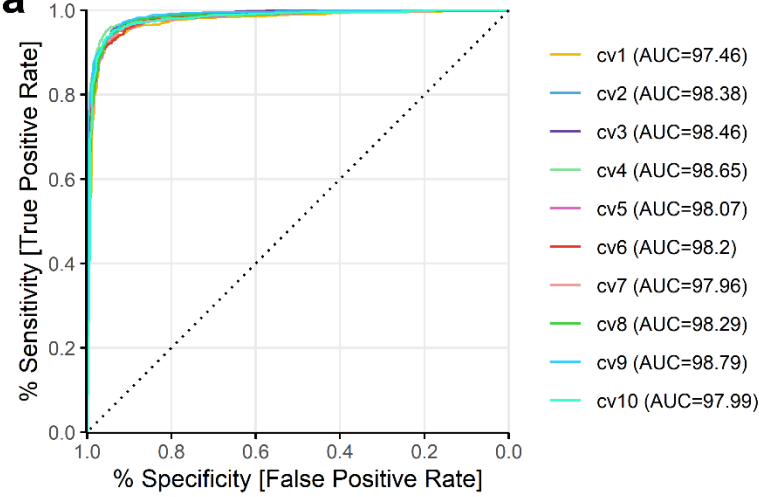


**Fig. S11**

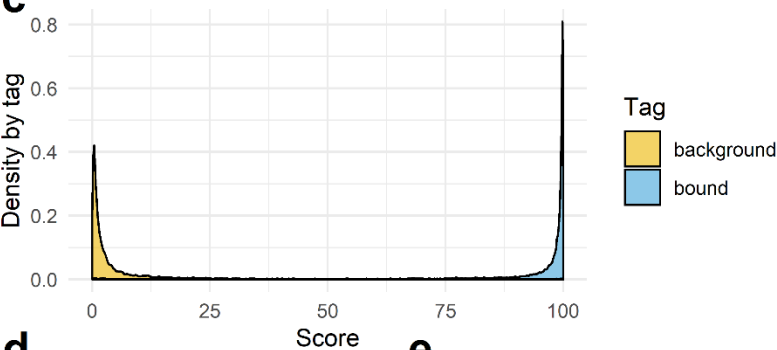


**Fig. S12**

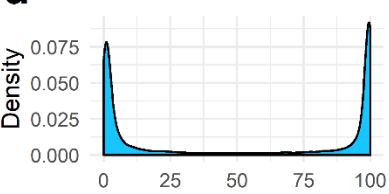
**a** ROC Curves: 10-fold cross-validation



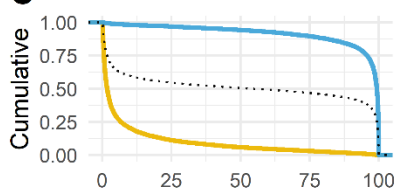
**c** Classification Model Results



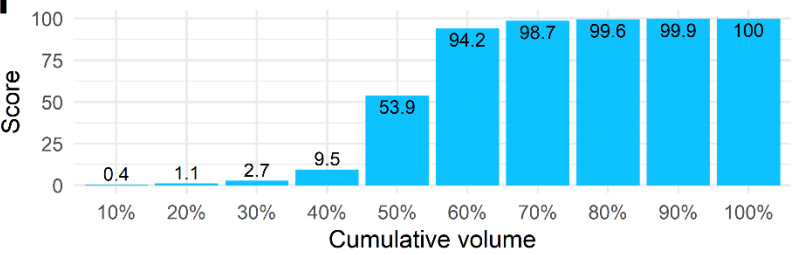
**d** Score



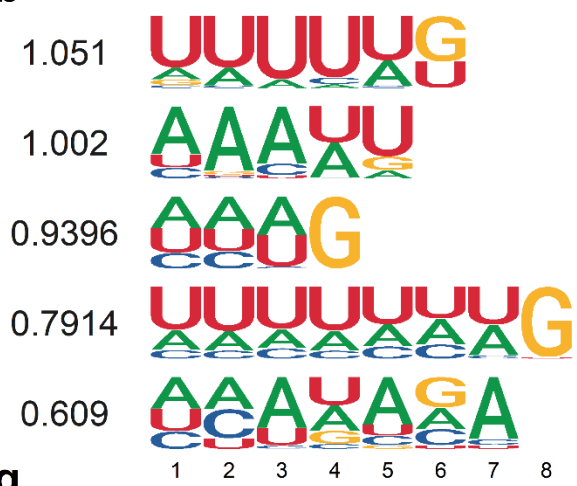
**e** Cumulative



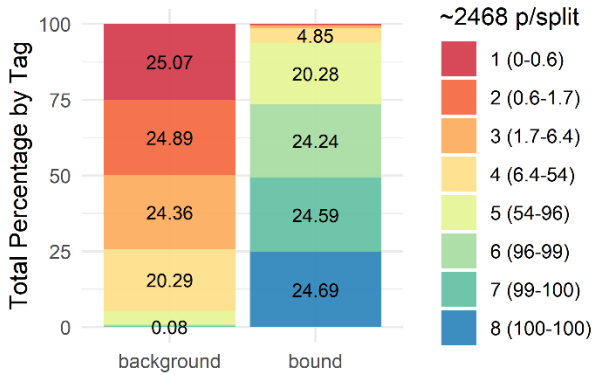
**f** Score cuts (10 equal-sized buckets)



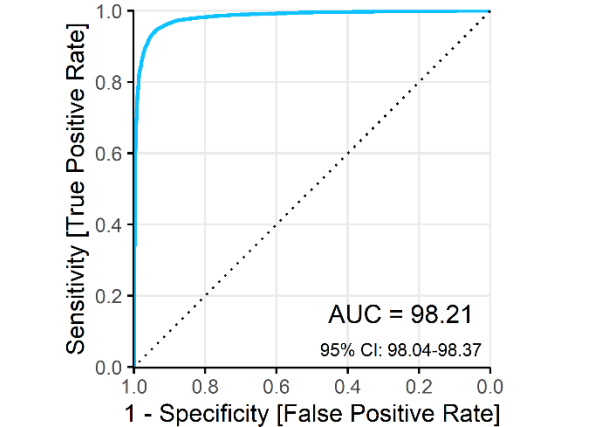
**b** ELAVL1A



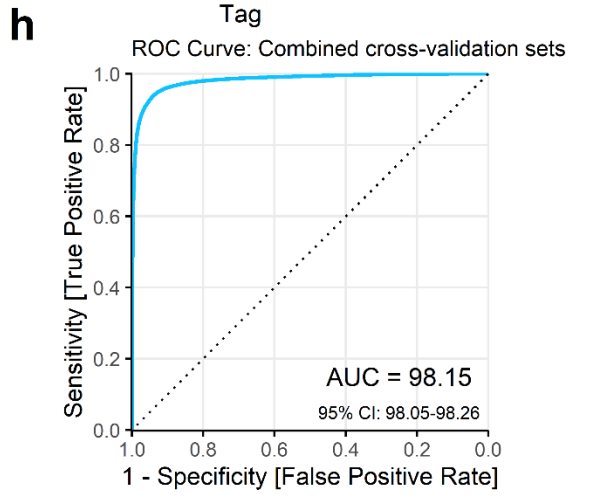
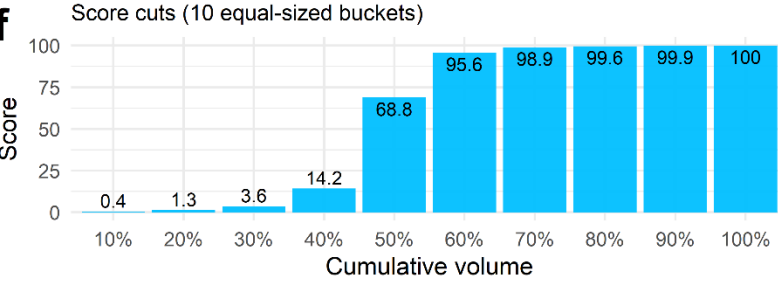
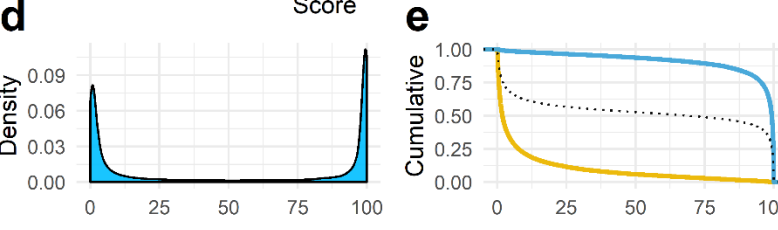
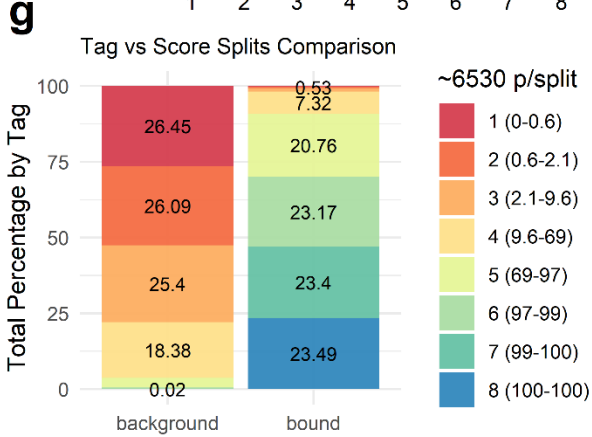
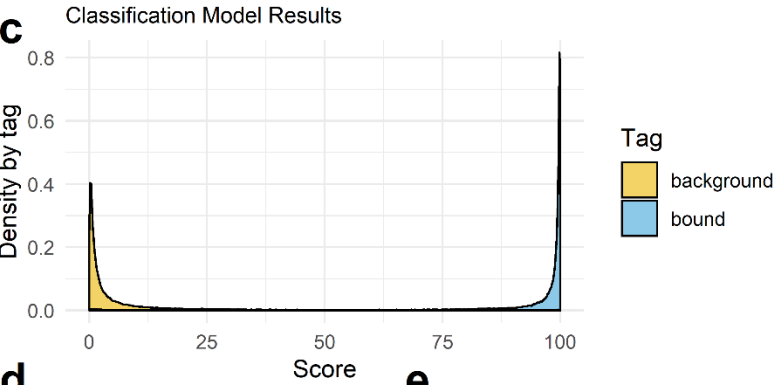
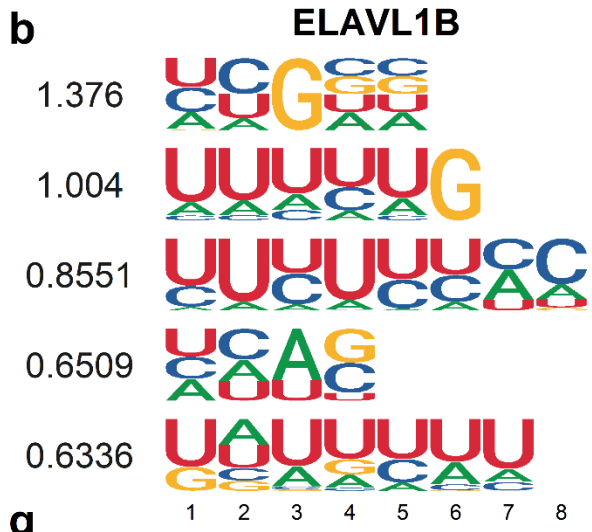
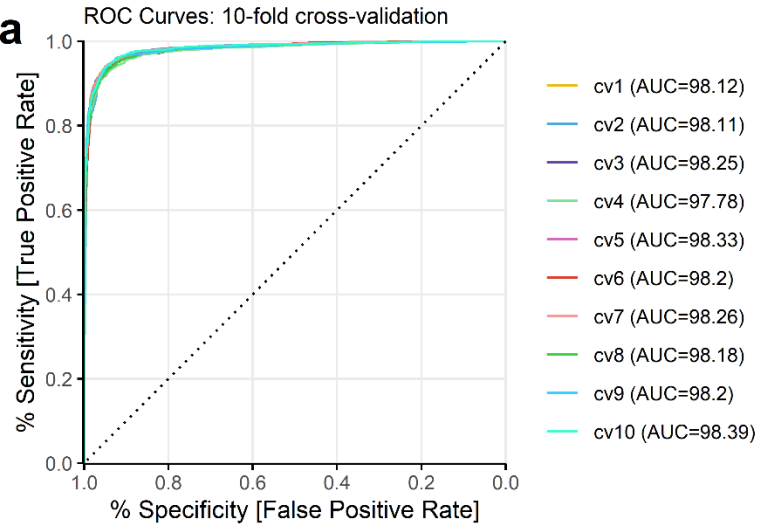
**g** Tag vs Score Splits Comparison



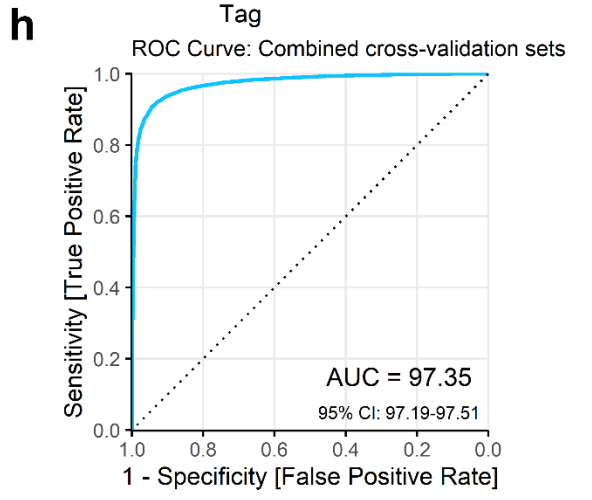
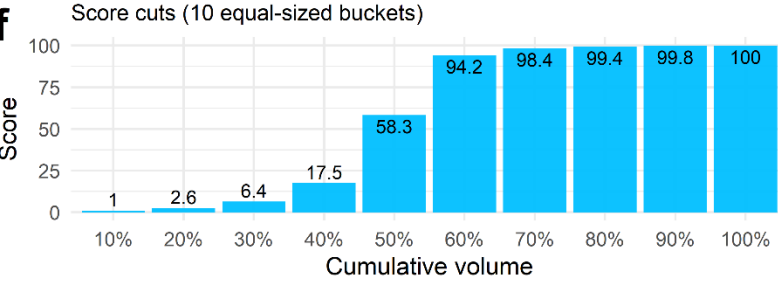
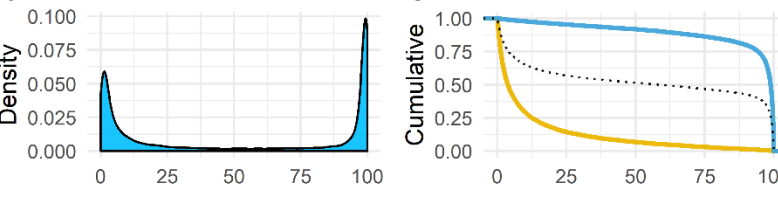
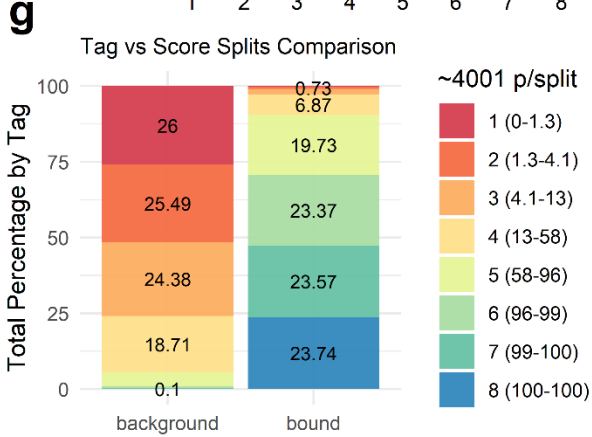
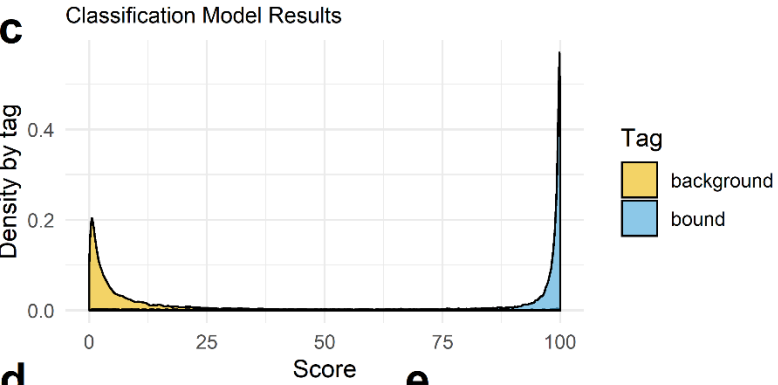
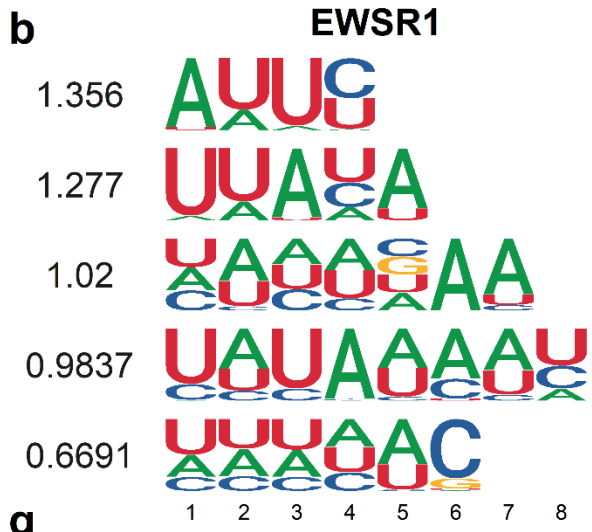
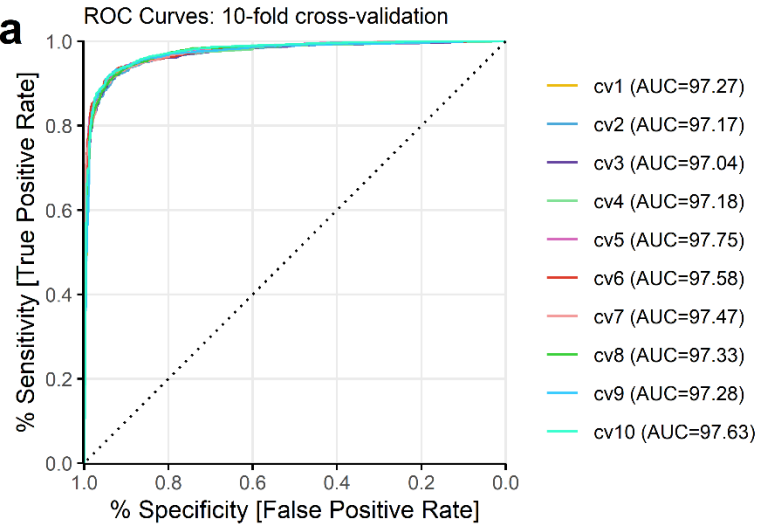
**h** Tag ROC Curve: Combined cross-validation sets



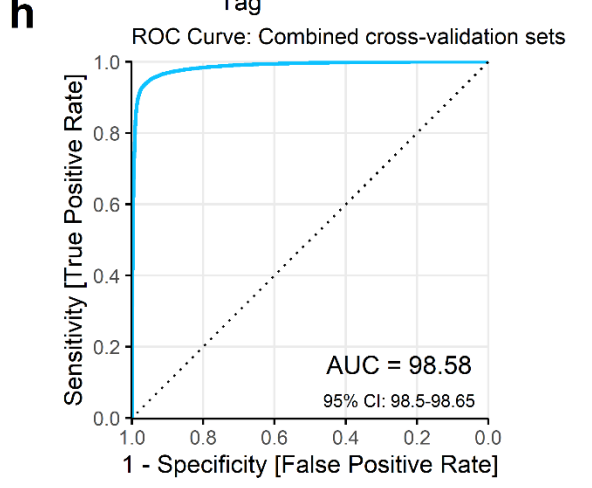
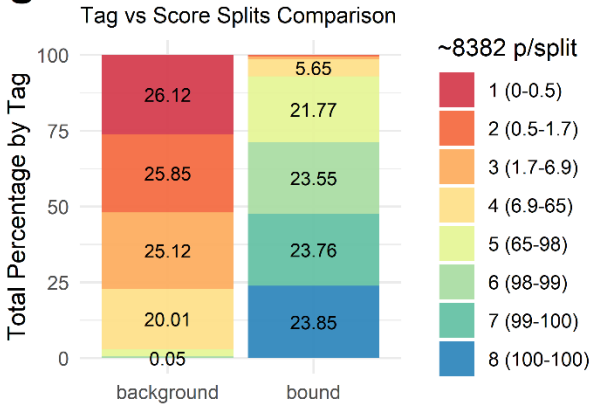
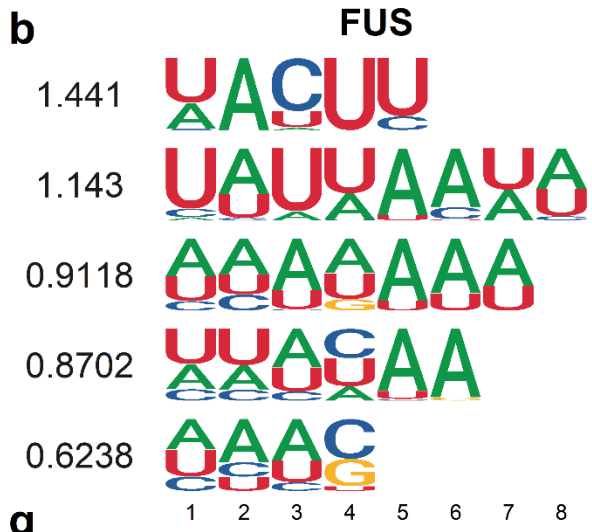
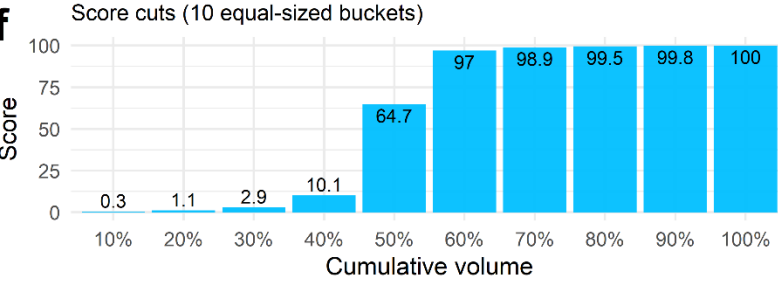
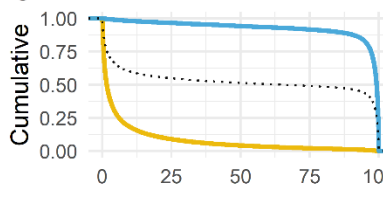
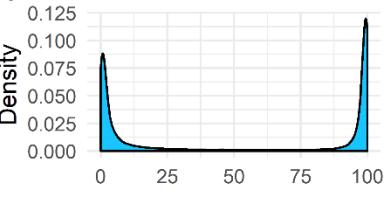
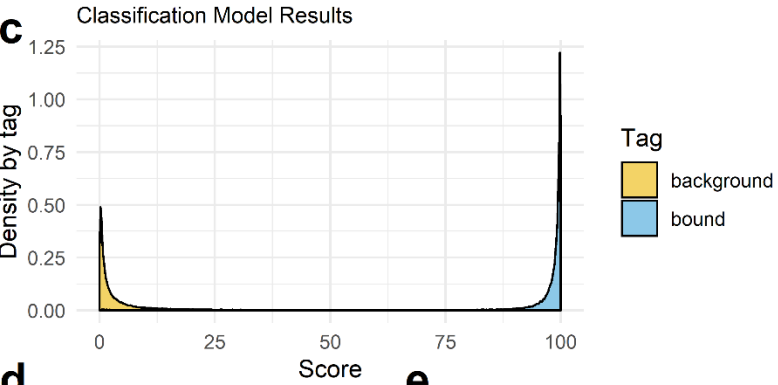
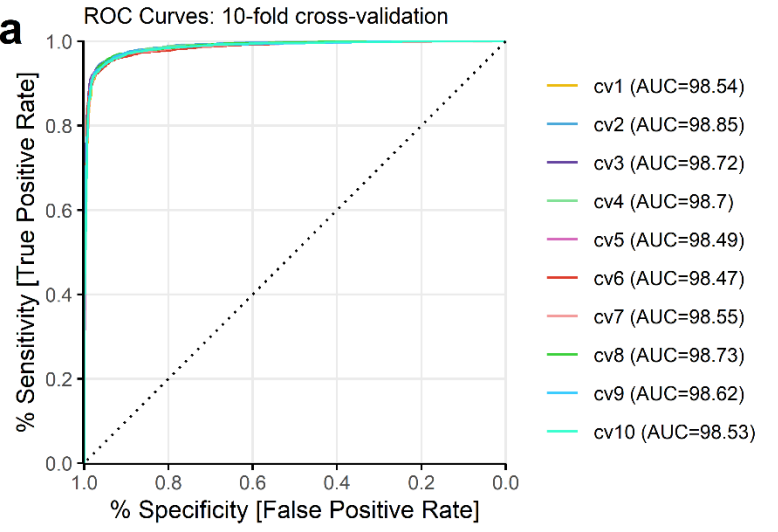
**Fig. S13**



**Fig. S14**

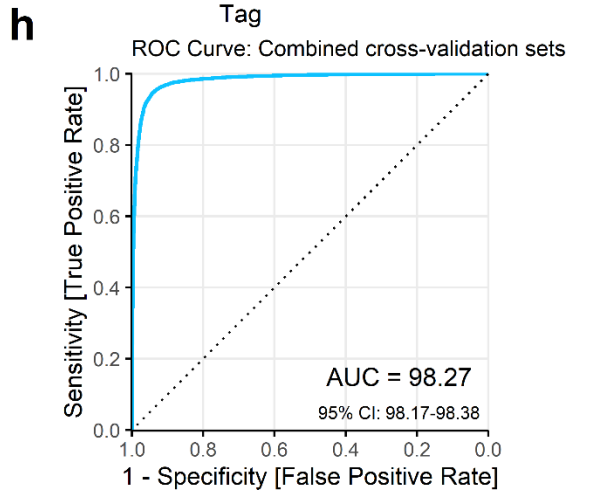
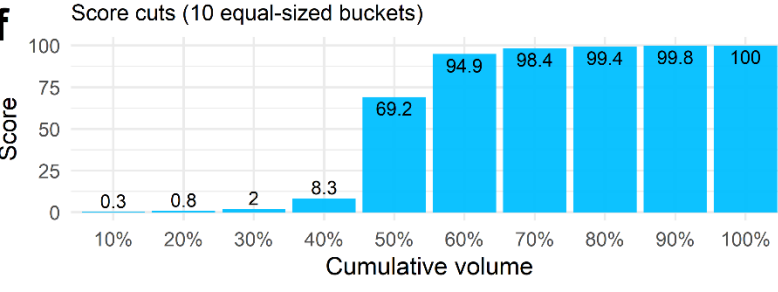
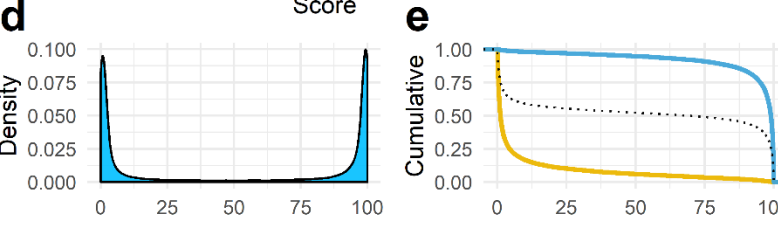
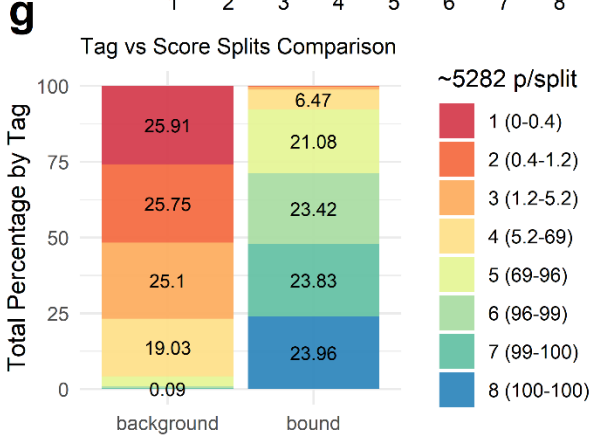
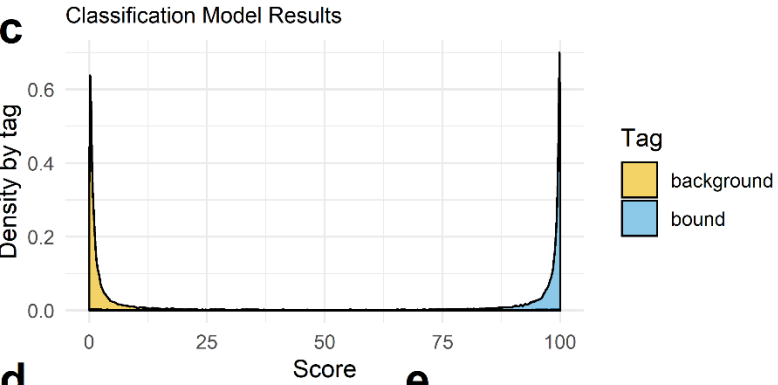
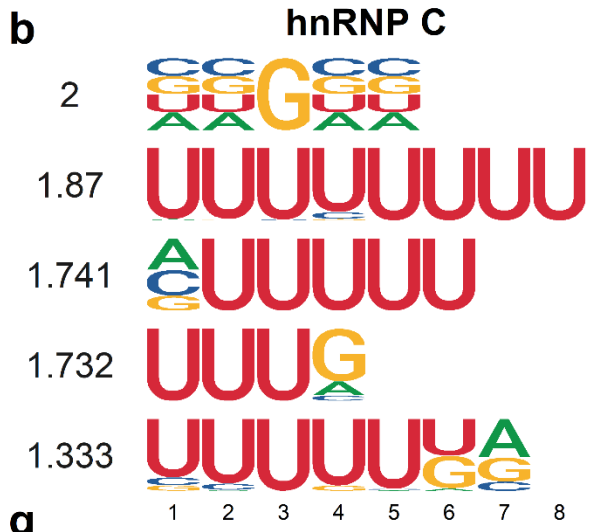
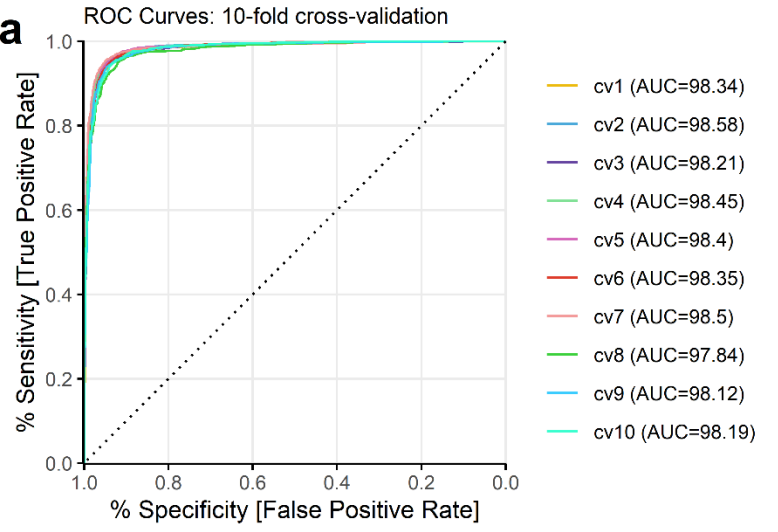


**Fig. S15**

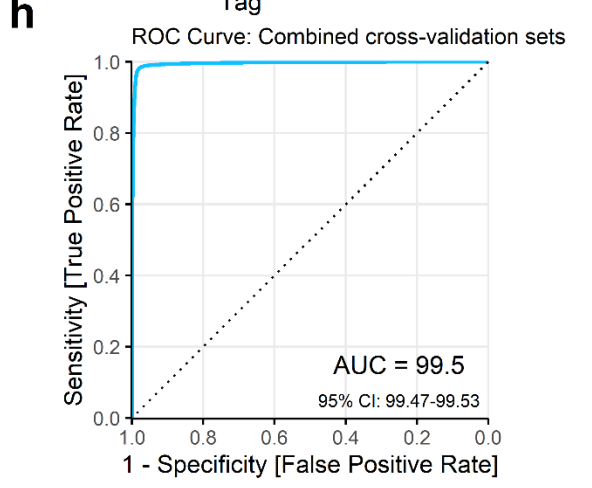
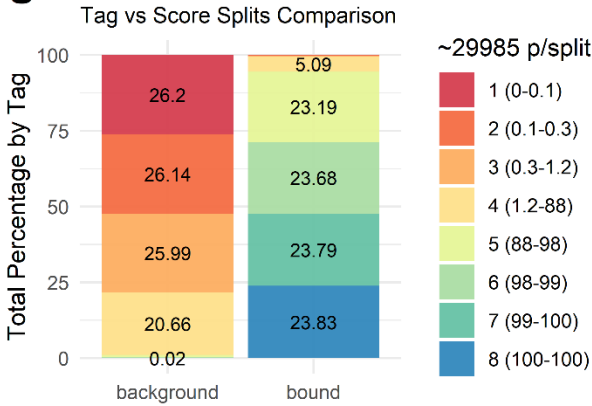
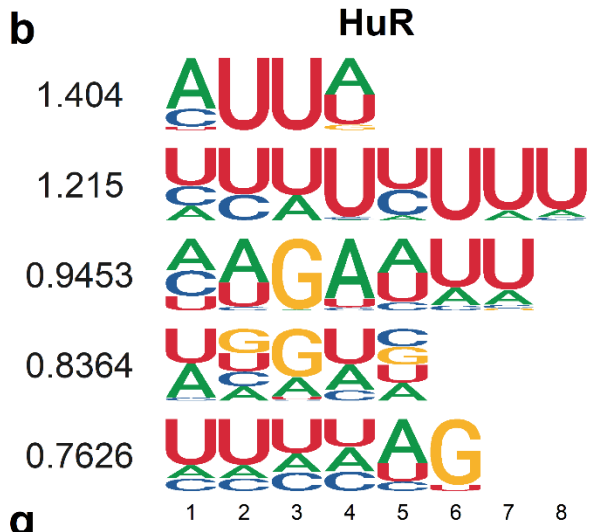
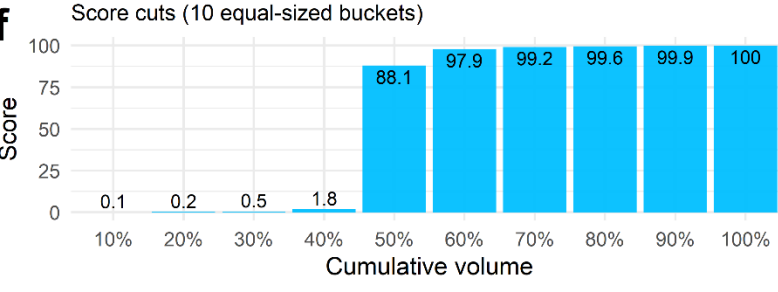
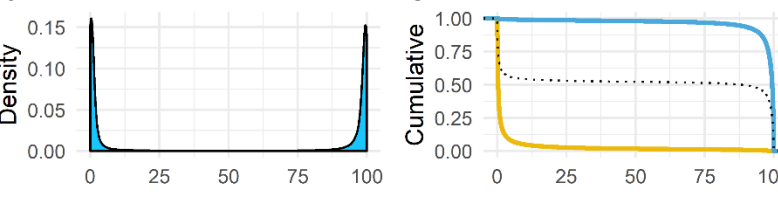
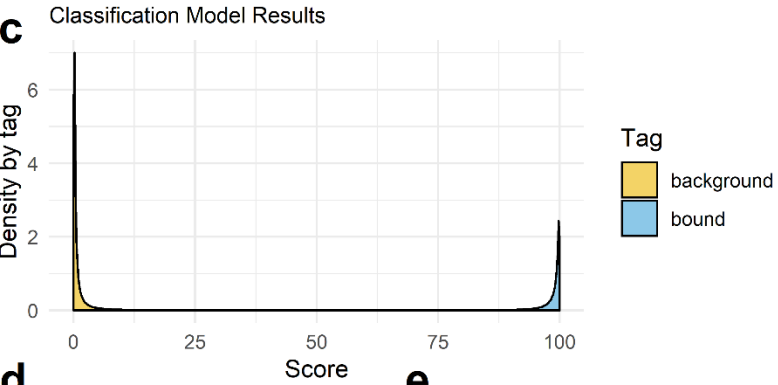
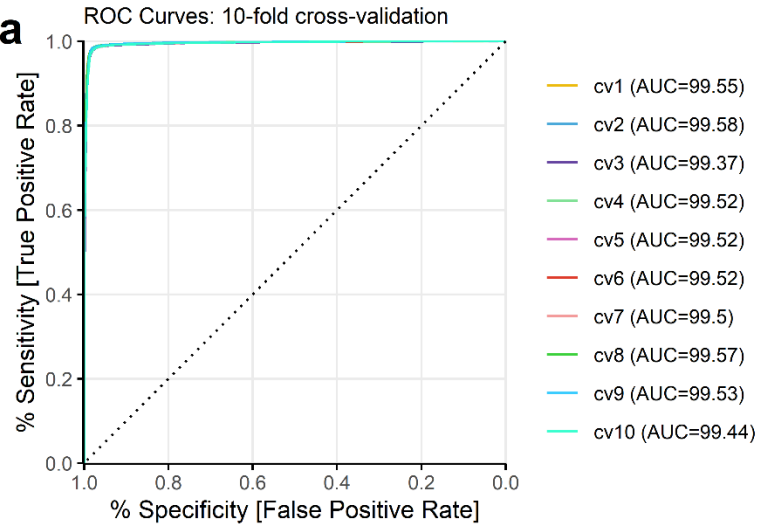




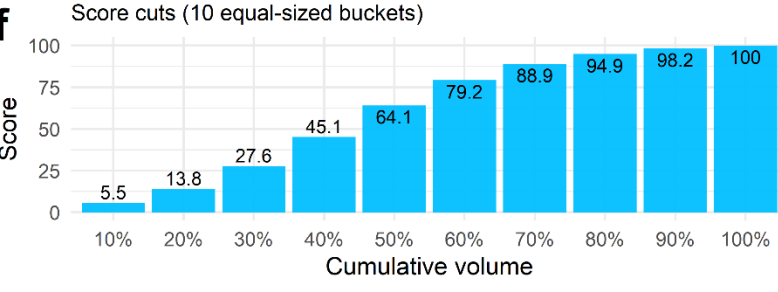
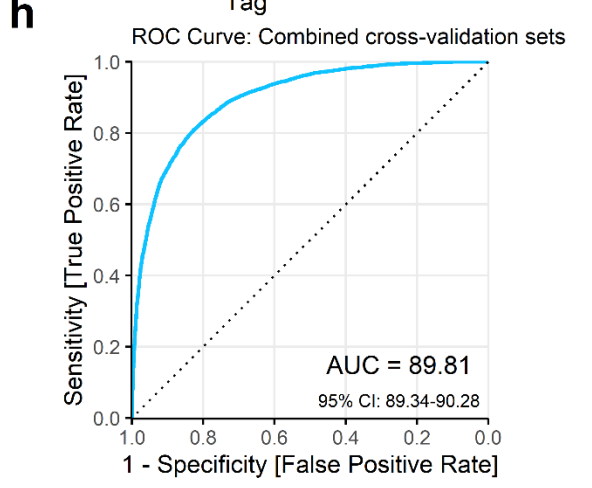
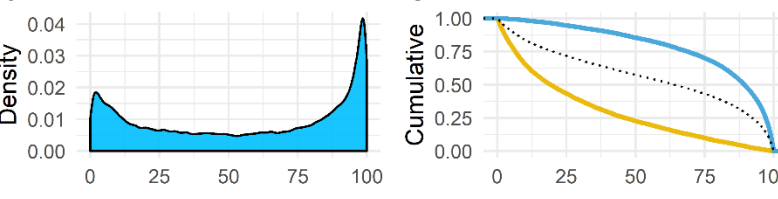
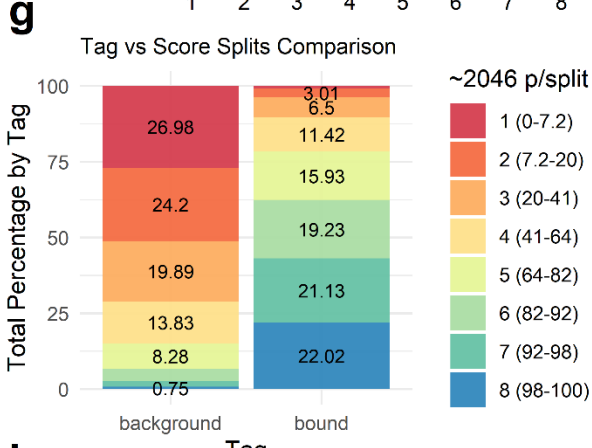
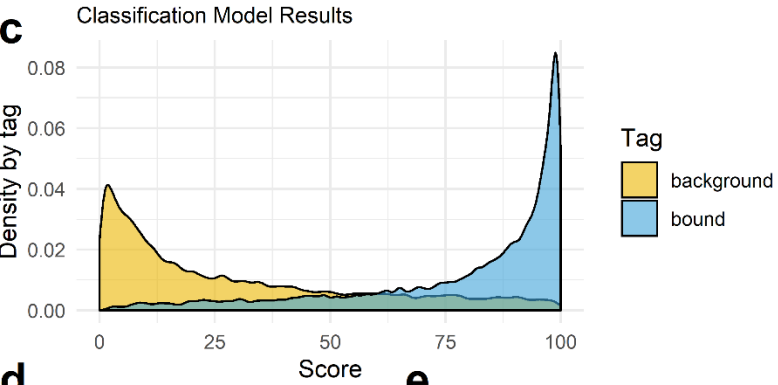
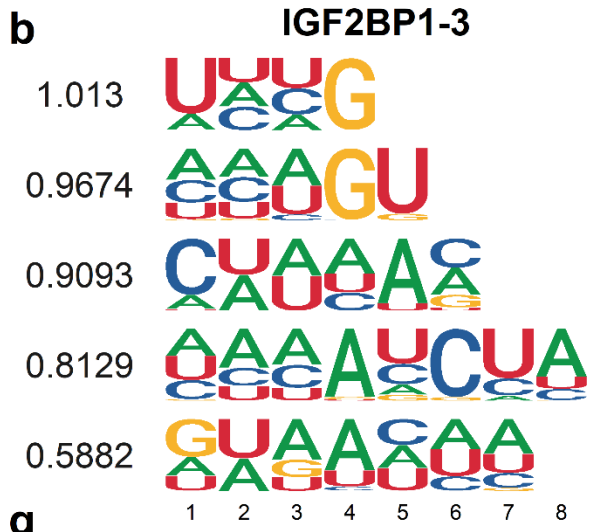
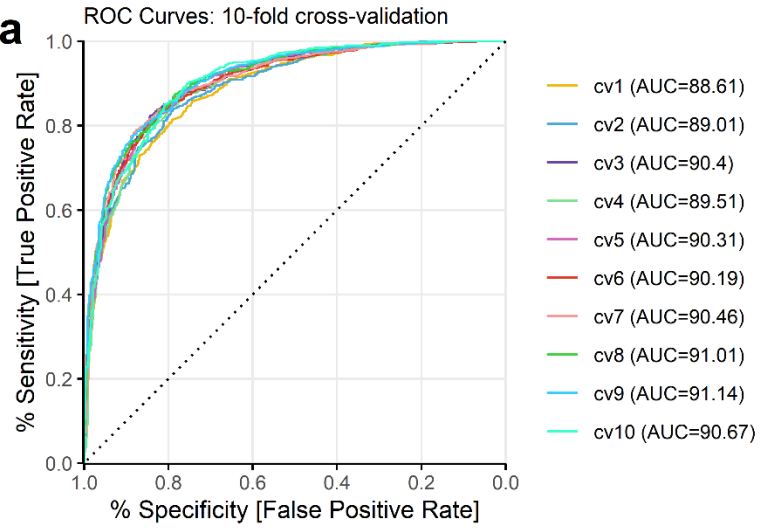
**Fig. S16**



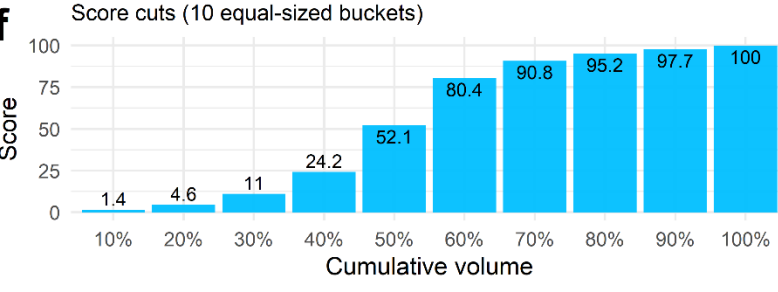
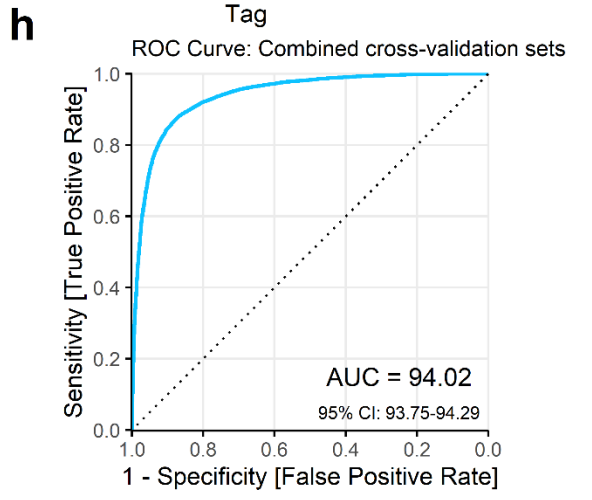
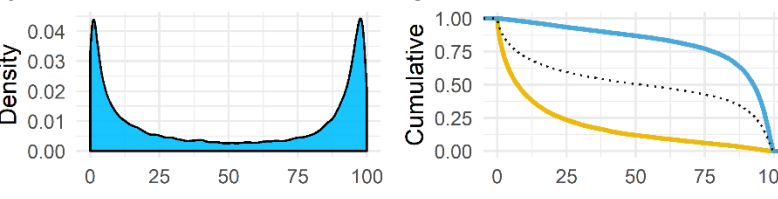
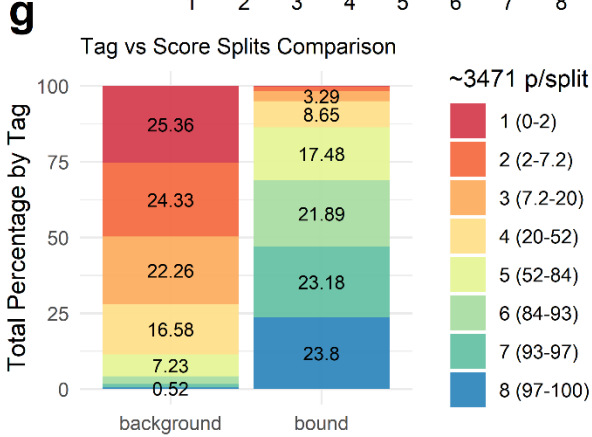
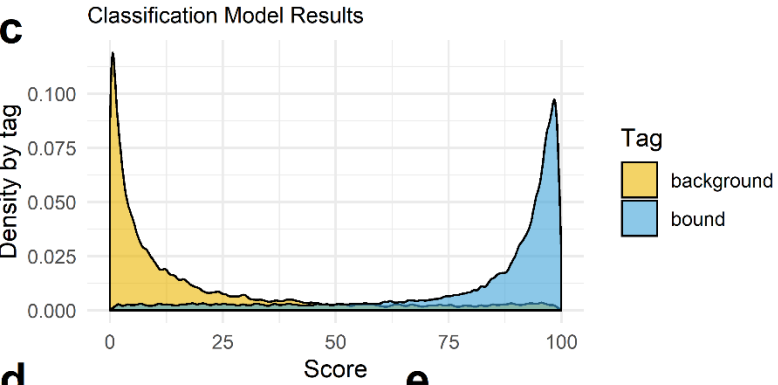
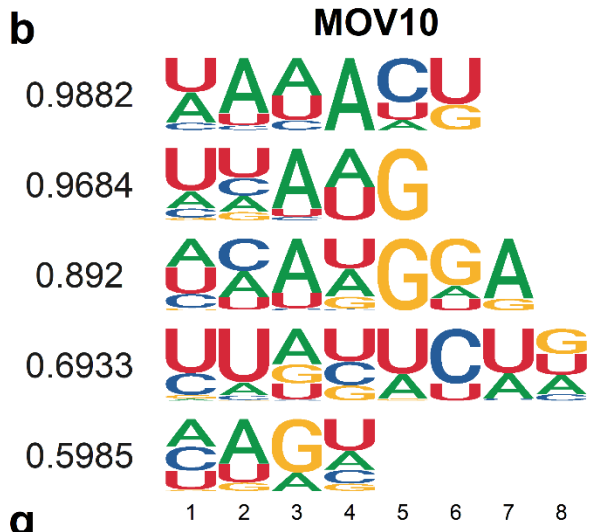
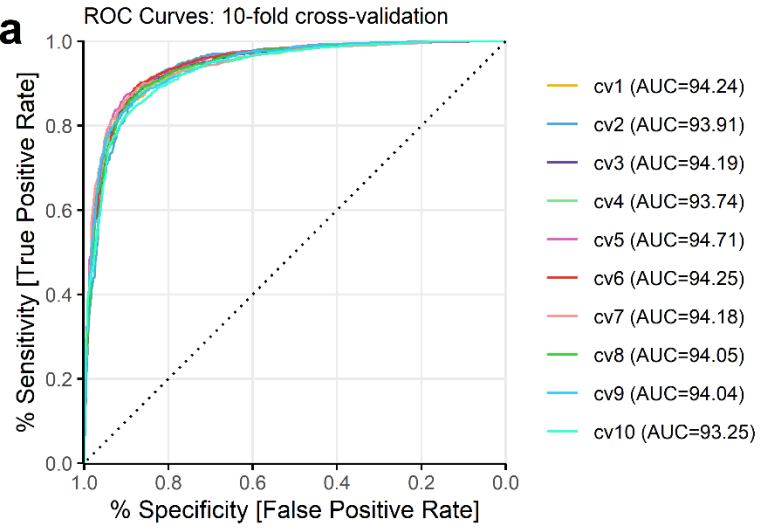
**Fig. S17**



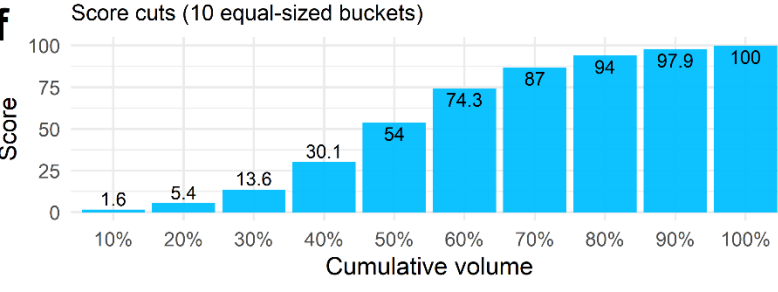
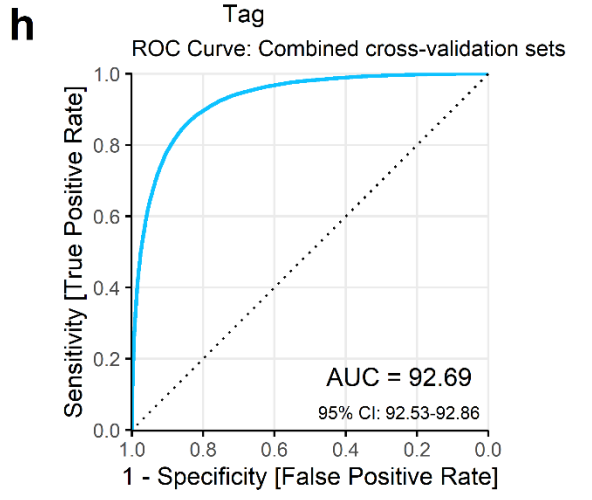
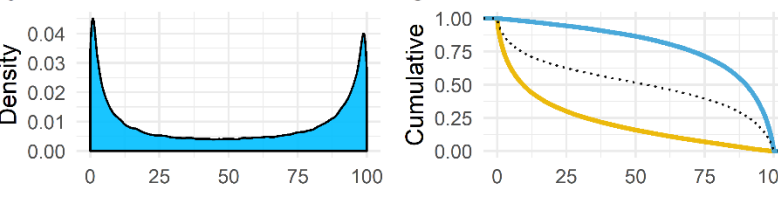
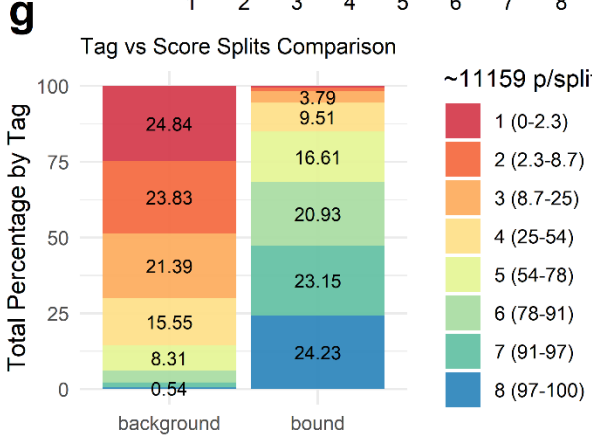
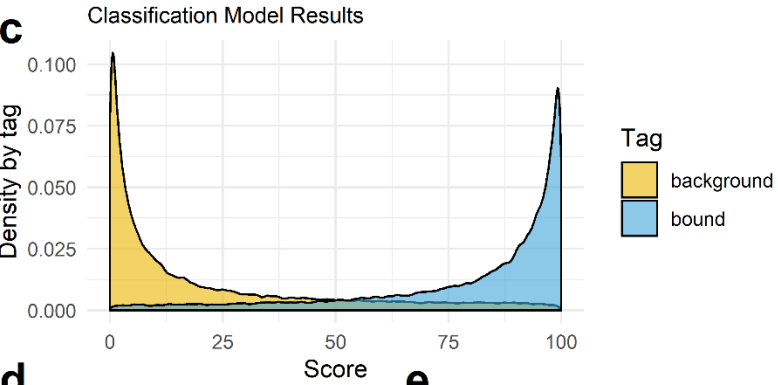
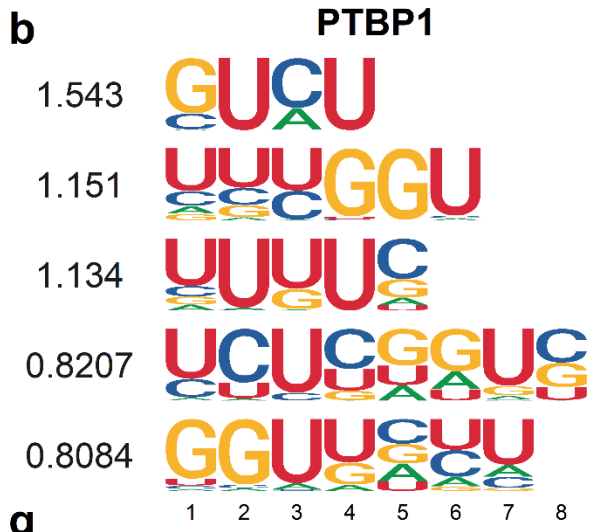
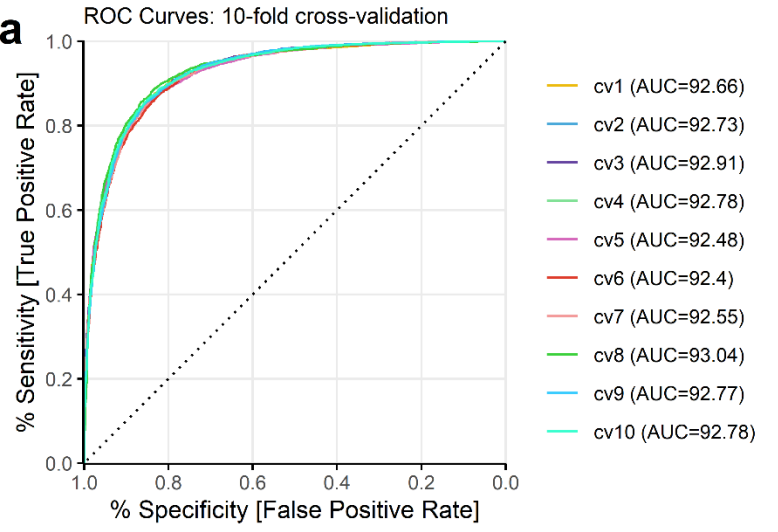
**Fig. S18**



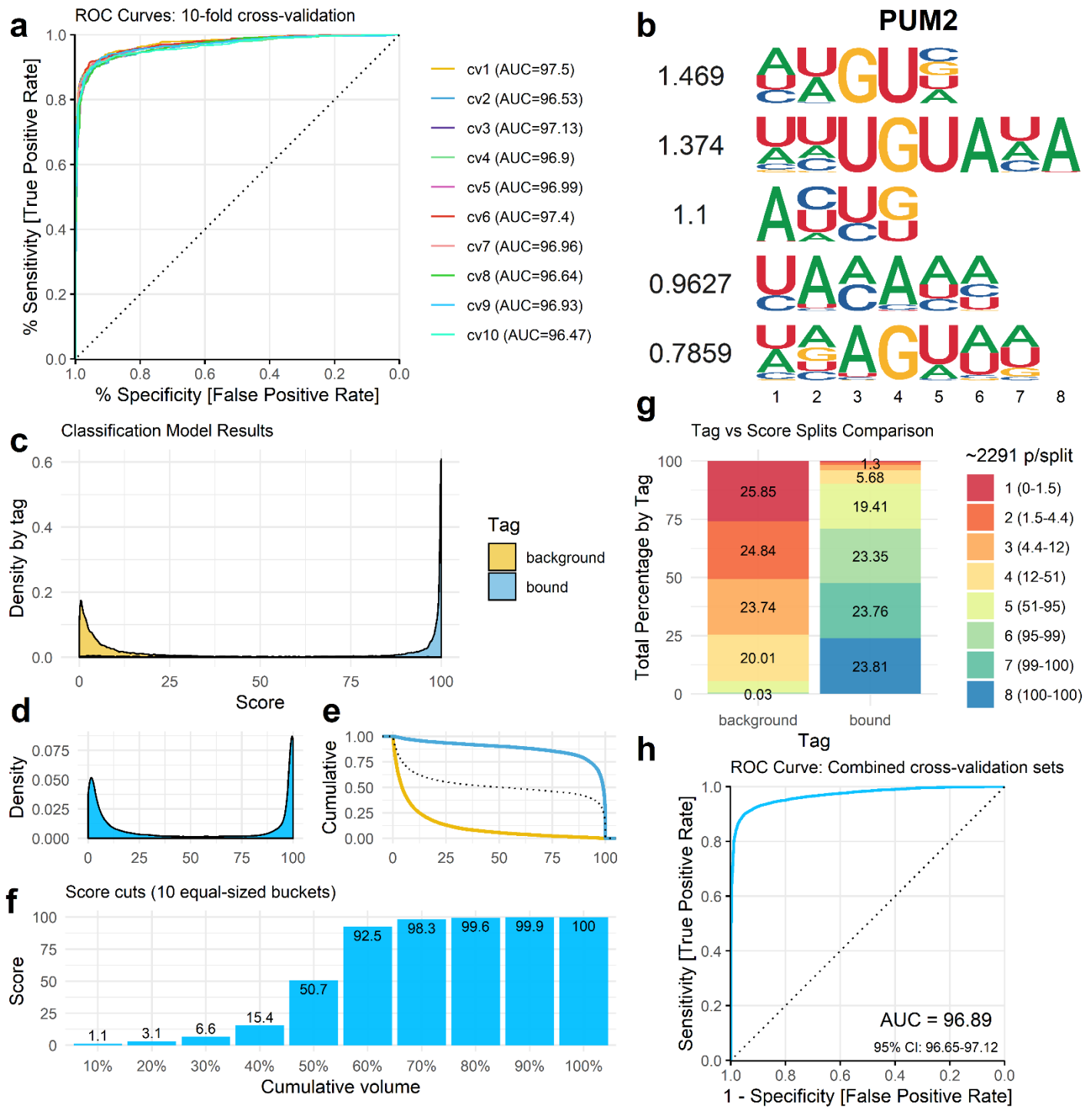
**Fig. S19**



**Fig. S20**

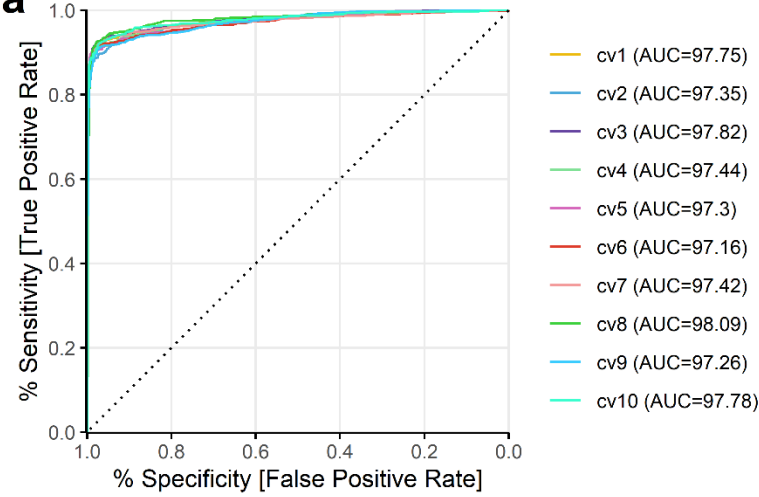


**Fig. S21**

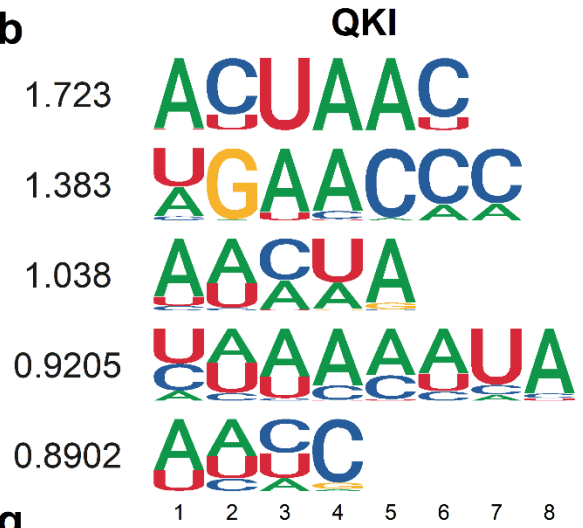


**Fig. S22**

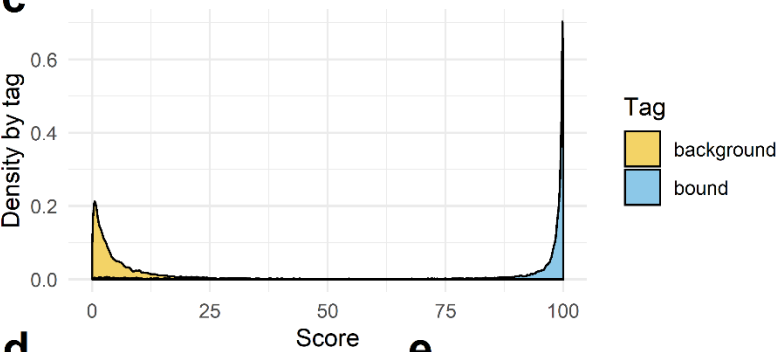
**a** ROC Curves: 10-fold cross-validation



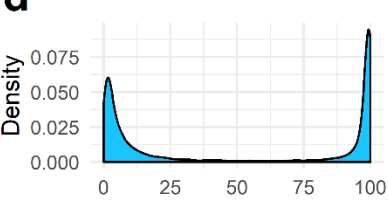
**b**



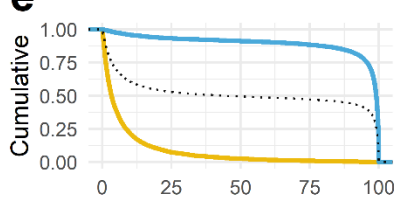
**c** Classification Model Results



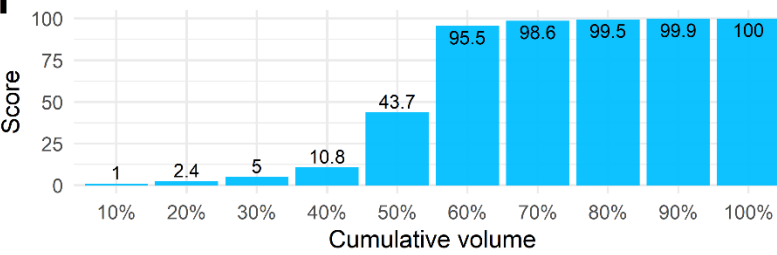
**d** Density



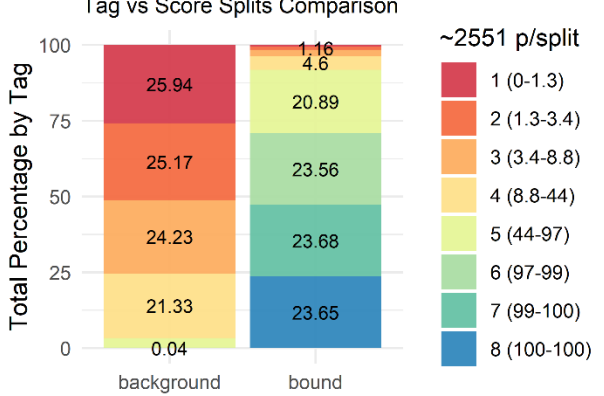
**e** Cumulative



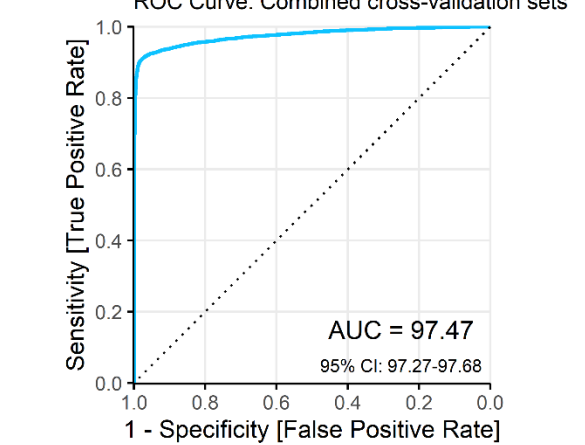
**f** Score cuts (10 equal-sized buckets)



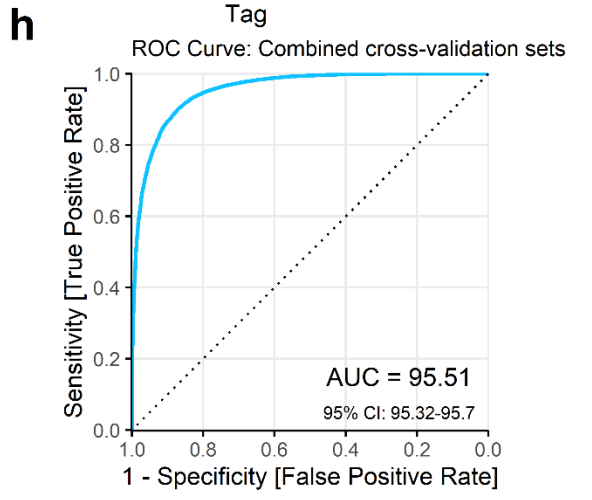
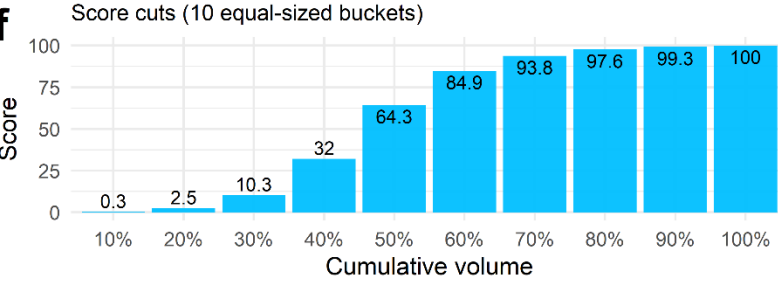
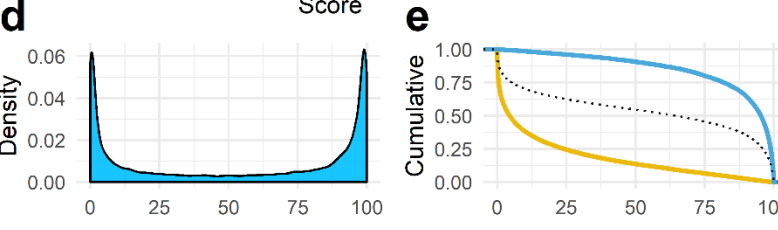
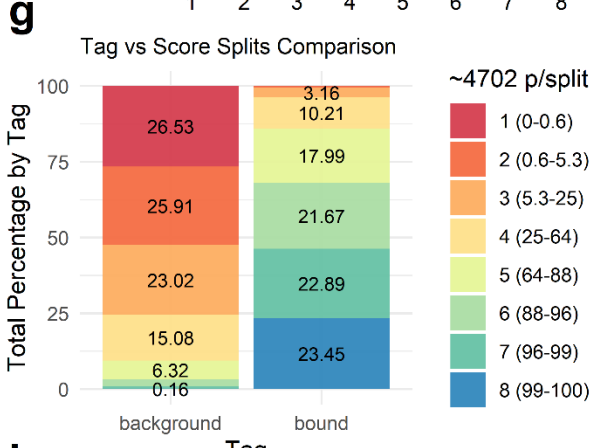
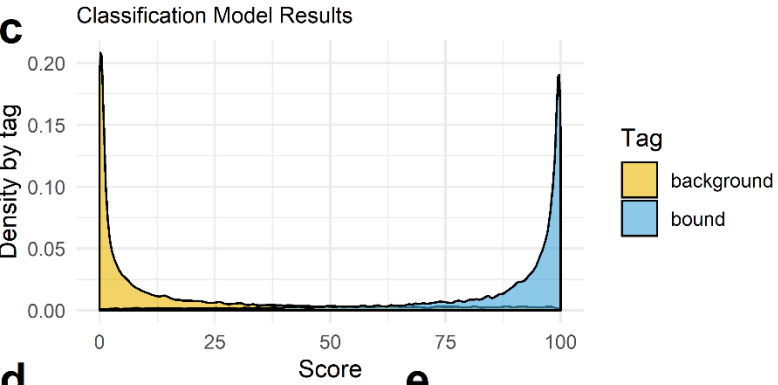
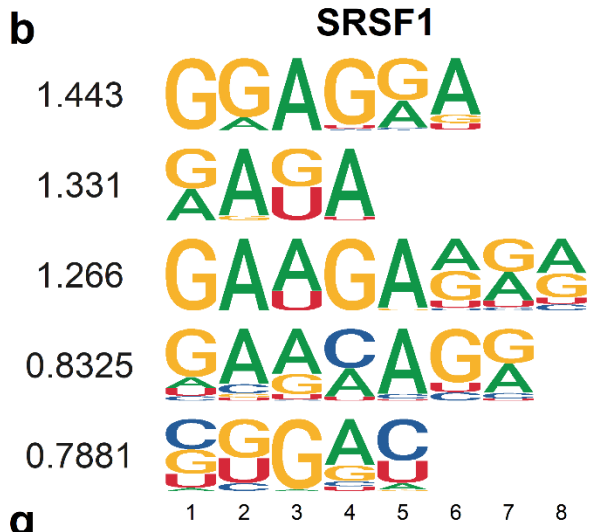
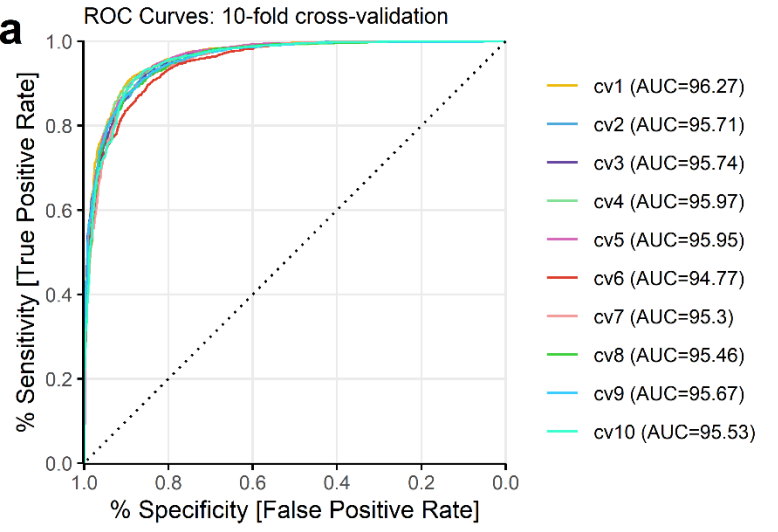
**g** Tag vs Score Splits Comparison



**h** Tag

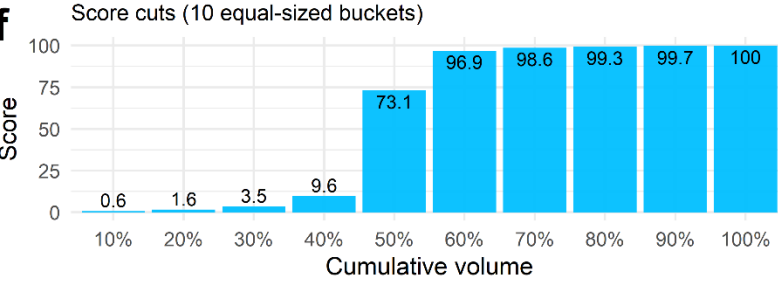
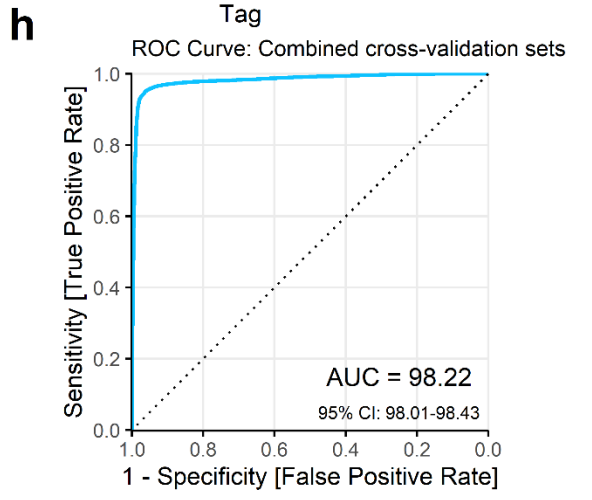
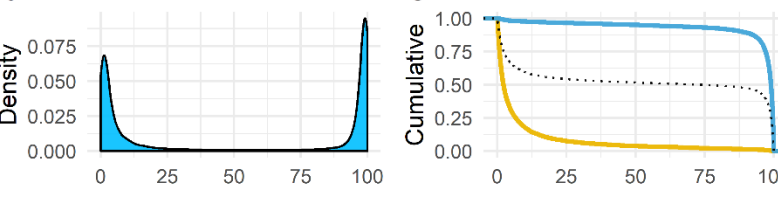
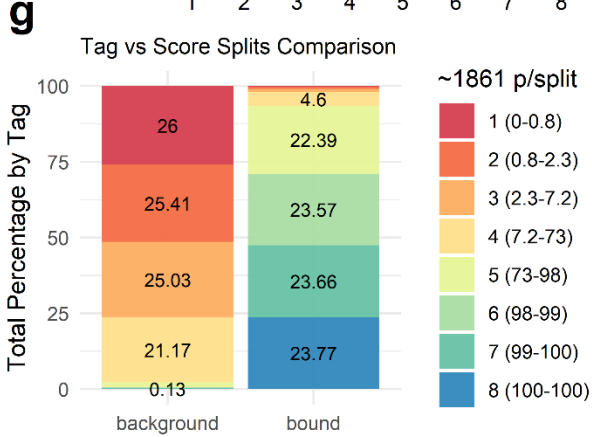
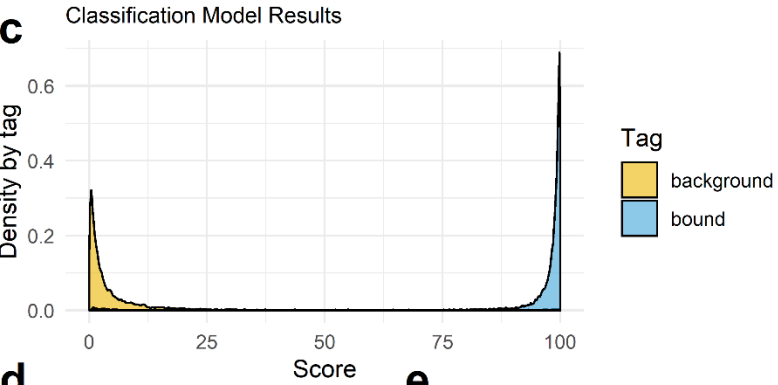
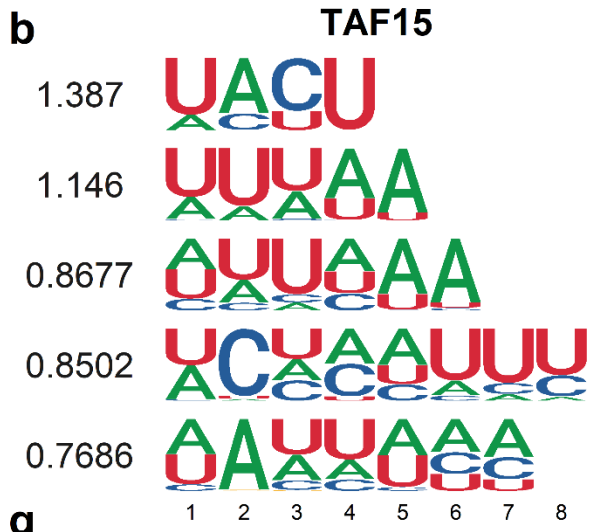
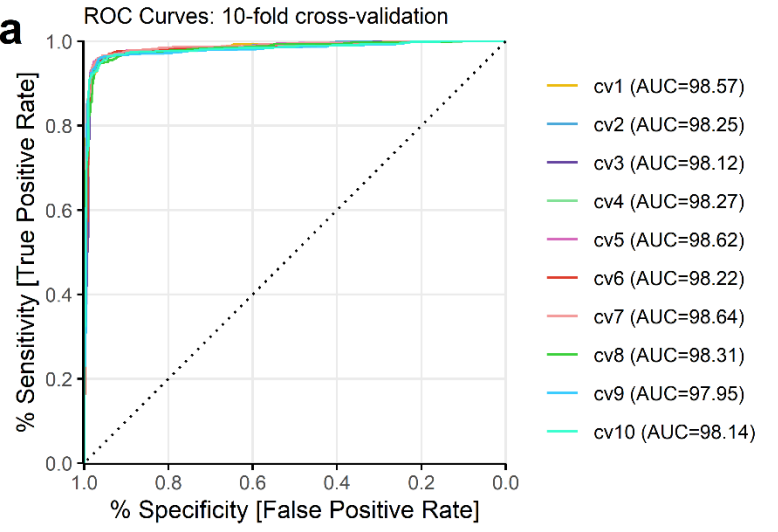


**Fig. S23**

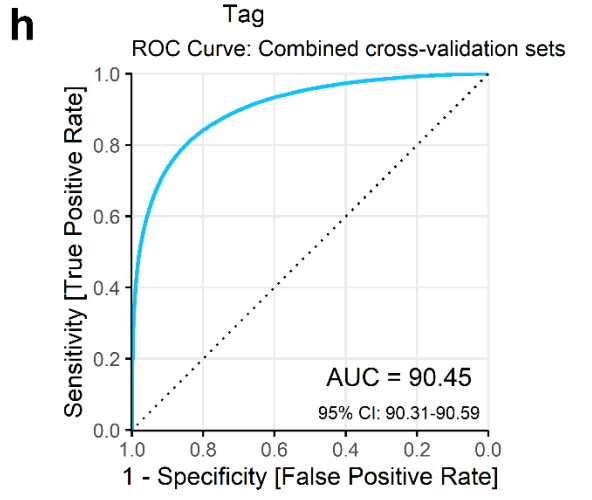
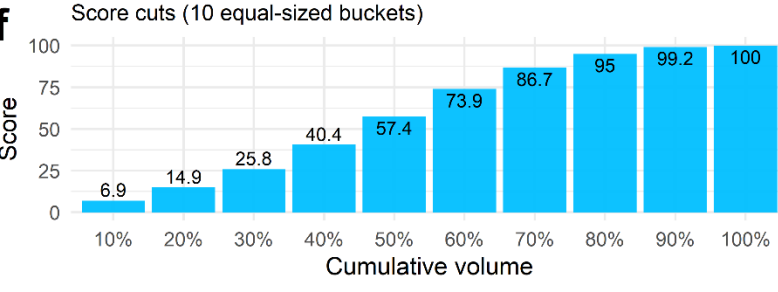
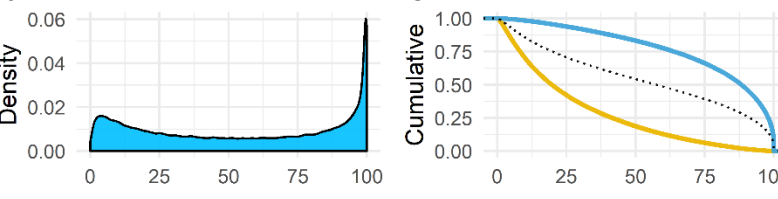
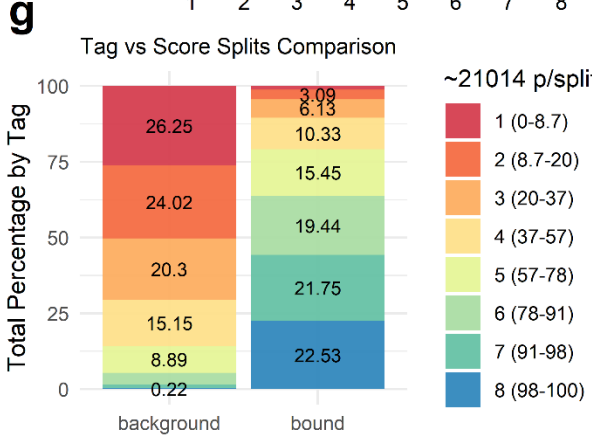
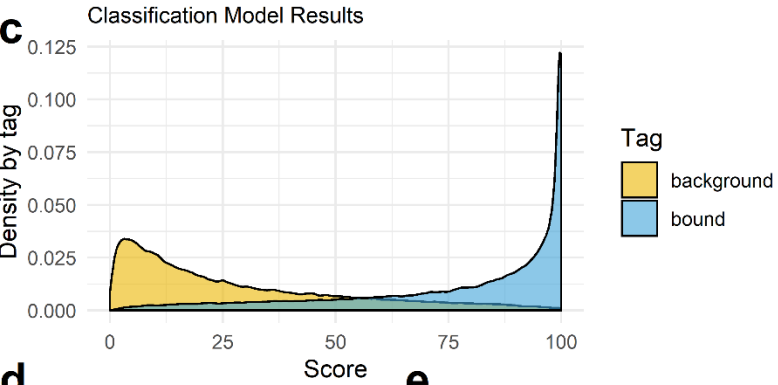
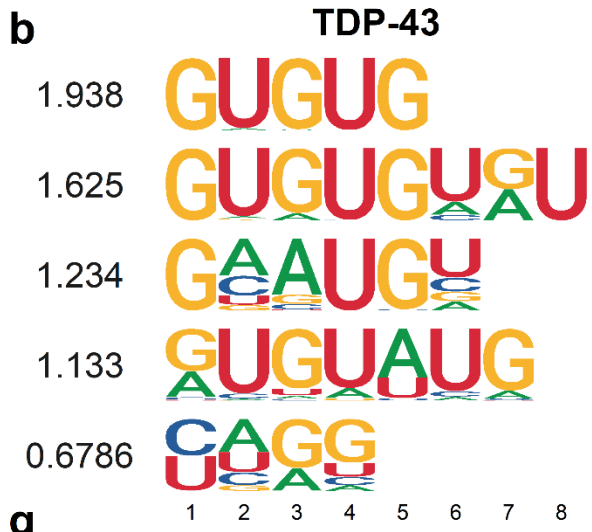
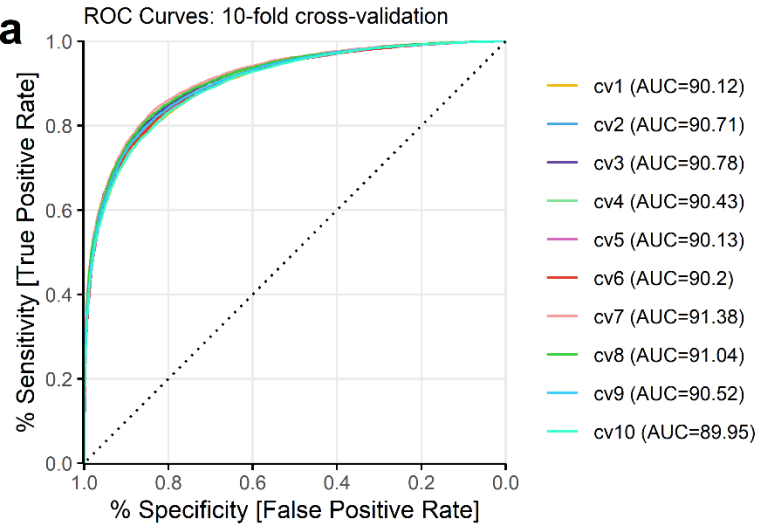




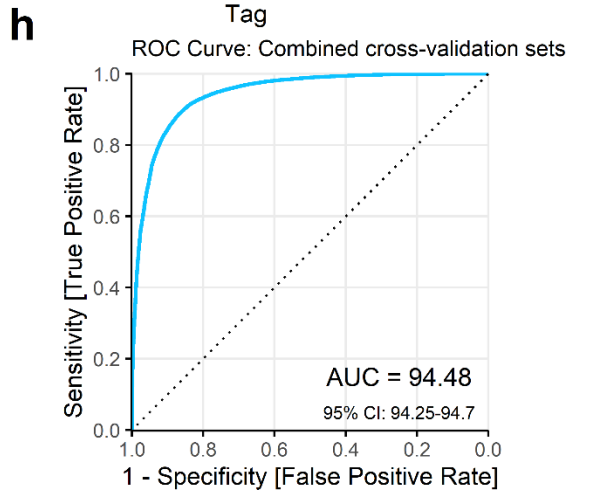
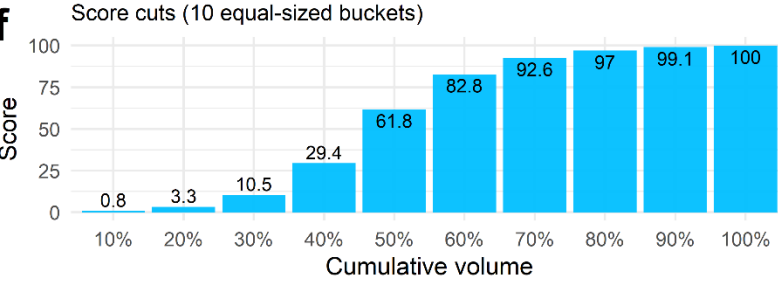
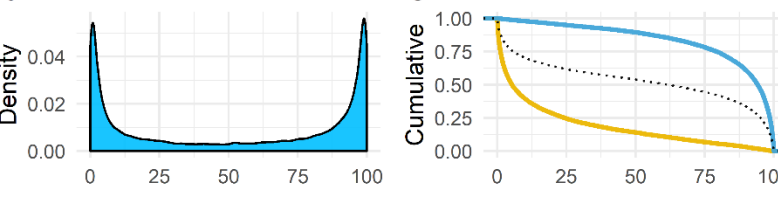
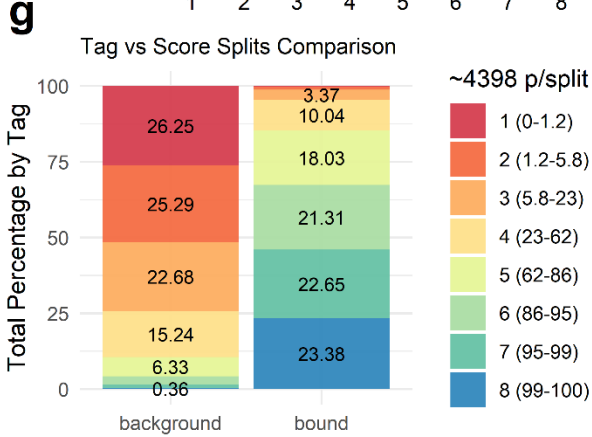
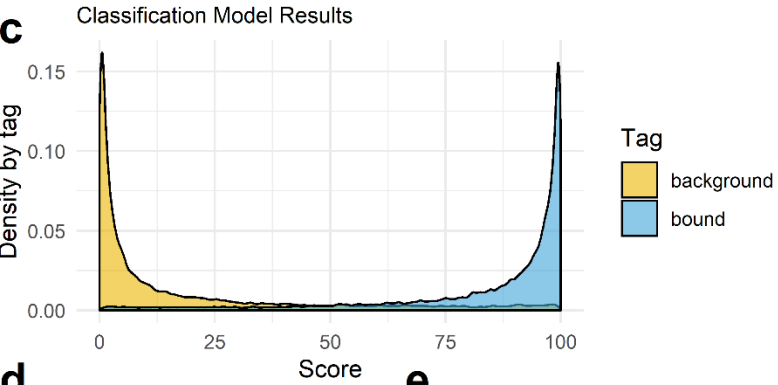
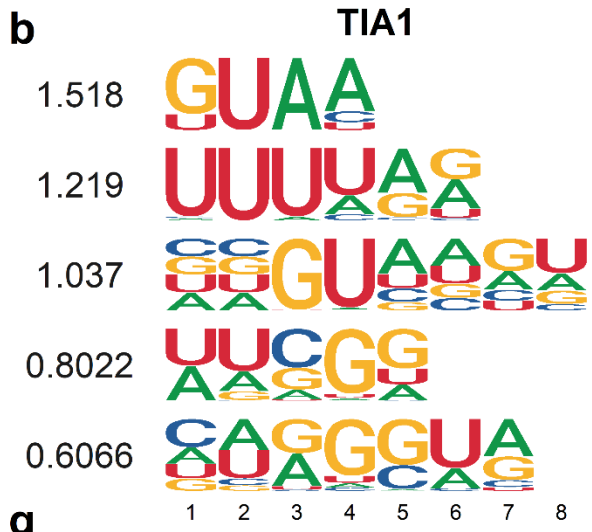
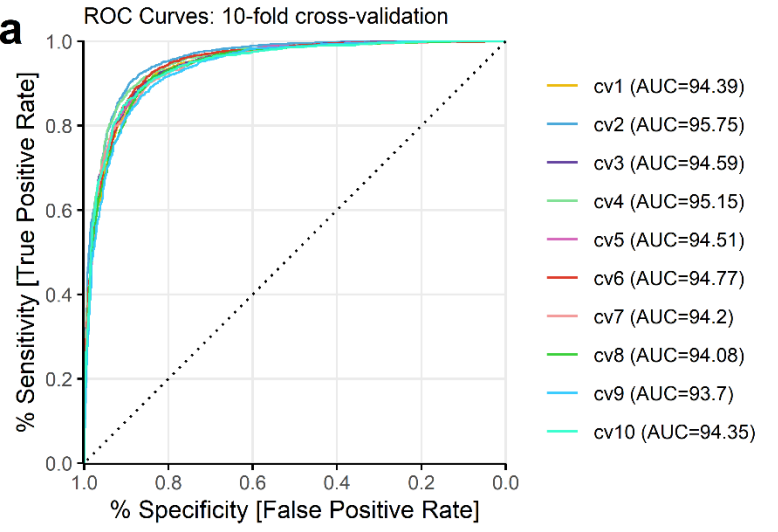
**Fig. S24**



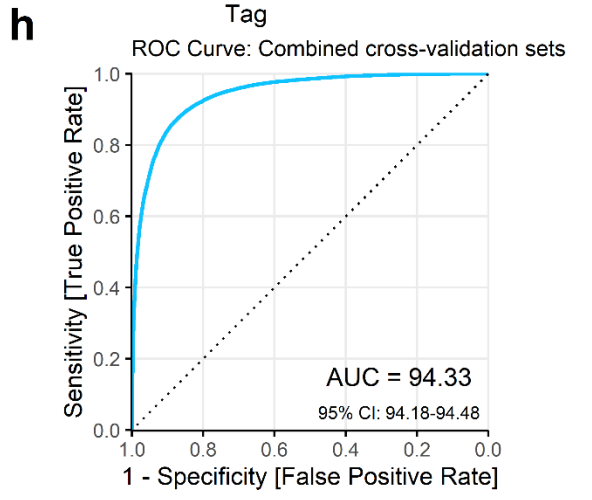
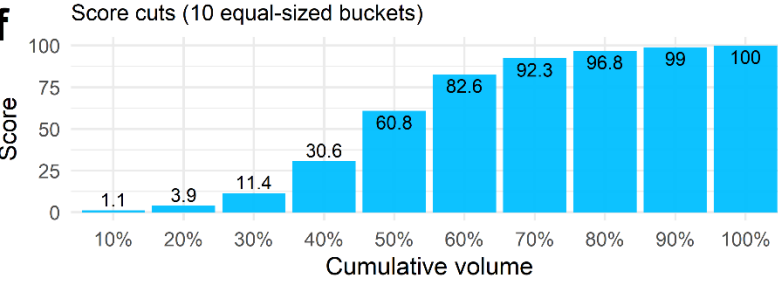
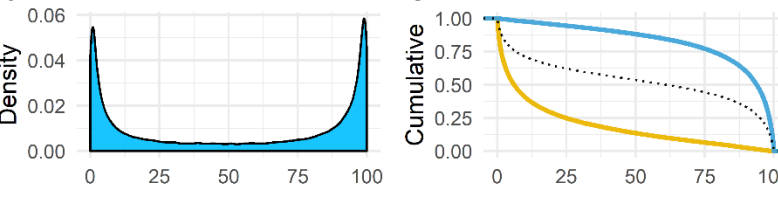
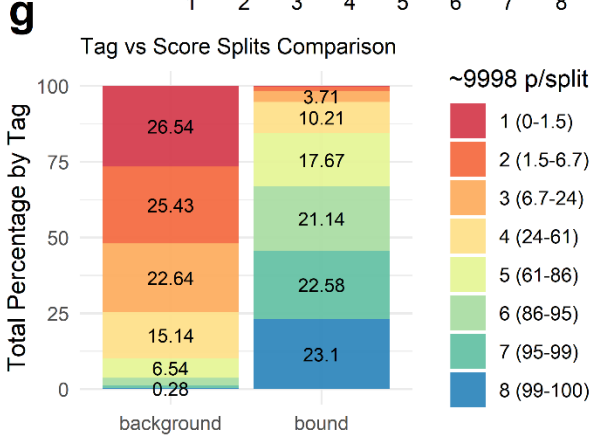
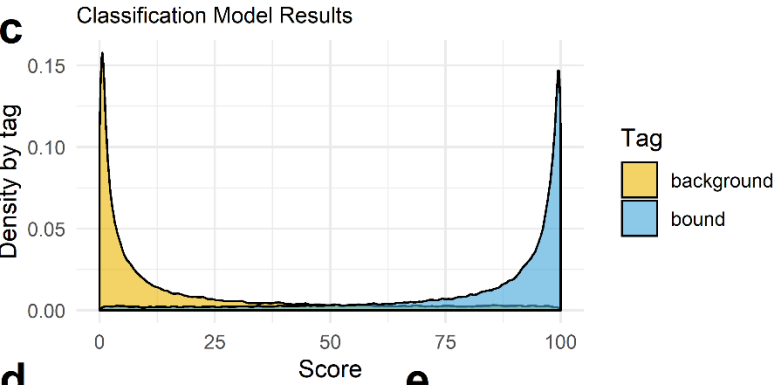
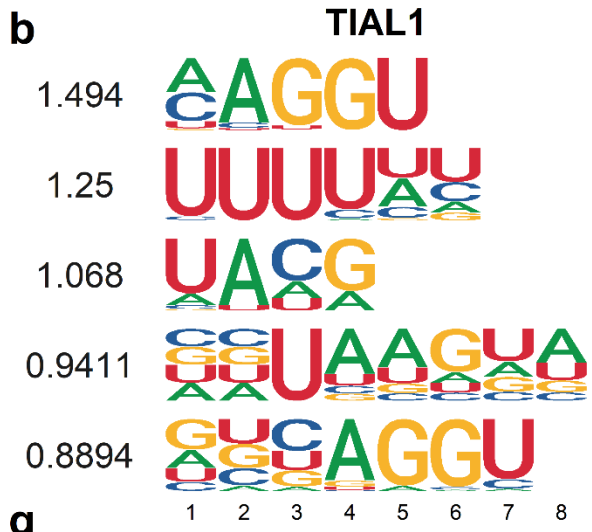
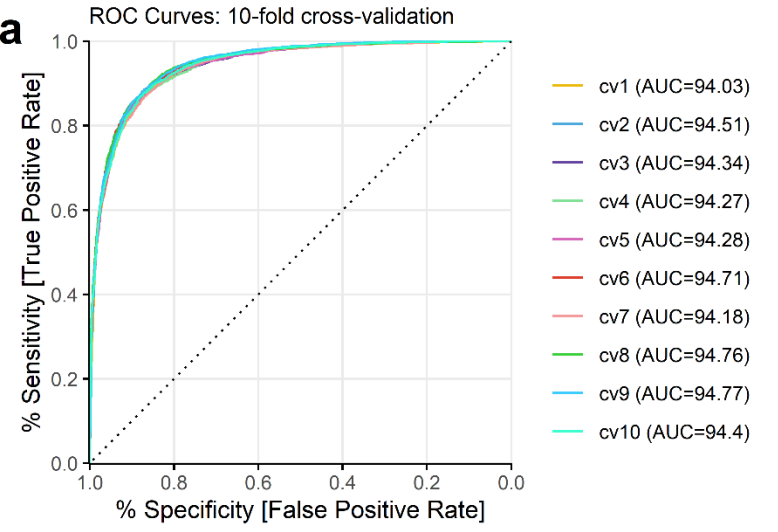
**Fig. S25**



**Fig. S26**

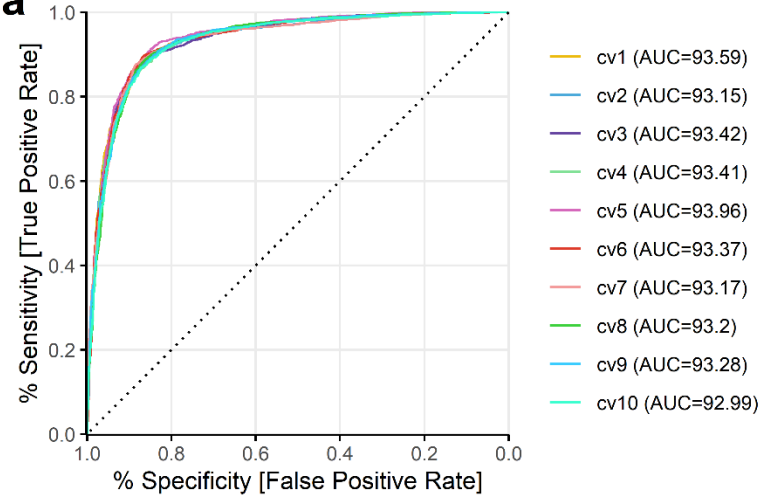


**Fig. S27**

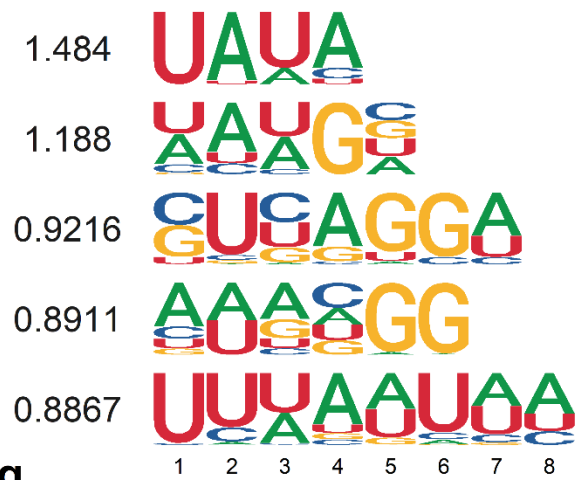


**Fig. S28**

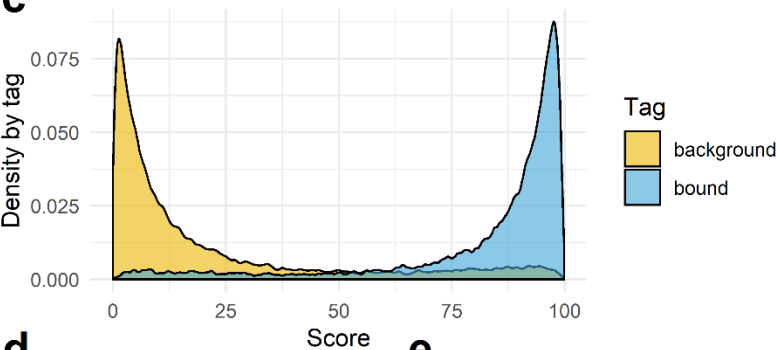
**a** ROC Curves: 10-fold cross-validation



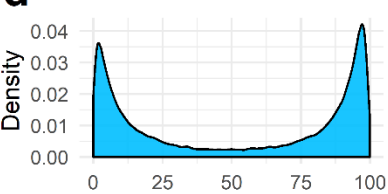
**b** ZC3H7B



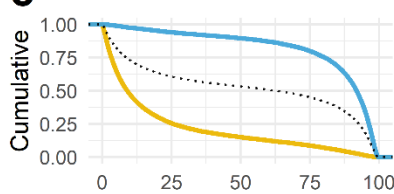
**c** Classification Model Results



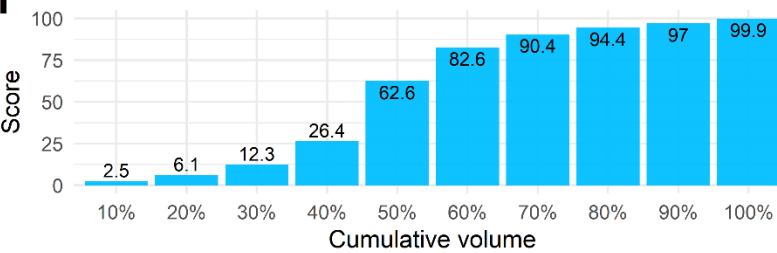
**d** Density vs Score



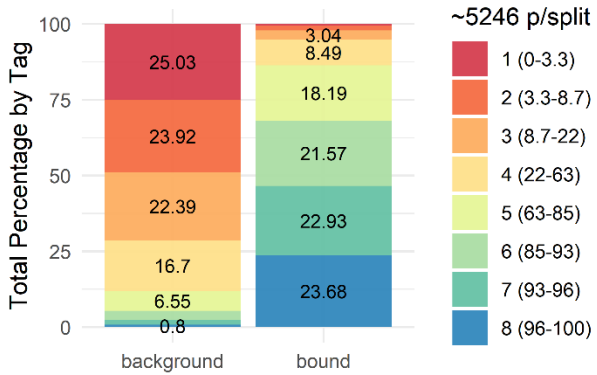
**e** Cumulative vs Score



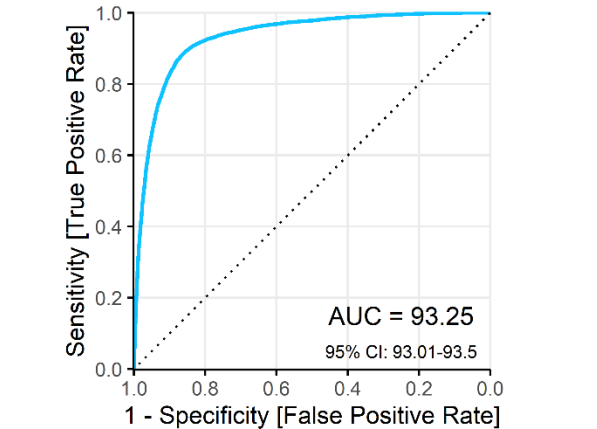
**f** Score cuts (10 equal-sized buckets)



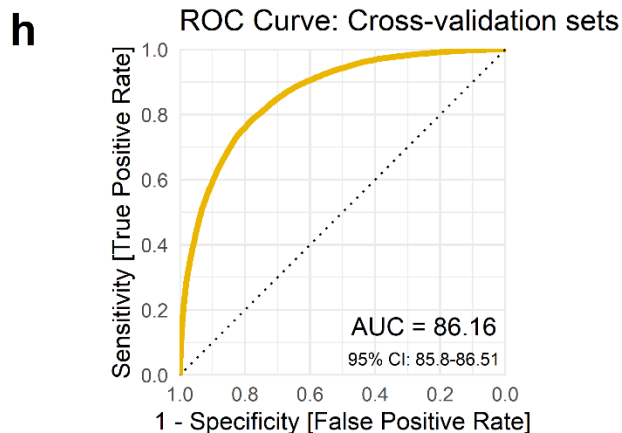
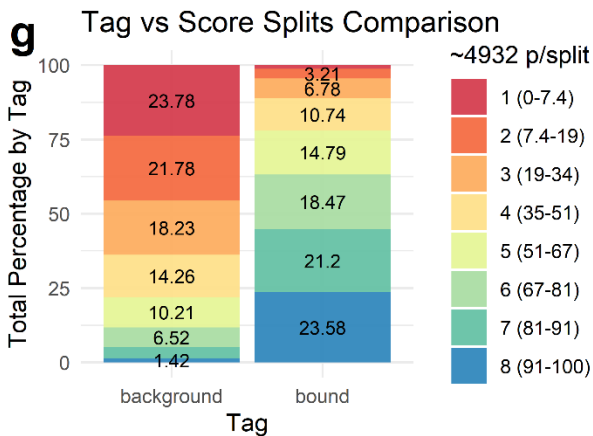
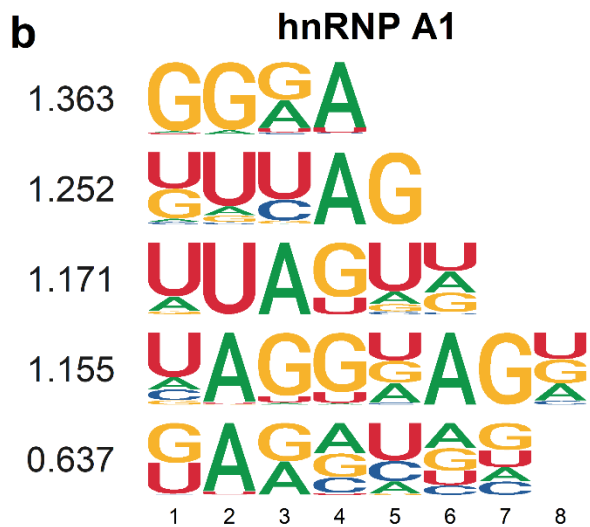
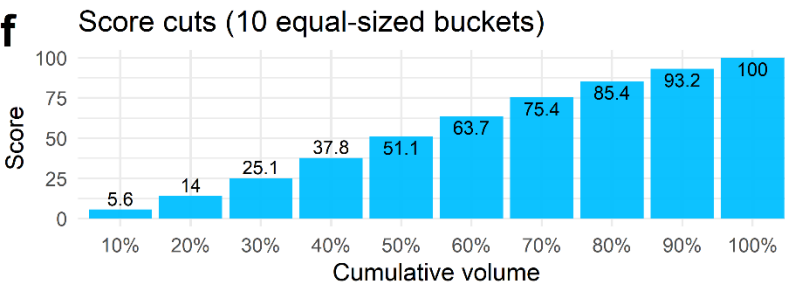
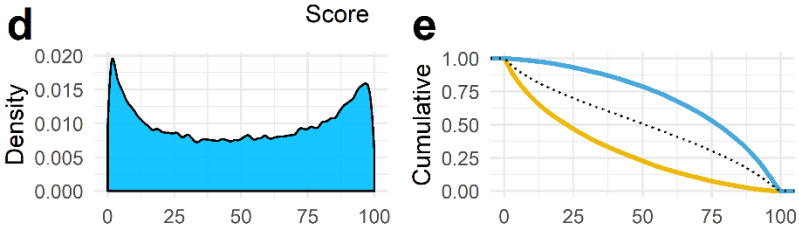
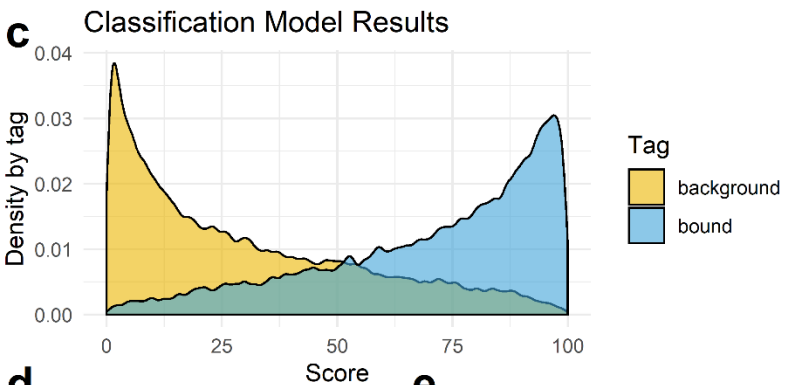
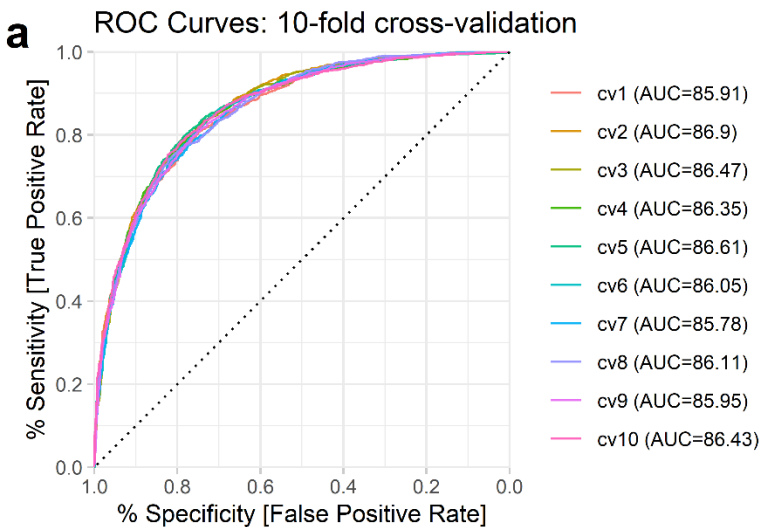
**g** Tag vs Score Splits Comparison



**h** ROC Curve: Combined cross-validation sets

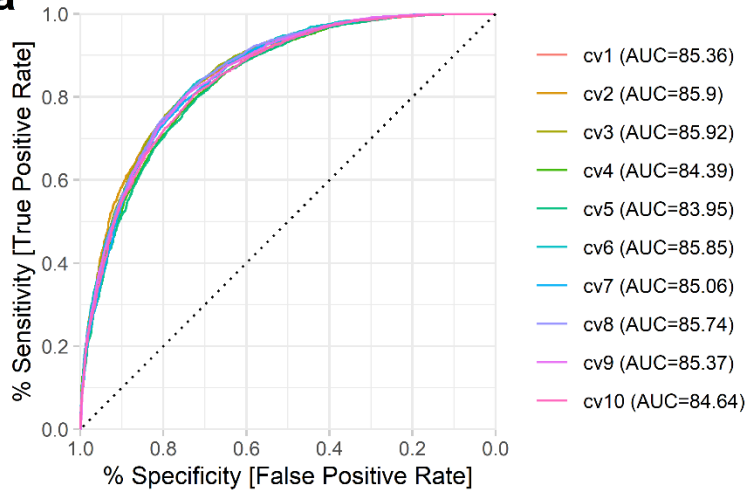


**Fig. S29**

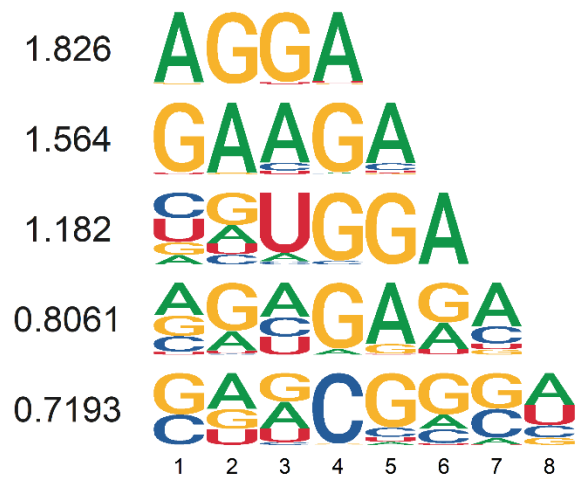


**Fig. S30**

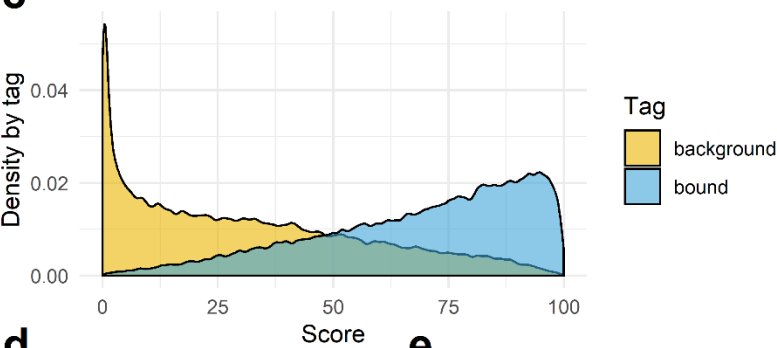
**a** ROC Curves: 10-fold cross-validation



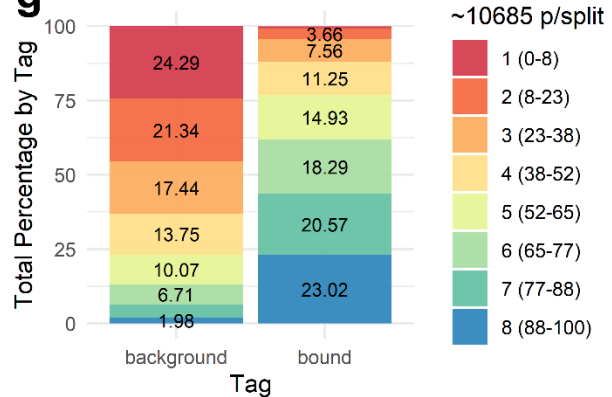
**b** SRSF1



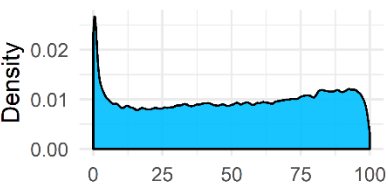
**c** Classification Model Results



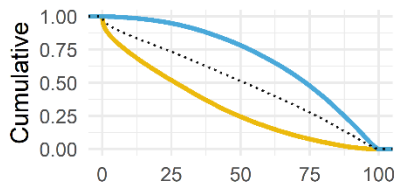
**g** Tag vs Score Splits Comparison



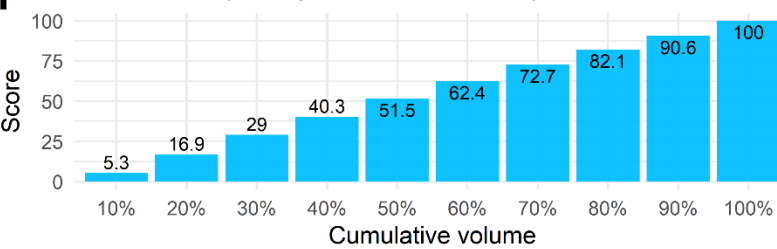
**d**



**e**



**f** Score cuts (10 equal-sized buckets)



**h** ROC Curve: Cross-validation sets

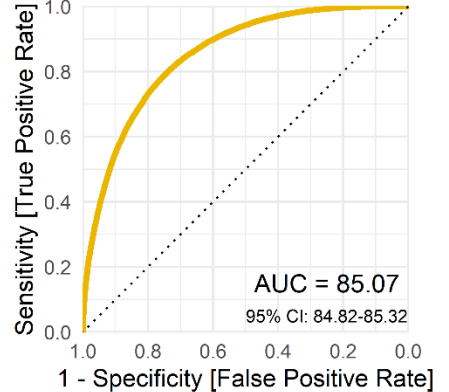


Fig. S31

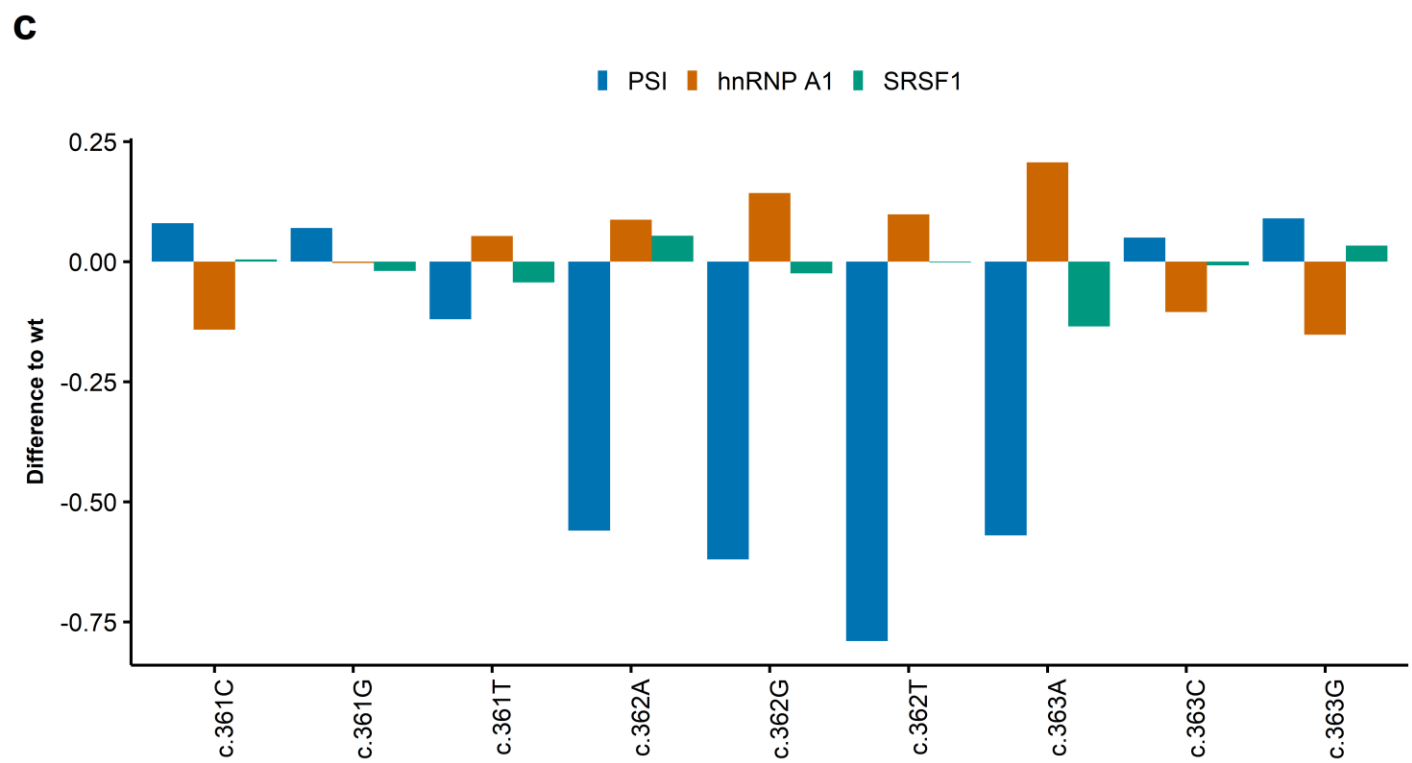
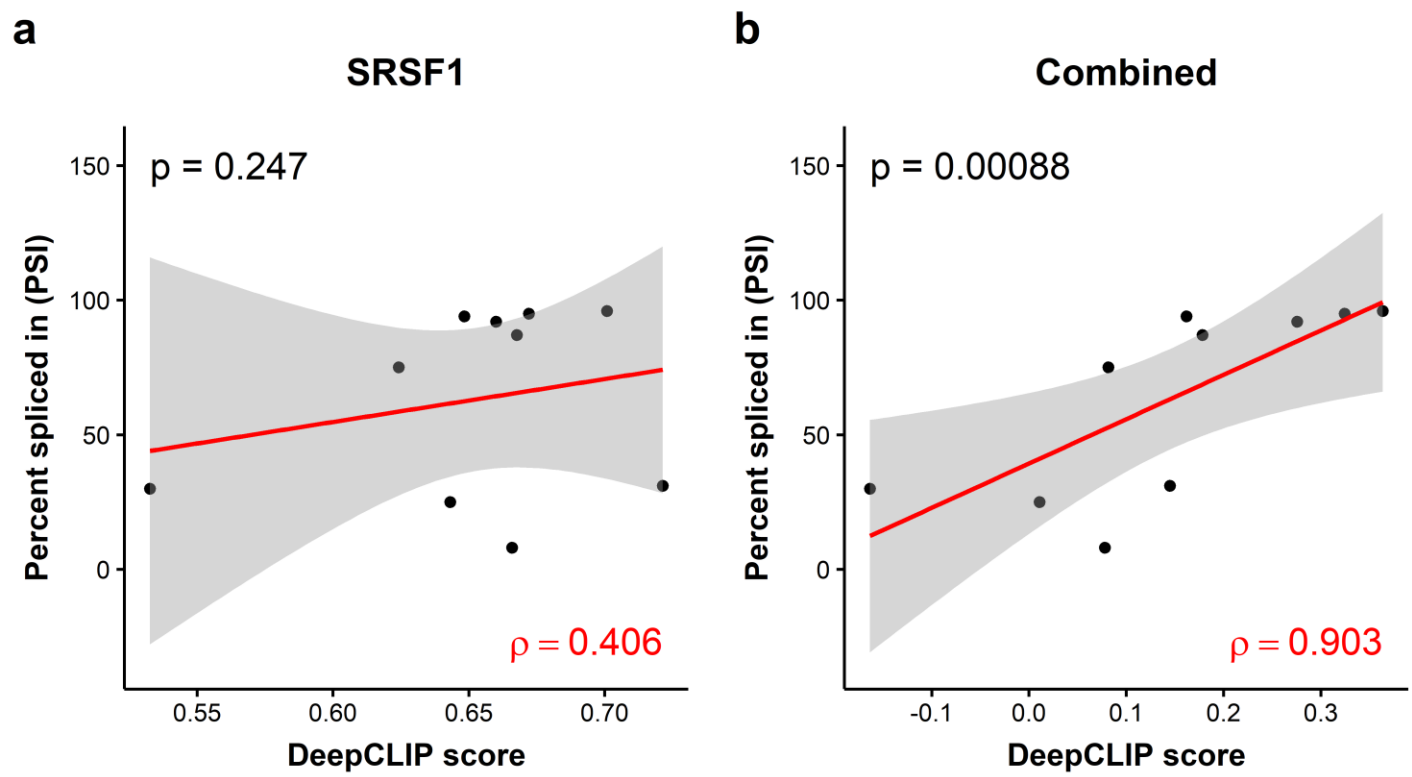
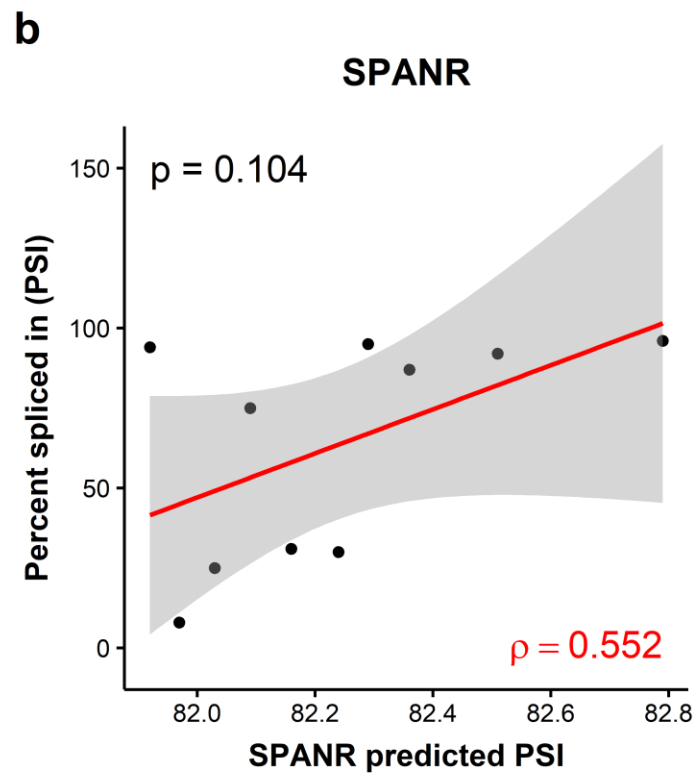
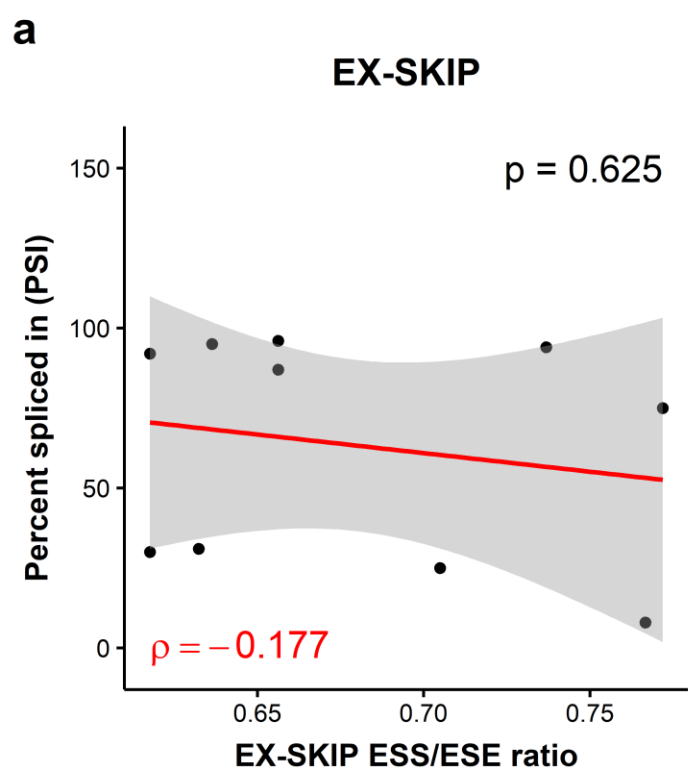




Fig. S32



**Fig. S33**

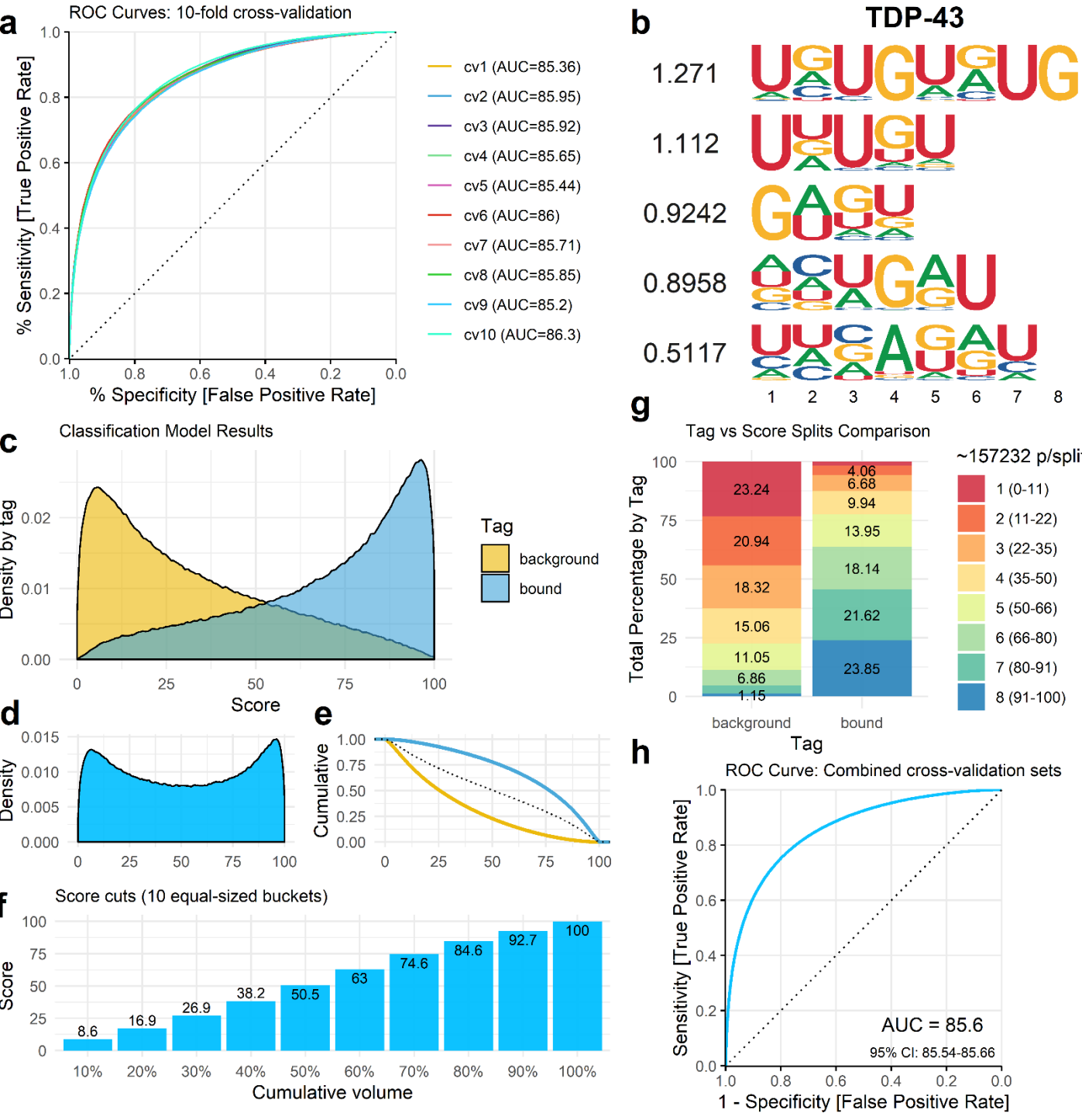


Fig. S34

hnRNP A1 SPRi binding plots with CLAMP model data

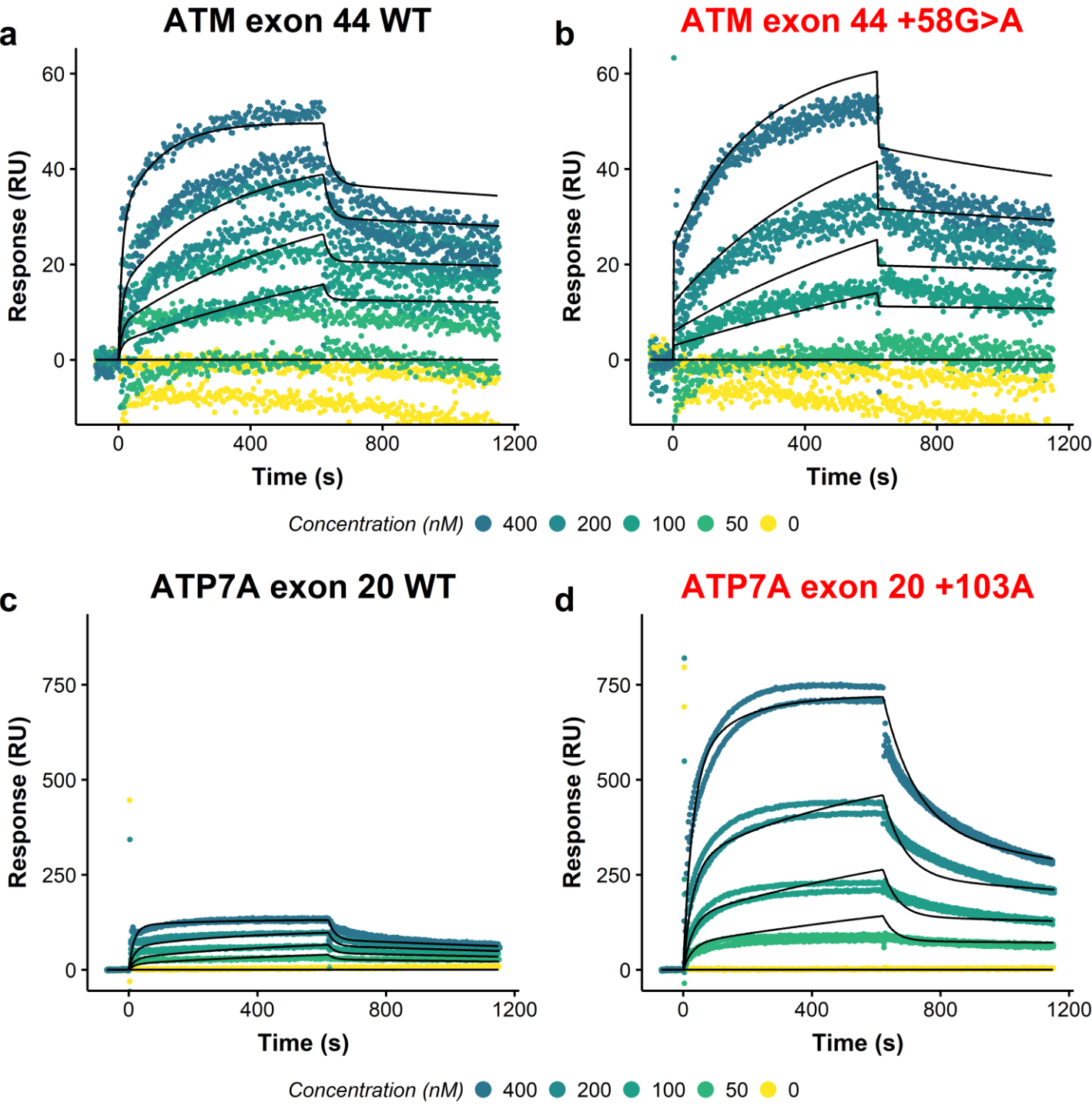


Fig. S35

hnRNP A1 SPRi binding plots with CLAMP model data

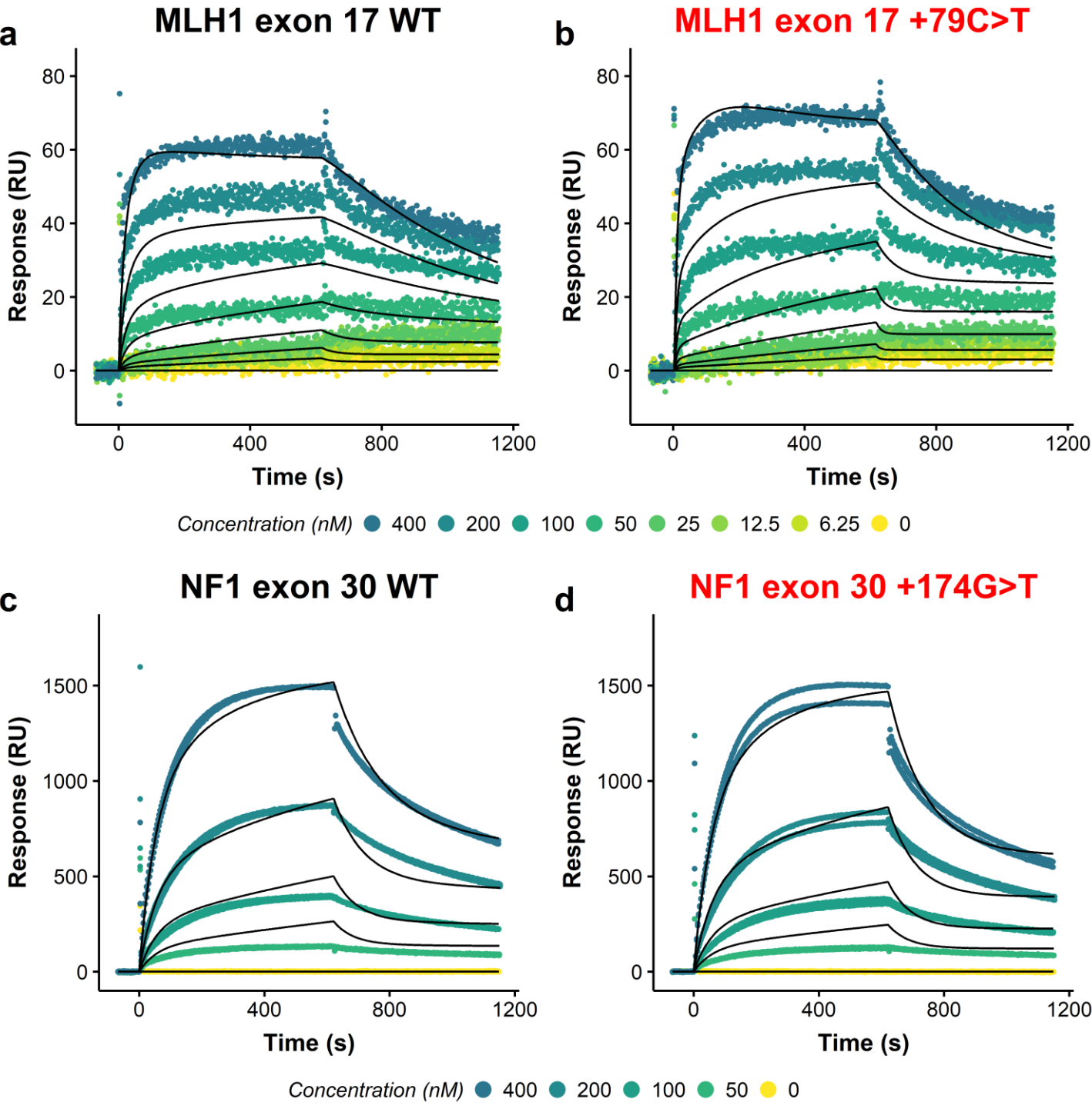


Fig. S36

hnRNP A1 SPRi binding plots with CLAMP model data

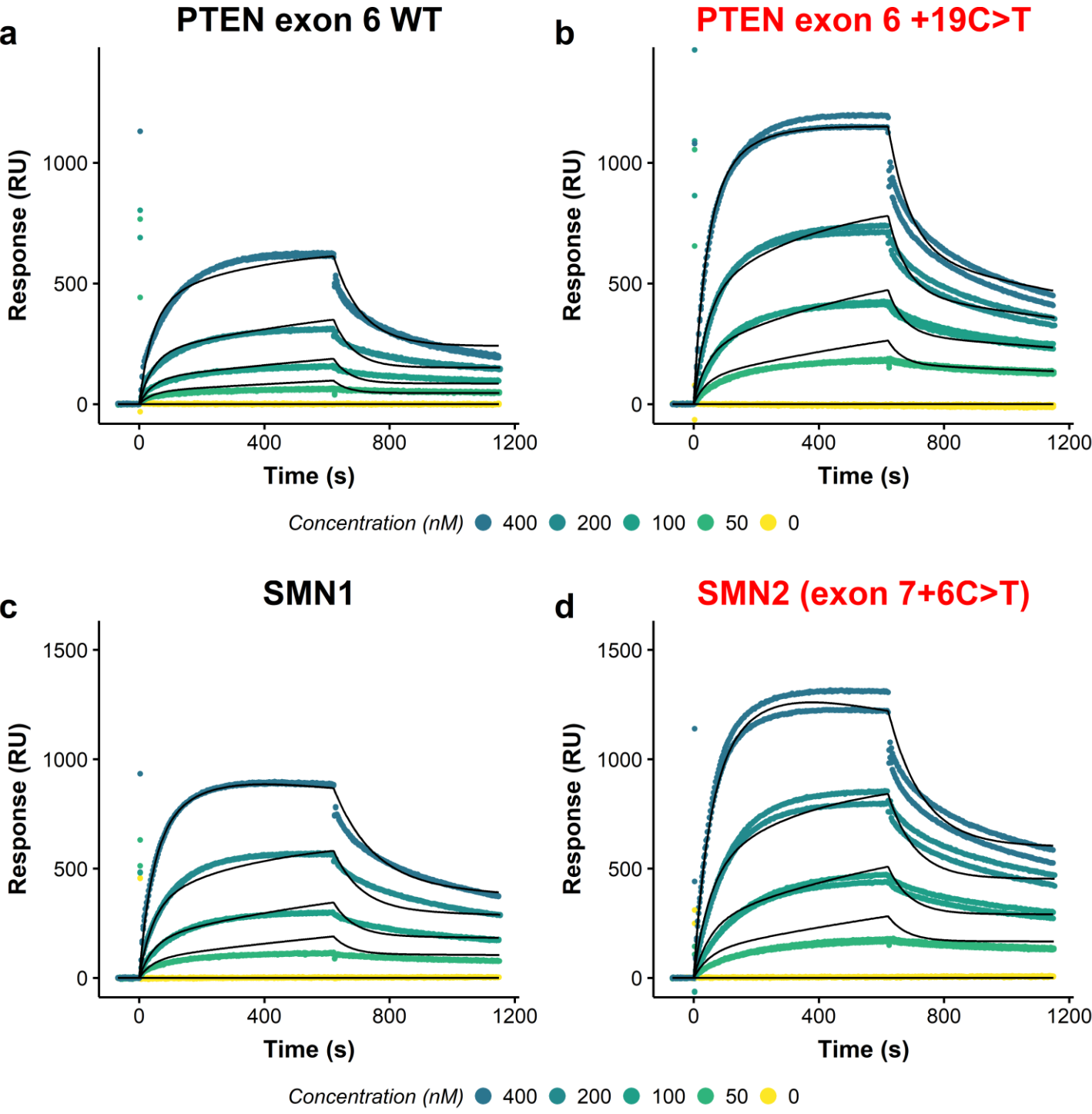


Fig. S37

SRSF1 SPRi binding plots with CLAMP model data

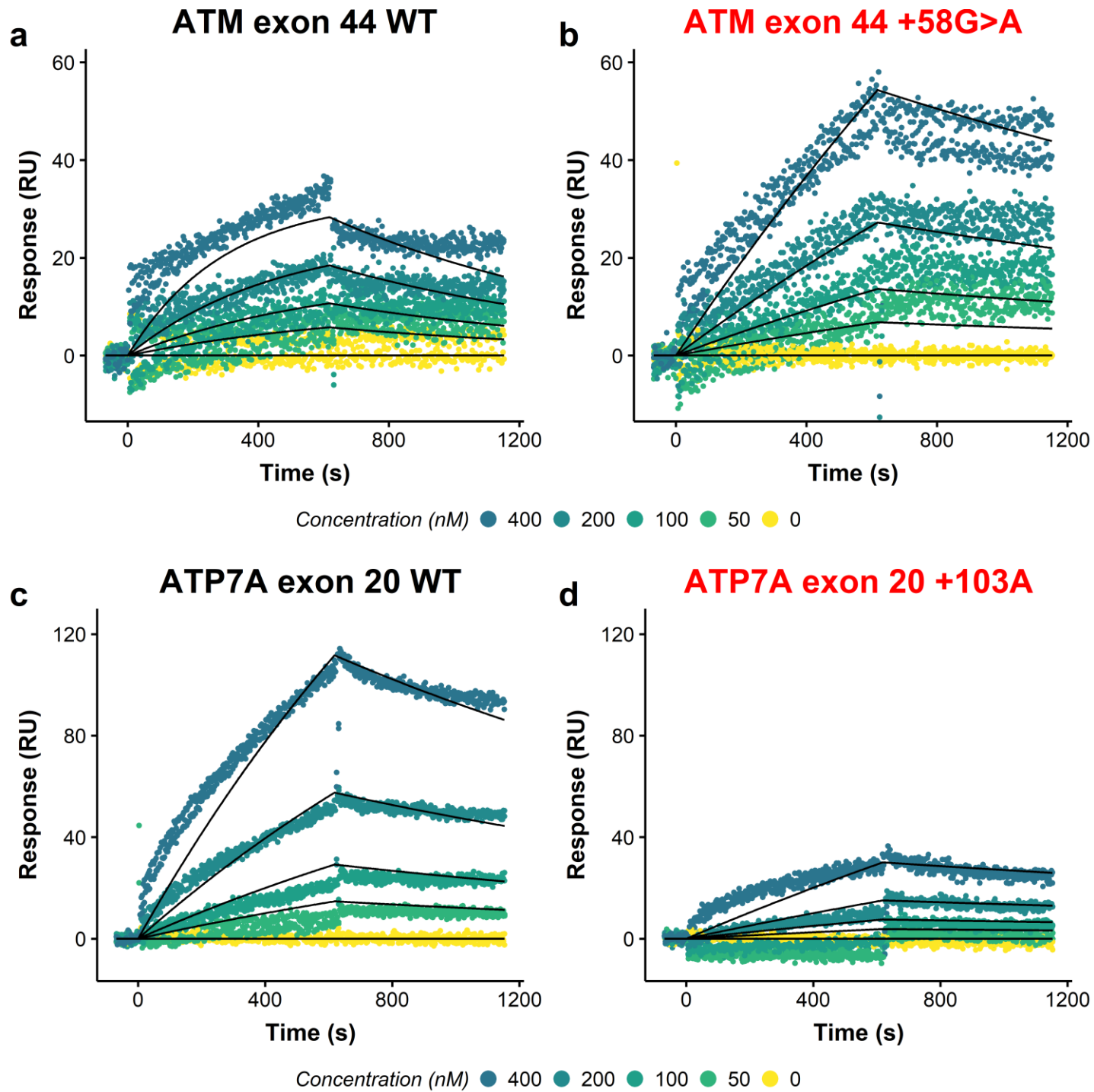


Fig. S38

SRSF1 SPRi binding plots with CLAMP model data

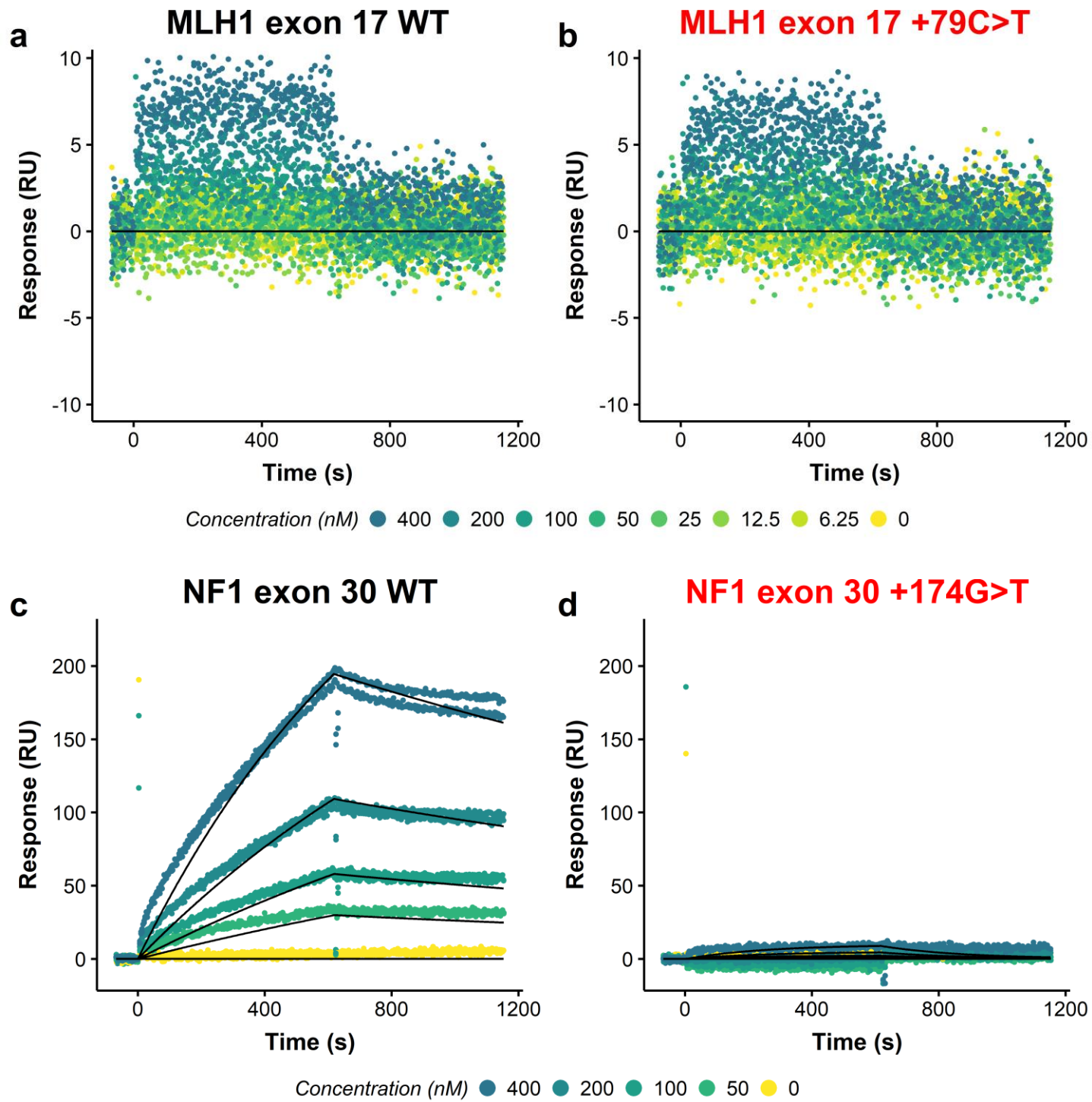


Fig. S39

SRSF1 SPRi binding plots with CLAMP model data

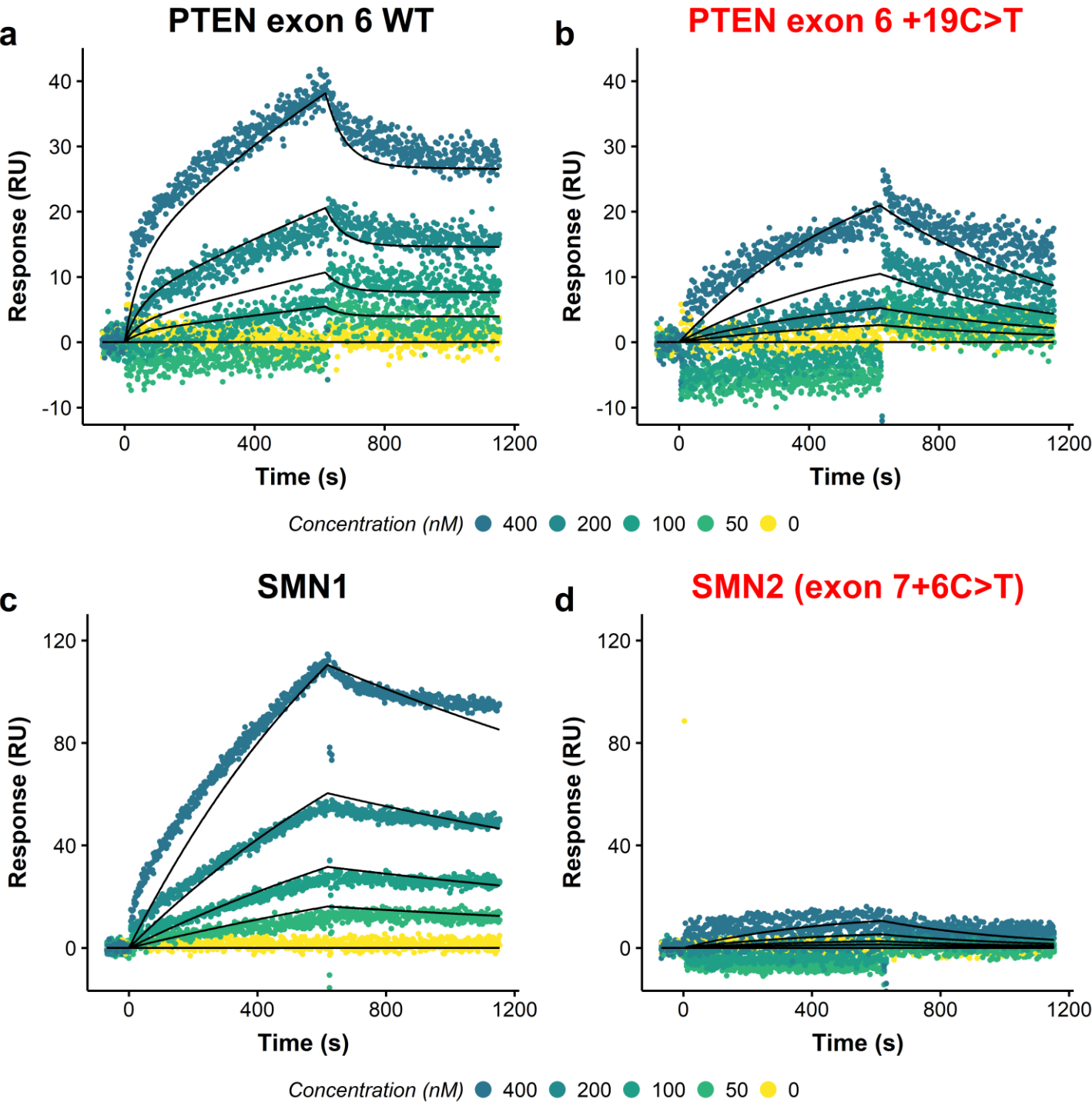
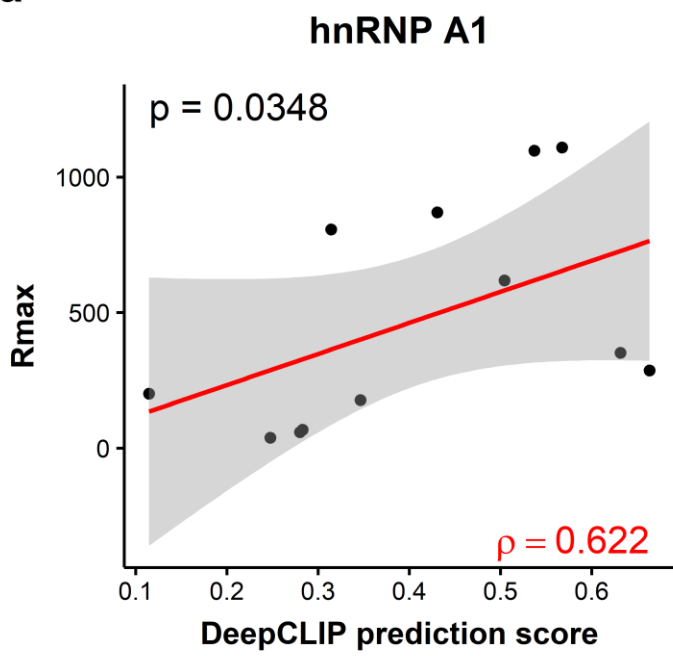




Fig. S40

**a**



**b**

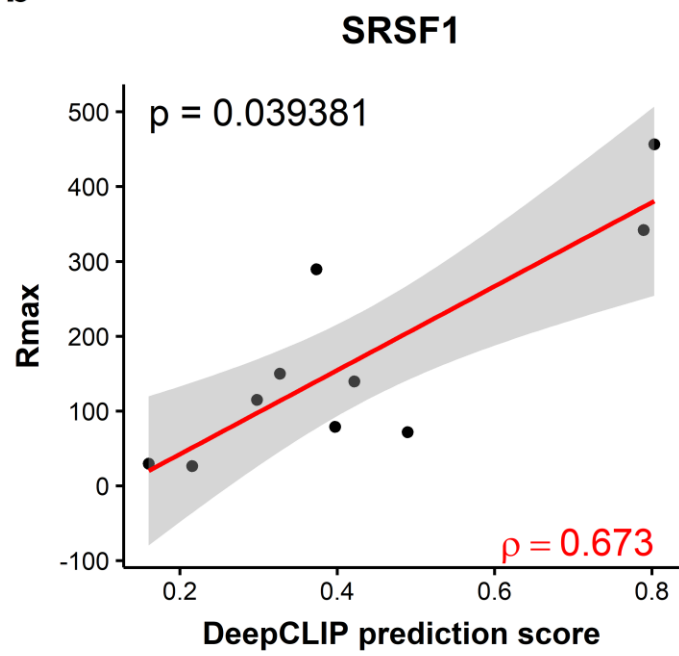


Fig. S41

