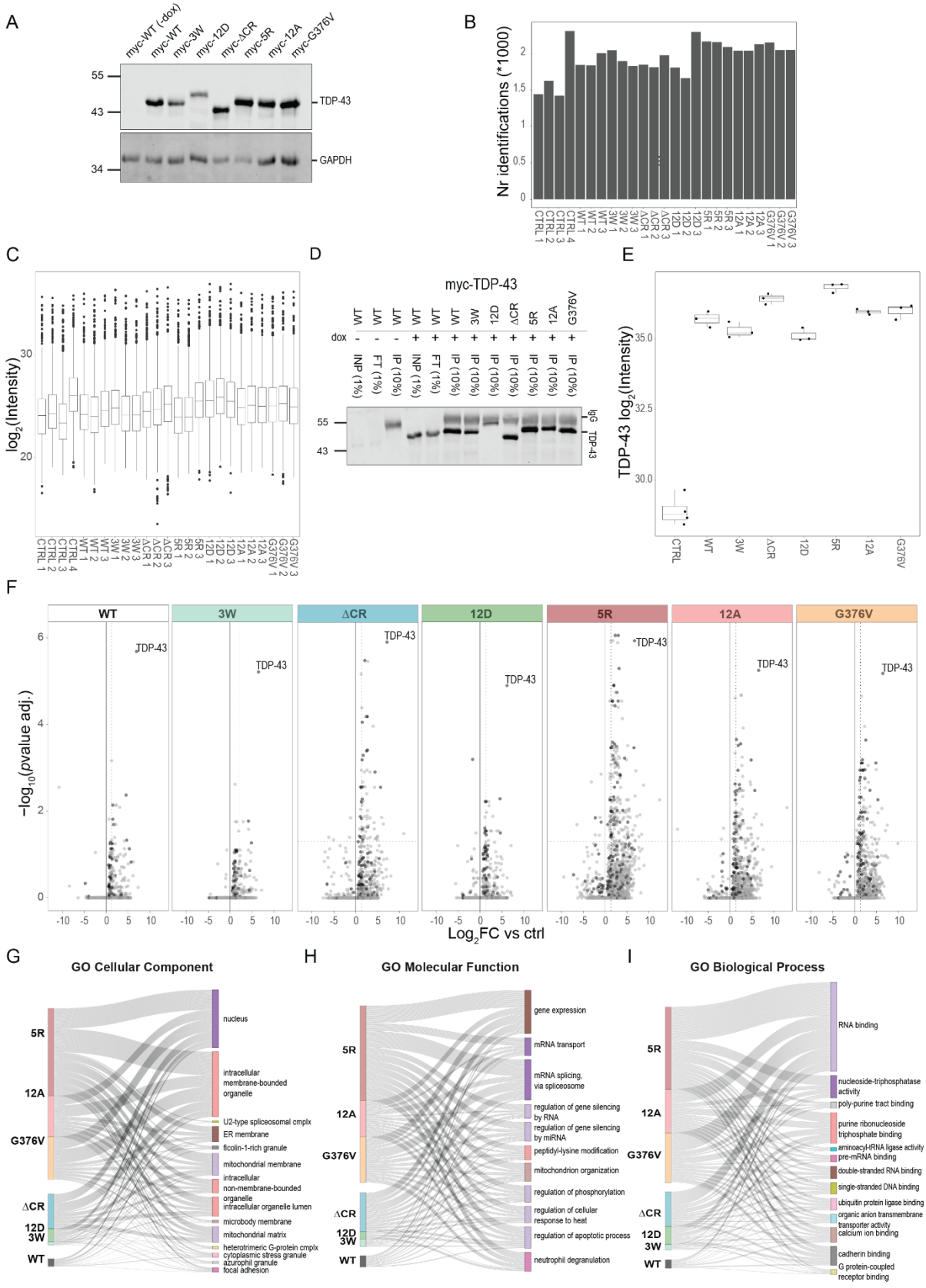
Supplementary information



Supplementary Figure 1

Supplementary Fig. 1

A. Western blot showing expression levels of the bait (TDP-43) across the analyzed variants. Antibodies used: mouse anti-TDP-43 (60019-2-Ig, Proteintech), rabbit anti-GAPDH (10494-1-AP, Proteintech).

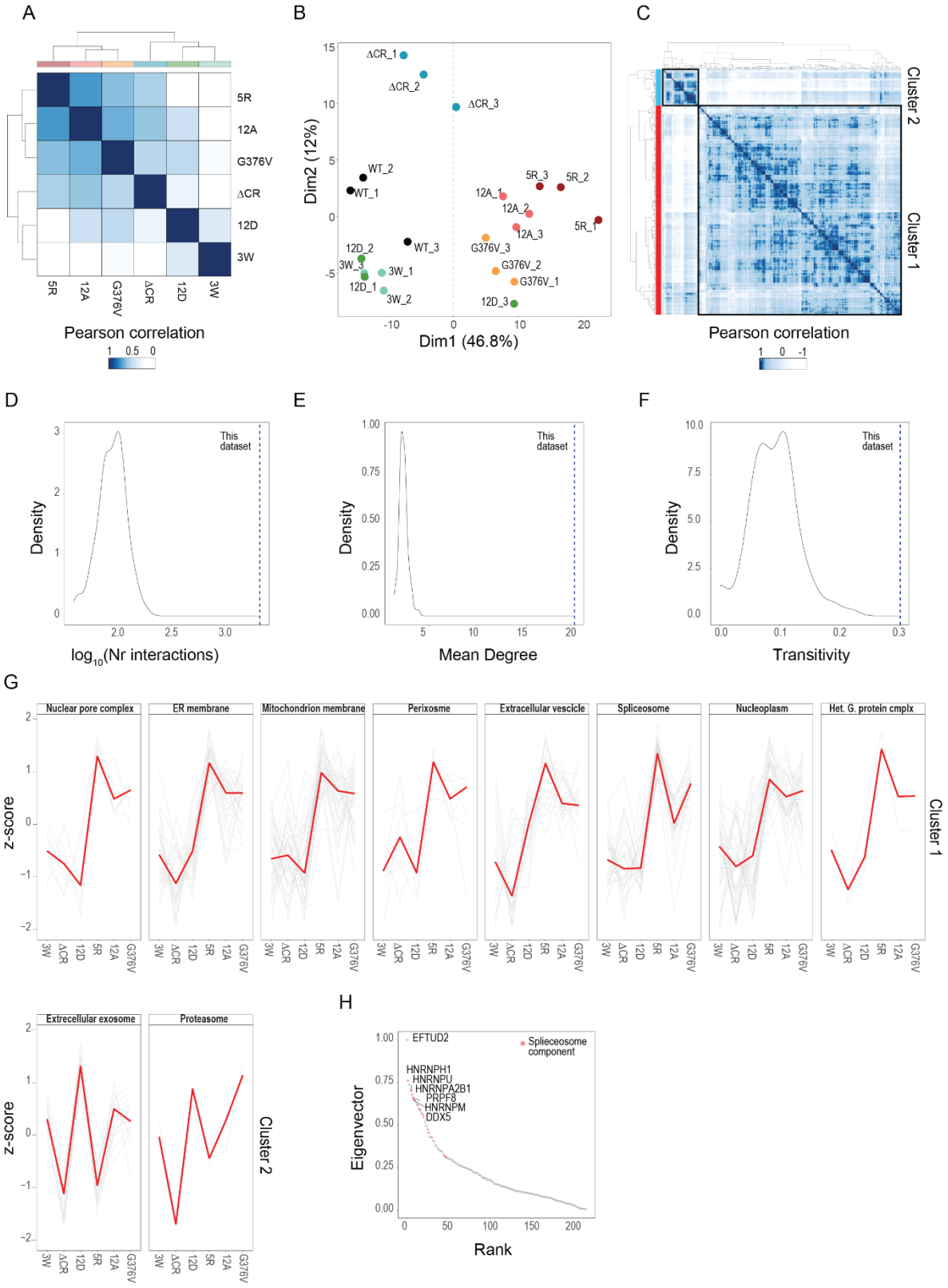
B, C. Number of protein identifications (B) and their abundance levels (C) across analyzed conditions.

D. Western blot showing the levels of the bait (TDP-43) across purification steps. INP: Input (1% of the total amount), FT: Flow-through (1% of the total amount), IP: Immunoprecipitation (10% of the total amount). Antibodies used: mouse anti-TDP-43 (60019-2-Ig, Proteintech), rabbit anti-GAPDH (10494-1-AP, Proteintech).

E. Abundance of the bait (TDP-43) purified in the different conditions.

F. Volcano plot showing the proteins enriched in the different purifications for the TDP-43 variants. High-confidence interactors (HCIs, TDP-43 interactors identified in at least three independent datasets) are shown in black, and abundance of the bait (TDP-43) is highlighted. Threshold for interactor definition: log2FC > 1 and *p*value adjusted < 0.05.

G, H, I. Flow diagram showing the GO enrichment for each purification (CC: Cellular Component (G); MF: Molecular Function (H); BP: Biological Process (I)). For the GO analysis, only terms with at least 5 (CC, MF) and 10 (BP) genes associated and identified as a parental term based on similarity are shown. For each gene only the term with the lowest *p*value is associated.



Supplementary Figure 2

Supplementary Fig. 2

A. Pearson correlation matrix of TDP-43 purifications. Euclidean distances between conditions were calculated from the interactor intensities normalized for TDP-43 levels to guide hierarchical clustering with the Ward.D2 method.

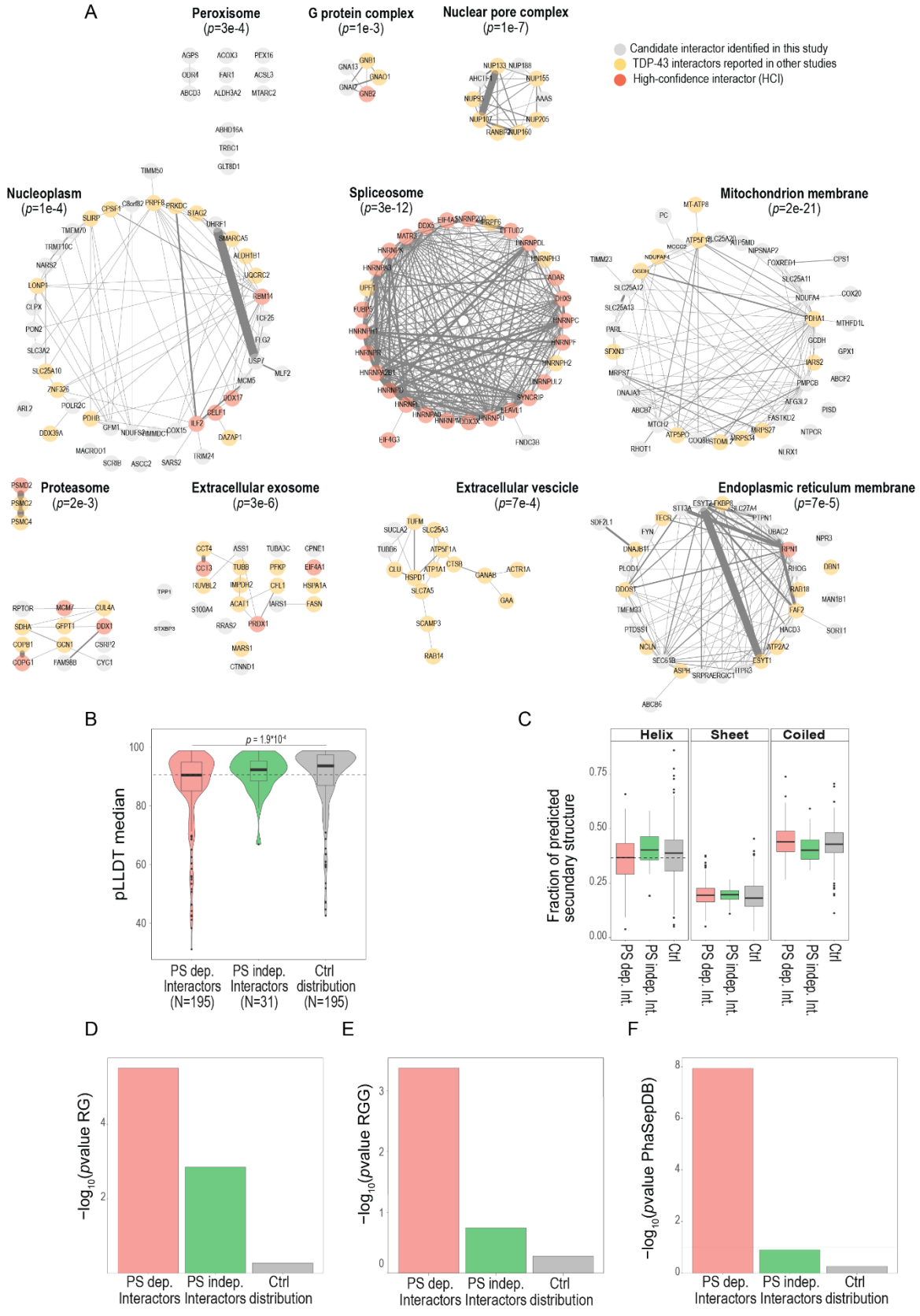
B. Principal component analysis of TDP-43 purifications. Eigenvector of the two principal components (Dim1 and Dim2) visualize the proportion of variance explained. TDP-43 variants are highlighted in different colors, and every dot represents a replicate condition.

C. Hierarchical clustering of TDP-43 interactors. The abundance of TDP-43 interactors was z-score-transformed and the Pearson correlation coefficients were calculated across all conditions.

D, E, F. Characterization of the network generated and comparison to a distribution of control networks generated by sampling the interactions of 226 random proteins from BioGRID 100 times. The plots report the number of interactions (D), the average node degree (E) and the transitivity (F).

G. z-score profile of interactors according to their module assignment. Profile of every protein is shown in grey, while the median profile is shown in red.

H. Organization of the network generated; proteins are ranked according to their eigenvector centrality. Proteins associated to the spliceosome complex (according to the QuickGO annotation) are highlighted in red.



Supplementary Figure 3

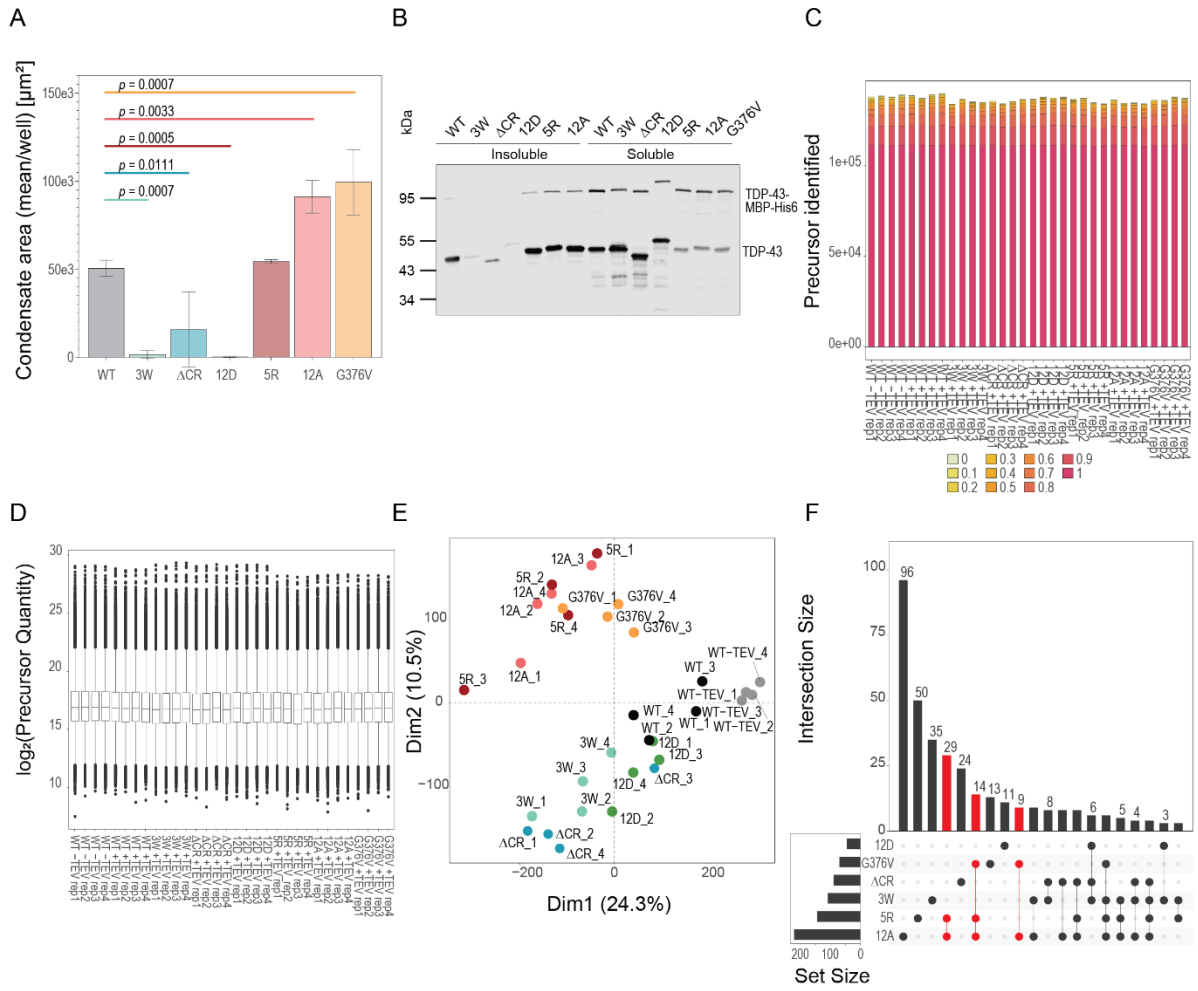
Supplementary Fig. 3

A. The protein-protein interaction network of TDP-43 is organized into modules based on reported interactions and protein localization. Proteins within each module are displayed using a circular layout. Edges between nodes represent interactions identified in at least two independent experiments (BioGRID v. 5.0.253). Nodes are color-coded to distinguish novel, reported and high-confidence interactors, represented in grey, yellow, and red, respectively.

B. Distribution of median of AlphaFold2 pLLDT confidence scores. The median-per-protein pLLDT scores were calculated for proteins identified as PS-dependent TDP-43 interactors (N = 195), PS-independent interactors (N = 31), and for proteins in the control group (N = 195, proteins that did not pass the interactor threshold). A statistical analysis was performed using an unpaired Wilcoxon test.

C. Fraction of the predicted secondary structure calculated for proteins identified as TDP-43 interactors: PS-dependent interactors (N =195), and PS-independent interactors (N = 31) or the control group (N = 195).

D, E, F. Enrichment analysis for PS-dependent and -independent TDP-43 interactors, and controls. Enrichment was evaluated for proteins containing at least one ‘RG’ motifs (D), ‘RGG’ motifs (E) and proteins annotated as associated to phase separation from PhaSepDB database (v. 2.1) (F). Reported values correspond to the enrichment calculated using a hypergeometric test.



Supplementary Figure 4

Supplementary Fig. 4

A. Quantification of the TDP-43 condensed fraction area per well for the images shown in Fig. 4B. 12 fields of view were quantified per condition. Mean values of 3 experimental replicates are shown. Error bars indicate standard deviation. Significance was determined using one-way ANOVA with a Dunnett's multiple comparisons test to WT. *p*values are indicated.

B. Western blot of TDP-43 in supernatant (soluble) and pellet (insoluble) fractions after centrifugation for the indicated TDP-43 variants. The band above 95 kDa represents TDP-43-His_6, while the band around 43 kDa corresponds to TDP-43 without the tag. Antibody used: rabbit anti-TDP-43 (80001-RR, Proteintech).

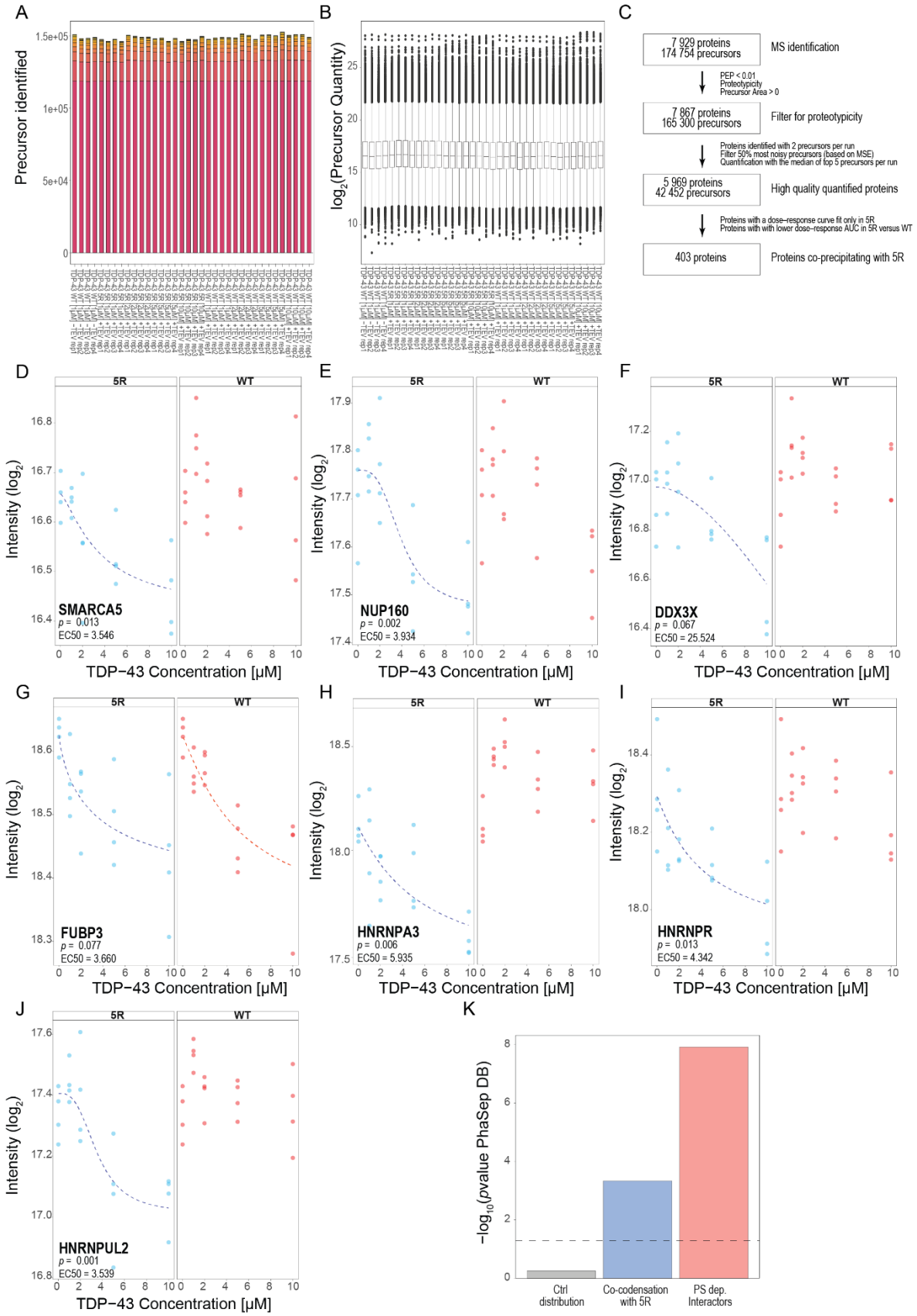
C. Precursors identified under the acquired conditions and their detection across all samples. The color scale indicates precursor completeness, expressed as the percentage of conditions in which each precursor was identified; red bars denote complete detection across the dataset.

D. Abundance of the precursors before normalization and filtering.

E. Principal component analysis (PCA) based on the abundance of identified proteins showing sample stratification between PS-deficient and solid-like conditions.

F. Overlap between proteins co-condensing ($log_2FC < -0.3$ and *p*values adjusted < 0.1) upon the spike-in of the different TDP-43 variants. The comparison was performed using TDP-43 WT + TEV as a reference. Proteins co-condensing in at least 2 conditions with solid-like TDP-43 condensates are highlighted in red.

G. Volcano plots showing the differential changes of protein levels upon the spike-in of the indicated TDP-43 variants into the HeLa cell lysate. Proteins showing a significant depletion ($log_2FC < -0.3$ and *p*values adjusted < 0.1) and identified previously as PS-dependent interactors are highlighted in red.



Supplementary Figure 5

Supplementary Fig. 5

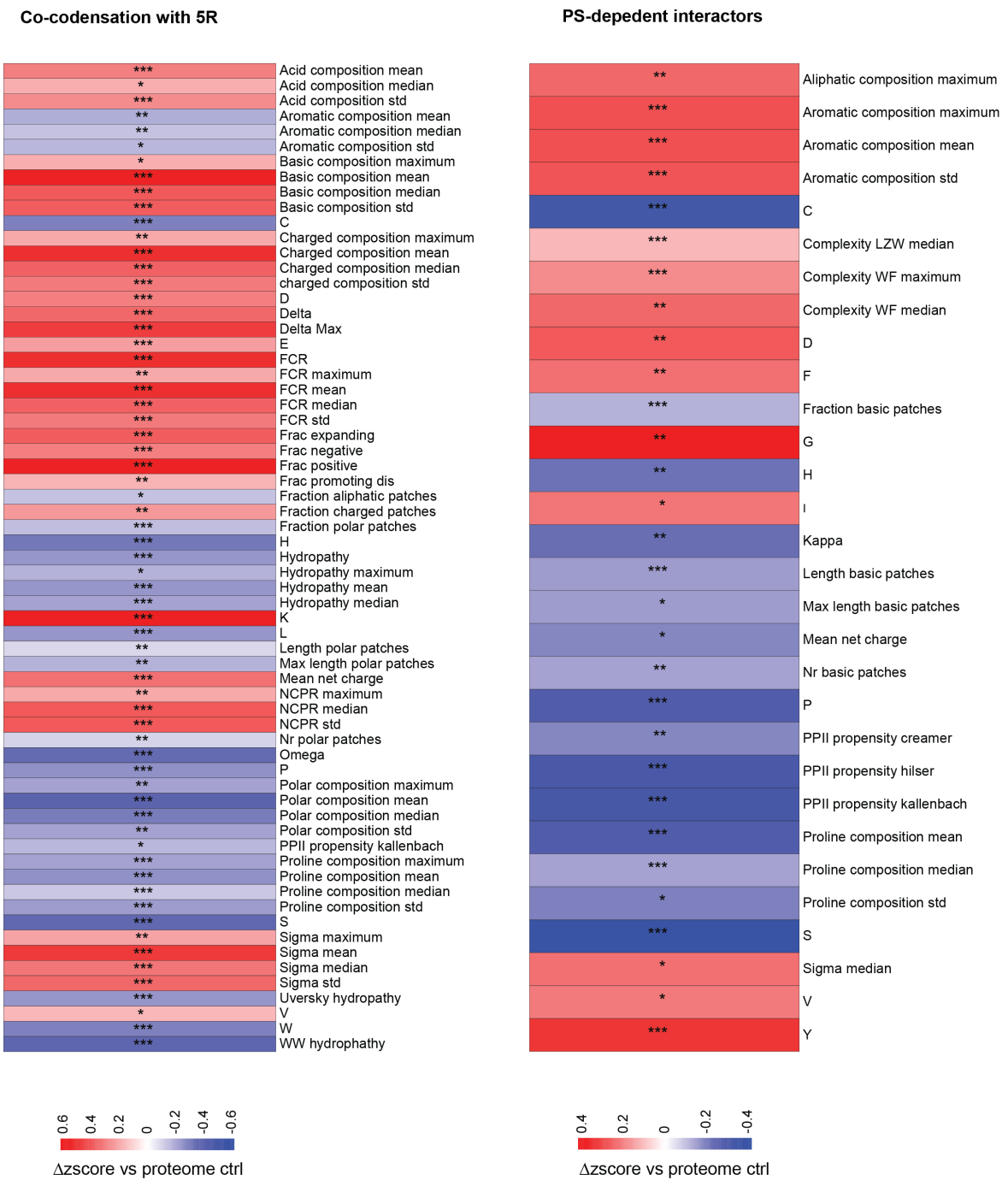
A. Precursors identified under the acquired conditions and their detection across all samples in the dose-response co-condensation profiling experiment. The color scale indicates precursor completeness, expressed as the percentage of conditions in which each precursor was identified; red bars denote complete detection across the dataset.

B. Abundance of the precursors before normalization and filtering.

C. Data analysis strategy for dose-response co-condensation profiling of protein levels upon the titration of TDP-43 WT or 5R. No imputation was applied, and only the 5 969 proteins identified with at least two precursors per condition were included in the analysis. Protein abundance was calculated as the median intensity of the top five precursors, following a filtering procedure designed to remove low-quality measurements, including precursors with poor identification confidence (PEP < 0.01), low intensity (> 0), and low reproducibility (excluding 50% of precursors per protein with the highest median standard error per condition). Dose-response curve fitting was applied to all quantified proteins, and proteins were retained if they exhibited a 5R-specific dose-response or a lower area under the curve (AUC) for the 5R condition compared to WT, yielding a final set of 403 proteins. Filtering criteria for curve fitting included the Hill coefficient (> 0.7), ANOVA *p*value (< 0.1), and correlation (> 0.7).

D - J. Abundance profiles for the indicated proteins upon titration of WT and 5R. Single points represent biological replicates, and the dashed line indicates the fitted dose-response curve (according to the criteria described above). The ANOVA *p*value and EC50 calculated for the 5R condition are reported for each profile. Representative profiles are shown for SMARCA5, NUP160, DDX3X, FUBP3, HNRNPA3, HNRNPR, and HNRNPUL2.

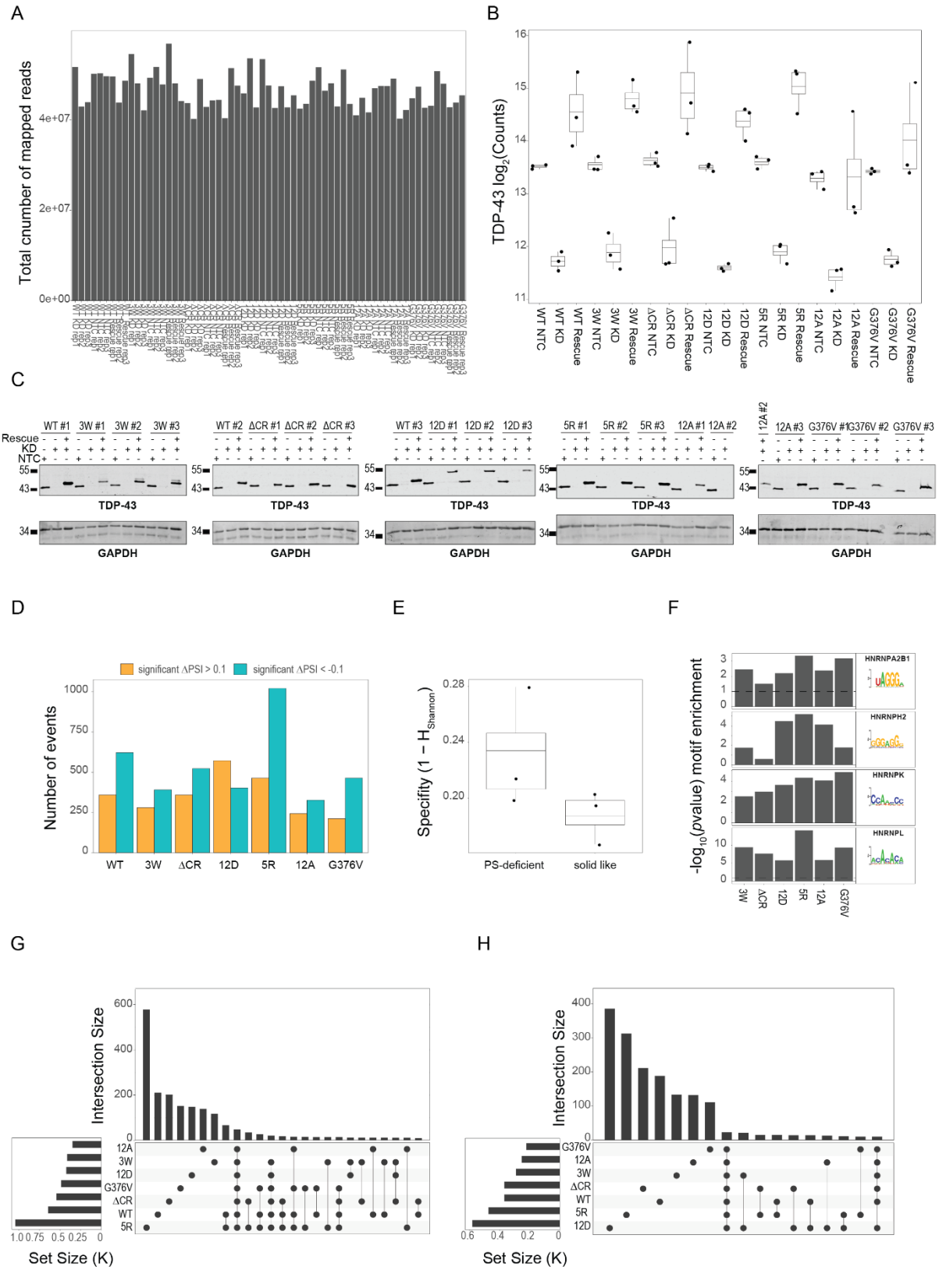
K. Enrichment analysis for proteins annotated to undergo phase separation from PhaSepDB database (v. 2.1). Reported values correspond to enrichment calculated using a hypergeometric test.



Supplementary Figure 6

Supplementary Fig. 6

Sequence features (“molecular grammar”) for proteins identified as co-condensing with 5R (left) or PS-dependent interactors (right). Only sequence features identified as significant (Welch test, *p*value adjusted < 0.1) are shown. A scale from blue (negative values) to red (positive values) displays the depletion and enrichment, respectively, of properties compared to the mean properties of the human proteome. The significance of the deviation from the human proteome is reported (Welch test, adjusted *p*values are as follows: *** < 0.001, ** < 0.01 and * < 0.1).



Supplementary Figure 7

Supplementary Fig. 7

A. Total number of reads assigned to genes for each condition analyzed.

B. TDP-43 read counts scaled by size factors as calculated by DESeq2 (log_2 scale). The boxplot reports the distribution of the values with the mean (central line), the interquartile range (1^{st} and 3^{rd} quartile) and whiskers (1.5XIQR, interquartile range).

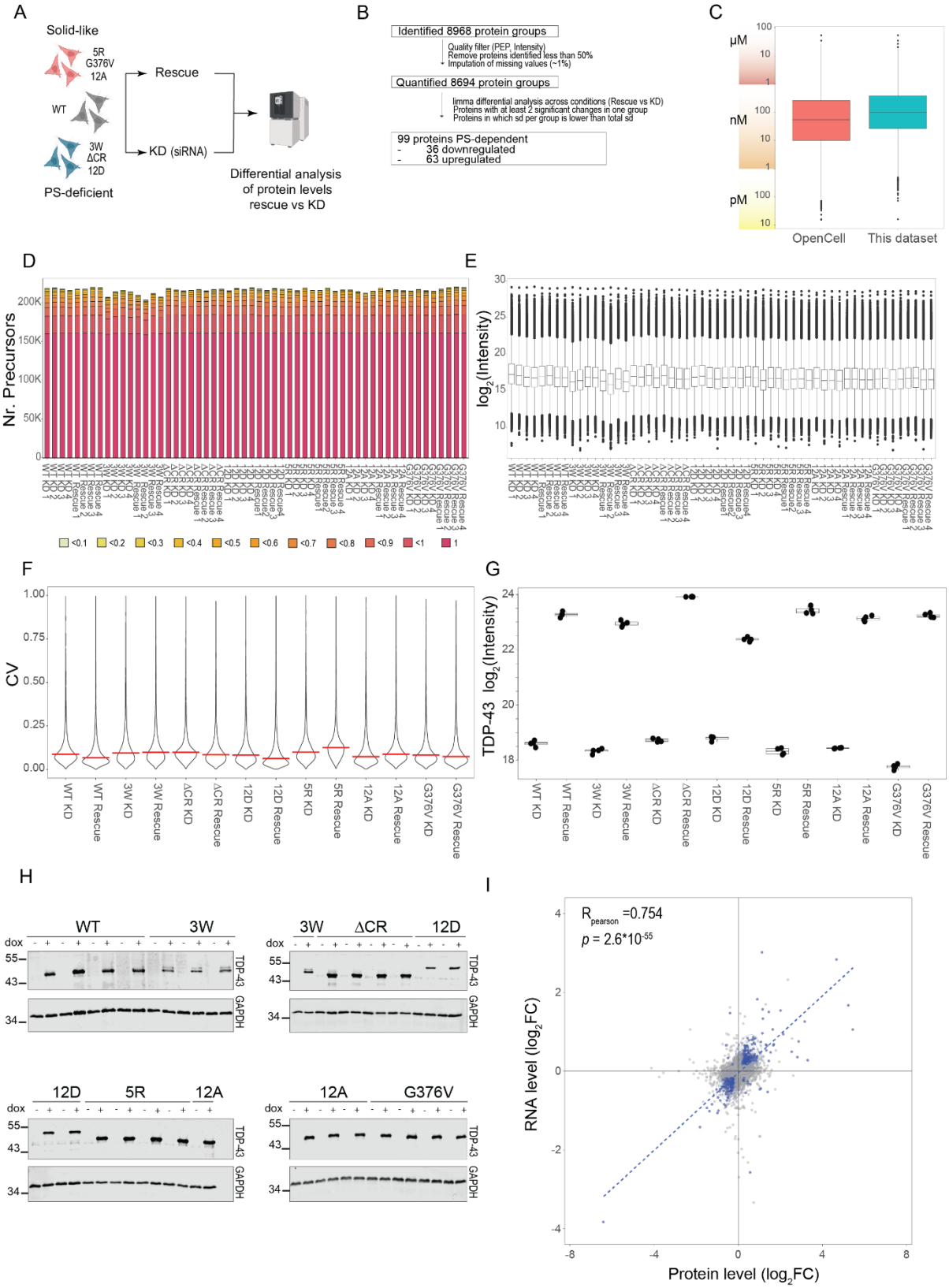
C. Western blots of TDP-43 levels for all examined cell clones (3 clones / TDP-43 variant, each in non-targeting control (NTC), knockdown (KD) and rescue condition). GAPDH is shown as loading control. Antibodies used: mouse anti-TDP-43 (60019-2-Ig, Proteintech), rabbit anti-GAPDH (10494-1-AP, Proteintech).

D. Number of significant (FDR < 0.05) alternatively spliced exons identified in the comparison rescue versus KD. Events with exon inclusion (ΔPSI > 0.1) and with exon skipping (ΔPSI < -0.1) are shown respectively in orange and cyan.

E. Specificity of RNA motifs identified based on the CISBP database. Specificity is calculated as $1\text{-}H_{Shannon}$. Shannon entropy was computed for significantly enriched RNA motifs (hypergeometric test, *p*value adjusted < 0.05) in each condition.

F. Significance (hypergeometric test) for RNA motif enrichment of proteins identified as PS-dependent TDP-43 interactors (HNRNPA2B1, HNRNPH2, HNRNPK, HNRNPL). The RNA motif enrichment for each single TDP-43 variant is shown as a barplot. The motif sequence logo (according to CISBP-RNA) is shown on the right.

G, H. Upset plots showing the overlap of alternatively spliced events that are significantly more included (ΔPSI>0.1) (G) or skipped (ΔPSI < 0.1) (H) in the comparison between rescue versus KD.



Supplementary Figure 8

Supplementary Fig. 8

A. Experimental workflow for proteome profiling of TDP-43 variants with distinct PS properties. Proteome profiling was performed in TDP-43 KD cells rescued with different TDP-43 variants: WT, solid-like variants (5R, 12A, G376V), and PS-deficient variants (3W, ΔCR, 12D). Differential analysis is performed between rescue and KD conditions. Created in part with BioRender.com.

B. Analysis pipeline used to identify proteins whose abundance is associated with the PS properties of TDP-43.

C. Comparison of the dynamic range of the protein abundance (expressed as protein concentration in the cell) in the acquired dataset with the OpenCell dataset (HEK293 cells). Boxplot reports the distribution of values with the median (central line), the interquartile range (1^{st} and 3^{rd} quartile), and whiskers (1.5XIQR, interquartile range).

D. Precursors identified under the acquired conditions and their detection across all samples. The color scale indicates precursor completeness, expressed as the percentage of conditions in which each precursor was identified; red bars denote complete detection across the dataset.

E. Distribution of precursor intensities across analyzed conditions. Boxplots represent the distribution of precursor abundance values, with the central line indicating the median, boxes representing the interquartile range (1^{st} and 3^{rd} quartile), and whiskers (1.5XIQR, interquartile range).

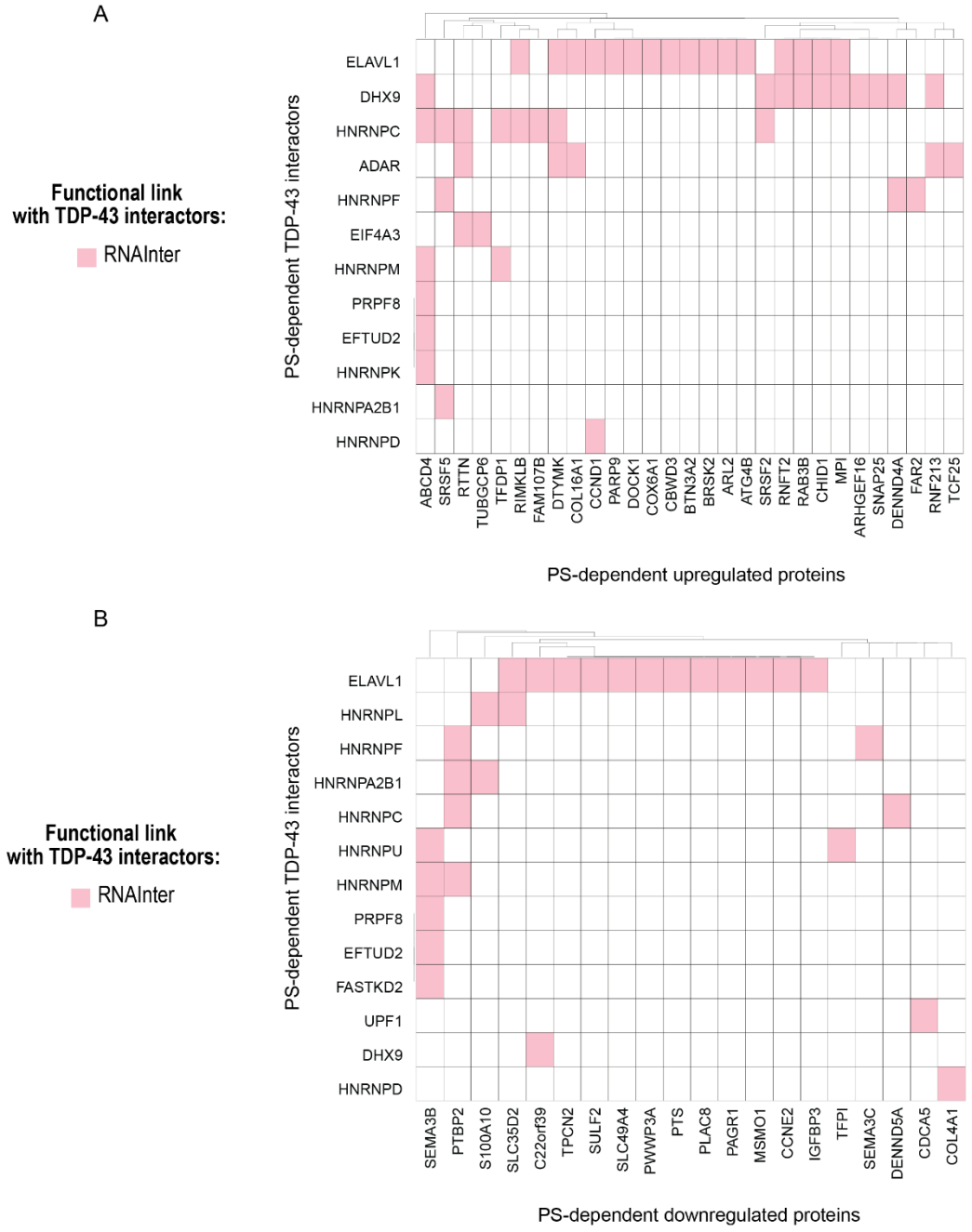
F. Coefficient of variation (CV) of protein abundance measurements across conditions. Violin plots show the distribution of CV values, with the median indicated by a red line.

G. Quantification of TDP-43 abundance across all analyzed conditions. Only peptide precursors detected in all conditions were used to determine TDP-43 abundance.

H. Western blots showing the level of expression of myc-TDP-43 across all conditions. Antibodies used: mouse anti-TDP-43 (60019-2-Ig, Proteintech), rabbit anti-GAPDH (10494-1-AP, Proteintech).

I. Correlation between RNA and protein abundance changes in TDP-43 WT rescue versus KD conditions. All genes and proteins are displayed in light gray, while those significantly regulated at both RNA and protein levels (*p*value adjusted < 0.05) are highlighted in blue. The Pearson correlation coefficient is calculated based on the subset of significantly regulated genes and

proteins.



Supplementary Figure 9

Supplementary Fig. 9

A-B. Functional links between TDP-43 PS-dependent interactors (rows) and their targets identified as PS-dependent proteins (columns). Functional links for PS-dependent up- and downregulated proteins are shown in A and B, respectively.