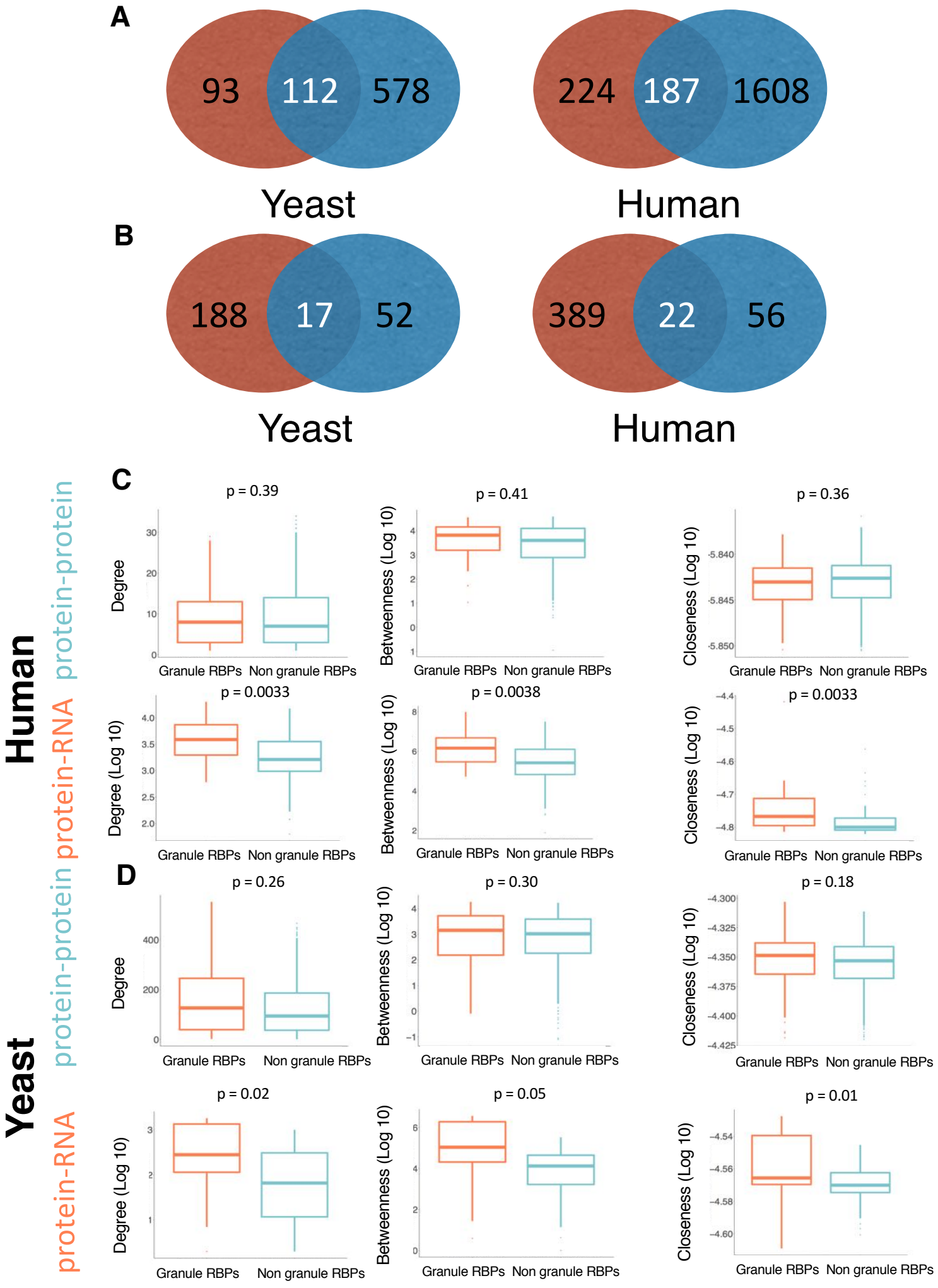
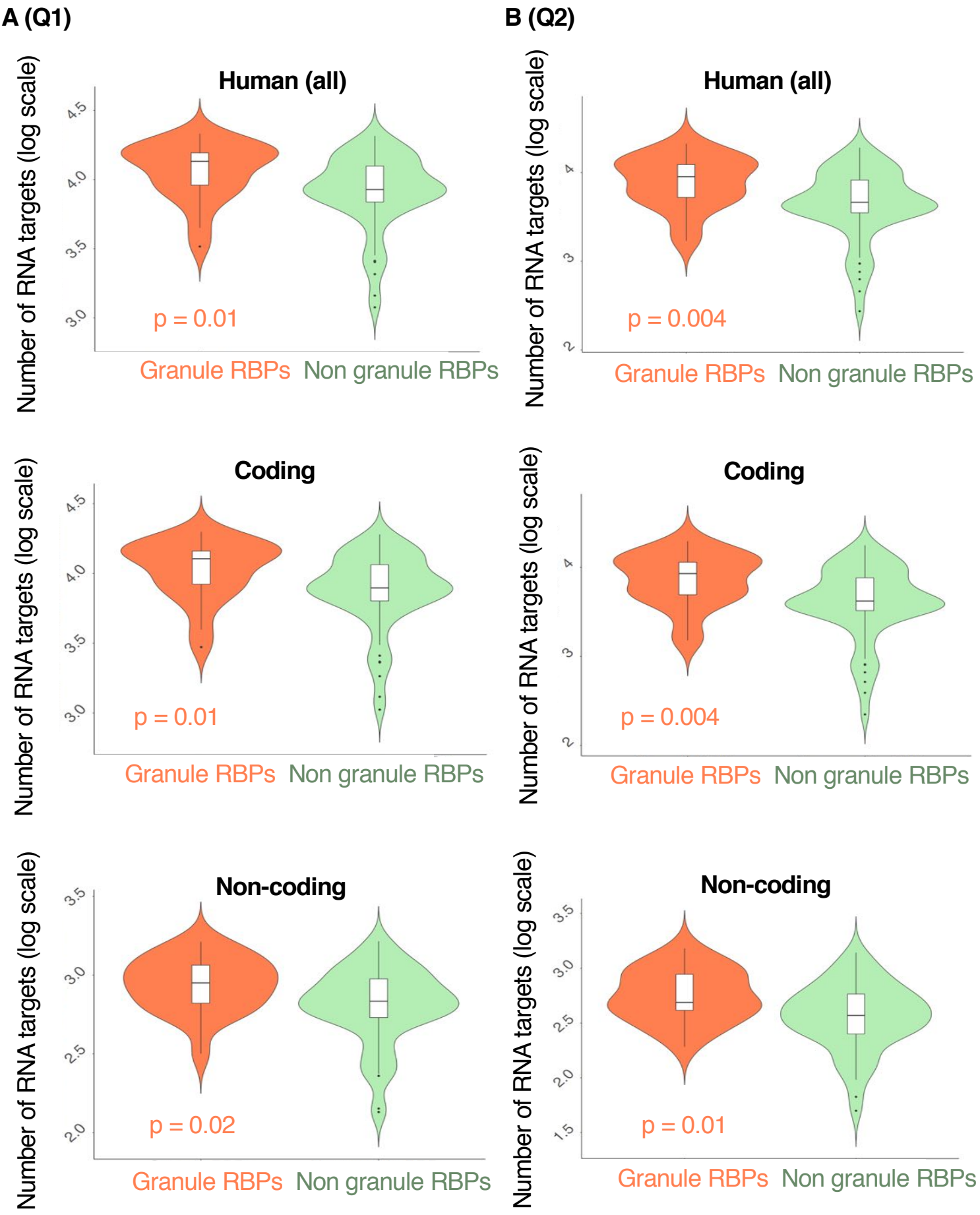


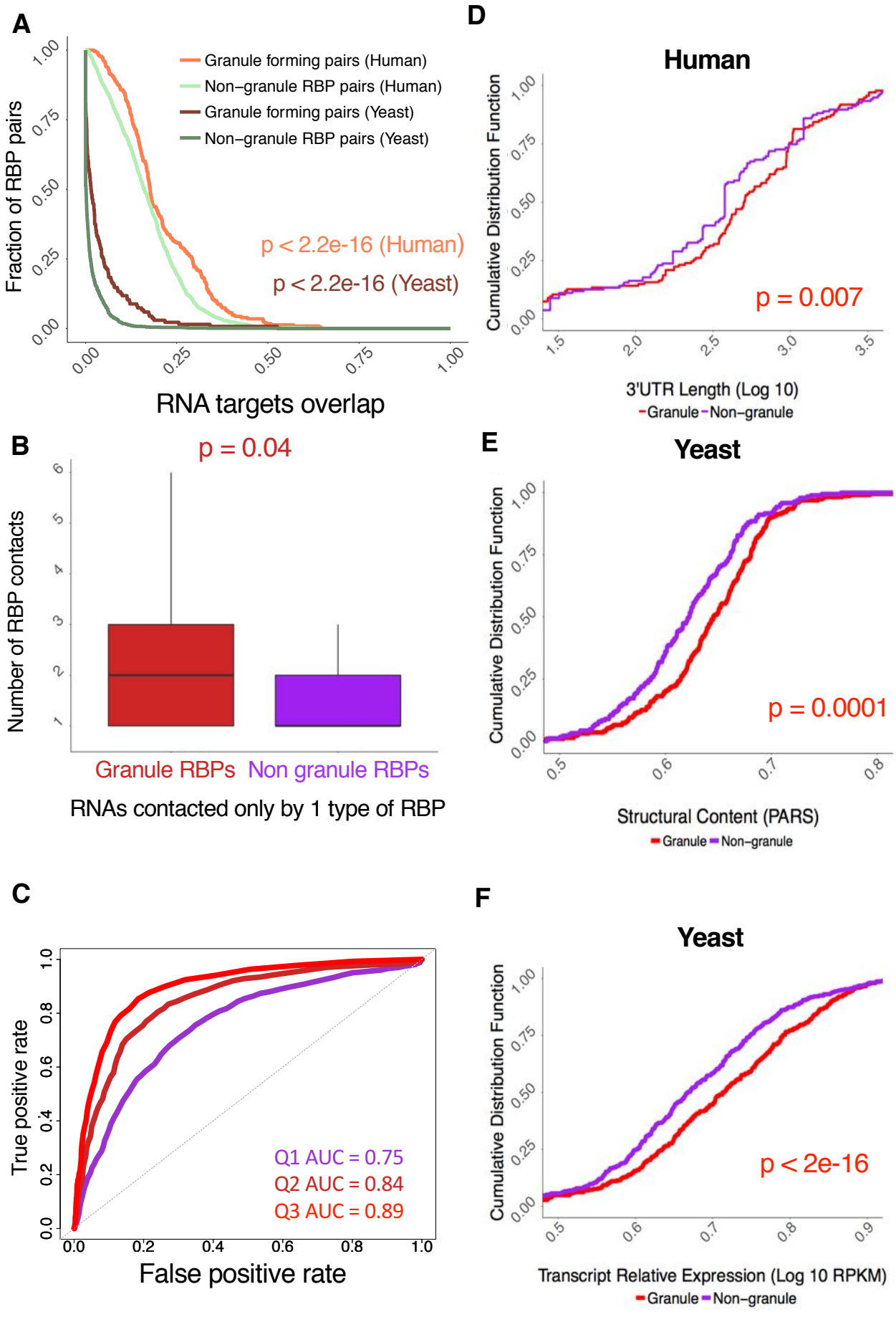
Supplementary Figure 1

Supplementary Figure 1 [related to Figure 1]. Datasets A) Granule RBPs Red circle: granule-forming proteins, Blue circle: RBPs, as defined in Gerstberger et al, 2014 (Gerstberger et al., 2014). Intersection represents granule RBPs. **B)** Number of interactions. Red circle: granule-forming proteins. Blue circle: RBPs with known targets. Intersection represents granule RBPs with known targets. **Distribution of centrality values of granule and non-granule RBPs in different interaction networks. C)** Centrality distributions for the human dataset. Up: Protein-protein network. (p-value (left) = 0.39, p-value (centre) = 0.41, p-value (right) = 0.36. Down: Protein-RNA network (p-value (left) = 0.003, p-value (centre) = 0.007, p-value (right) = 0.01. **D)** Centrality distributions for the yeast dataset. Up: Protein-protein network. (p-value (left) = 0.26, p-value (centre) = 0.30, p-value (right) = 0.18. Down: Protein-RNA network (p-value (left) = 0.02, p-value (centre) = 0.05, p-value (right) = 0.01.



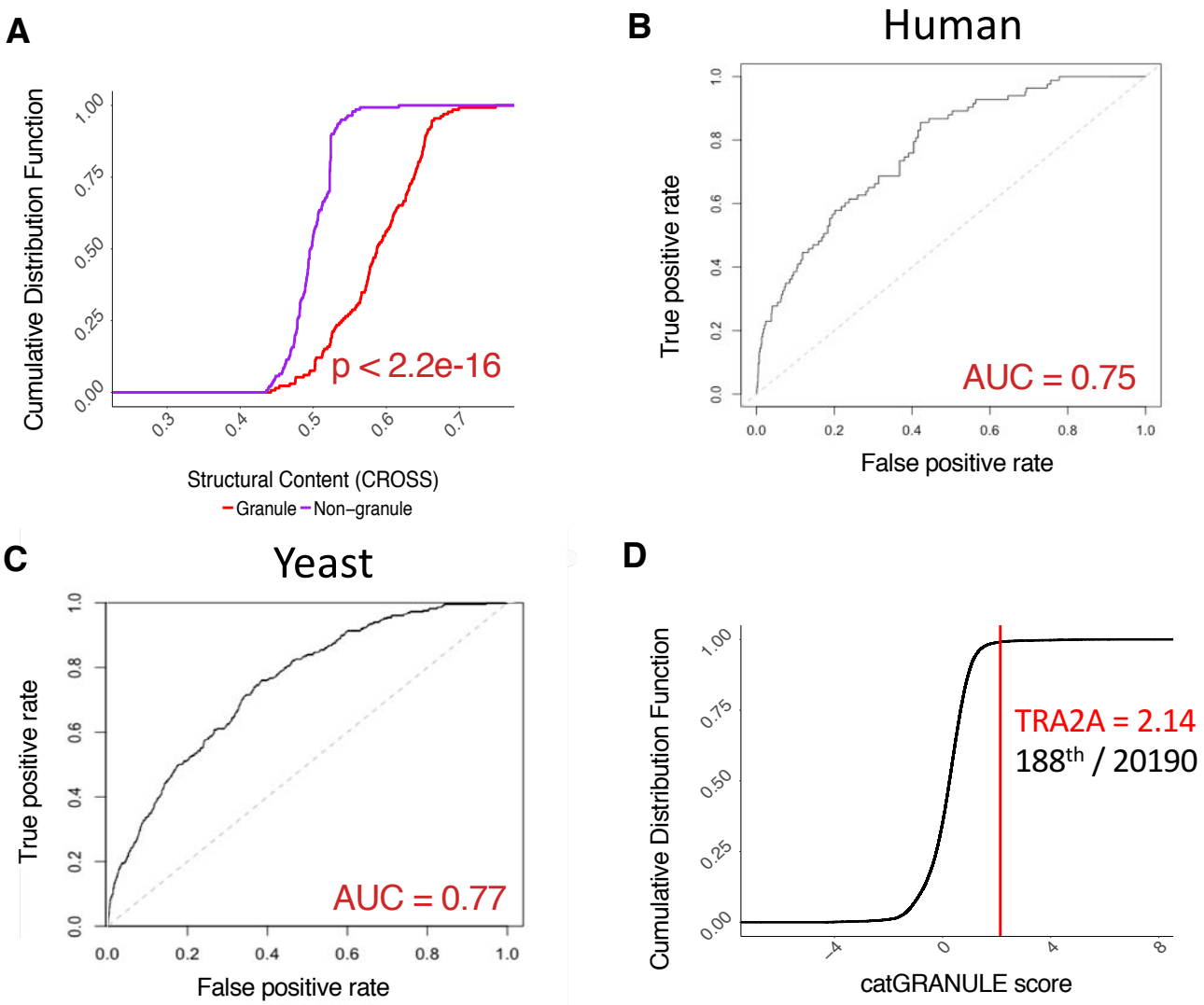
Supplementary Figure 2

Supplementary Figure 2 [related to Figure 1]. Number of RNA targets of granule and non-granule RBPs: A) First quartile of the reads/expression distribution (Q1). **B)** Second quartile (Q2).



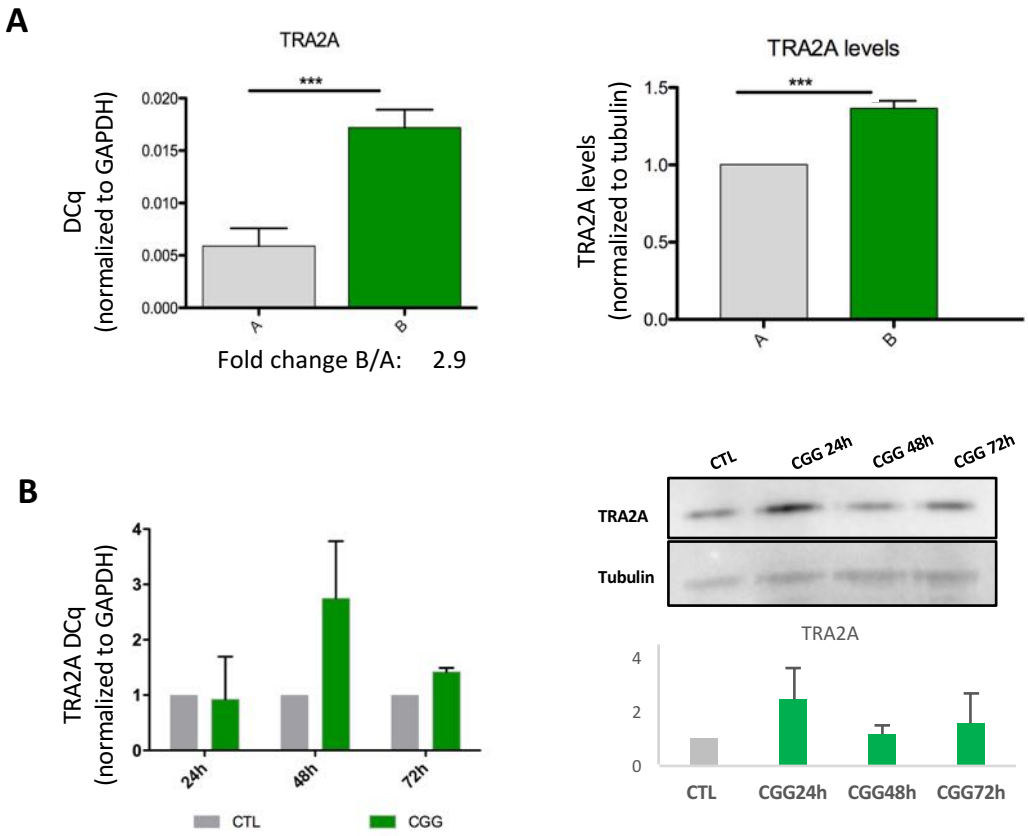
Supplementary Figure 3

Supplementary Figure 3 [related to Figure 1,2]. Properties of granule RNAs. A) RNAs interacting exclusively with granule forming RBPs have higher number of protein contacts (p-value = 0.04, Wilcoxon test). Human transcripts: **B)** Granule RNAs have more structured UTRs (p-value = 0.007; KS test). PARS analysis on 3'UTR of granule and non-granule RNAs**.** Yeast granule RNA are **C)** structured (p-value = 0.001; KS test; PARS data), and **D)** more abundant (p-value = 2.2e-16; KS test) than non-granule RNAs. The UTR analysis was not performed due to the lack of annotation.



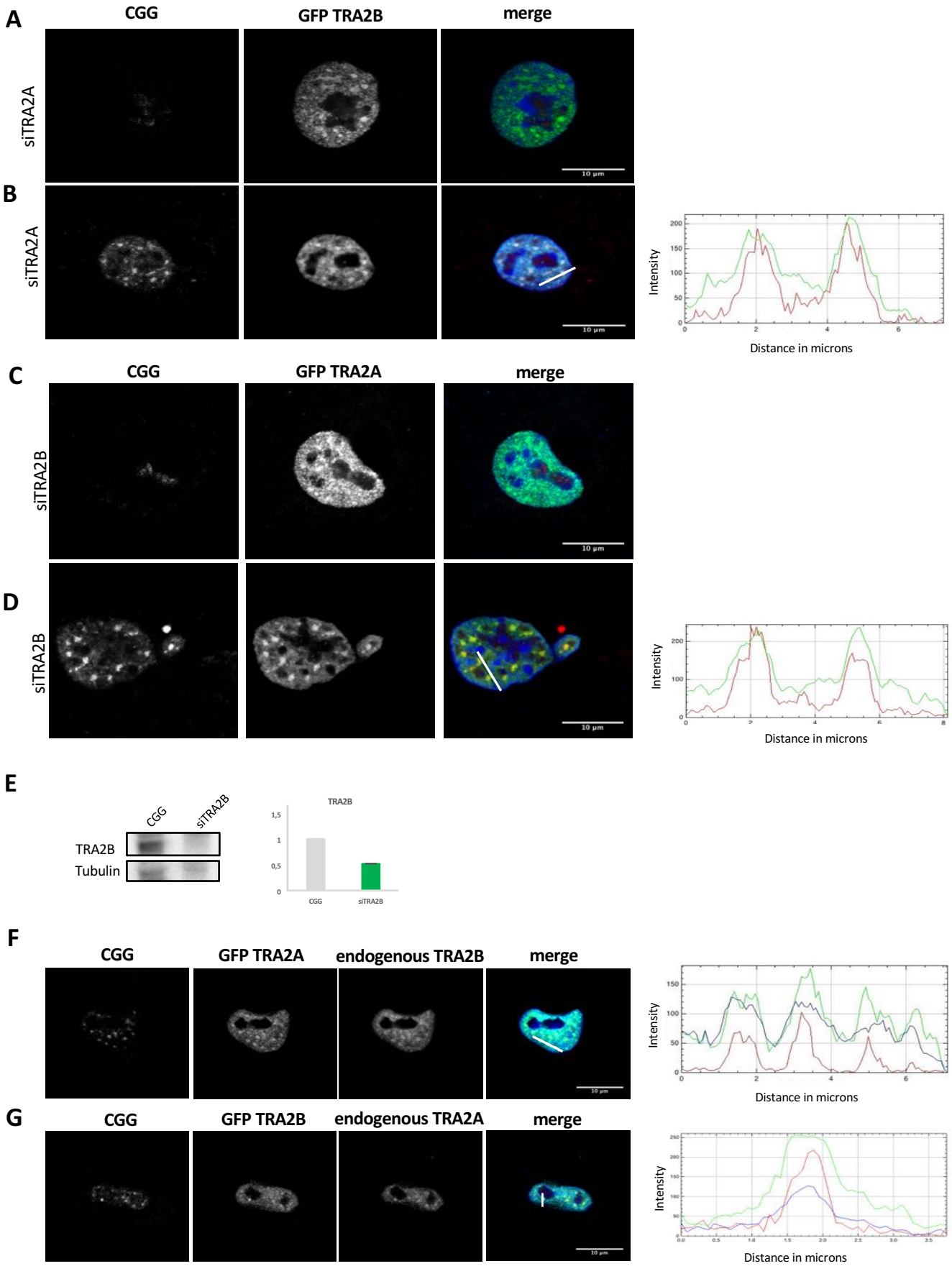
Supplementary Figure 4

Supplementary Figure 4 [related to Figure 2,3]. Computational predictions of granule-forming components. A) Granule transcripts are predicted to be more structured (structural content according is measured using CROSS; p-value < 2.2e-16, KS test). **B** and **C)** *cat*GRANULE performances on human and yeast experimentally described granule-forming proteins. AUC (Area under the ROC curve) is used to measure the discriminative power of the method. **D)** Distribution of *cat*GRANULE scores for the whole human proteome. TRA2A (*cat*GRANULE score = 2.14) ranks 188th out of 20190 human proteins (i.e. 1% of the distribution).



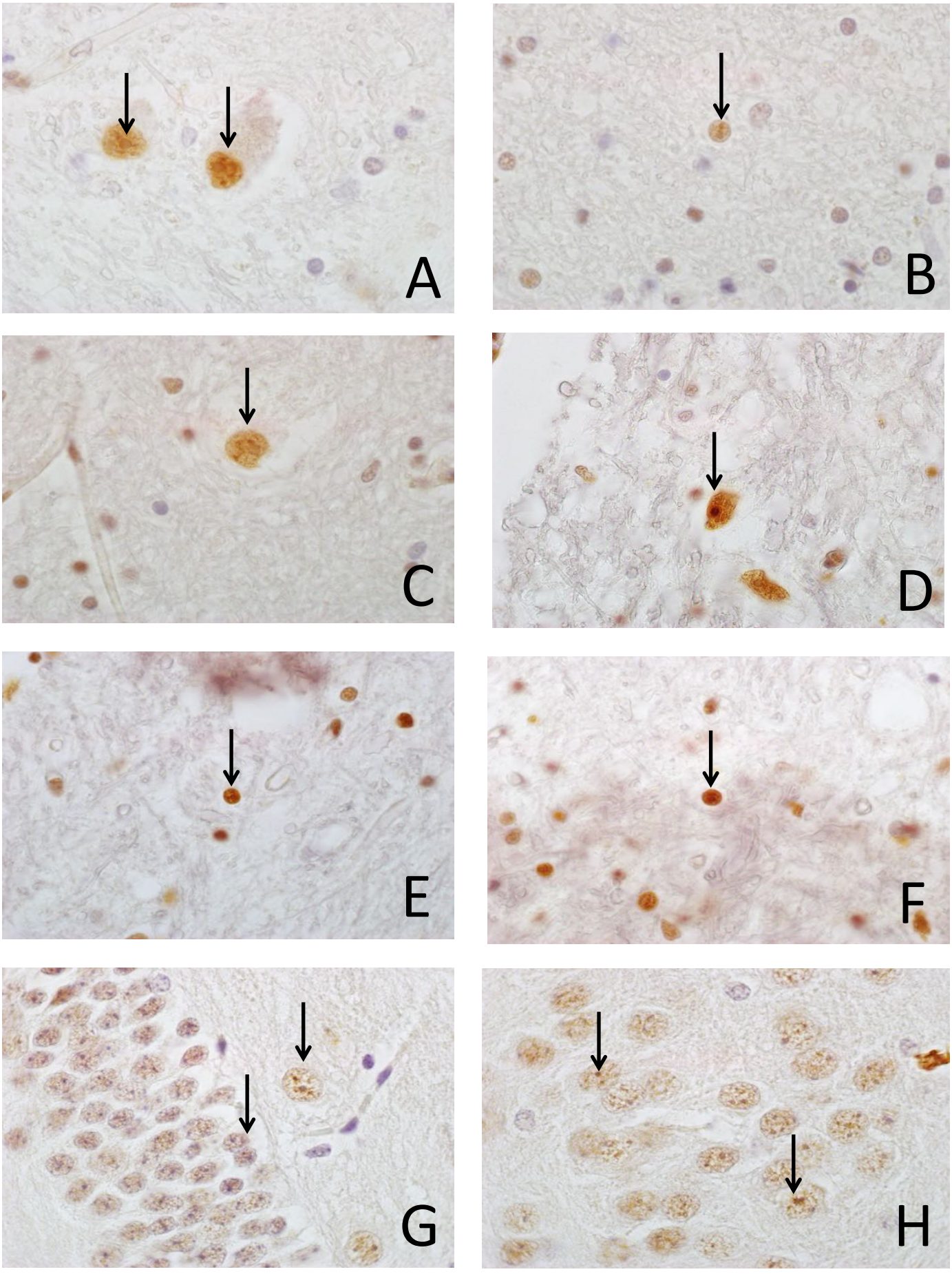
Supplementary Figure 5

Supplementary Figure 5 [related to Figure 4]. TRA2A levels in human lymphocytes and COS-7 cell model. A) Human lymphocytes from control (A) or pre mutation-carrier (B) were lysated and both RNA and protein were isolated (*** p-value < 0.01). Relative TRA2A RNA expression (left panel) and TRA2A protein (right panel) are represented. **B)** COS-7 cells were transfected with CGG(60X) and compared to controls. After 24h, 48h or 72h of transfection cells were pelleted and RNA and protein extraction was performed. Relative TRA2A RNA expression (left panel) and TRA2A protein (right panel) are represented.



Supplementary Figure 6

Supplementary Figure 6 [related to Figure 6]. TRA2B over-expression and TRA2A knock-down. **A)** Control COS-7 cells (without CGG(60X) transfection) were transfected with GFP-TRA2B and siTRA2A. **B)** COS-7 cells were transfected with CGG(60X), GFP-TRA2B and siTRA2A. In both A and B, after 48 hours, cells were hybridized with Cy3-GGC(8X) probe and immunostained with an antibody against TRA2B. The graph represents TRA2B/CGG levels**. TRA2A over-expression and TRA2B knock-down. C)** Control COS-7 cells were transfected with siTRA2B and GFP-TRA2A (in absence of CGG(60X) transfection). **D)** COS-7 cells were transfected with CGG(60X), siTRA2B and GFP-TRA2A. In both A and B, after 48 hours of transfection cells were hybridized with Cy3-GGC(8X) probe and immunostained with antiGFP. The graphs represent TRA2A/CGG levels. **E)** TRA2B protein levels in COS-7 cells treated as in B. **TRA2A and TRA2B over-expression** COS-7 cells were transfected with GFP-TRA2A **F)** or GFP-TRA2B **G)** and CGG(60X). After 48 hours, cells were hybridized with Cy3-GGC(8X) probe and immunostained with an antibody against either TRA2A or TRA2B. Graphs represent TRA2A/TRA2B/CGG levels.



Supplementary Figure 7

Supplementary Figure 7 [related to Figure 10]. A-F) TRA2A immunohistochemistry in human hippocampus from FXTAS. **G-H)** TRA2A immunohistochemistry in premutated mouse model (counterstaining is done with haematoxylin; the arrows points to the inclusions).

Table S1 [related to Figure 1]. Protein and RNA interactions. A) list of human and yeast granule proteins. **B**, **C, D, E**) Granule and non-granule RBPs, RNA partners of human RBPs identified at different cut-offs of the reads/expression distribution (first, second and third quartile are indicated with Q1, Q2 and Q3). Names starting by NM indicate coding transcripts and names starting by NR indicate non-coding transcripts. **F)** RNA partners of yeast RBPs. **G, H, I)** Overlap between interactomes of human RBPs calculated using the Jaccard index (first, second and third quartile are indicated with Q1, Q2 and Q3) and **J)** Overlap between yeast RBP interactomes. K) Intersection between Q3 RNA interactome and granule transcripts reported in (Khong et al., 2017).

Table S2 [related to Figure 2]. A) RBP contacts of human RNAs. Names starting by NM indicate coding transcripts and names starting by NR indicate non-coding transcripts. **B)** Number of total, granule and non-granule contacts, structural content, length and UTR size of human transcripts; **C)** Number of total, granule and non-granule contacts, structural content, length of yeast transcripts.

Table S3 [related to Figure 3]. *in silico* predictions of CGG interactions. *cat*RAPID scores (discriminative power *DP* or interaction score, interaction strength *IS* or specificity) (Agostini et al., 2013), name of the gene, catGRANULE score (Bolognesi et al., 2016), granule ability (predicted / validated) and empirical p-value indicating the ability of proteins to interact with CGG repeats (calculated on 3340 DNA-binding, RNA-binding and structurally disordered proteins).

Table S4 [related to Figure 3]. *in vitro* validation of CGG interactions. We employed protein arrays to perform a large *in vitro* screening of RBP interactions with the first *FMR1* exon (Cirillo et al., 2017; Marchese et al.). We probed both expanded ("PRE"; 79 CGG) and normal ("WT"; 21 CGG) on three independent arrays, obtaining highly reproducible results.

Table S5 [related to Figure 7]. Microarray and RNA-seq analysis of splicing events. A) CGG over-expression vs CTL; B) CGG over-expression and TRA2A knock-down vs control. Fold Changes, significance and (sub)-exon names are reported. Microarray: Transcriptome Analysis Console (TAC) 4.0 software (ThermoFisher) was used to identify splicing events. RNA-seq: Statistical analysis of alternative splicing events was done using EventPointer v1.0.0 and DEXSeq v 1.24.2 (GO analysis: http://www.tartaglialab.com/GO_analyser/render_GO_universal/2105/64ce4f8d1d/, http://www.tartaglialab.com/GO_analyser/render_GO_universal/2108/eef220536a/).