

Figure S1

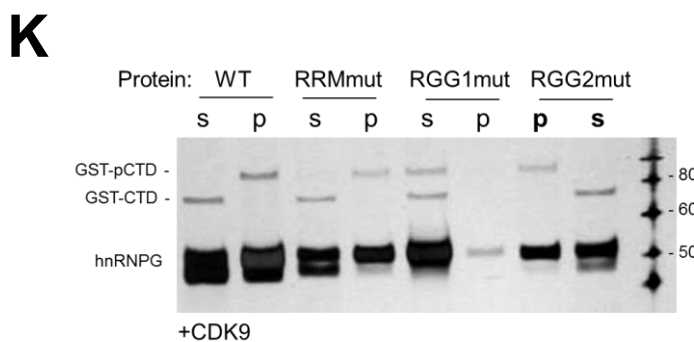
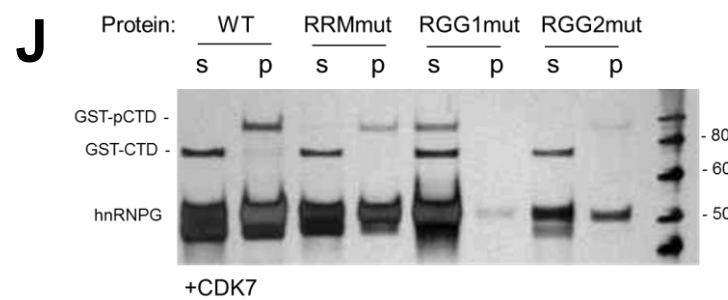
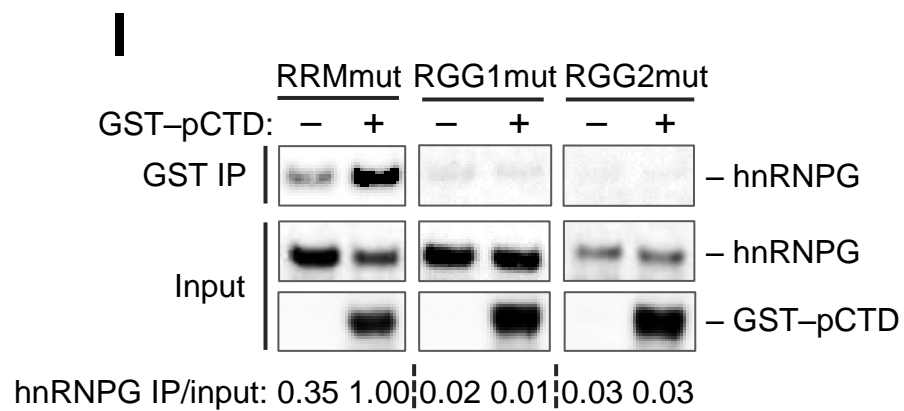
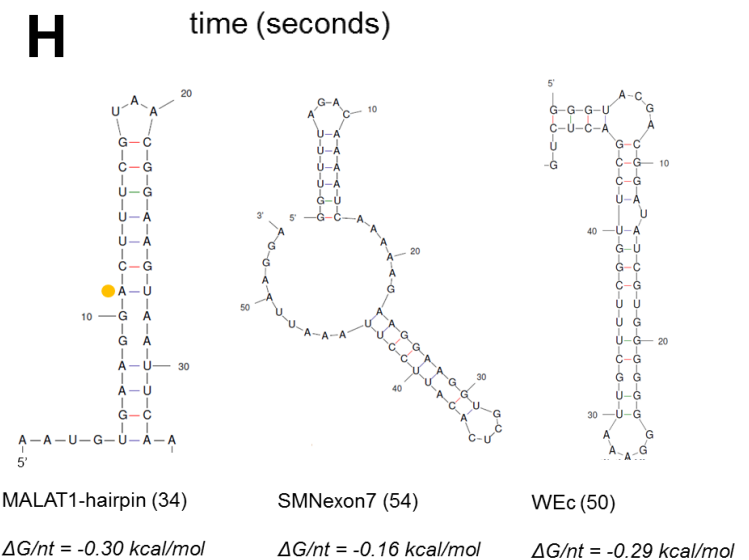
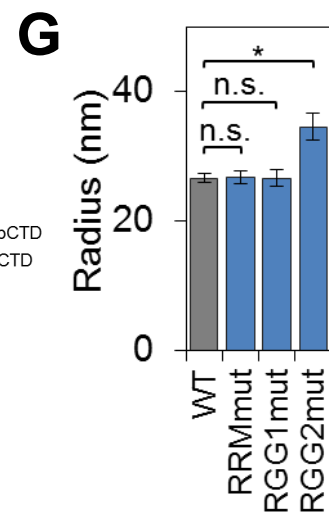
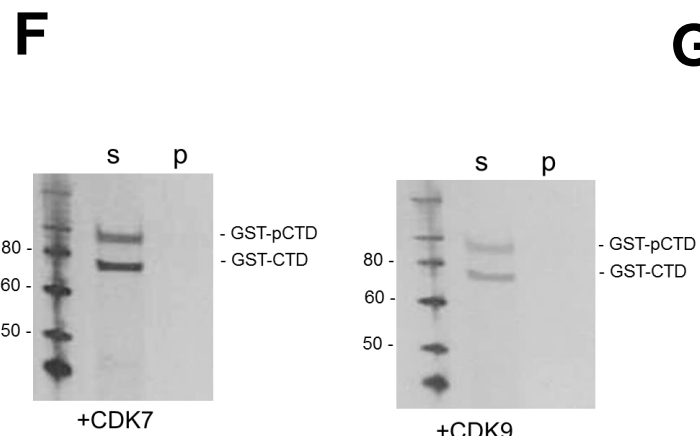
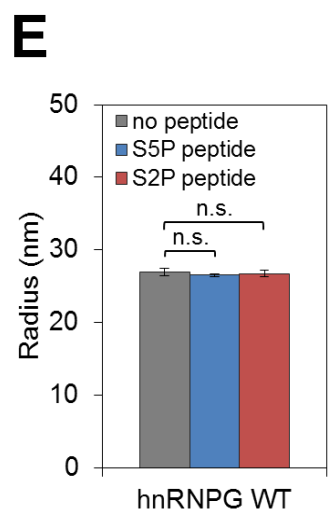
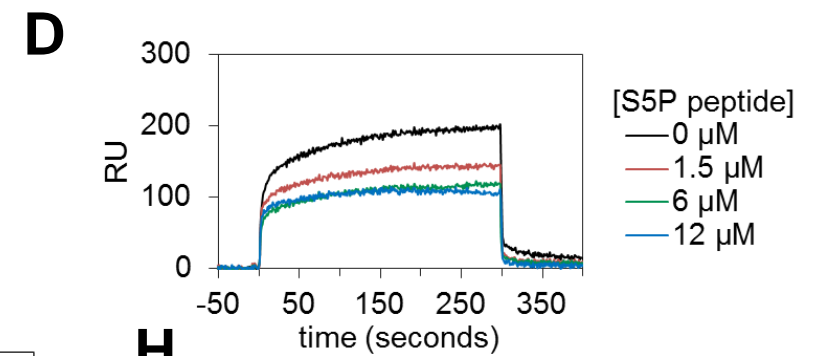
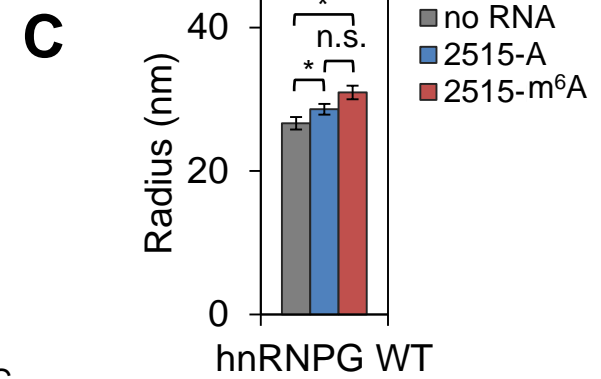
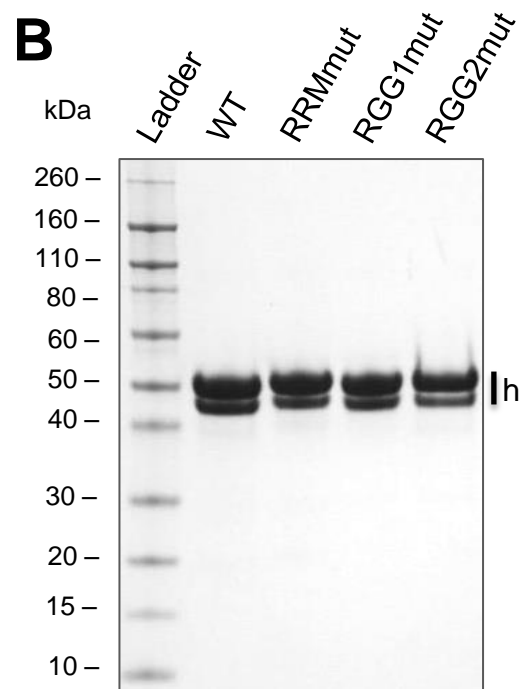
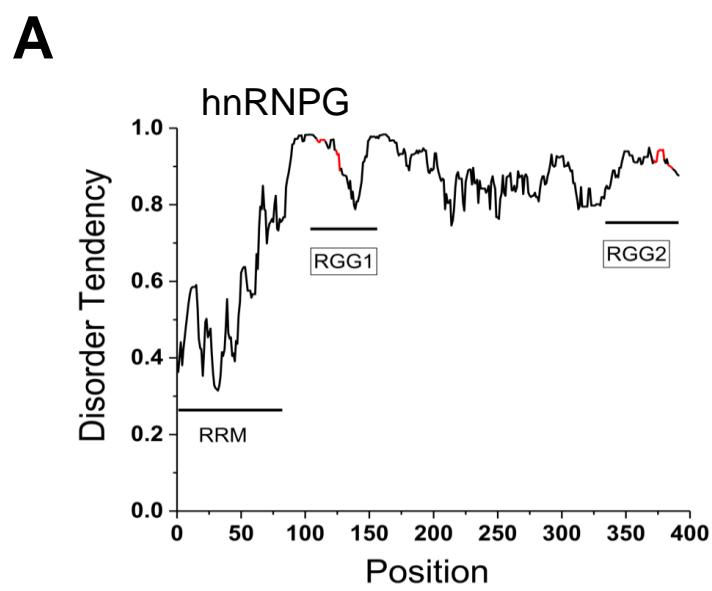


Figure S2

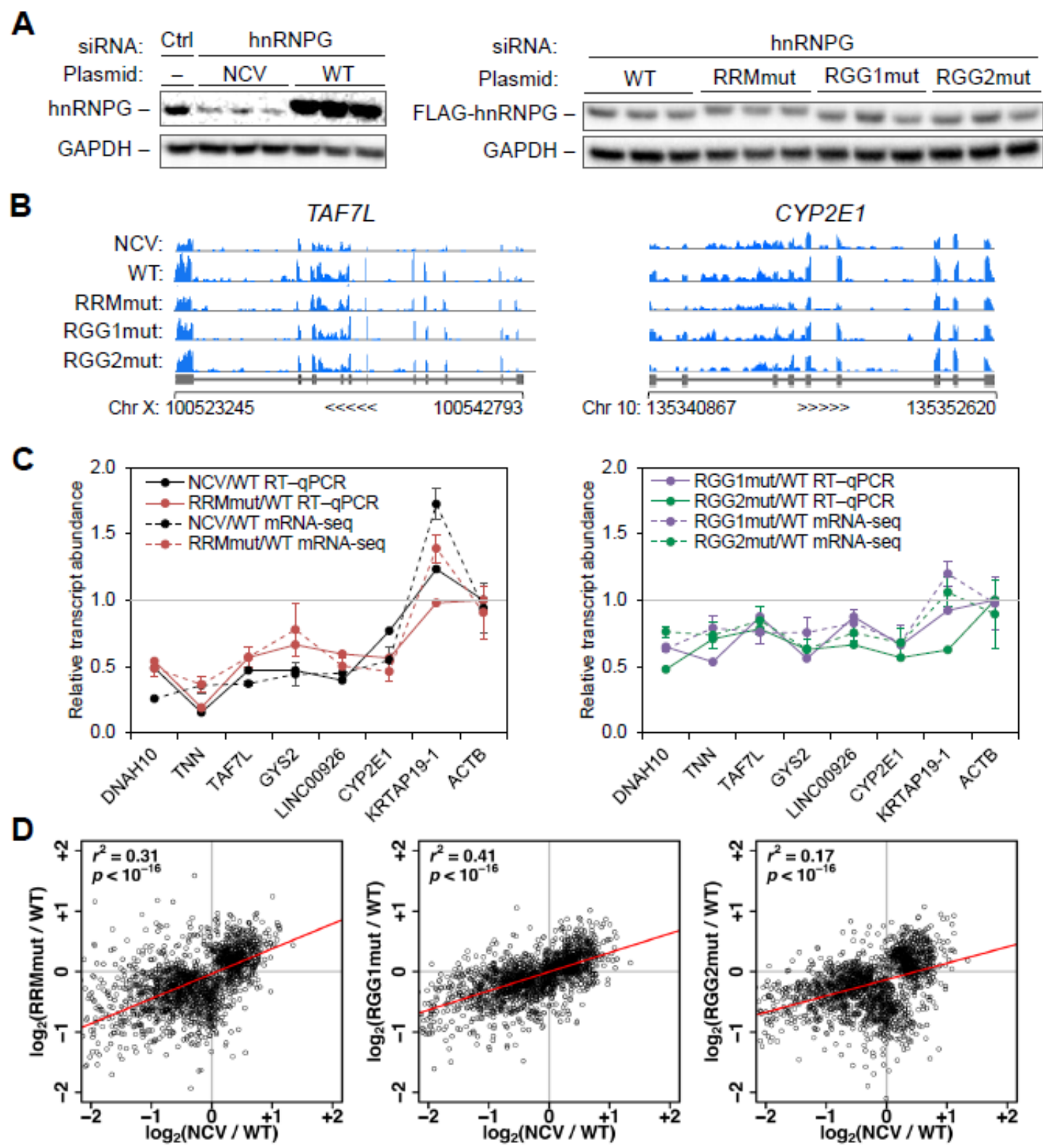
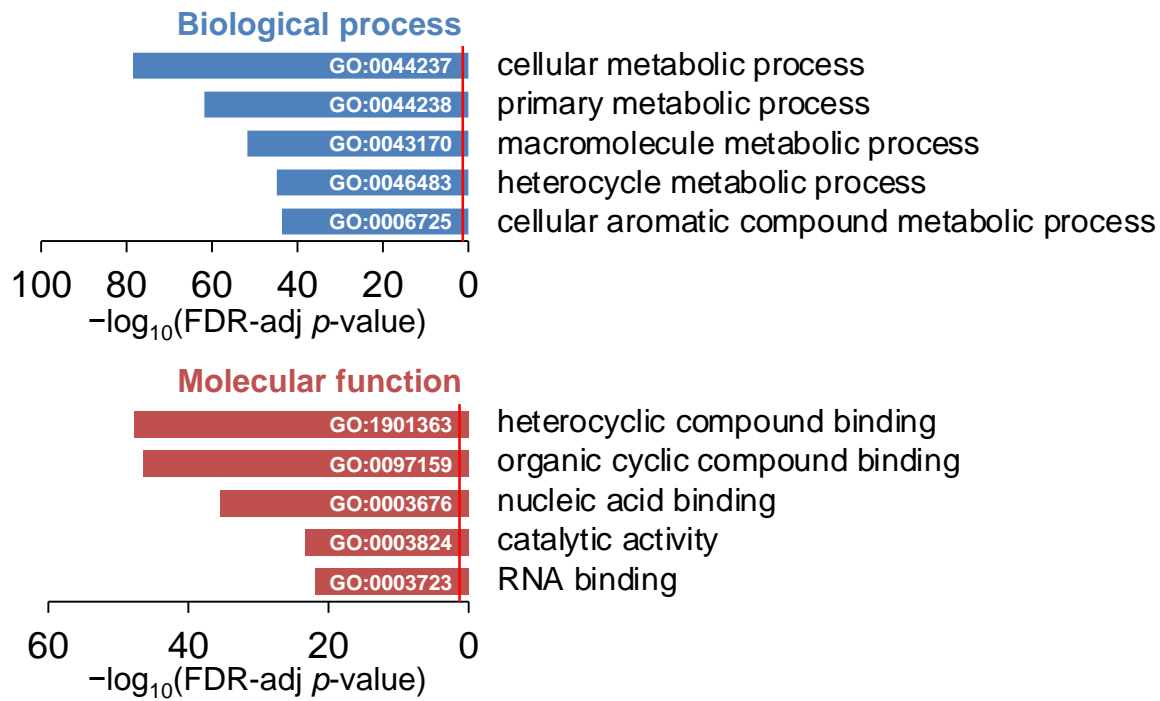
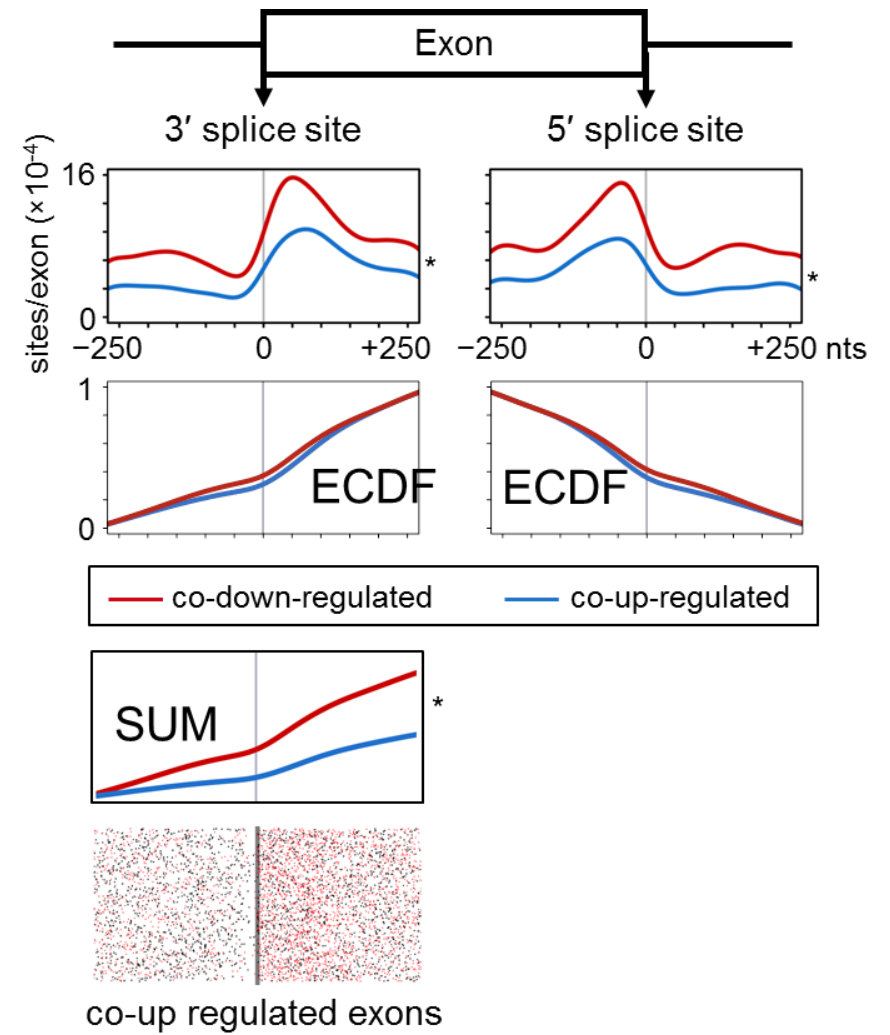


Figure S3

A Enriched in hnRNPG KD / Control: GO analysis



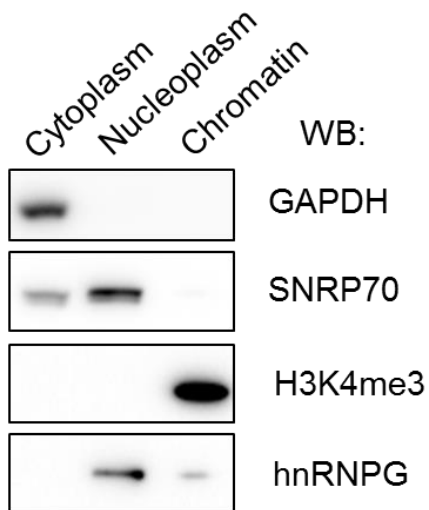
B



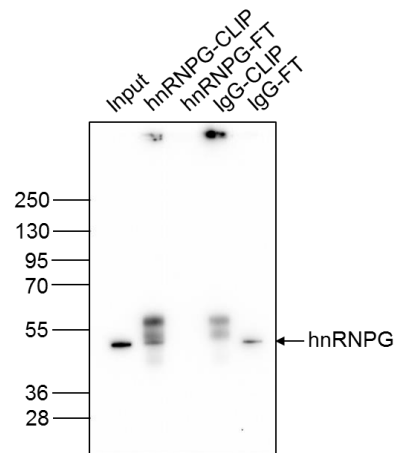
C

	co-down-regulated exons	co-up-regulated exons
exons with hnRNPG-bound m ⁶ A site at 3' ss \pm 300 nt	36%	22%
exons with hnRNPG-bound m ⁶ A site at 5' ss \pm 300 nt	36%	21%

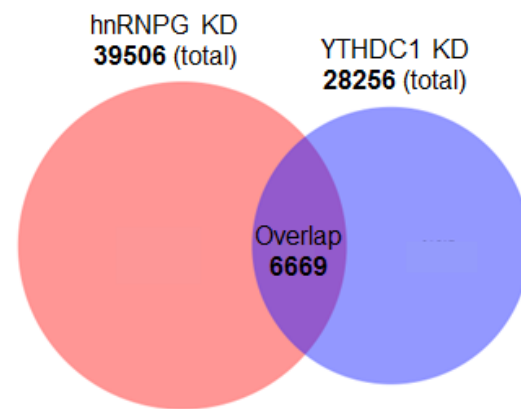
D



E



F



G

	YTHDC1 KD	
	down-regulated exons	up-regulated exons
hnRNPG KD down-regulated exons	1909	1291
hnRNPG KD up-regulated exons	1619	1850

H

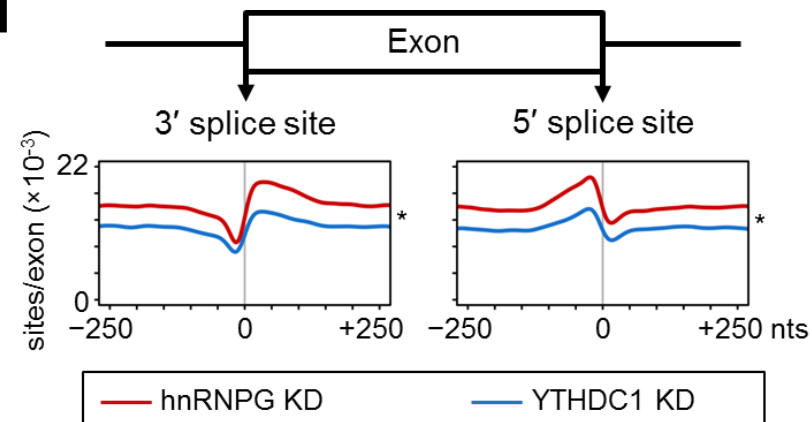


Figure S4

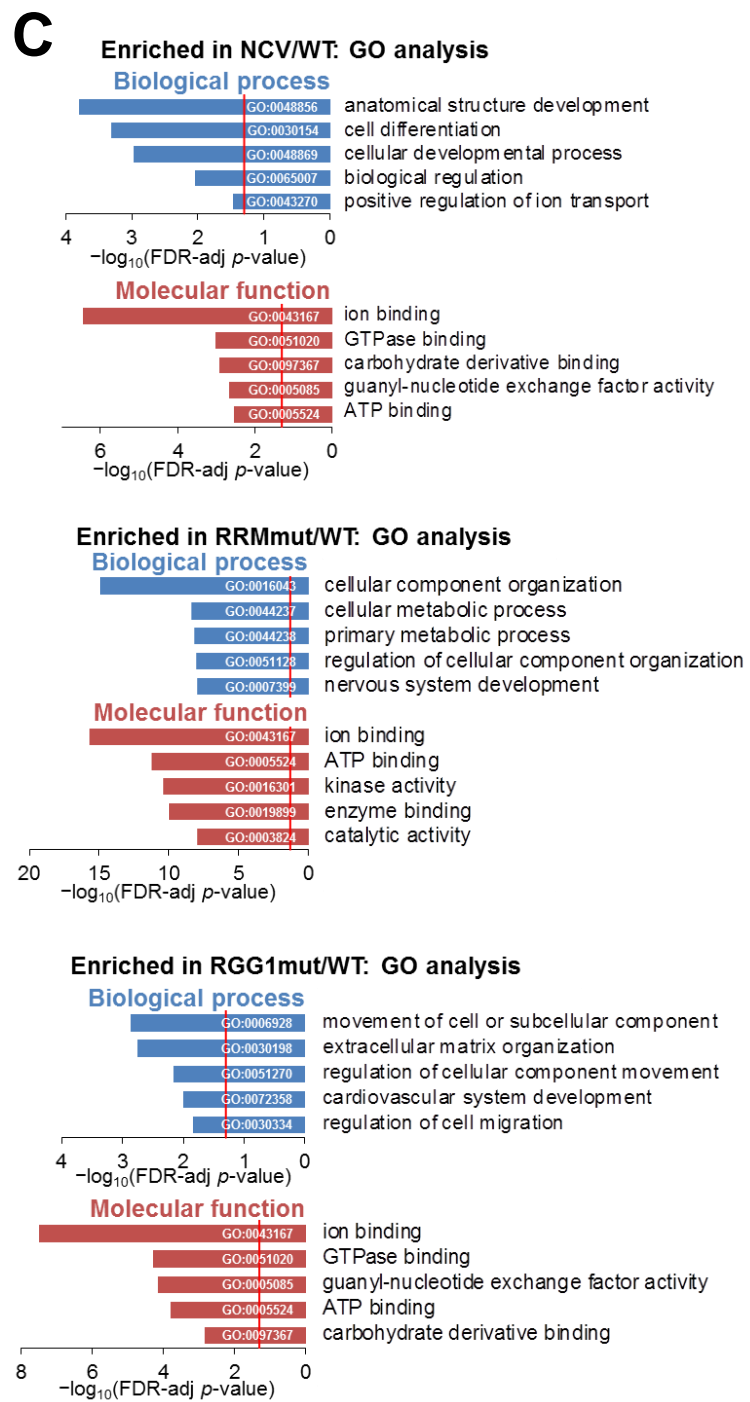
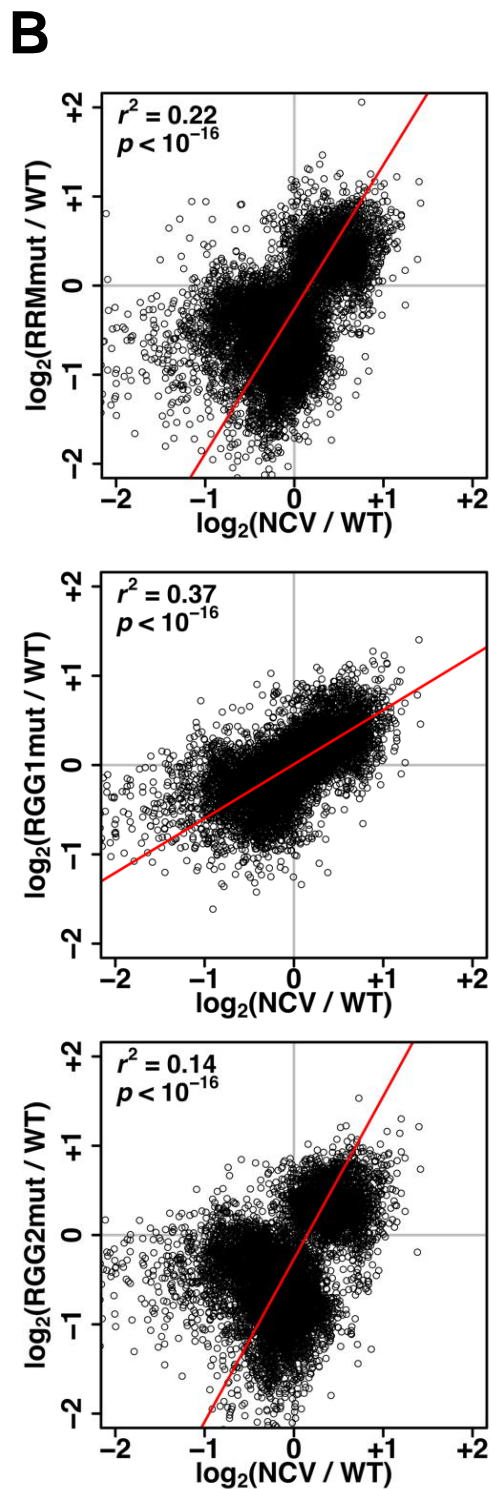
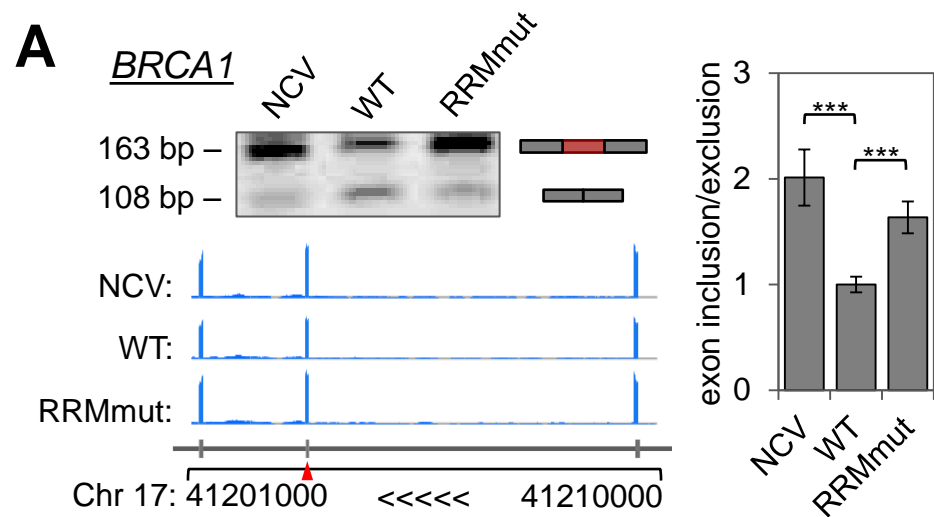
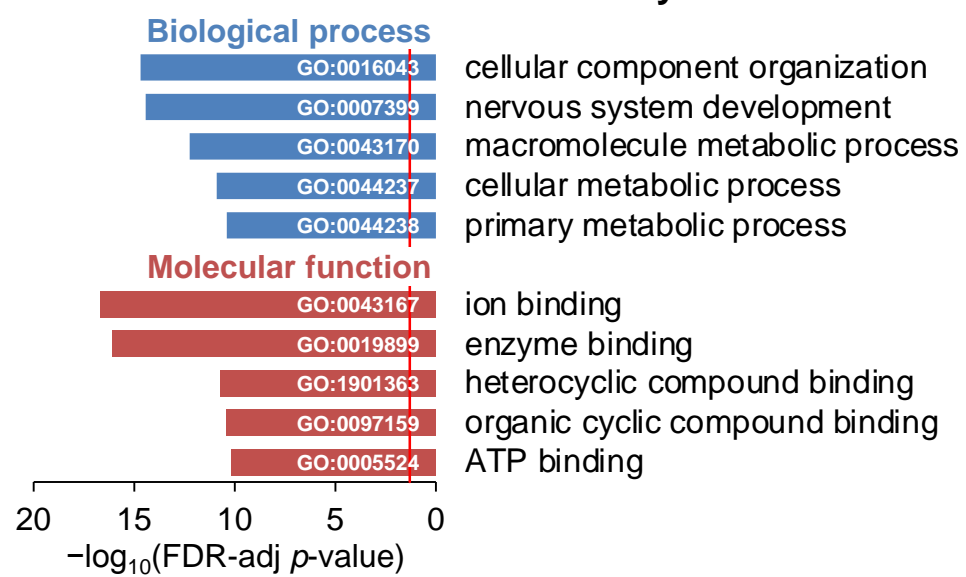
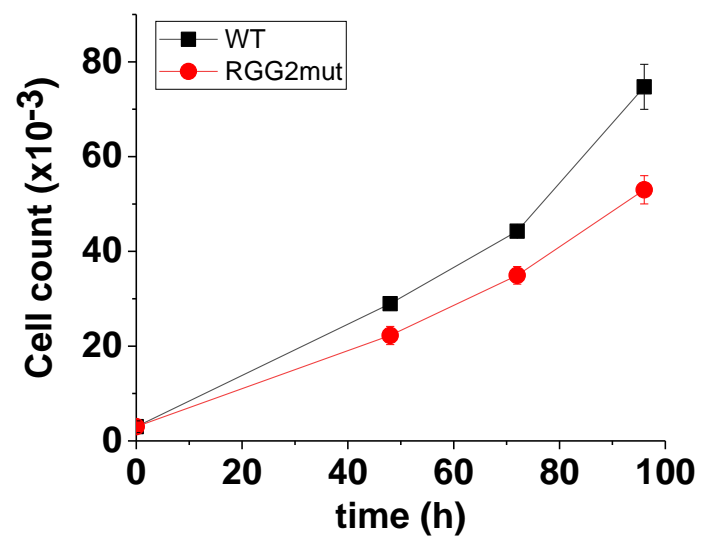


Figure S5

A**Enriched in RGG2mut/WT: GO analysis****B****Figure S6**

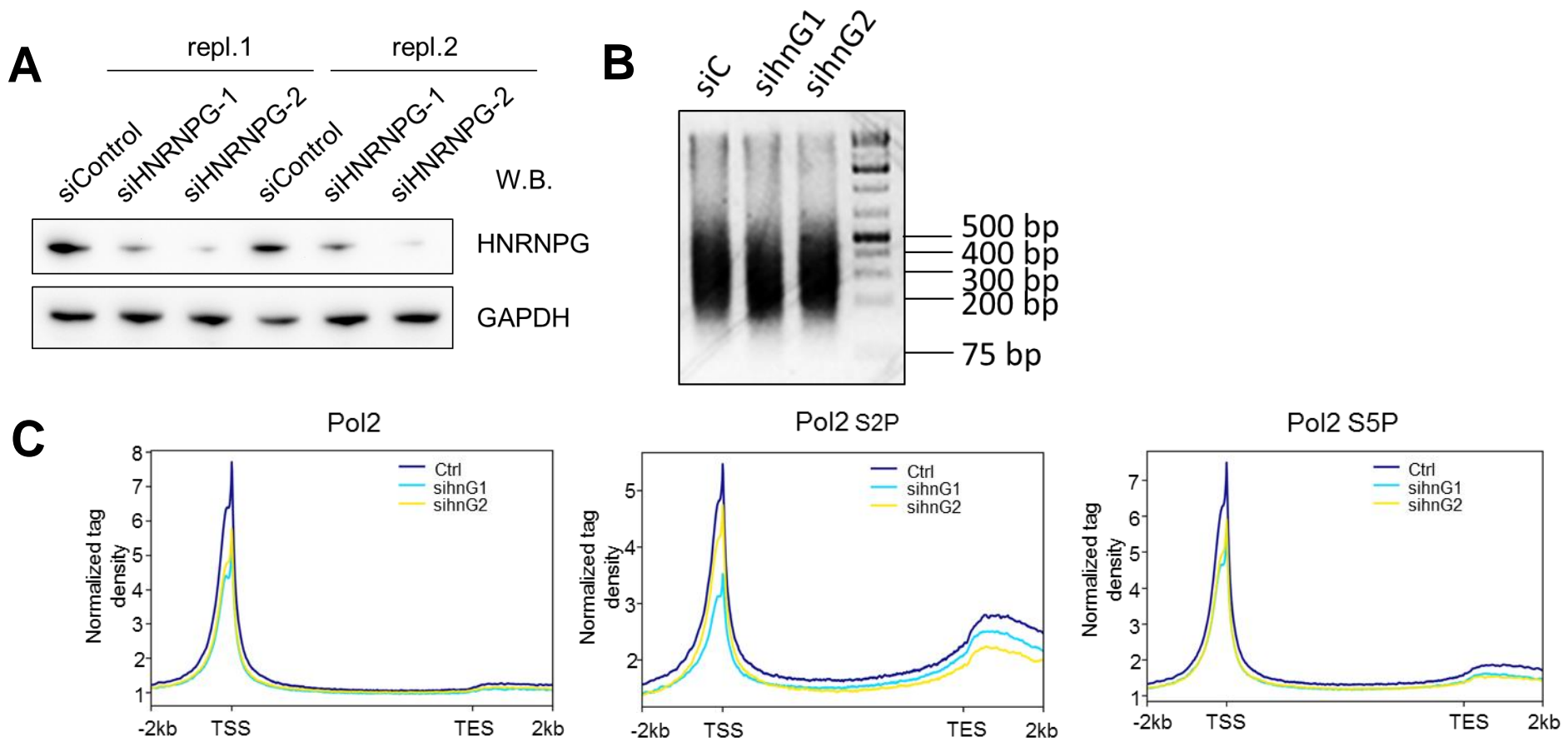


Figure S7

Supplemental figure legends

Figure S1: Fractionation, immunoprecipitation, and domain structure of hnRNPG. Related to Figure 1.

- A. Western blot showing co-immunoprecipitation (co-IP) of RNAPII S5P with hnRNPG in chromatin extracts of cells transfected with control siRNA (ctrl) or one of two different hnRNPG siRNAs (G1 and G2).
- B. Western blot showing co-IP of RNAPII S5P in whole cell lysates upon immunoprecipitation of hnRNPG using three different antibodies (Ab #1–3). in: input (whole cell lysate); un: unbound; IP: hnRNPG IP.
- C. Immunofluorescence staining of hnRNPG in cells treated with 2% v/v H₂O for 9 hours, 20 µg/mL α-amanitin for 9 hours, or 5 µg/mL actinomycin D for 2 hours. Arrowheads: dense clusters of hnRNPG in nuclei of cells treated with α-amanitin or actinomycin D. Scale bars: 10 µm.
- D. Diagram showing the RRM, RGG1, RGG2, and low-complexity regions of full-length hnRNPG, as well as the mutations introduced in the RRM, RGG1, and RGG2 regions to generate RRMmut, RGG1mut, and RGG2mut.
- E. Western blot showing co-IP of METTL3 with RNAPII S5P in the chromatin extract. in: input (chromatin extract); IP: RNAPII S5P IP.

Figure S2: Disorder prediction, purity, self-assembly of hnRNPG, and pCTD binding *in vitro*. Related to Figure 2.

- A. Prediction of intrinsic disorder tendency based on the primary sequence of full-length human hnRNPG (Dosztányi et al., 2005a, 2005b). The RRM, RGG1, and RGG2 regions are identified by horizontal bars. Red color of the disorder tendency curve identifies the locations of Arg-Gly-Gly (RGG) sequences.

- B. Denaturing protein gel showing the full-length WT, RRMmut, RGG1mut, and RGG2mut forms of the hnRNPG protein purified from insect cells. The protein runs as two bands due to variable N-glycosylation in insect cells. Ladder: Novex Sharp Pre-stained Protein Standard (LC5800, Thermo Fisher).
- C. Radius of wild-type hnRNPG assemblies measured by dynamic light scattering at 4 °C with or without pre-binding to MALAT1 34mer RNA hairpin (2515-A/m⁶A). Error bars: ± 1 standard deviation; *n* = 4 replicates; n.s., not significant by two-sample t-test; * *p* < 0.05 by two-sample t-test.
- D. Surface plasmon resonance (SPR) showing binding of 3 μM hnRNPG to immobilized MALAT1 2515-A-Biotin RNA in the presence of 0–12 μM S5P peptide. RU = response units.
- E. Radius of wild-type hnRNPG assemblies, with or without pre-binding of hnRNPG to four molar equivalents of S5P or S2P peptide, measured by dynamic light scattering at 4 °C. Error bars: ± 1 standard deviation for three consecutive measurements of a single sample; n.s., not significant by two-sample t-test.
- F. Spin-down assay for GST-CTD and phosphorylated GST-CTD in the absence of the hnRNPG protein. CDK7 phosphorylates S5, and CDK9 phosphorylates S2 of CTD.
- G. Radius of mutant hnRNPG assemblies measured by dynamic light scattering at 4 °C. Error bars: ± 1 standard deviation; *n* = 3 replicates; n.s., not significant by two-sample t-test; * *p* < 0.05 by two-sample t-test.
- H. Mfold predicted secondary structure of the RNAs used in the binding assay. The size of the RNA is in parentheses. The MALAT1 hairpin can also contain m⁶A (orange dot next to A). The structure of the MALAT1 hairpin has been validated experimentally by in-solution probing (Liu et al., 2017). The delta free energy of the predicted RNA structure is indicated by ΔG per nucleotide (ΔG/nt): higher negative values correspond to more stable RNA structure.
- I. Western blot showing co-IP of recombinant hnRNPG mutants RRMmut, RGG1mut, and RGG2mut with GST-CTD pre-phosphorylated with CDK7 kinase (GST-pCTD).

J. Silver stained denaturing protein gel showing pCTD binding (S5P by CDK7) to wild-type and mutant hnRNPG proteins by the spin-down assay. Final NaCl concentration was 0.25 M; final hnRNPG concentration was 1.7 μ M. s: supernatant fraction; p: pellet fraction.

K. Silver stained denaturing protein gel showing pCTD binding (S2P by CDK9) to wild-type and mutant hnRNPG proteins by the spin-down assay. Final NaCl concentration was 0.25 M; final hnRNPG concentration was 1.7 μ M. s: supernatant fraction; p: pellet fraction.

Figure S3: Knockdown and expression of hnRNPG for mRNA-seq and validation of changes in transcript abundance. Related to Figure 3.

A. Western blot showing knockdown of endogenous hnRNPG and transfection with pCMV3-Flag negative control vector (NCV) or pCMV3-Flag-RBMX encoding wild-type (WT) or mutant (RRMmut, RGG1mut, or RGG2mut) hnRNPG in whole cell lysates. Ctrl: control siRNA; GAPDH: loading control.

B. mRNA-seq reads for *TAF7L* and *CYP2E1* genes upon knockdown of endogenous hnRNPG and transfection of NCV, WT, RRMmut, RGG1mut, or RGG2mut plasmid.

C. RT-qPCR validation of differential transcript abundance upon knockdown of endogenous hnRNPG and transfection of NCV, WT, RRMmut, RGG1mut, or RGG2mut plasmid. Error bars: \pm 1 standard deviation; $n = 3$ biological replicates. Gene ontologies corresponding to analysis in Figure 3C include: TNN, extracellular region; *TAF7L*, cell differentiation; *GYS2*, cell periphery; *KRTAP19-1*, cellular developmental process; *ACTB*, plasma membrane.

D. Correlated changes in transcript abundance, quantified as $\log_2(\text{fold change relative to WT})$, in mRNA sequencing data for NCV, RRMmut, RGG1mut, and RGG2mut. Each point is a differentially expressed gene. r , Pearson correlation coefficient; p , p -value using Fisher transformation; red line, model II major axis linear regression.

Figure S4: hnRNPG and m⁶A methyltransferase co-regulate exon splicing. Related to Figure 4.

A. Gene ontology (GO) analysis showing the false discovery rate (FDR) adjusted p -value ($-\log_{10}$) for biological processes (blue) and molecular functions (red) enriched among genes containing exons that were differentially expressed upon hnRNPG KD relative to Control KD. Red line: FDR threshold ($p = 0.05$).

B. Additional analysis of m⁶A sites and splicing. For the Figure 4D graphs (top) we added empirical distribution function (ECDF), cumulative sum of the affected exons (SUM), and individual exon (each dot) distributions for the co-up-regulated exons.

C. Percentage of exons with hnRNPG-bound m⁶A sites within 300 nucleotides of 3' and 5' splice sites, among exons co-regulated by hnRNPG KD and either METTL3 KD or METTL14 KD.

D. Western blots of GAPDH (a marker of cytoplasm), SNRP70 (a marker of nucleoplasm), Histone3 tri methyl K4 (H3K4me3, a marker of chromatin), and hnRNPG in different cell fractions extracted in the assay of chromatin-associated hnRNPG PAR-CLIP. hnRNPG is present in both nucleoplasm and chromatin. Fractionated lysates corresponding to same numbers of cells were loaded.

E. Western blot of hnRNPG showing its specific pull down in the assay of hnRNPG chromatin PAR-CLIP. The "Input" is the chromatin extract. "FT" stands for the "Flow-Through". Rabbit isotype control antibody (IgG) pull down was performed in parallel as a negative control. The band of hnRNPG is indicated. Protein samples corresponding to same numbers of cells were loaded.

F. Venn diagram showing the overlap between exons that were differentially expressed upon hnRNPG KD (red, 39506 exons) or upon YTHDC1 KD (blue, 28256 exons). The overlapping exons (6669) were regulated upon both hnRNPG KD and YTHDC1 KD.

G. Number of exons up- or down-regulated upon hnRNPG KD or YTHDC1 KD, among the 6669 overlapping exons from Figure S4F. The overlapping exons include 2910 co-regulated (1619 + 1291) and 3759 anti-co-regulated (1909 + 1850) exons.

H. Number of RRACH sites per regulated exon at each site in the -250 to +250 nucleotide region around the 3' and 5' splice sites of exons that were differentially expressed upon hnRNPG KD (red) or

YTHDC1 KD (blue). Exons differentially expressed upon hnRNPG KD (Liu et al., 2017) or YTHDC1 KD (Xiao et al., 2016) were found based on previously published sequencing data. * $p < 10^{-16}$ based on paired t-test between curves for down-regulated versus up-regulated exons.

Figure S5: The RRM, RGG1, and RGG2 regions function in the regulation of alternative splicing by hnRNPG. Related to Figure 5.

A. Native polyacrylamide gel and quantification of RT-PCR validation; mRNA-seq reads showing differential exon usage in *BRCA1* RNA for NCV, WT, and RRMmut. Red arrowhead: alternatively spliced exon. Error bars: ± 1 standard deviation; $n = 3$ biological replicates; *** $p < 0.001$ by two-sample t-test.

B. Correlated changes in exon expression, quantified as $\log_2(\text{fold change relative to WT})$, in mRNA sequencing data for NCV, RRMmut, RGG1mut, and RGG2mut. Each point is a differentially expressed exon. r , Pearson correlation coefficient; p , p -value using Fisher transformation; red line, model II major axis linear regression.

C. Gene ontology (GO) analysis showing the false discovery rate (FDR) adjusted p -value ($-\log_{10}$) for biological processes (blue) and molecular functions (red) enriched among genes containing exons that were differentially expressed in NCV, RRMmut, or RGG1mut relative to WT. Red line: FDR threshold ($p = 0.05$).

Figure S6: The RGG2 mutation affects the splicing of many genes and cell proliferation. Related to Figures 5 and 6.

A. Gene ontology (GO) analysis showing the false discovery rate (FDR) adjusted p -value ($-\log_{10}$) for biological processes (blue) and molecular functions (red) enriched among genes containing exons that were differentially expressed in RGG2mut relative to WT. Red line: FDR threshold ($p = 0.05$).

B. Cell proliferation of HEK293T cells containing WT or RGG2mut hnRNPG protein. Data are mean \pm s.d.; $n = 8$ biological replicates.

Figure S7: m⁶A-dependent regulation of alternative splicing. Related to Figure 7.

A. Western blot of hnRNPG knockdown using two different siRNAs for the RNAPII ChIP-seq experiments. GAPDH is the loading control.

B. Fragmented DNA from cells treated with control RNA (siC) or hnRNPG siRNA (sihnG1, sihnG2), which was used for library preparation for ChIP-seq.

C. Overall pattern of RNAPII ChIP-seq densities at the transcription start site (TSS), gene body, and transcription end site (TES) after normalization to the spiked-in *Drosophila* chromatin.

Table S1: Summary of the mRNA sequencing samples for splicing analysis (100 bp paired-end). Related to Figures 3 and 5.

Sample (sum of 3 replicates)	Total reads	Uniquely mapped reads	Uniquely mapped rate	Correlation (r^2) in \log_2 (reads) between replicates	
				Transcript abundance	Exon expression
NCV	289 065 456	263 566 171	91.2%	0.985–0.994	0.938–0.984
WT	284 794 332	261 018 028	91.7%	0.977–0.989	0.902–0.961
RRMmut	289 064 553	261 053 178	90.3%	0.976–0.991	0.920–0.966
RGG1mut	298 419 121	273 629 200	91.7%	0.980–0.993	0.905–0.981
RGG2mut	294 780 703	271 835 951	92.2%	0.976–0.992	0.907–0.976
Individual sample	Input (pairs)	Uniquely mapped (pairs)	Uniquely mapped (%)		
NCV_1	49,728,476	44,985,207	90.46%		
NCV_2	44,724,346	40,955,949	91.57%		
NCV_3	53,127,004	48,606,533	91.49%		
WT_1	46,063,721	42,410,275	92.07%		
WT_2	43,507,198	39,309,288	90.35%		
WT_3	56,752,091	52,350,145	92.24%		
RRM_1	43,459,345	39,709,083	91.37%		
RRM_2	45,506,769	39,889,394	87.66%		
RRM_3	57,128,232	52,422,036	91.76%		
RGG1_1	48,783,884	44,701,286	91.63%		
RGG1_2	47,462,619	42,693,400	89.95%		
RGG1_3	50,074,884	46,736,044	93.33%		
RGG2_1	46,544,832	43,322,431	93.08%		
RGG2_2	49,128,789	44,851,108	91.29%		
RGG2_3	56,450,285	51,975,846	92.07%		

For each sample, endogenous hnRNPG was knocked down, and either FLAG (NCV, negative control vector) or FLAG–hnRNPG (WT, RRMmut, RGG1mut, or RGG2mut) was expressed in HEK293T cells, prior to RNA extraction for mRNA sequencing. The square of the Pearson correlation coefficient (r^2) is reported as the range of r^2 values obtained by pairwise comparison of gene or exon expression in three biological replicates (three pairwise comparisons: replicates 1 and 2, replicates 2 and 3, replicates 1 and 3).

Table S2: Summary of the chromatin hnRNPG PARCLIP sequencing samples (84 bp single-end). Related to Figure 4.

Sample	Total reads	Uniquely mapped reads	Uniquely mapped rate	Average fragment mapped length
chromatin-hnG-PAR-CLIP-R1	57,095,456	40,395,102	70.8%	38.9 nt
chromatin-hnG-PAR-CLIP-R2	57,914,425	41,864,901	72.3%	38.9 nt
chromatin-hnG-PAR-CLIP-R3	64,480,814	46,427,736	72.0%	48.1 nt

WT HEK293T cells were fed with 200 μ M 4SU for 14 hours before crosslinking by 365 nm UV light. The chromatin fraction was separated and then solubilized by micrococcal nuclease (MNase) digestion. MNase also digests RNAs that are not protected by protein binding. The chromatin extract was then subject to hnRNPG pull down using Rabbit anti-hnRNPG antibody (ab190352, Abcam) and Protein A beads. The RNAs crosslinked with hnRNPG were end-repaired on beads, and size-selected using SDS-PAGE gels. The RNAs were recovered by reverse-crosslinking with Protease K, Acid-Phenol:Chloroform extraction, and overnight ethanol precipitation. RNA libraries were generated using NEBNext multiplex small RNA library preparation kit (NEB, E7300S) and sequenced on NextSeq500. Three biological replicates were sequenced.

Table S3: Summary of the Pol II Chip-seq samples upon hnRNPG knockdown (84 bp single-end). Related to Figure 7.

Sample	Total reads	Mapped reads	Spike-in reads
siC-1-Input	30,258,545	27,774,266	14,110
siC-1-Pol2	32,948,392	30,398,565	11,564
siC-1-S2P	33,332,882	30,897,131	20,640
siC-1-S5P	34,057,394	31,486,988	19,197
sihnG3-1-Input	29,972,085	27,405,787	13,892
sihnG3-1-Pol2	35,190,065	32,348,844	12,681
sihnG3-1-S2P	34,104,827	31,555,444	22,463
sihnG3-1-S5P	33,663,519	31,053,850	20,862
sihnG4-1-Input	30,563,909	28,222,086	12,693
sihnG4-1-Pol2	34,639,285	31,831,575	12,455
sihnG4-1-S2P	33,158,581	30,683,436	22,398
sihnG4-1-S5P	33,061,182	30,476,045	19,255
siC-2-Input	31105231	28375292	14390
siC-2-Pol2	33330849	30469146	13101
siC-2-S2P	33794356	29291014	32397
siC-2-S5P	33572315	30855436	23159
sihnG3-2-Input	31579658	28790560	14152
sihnG3-2-Pol2	34286055	31391392	13543
sihnG3-2-S2P	36457580	33068363	37083
sihnG3-2-S5P	34333570	31549394	24111
sihnG4-2-Input	32934383	30239511	13138
sihnG4-2-Pol2	35250092	32362805	13021
sihnG4-2-S2P	35142812	32326815	28670
sihnG4-2-S5P	34184457	31472522	21280

Endogenous hnRNPG was knocked down by two different siRNAs respectively in two independent biological samples to avoid potential off-target effects. For each of the ChIP antibodies, the percentages of Spike-in reads in the siControl and the two si-hnRNPG ChIP libraries are very similar, indicating that the ChIP procedure was able to be performed uniformly for all three biological samples.