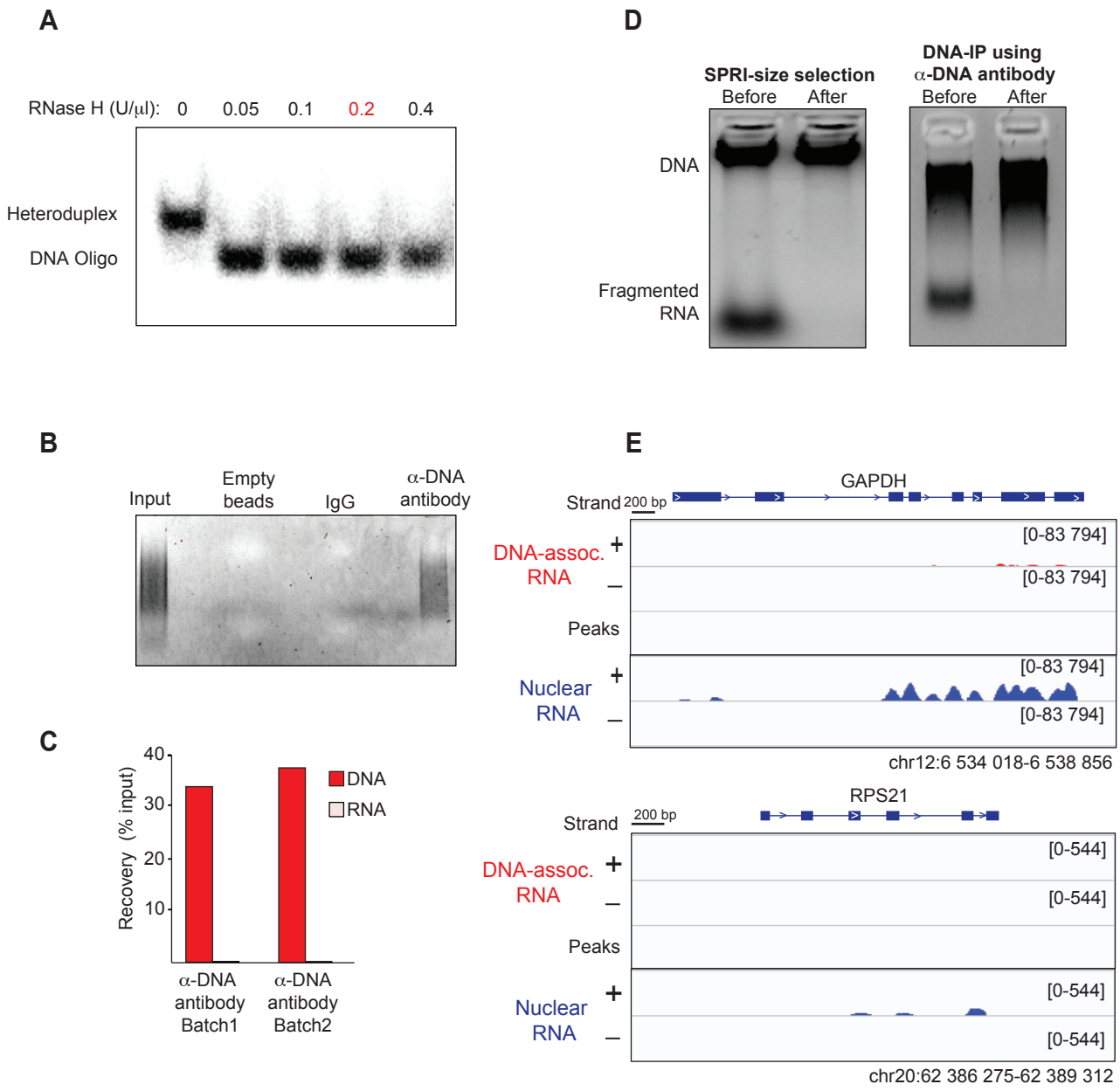


Senturk Cetin et al._Fig. S1

Figure S1.

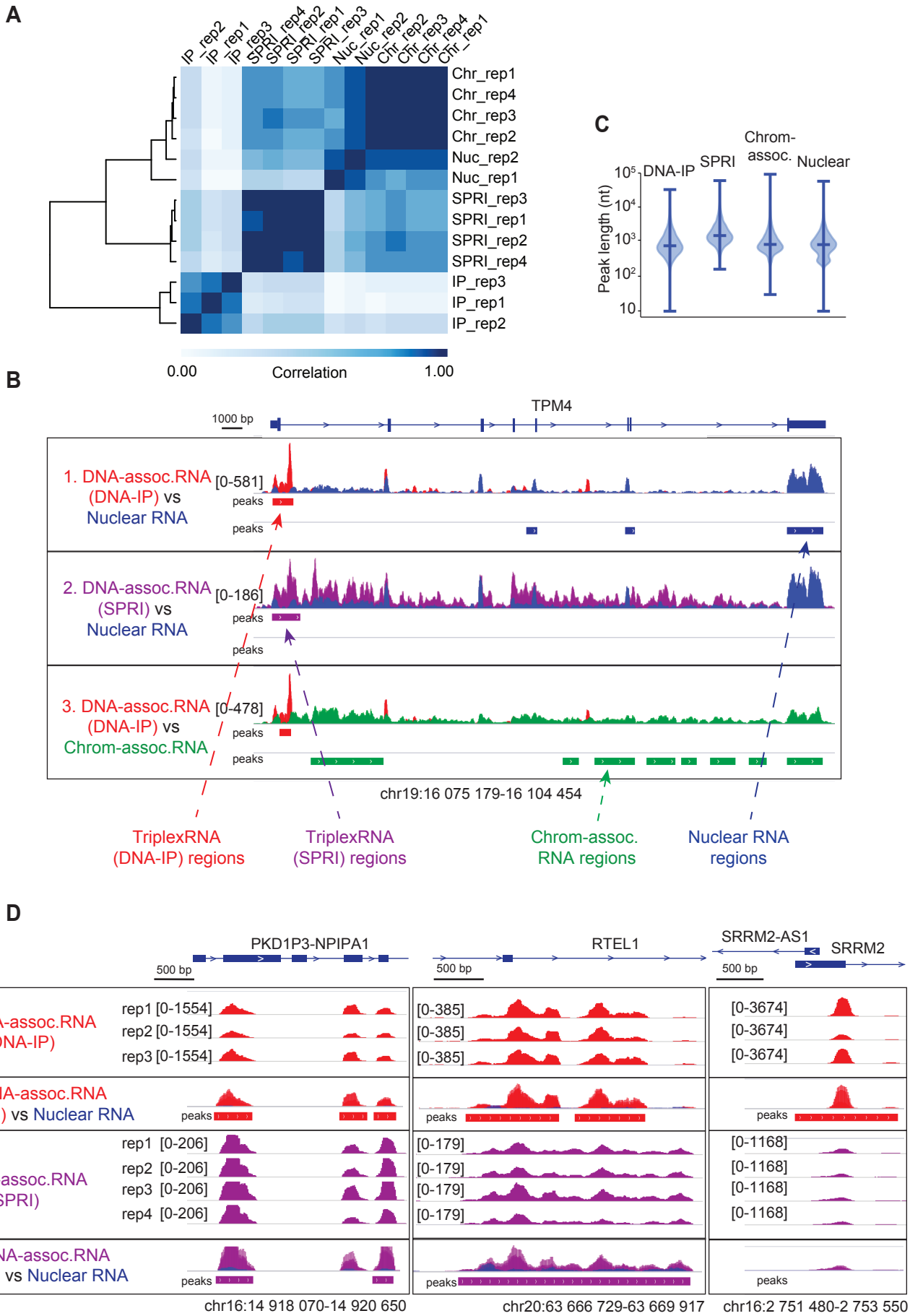
(A) Polyacrylamide gel electrophoresis showing digestion of heteroduplexes in the fractionation protocol. Synthetic heteroduplexes formed with a radiolabeled DNA and an unlabeled RNA oligo were spiked into chromatin-associated nucleic acids and digested with different concentrations of RNase H. 0.2 U/μl of RNase H were used for all further experiments.

(B) Gel electrophoresis validating the recovery of DNA after purification with anti-DNA antibody (DNA-IP) but not with empty beads or IgGs.

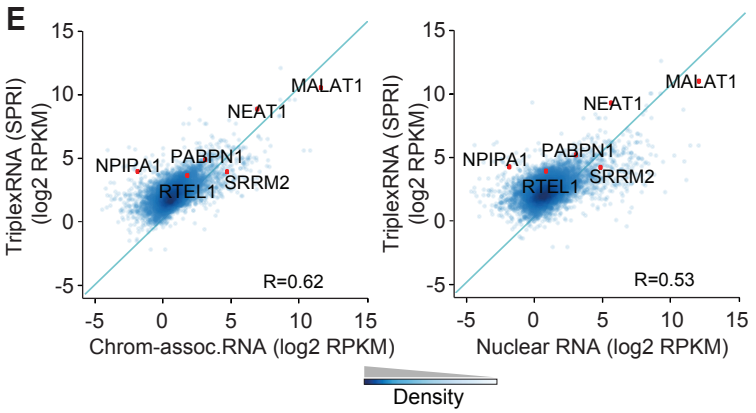
(C) Comparison of DNA and RNA recovery by anti-DNA antibody immunopurification. Radiolabeled RNA was spiked into chromatin-associated nucleic acids before DNA was immunopurified with anti-DNA antibody. DNA and spiked-in RNA were quantified by QUBIT fluorometry and scintillation counting, respectively.

(D) Agarose gel electrophoresis showing separation of RNase I-fragmented RNA (average below 150 nt) from long DNA by SPRI-size selection (left) and DNA-IP (right).

(E) RNA-seq profiles for RPS21and GAPDH in DNA-associated RNA (DNA-IP) and nuclear RNA from U2OS cells. Minus (-) and plus (+) strands are shown.

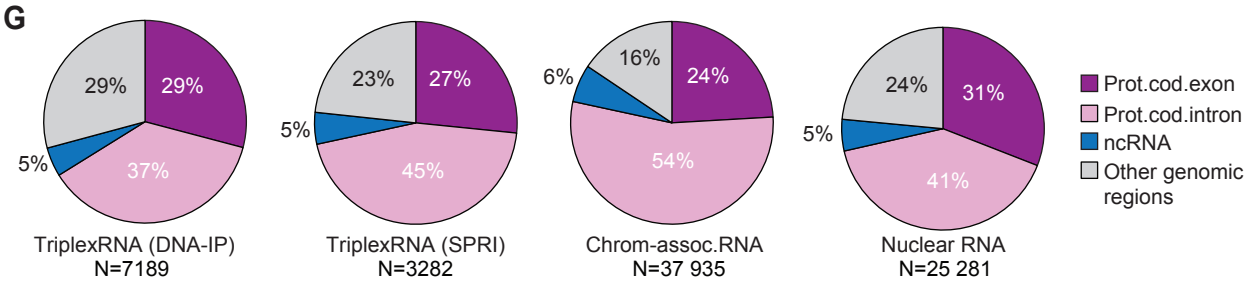


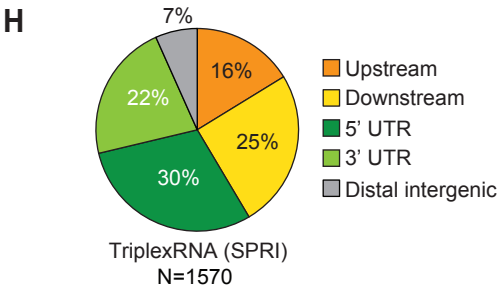
Senturk Cetin et al._Fig. S2_1

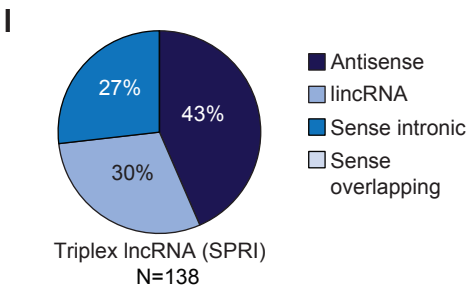


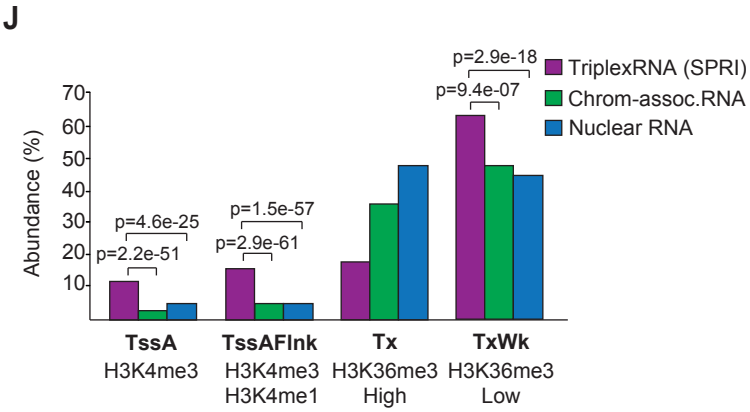
F

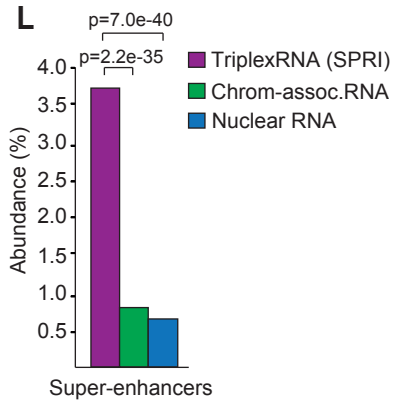
	Total Peaks	Peaks overlapping DRIPc-seq
TriplexRNA (DNA-IP)	7189	1524
TriplexRNA (SPRI)	3282	745

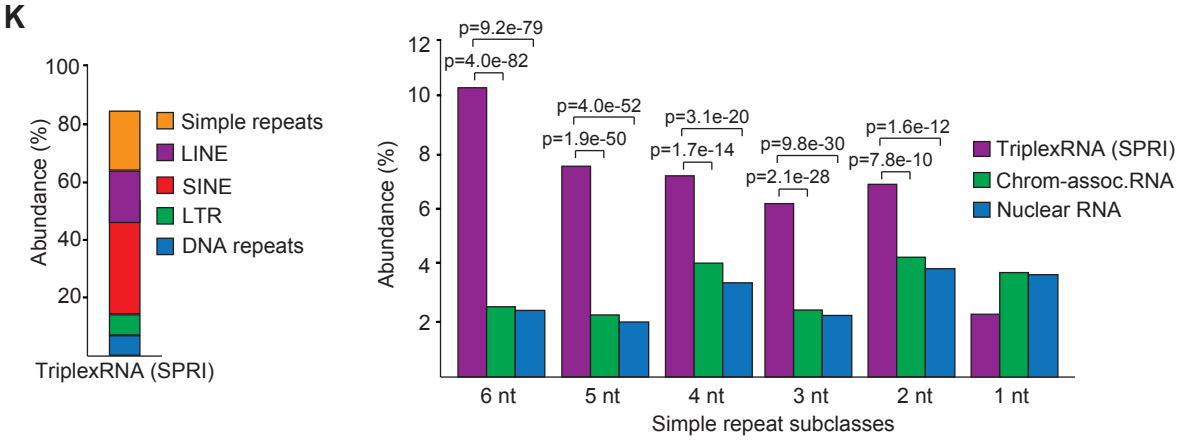












Senturk Cetin et al._Fig. S2_2

Figure S2.

(A) Heatmap demonstrating the correlation of the signals among biological replicates of TriplexRNA-seq libraries from HeLa S3 cells. Pearson's correlation coefficient R is shown.

(B) Computational pipeline to detect enriched RNA regions in DNA-associated, chromatin-associated and nuclear RNA exemplified at the *TPM4* locus. Differential peak calling considering replicates was performed for pairs of RNA fractions. Differential peak regions are indicated below the signal tracks. Nuclear RNA regions are blue (track 1, 2), chromatin associated regions green (track 3), DNA-associated RNA regions from DNA-IP are red (track 1) and from SPRI-size selection are purple (track 2). For simplicity, only positive strands are shown.

(C) Violin plot showing the length of regions enriched in TriplexRNA (DNA-IP & SPRI-size selection), chromatin-associated and nuclear RNA. y-axis is in log10 scale. Whiskers represent 100%, 50% and 0% quantiles of the distribution.

(D) Screenshot of representative TriplexRNA-seq regions, comparing signals from DNA-IP and SPRI-size selection. Signals in biological replicates are shown in Track 1 and 3. Differential peaks identified from the comparison of DNA-associated RNA and nuclear RNA are demonstrated in Track 2 and 4. For simplicity, only positive strands are shown.

(E) Scatter plots showing the correlation of TriplexRNA (SPRI) with chromatin-associated RNA and nuclear RNA across 7148 genes overlapping peaks. Pearson correlation coefficient (R) is shown. Green diagonal line x=y. Some representative genes that overlap TriplexRNAs and control RNAs are highlighted.

(F) Abundance of TriplexRNA peaks in HeLa S3 cells overlapping R-loop regions identified by DRIPc-seq in Ntera2 cells (40).

(G) Pie charts depicting genomic distribution of peaks from TriplexRNA, chromatin-associated and nuclear RNA.

(H) Pie chart depicting the genomic distribution of TriplexRNA (SPRI), excluding intronic and exonic gene regions. Upstream and downstream regions are defined within 2.5 kb proximity of the closest gene.

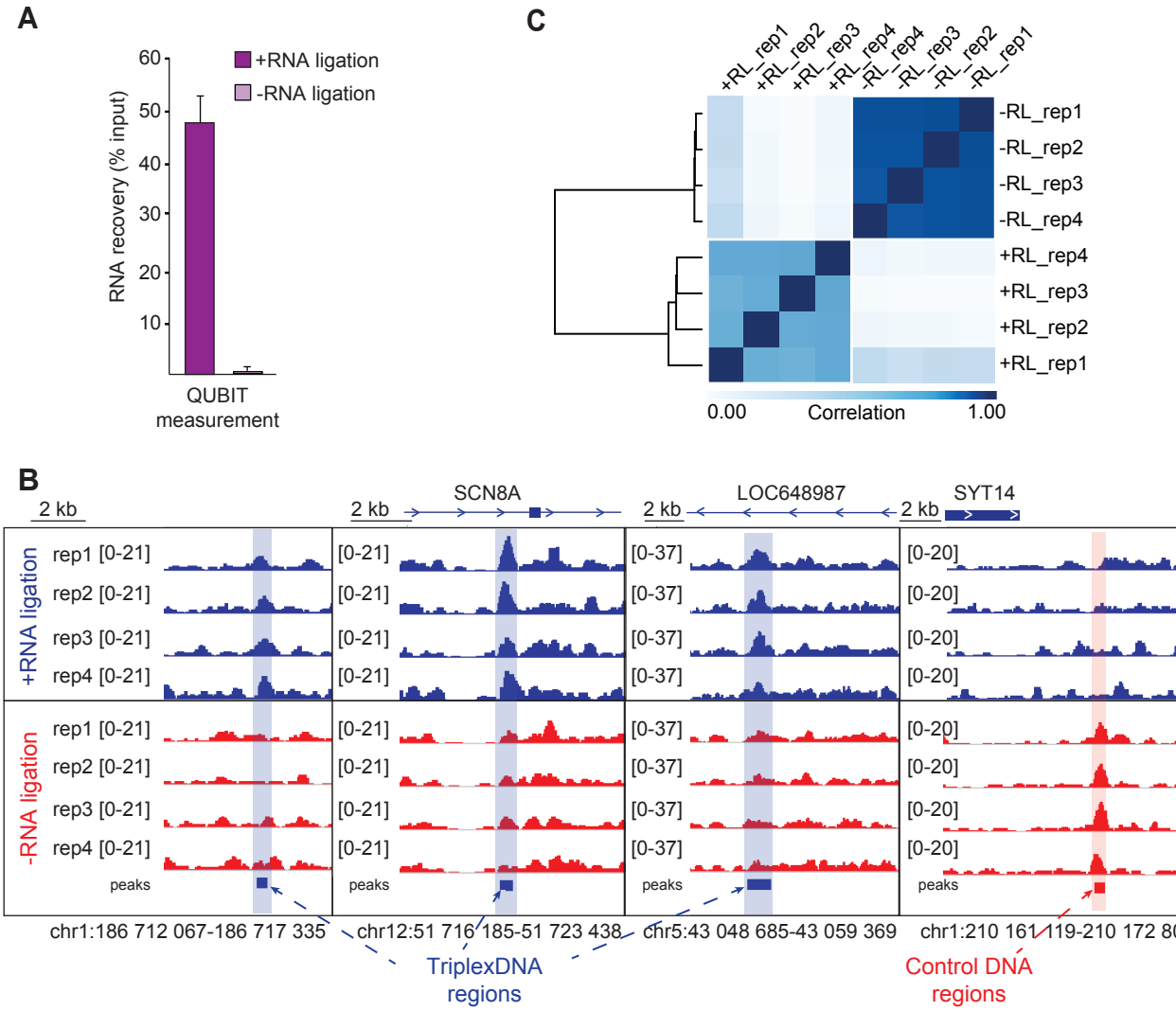
(I) Pie chart showing the classification of lncRNAs that overlap peaks in TriplexRNA (SRPI).

(J) Association of TriplexRNA (SPRI) and control RNAs with ChromHMM promoter states and transcribed states. Active transcription start site (TssA), flanking active TSS (TssAFlnk), strong (Tx) and weak (TxWk) transcription regions are shown.

(K) Left: TriplexRNA (SPRI) overlapping different classes of repeat elements. Right: Abundance of simple repeat subclasses.

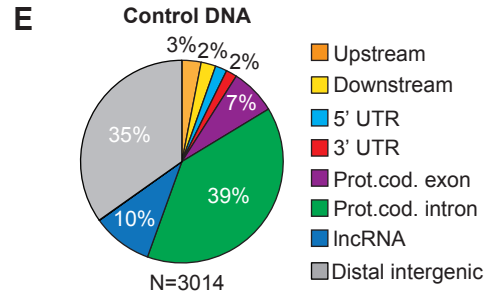
(L) Abundance of TriplexRNA peaks (SPRI) overlapping super-enhancers in HeLa S3 cells.

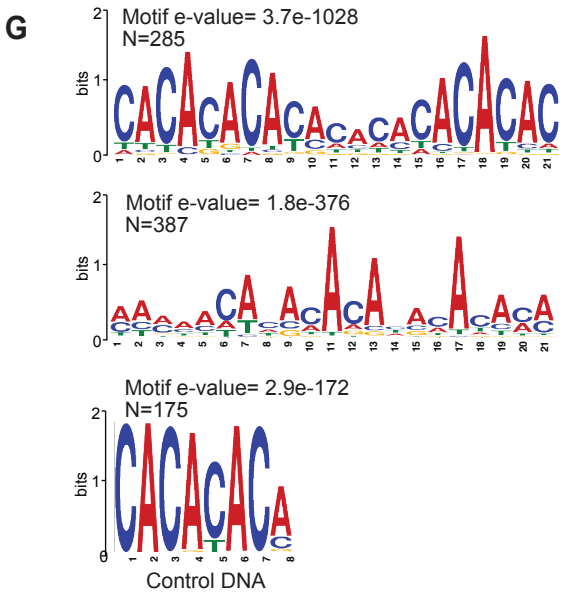
Adjusted p-values <0.05 in panels (J-L) were obtained from one-tailed Fisher's exact test using chromatin-associated and nuclear RNA as control.

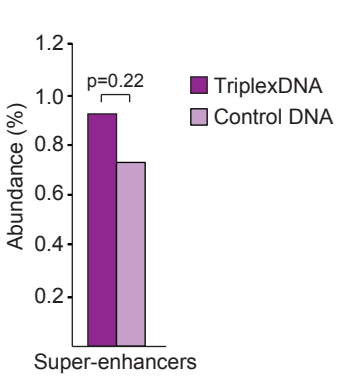


D

	Total Peaks	Peaks overlapping DRIPc-seq
TriplexDNA-seq	2547	331







Senturk Cetin et al._Fig. S3

Figure S3.

(A) QUBIT fluorometric quantification of RNAs recovered from ligation-mediated pull-down of chromatin-associated RNAs as compared to control samples without ligation. Values are presented relative to input RNA (±SD, N=4).

(B) Screenshots of representative TriplexDNA-seq regions, displaying signals in biological replicates. Enriched regions identified by differential peak calling are shaded.

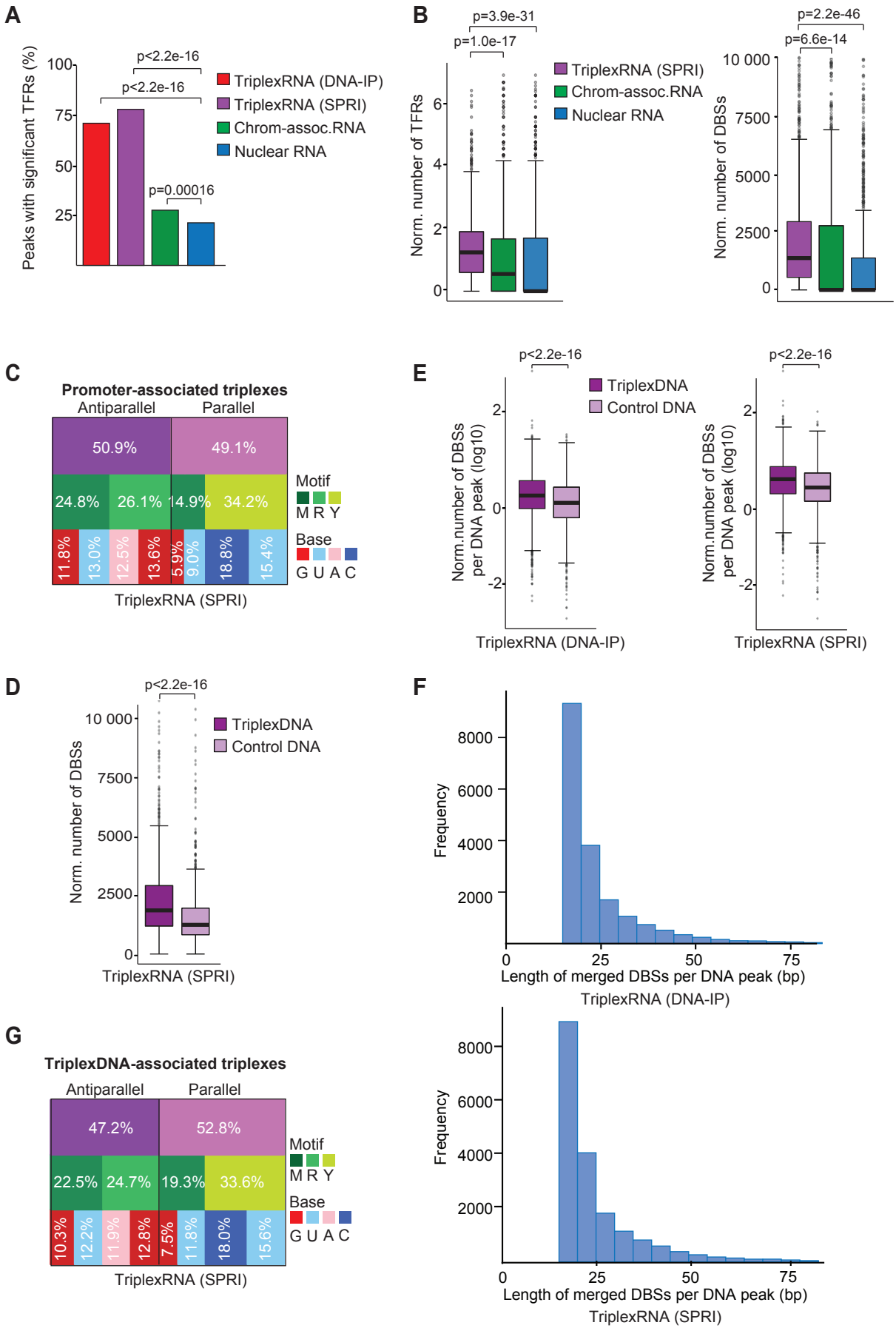
(C) Heatmap depicting the correlation of the signals among biological replicates of TriplexDNA-seq libraries from HeLa S3 cells. Spearman's correlation coefficient (r_s) is shown.

(D) Abundance of TriplexDNA peaks in HeLa S3 cells overlapping R-loop regions identified by DRIPc-seq in Ntera2 cells (40).

(E) Pie charts depicting the genomic distribution of control DNA regions. Upstream and downstream regions are defined within 2.5 kb proximity of the closest gene.

(F) Abundance of TriplexDNA-seq peaks overlapping super-enhancers in HeLa S3 cells. Adjusted p-value was obtained from one-tailed Fisher's exact test.

(G) MEME motif analysis identifying consensus motifs in randomly selected 500 control DNA peaks.

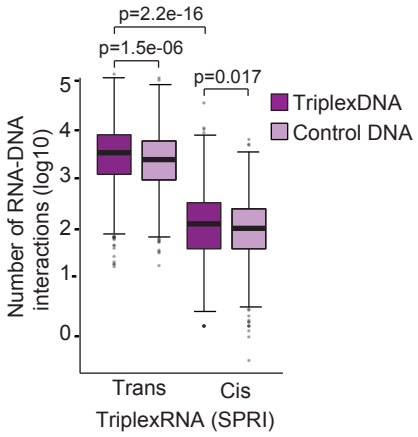

A
Peaks with significant TFRs (%)
100
75
50
25
p<2.2e-16
p<2.2e-16
p=0.00016
TriplexRNA (DNA-IP)
TriplexRNA (SPRI)
Chrom-assoc.RNA
Nuclear RNA
B
p=3.9e-31
p=1.0e-17
Norm. number of TFRs
6
4
2
0
TriplexRNA (SPRI)
Chrom-assoc.RNA
Nuclear RNA
p=2.2e-46
p=6.6e-14
Norm. number of DBSs
10 000
7500
5000
2500
0
C
Promoter-associated triplexes
Antiparallel
Parallel
50.9%
49.1%
24.8%
26.1%
14.9%
34.2%
11.8%
13.0%
12.5%
13.6%
5.9%
9.0%
18.8%
15.4%
Motif
M R Y
Base
G U A C
TriplexRNA (SPRI)
E
p<2.2e-16
TriplexDNA
Control DNA
Norm.number of DBSs per DNA peak (log10)
2
0
-2
TriplexRNA (DNA-IP)
p<2.2e-16
Norm.number of DBSs per DNA peak (log10)
TriplexRNA (SPRI)
D
p<2.2e-16
TriplexDNA
Control DNA
Norm. number of DBSs
10 000
7500
5000
2500
0
TriplexRNA (SPRI)
F
Frequency
8000
6000
4000
2000
0
25
50
75
Length of merged DBSs per DNA peak (bp)
TriplexRNA (DNA-IP)
Frequency
8000
6000
4000
2000
0
25
50
75
Length of merged DBSs per DNA peak (bp)
TriplexRNA (SPRI)
G
TriplexDNA-associated triplexes
Antiparallel
Parallel
47.2%
52.8%
22.5%
24.7%
19.3%
33.6%
10.3%
12.2%
11.9%
12.8%
7.5%
11.8%
18.0%
15.6%
Motif
M R Y
Base
G U A C
TriplexRNA (SPRI)

H

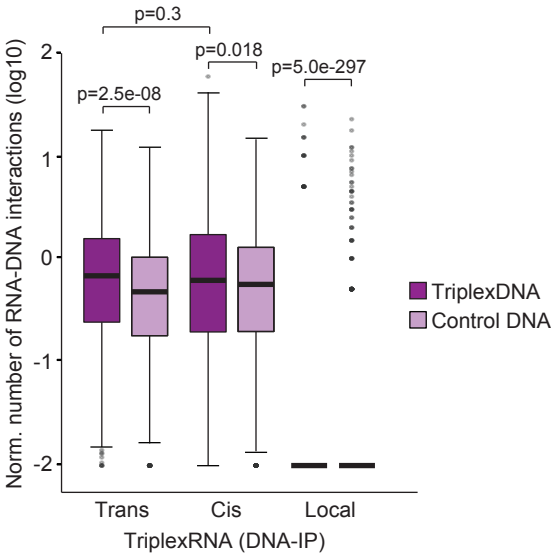
Triplex interactions between TriplexRNA and TriplexDNA

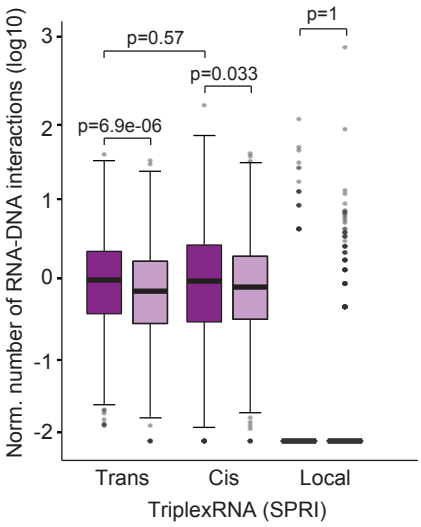
	Trans	Cis	Local
DNA-IP	3 662 965	162 345	45
SPRI	6 760 593	357 172	98

I



J





K

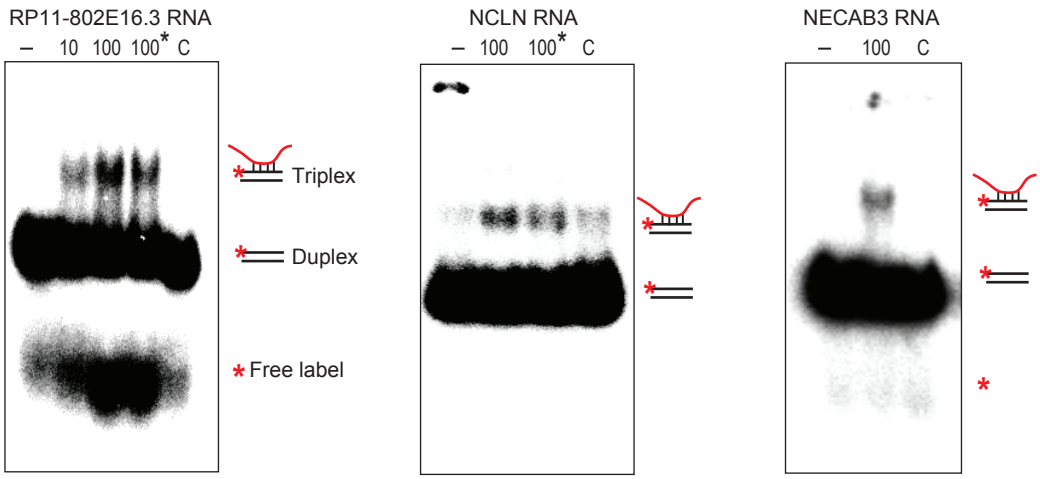


Figure S4.

(A) TDF analysis showing the number of peaks in TriplexRNA, chromatin-associated RNA and nuclear RNA that have potential to form triplex with active promoters defined by ChromHMM. Adjusted p-values were obtained from one-tailed Fisher's exact test.

(B) TDF analysis predicting the potential of top 1000 enriched TriplexRNA (SPRI) regions (ranked by peak p-value) to bind to active promoters. Left: Number of TFRs in RNA (per kilobase of RNA). Right: Number of putative DBSs at promoters (per kilobase of RNA).

(C) Motif analysis of triplexes formed between TriplexRNA (SPRI) and active promoters. The diagram depicts the fraction of antiparallel and parallel triplexes with the respective motif and nucleotide composition of TFRs in TriplexRNA.

(D) TDF analysis comparing the triplex-forming potential of top 2000 TriplexDNA-seq regions with top 1000 TriplexRNA (SPRI) (ranked by peak p-value). The number of putative DBSs (per kilobase of RNA) is shown.

(E) Box plot showing the number of DBSs per TriplexDNA peak normalized by the peak length.

(F) Frequency and length of DBSs in TriplexDNA peaks. Merged length is shown for peaks with overlapping DBSs.

(G) Motif analysis of predicted triplexes formed between TriplexRNA (SPRI) and TriplexDNA. The diagram depicts the fraction of antiparallel and parallel triplexes and the respective motif and nucleotide composition of TFRs in TriplexRNA.

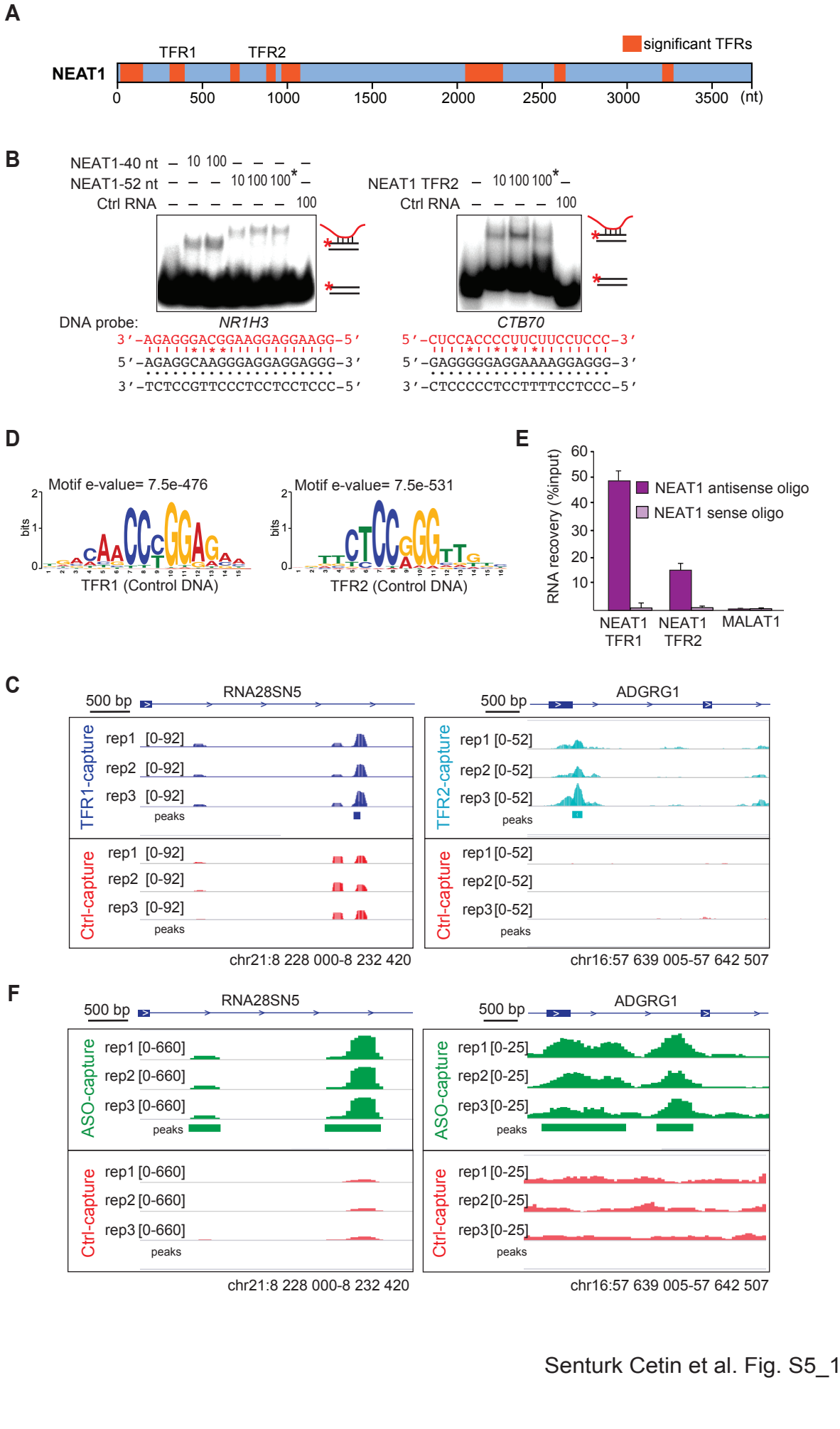
(H) Table showing the numbers of predicted triplexes between TriplexRNA and TriplexDNA regions, classified as local (within 10 kb distance), *cis* (>10 kb in the same chromosome) and *trans* (at different chromosomes).

(I) Box plot classifying triplex interactions between TriplexRNAs (SPRI) and TriplexDNA-seq regions as *cis* (>10 kb in the same chromosome) and *trans* (at different chromosomes) interactions, excluding underrepresented local interactions (within 10 kb distance).

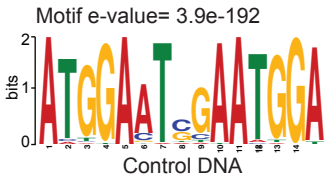
(J) Boxplot showing the number of local, *cis* and *trans* triplex interactions between TriplexRNA and TriplexDNA regions normalized by the expected genomic region size.

(K) Full gel images of the EMSAs shown in **Figure 4F**. Positions of triplex, duplex and free radiolabel are indicated.

Adjusted p-values <0.05 in panels (B,D,E,I,J) are based on one-tailed Mann-Whitney test.

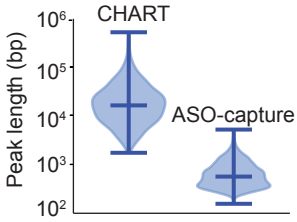

A
TFR1
TFR2
significant TFRs
NEAT1
0
500
1000
1500
2000
2500
3000
3500
(nt)
B
NEAT1-40 nt
NEAT1-52 nt
Ctrl RNA
NEAT1 TFR2
Ctrl RNA
DNA probe:
NR1H3
CTB70
3′-AGAGGGACGGAAGGAGGAAGG-5′
5′-AGAGGCAAGGGAGGAGGAGGG-3′
3′-TCTCCGTTCCCTCCTCCTCCC-5′
5′-CUCCACCCCUUCUUCCUCCC-3′
5′-GAGGGGGAGGAAAAGGAGGG-3′
3′-CTCCCCCTCCTTTTCCTCCC-5′
D
Motif e-value= 7.5e-476
TFR1 (Control DNA)
Motif e-value= 7.5e-531
TFR2 (Control DNA)
bits
E
RNA recovery (%input)
NEAT1 antisense oligo
NEAT1 sense oligo
NEAT1 TFR1
NEAT1 TFR2
MALAT1
C
500 bp
RNA28SN5
ADGRG1
TFR1-capture
TFR2-capture
Ctrl-capture
rep1 [0-92]
rep2 [0-92]
rep3 [0-92]
rep1 [0-52]
rep2 [0-52]
rep3 [0-52]
peaks
chr21:8 228 000-8 232 420
chr16:57 639 005-57 642 507
F
ASO-capture
rep1 [0-660]
rep2 [0-660]
rep3 [0-660]
rep1 [0-25]
rep2 [0-25]
rep3 [0-25]

G



H

	Total Peaks	Peaks overlapping ASO-capture-seq	Peaks overlapping CHART-seq
ASO-capture-seq	3780	3780	476
CHART-seq	1244	233	1244



I

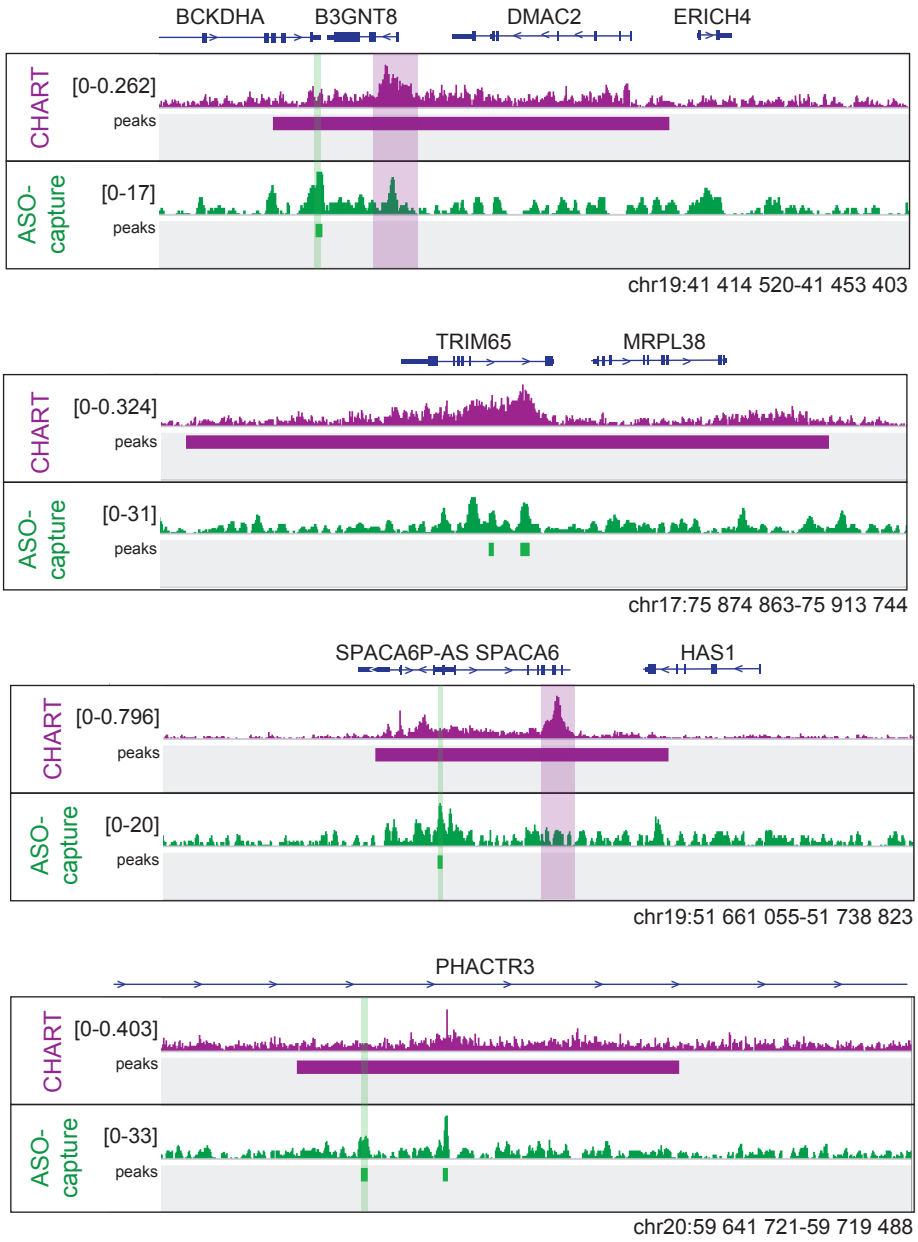


Figure S5.

(A) TDF-mediated prediction of TFRs in NEAT1 which bind to DNA regions that have been identified by CHART-seq (7). TFR1 and TFR2 positions are indicated.

(B) EMSAs using 10 or 100 pmol of synthetic NEAT1 versions comprising TFR1 (40 or 52 nt) or TFR2 incubated with 0.25 pmol of double–stranded ^{32}P-labeled oligonucleotides (**Supplementary Table S2**). NEAT1 targets were from CHART-seq data (7). Reactions marked with an asterisk (*) were treated with 0.5 U RNase H. RNA without a putative TFR was used as a control. Potential Hoogsteen base pairing between motifs and respective TFR sequences are shown; mismatches are marked (*).

(C) Screenshots of representative regions enriched by NEAT1-TFR capture. Signals from biological replicates are displayed.

(D) MEME motif analysis identifying consensus motifs in control DNA regions from TFR1-based (in 376 of top 500 peaks) and TFR2-based (in 364 of top 500 peaks) capture experiments. Ranking is based on peak p-values.

(E) RT-qPCR analysis of RNA recovered by ASO-based capture of NEAT1. As control, a biotinylated sense oligo with the same sequence as NEAT1 was used. Data are presented relative to input (±SD, N=4).

(F) Screenshots of representative regions enriched by NEAT1-ASO capture, displaying signals in biological replicates.

(G) MEME motif analysis identifying a consensus motif in 61 of top 500 peaks control DNA regions identified by ASO-based NEAT1 capture experiments. Ranking is based on peak p-values.

(H) Top: Overlap between data from CHART-seq (7) and ASO-capture-seq in MCF7 cells. Bottom: Violin plots showing the peak lengths in both approaches. y-axis is in log10 scale. Whiskers represent 100%, 50% and 0% quantiles of the distribution.

(I) Screenshots of representative regions that are co-enriched by ASO-capture-seq and CHART-seq, Identified peaks in both approaches are indicated below the signal tracks. The regions shaded green show ASO-capture signals underrepresented in CHART-seq. The purple shaded regions demonstrate the CHART-seq signals which are not detected as peak in ASO-capture-seq, presumably protein-mediated interactions. For simplicity, control libraries are not included.

Data are from HeLa S3 cells unless otherwise stated.

Supplementary Table S1.
Sequences of oligos used in RT-qPCR analyses

Name	Sequence (5'-3')
18S, for	GCGACCTCAGA TCAGACGTGG
18S, rev	CTGTTCACTCGCCGTTACTGAG
GAPDH, for	GAGTCAACGGATTTGGTCGT
GAPDH, rev	TTGATTTTGGAGGGATCTCG
pre-GAPDH, for	GAGCTGGGGAATGGGACT
pre-GAPDH, rev	TGATGGCATGGACTGTGG
KHPS1, for	CAGGATCCGACCTTTCTGG
KHPS1, rev	ACTCAAGGCTGGTGGTAGTGG
MALAT1, for	AAATGTGAAGGACTTTCGTAACG
MALAT1, rev	CACCTGGGTCAGCTGTCAAT
NEAT1 TFR1, for	TTGCATAGCTGAGCGAGCC
NEAT1 TFR1, rev	TGCCTGCCTTCCTGATCATTT
NEAT1 TFR2, for	GGCGCTAAACTCTTCTTGTGAG
NEAT1 TFR2, rev	CTTACTGTCCCGGGCTTACCA
RPS21, for	TTTCATTCTTACCGCCCAAGTCCC
RPS21, rev	AGCCTCTGCCTGTACCTTCACCC

Supplementary Table S2.
Sequences of RNAs and DNA oligos used in EMSAs

Name	Sequence (5'-3')
RP11-802E16.3 RNA	UGTGCUGGGGAAGAAGAGGAGGAGGUGGAGGAGGGCUGAUG AUGGGGCCACAUGCUGCUGGG
NCLN RNA	GUGGUGGGGAAAGAUGGGAGUUGCCGCUGUGGAGAGUGAGG AGAAA
NECAB3 RNA	GGAGACAGGGAGAAGAGGAUGGAGACAGGGAGGCUGGGGG GCAGGUUGGCGG
NEAT1 TFR1 RNA 40 nt	UGUGGAAGGAGGAAGGCAGGGAGAGGUAGAAGGGGUGGAG
NEAT1 TFR1 RNA 52 nt	UGUGGAAGGAGGAAGGCAGGGAGAGGUAGAAGGGGUGGAG GAGUCAGGAGGA
NEAT1 TFR2 RNA 55 nt	GAGCUCACUCCACCCCUUCUUCCUCCCUUUAACUUAUCCAUU CACUUAAAACAUU
Negative control RNA	GUCGGUCUUAUCAGUUCUCCGGGUUGUCAG
RAB6A, for	GGAGGGGAAGGAAAGAAAGGGGAGGGGGAGAAGGGTGGGG
RAB6A, rev	CCCCACCCTTCTCCCCCTCCCCTTTCTTTCCTTCCCCTCC
NCLN, for	CTGTCCCCACCTTTCGCCCCTCACCCTAGC
NCLN, rev	GCTAGGGTGAGGGGCGAAAGGTGGGGACAG
KLHL21, for	TGGGAGGCTGAGGCAGGAGAATCGCTTGAAC
KLHL21, rev	GTTCAAGCGATTCTCCTGCCTCAGCCTCCCA
FLI1, for	GGGAGAGAGAGAGAGGGAAGGAAGGGAAGG
FLI1, rev	CCTTCCCTTCCTTCCCTCTCTCTCTCCCC
GRIK4, for	GGGAGGAGGAAGAGGGAGAGGGAGGAGGAA

GRIK4, rev	TTCCTCCTCCCTCTCCCTCTTCCTCCTCCC
NR1H3, for	TCCCAGAGGCAAGGGAGGAGGAGGGAGGCT
NR1H3, rev	AGCCTCCCTCCTCCTCCCTTGCCTCTGGGA
CTB70, for	GAAGGAGAGGGGGAGGAAAAGGAGGGTGGA
CTB70, rev	TCCACCCTCCTTTTCCTCCCCCTCTCCTTC
CYP4F22, for	AGAAGGAGGAGGAGAAGAGGGAGGAAGAAG
CYP4F22, rev	CTTCTTCCTCCCTCTTCTCCTCCTCCTTCT

Supplementary Table S3.

Sequences of oligos used in RNA-associated DNA experiments

Name	**Sequence (5'-3')**	**Application**
NEAT1 TFR1 RNA oligo	Pso6C/AAGGAGGAAGGCAGGGAGAGGUAG AAGGGGUGGAGGAGUCAGGAGGA/ TEGbiotin	Capture of NEAT1-assoc. DNA
NEAT1 TFR2 RNA oligo	Pso6C/UCACUCCACCCCUUCUUCCUCCCUU UAACUUAUCCA/TEGbiotin	Capture of NEAT1-assoc. DNA
RNA control oligo	Pso6C/GUCGGUCUUAUCAGUUCUCCGGGU UGUCAG/TEGbiotin	Capture of NEAT1-assoc. DNA
NEAT1 antisense oligo	TGTCTGTCCCCTGAAGCCCTG/TEGbiotin	ASO-based capture of NEAT1-assoc. DNA
NEAT1 sense oligo	GCTGGGGGTGCGAGAAGGAA/TEGbiotin	ASO-based capture of NEAT1-assoc. DNA
Ligation linker oligo	5Phos/GTTGGAGTTCGGTGTGTGGG/ TEGbiotin	Isolation of RNA-assoc. DNA

Supplementary Table S4.

Read/ peak numbers in TriplexRNA-seq

Samples	**Library average size (bp)**	**Uniquely mapped reads**	**Total reads**	**% of uniquely mapped reads**
Nuclear_rep1	330	39 081 484	49 027 136	79.71
Nuclear_rep2	346	29 714 487	36 158 302	82.18
Chromatin_rep1	383	38 461 945	46 939 070	81.94
Chromatin_rep2	364	39 710 878	47 898 697	82.91
Chromatin_rep3	387	33 644 670	39 119 495	86.00
Chromatin_rep4	384	37 304 975	43 641 809	85.48
SPRI_rep1	429	35 871 152	43 745 808	82.00
SPRI2_rep2	413	37 093 026	44 458 085	83.43
SPRI_rep3	433	35 327 259	42 523 924	83.08
SPRI_rep4	421	35 286 452	43 094 668	81.88
IP_rep1	336	28 510 682	60 801 254	46.89
IP_rep2	368	23 752 023	45 413 818	52.30
IP_rep3	351	32 159 970	68 273 091	47.10

Data sets	**Comparison**	**Number of peaks**
TriplexRNA (SPRI)	SPRI vs. Nuclear	3282
TriplexRNA (DNA-IP)	DNA-IP vs. Nuclear	7189
Chrom-assoc.RNA	Chrom-assoc. vs DNA-IP	37 935
Nuclear RNA	Nuclear vs. DNA-IP	25 281

Supplementary Table S5.
Read/ peak numbers in TriplexDNA-seq

Samples	Library average size (bp)	Uniquely mapped reads	Total reads	% of uniquely mapped reads
-RL_rep1	350	28 250 630	33 672 339	83.90
+RL_rep1	348	28 131 092	34 159 767	82.35
-RL_rep2	364	25 134 249	30 632 181	82.05
+RL_rep2	364	23 459 098	27 595 132	85.01
-RL_rep3	357	26 968 596	31 906 908	84.52
+RL_rep3	376	23 547 362	27 698 628	85.01
-RL_rep4	380	28 751 651	34 463 848	83.43
+RL_rep4	364	23 515 829	27 735 597	84.79
Data sets	**Comparison**			**Number of peaks**
TriplexDNA	+RNA ligase vs. –RNA ligase			2547
Control DNA	-RNA ligase vs. +RNA ligase			3014

Supplementary Table S6.
Read/ peak numbers in NEAT1-TFR-capture-seq

Samples	Library average size (bp)	Uniquely mapped reads	Total reads	% of uniquely mapped reads
Control_rep1	311	49 707 873	53 256 870	93.34
Control_rep2	331	35 594 218	38 239 800	93.08
Control _rep3	330	38 038 763	40 927 647	92.94
TFR1_rep1	330	46 966 681	51 630 284	90.97
TFR1_rep2	308	40 827 126	44 675 435	91.39
TFR1_rep3	329	45 171 398	49 180 705	91.85
TFR2_rep1	301	48 770 794	53 942 991	90.41
TFR2_rep2	346	38 557 848	42 932 097	89.81
TFR2_rep3	320	94 597 187	104 703 363	90.35
Data sets	**Comparison**			**Number of peaks**
TFR1-assoc.DNA	TFR1 vs. Control oligo			622
TFR1 control DNA	Control vs. TFR1 oligo			4001
TFR2-assoc.DNA	TFR2 vs. Control oligo			4423
TFR2 control DNA	Control vs. TFR2 oligo			1594

Supplementary Table S7.
Read/ peak numbers in ASO-mediated-capture-seq

Samples	**Library average size (bp)**	**Uniquely mapped reads**	**Total reads**	**% of uniquely mapped reads**
Sense_rep1	364	31 547 105	36 865 879	85.57
Antisense_rep1	367	26 631 756	31 011 185	85.88
Sense_rep2	372	33 532 439	38 621 482	86.82
Antisense_rep2	373	37 575 051	43 419 553	86.54
Sense_rep3	289	44 050 048	50 702 789	86.88
Antisense_rep3	308	36 859 683	43 023 472	85.67
Sense_MCF7	285	40 345 941	46 534 218	86.70
Antisense_MCF7	335	40 516 315	47 114 634	86.00
Data sets	**Comparison**			**Number of peaks**
NEAT1-assoc.DNA HeLa S3	Antisense vs. Sense oligo			3692
Control DNA HeLa S3	Sense vs. Antisense oligo			2066
NEAT1-assoc.DNA MCF7	Antisense vs. Sense oligo			3780
Control DNA MCF7	Sense vs. Antisense oligo			2240