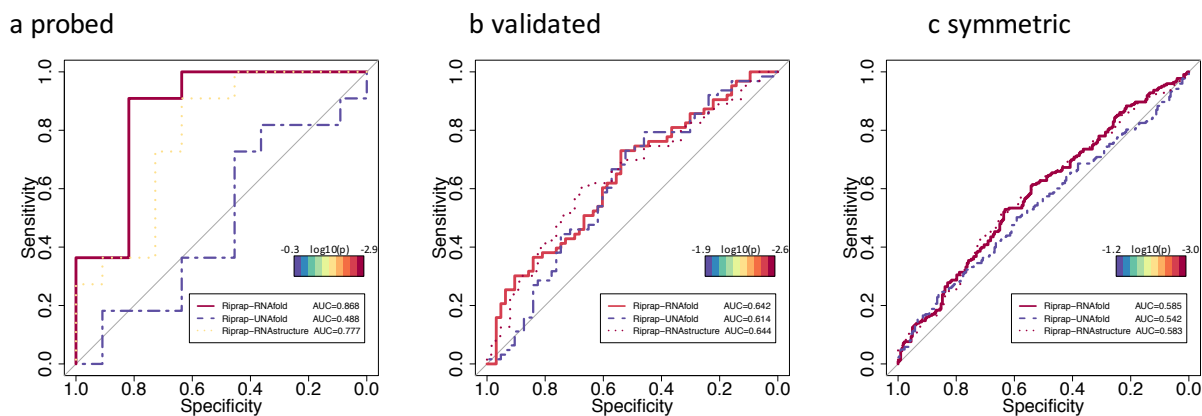
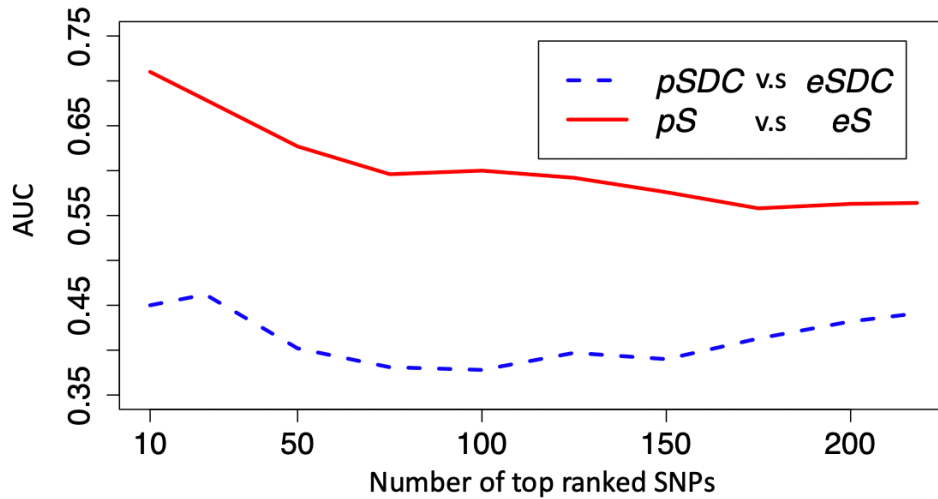


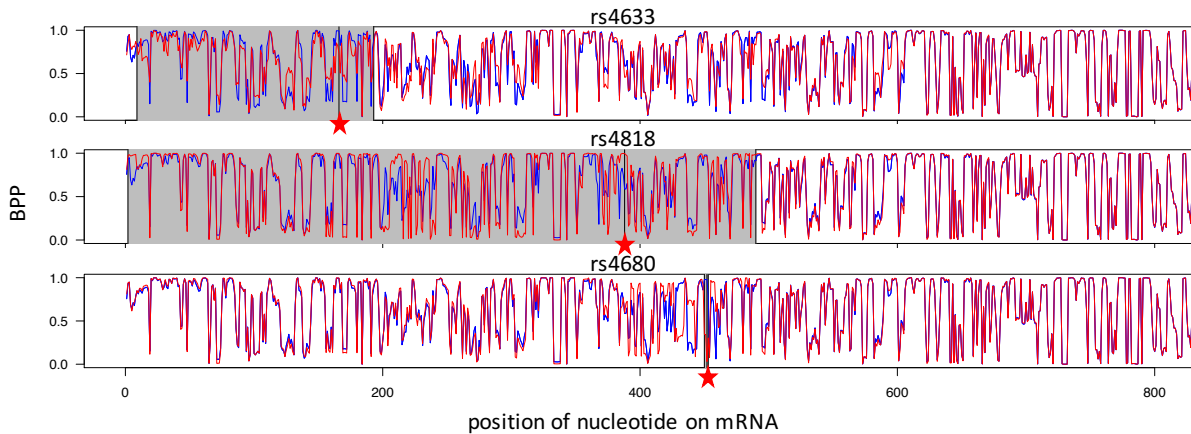
**Supplementary Figure 1. Diagram of Riprap and RNAsnp.** (a) Diagram, structural measuring unit and distance metric used in Riprap algorithm. Please see details in the Methods section. (b) Diagram, structural measuring unit and distance metric used in RNAsnp algorithm (3).



**Supplementary Figure 2. ROC analysis of Riprap-RNAfold, Riprap-UNAFold, and Riprap-RNAstructure on riboSNitch benchmarks.** (a) to (c) show the ROC curves for Riprap-RNAfold, Riprap-UNAFold and Riprap-RNAstructure applied to “probed”, “validated”, and “symmetric” datasets, respectively. The corresponding ROC curves for Riprap-RNAfold, Riprap-UNAFold and Riprap-RNAstructure are indicated by solid, dashed, and dotted lines, respectively.



**Supplementary Figure 3. Comparisons between the local metric Riprap score  $S$  and the local SDC and assessments on SHAPE data.** AUC values of ROC analysis on the top and bottom  $n$  SNVs ranked by  $eSDC$  and  $eS$  scores, respectively. The red solid curve and blue dashed curve represent the AUCs of  $pSDC$  on  $eSDC$ , and  $pS$  on  $eS$ , respectively.



**Supplementary Figure 4. BPP curves of rs4633, rs4818, and rs4680 in S-COMT.** The X-axis represents the nucleotide position on the mRNA. The Y-axis represents the BPPs. The blue curve is for the WT RNA, while the red curve is for the mutant RNA. The SNV position is highlighted by a red star and a black vertical line. The finally selected region by Riprap is highlighted in grey. The three SNVs were studied in the Human COMT haplotypes (1).

**a**

RBP miRNA Methylation sites ClinVar

RBP binding sites

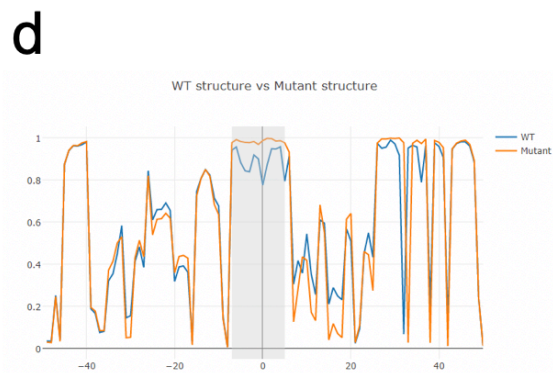
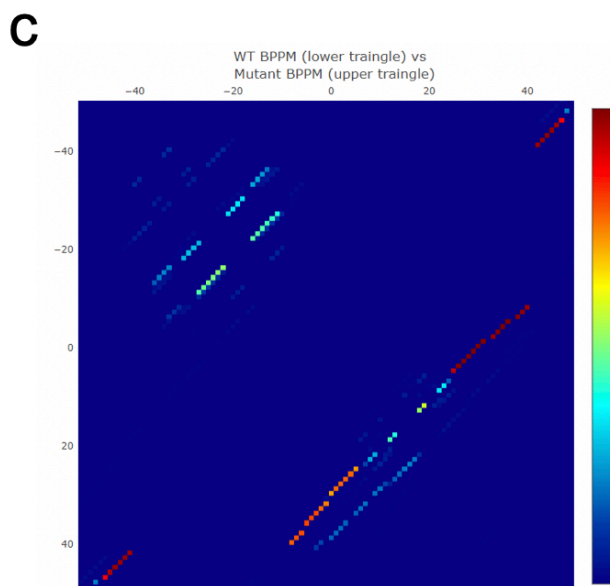
Chromosome	Start	End	Strand	RBP name	CLIP Method	Cell Line	Data source	Score
chr2	217070030	217070254	+	CSTF2T	PAR-CLIP,PARalyzer	HEK293	GSE37401,GSM917677	0.831945
chr2	217070110	217070244	+	CSTF2	PAR-CLIP,PARalyzer	HEK293	GSE37401,GSM917676	0.777442
chr2	217070111	217070291	+	ATXN2	PAR-CLIP,PARalyzer	HEK293T	DRA001158,DRS012390	0.989994
chr2	217070130	217070319	+	TARDBP	PAR-CLIP,PARalyzer	HEK293T	DRA001158,DRS012387	0.906124
chr2	217070196	217070244	+	CPSF7	PAR-CLIP,PARalyzer	HEK293	GSE37401,GSM917663	0.586511
chr2	217070211	217070244	+	CPSF1	PAR-CLIP,PARalyzer	HEK293	GSE37401,GSM917672	0.596587
chr2	217070211	217070301	+	HNRNPC	PAR-CLIP,PARalyzer	HEK293T	GSE56010,GSM1350199	0.739263
chr2	217070212	217070253	+	FUS	PAR-CLIP,PARalyzer	HEK293T	DRA001158,DRS012384	0.965826
chr2	217070214	217070245	+	HNRNPC	PAR-CLIP,PARalyzer	HEK293T	GSE56010,GSM1350200	0.62185
chr2	217070216	217070291	+	HNRNPC	PAR-CLIP,PARalyzer	HEK293T	GSE56010,GSM1350201	0.712915

Showing 1 to 10 of 35 rows 10 rows per page

**b**

WT RNA sequence  
 TTACATACATGCTTTGAAGTTTCTGAAAAGTAGATCTTTTCTTGACCTAGTATATCAGTGACAGTTGCAGCCCTTGATGTGATTAGTGCTCATGTGGA

Mutant RNA sequence  
 TTACATACATGCTTTGAAGTTTCTGAAAAGTAGATCTTTTCTTGACCTAGCATAATCAGTGACAGTTGCAGCCCTTGATGTGATTAGTGCTCATGTGGA



**Supplementary Figure 5. Example riboSNitch rs1051677 in riboSNitchDB.** (a) RBPs' binding peaks in the 201 nt region (100 nt upstream to 100 nt downstream) around the riboSNitch rs1051677. (b) The wild-type and mutant sequence of RNA. (c) BPPM heatmap of both the wild-type (upper triangle) and mutant RNA (lower triangle). (d) The BPP curve of both the wild-type (highlighted by blue) and mutant RNA (highlighted by orange).

**Supplementary Table 1. Threshold, accuracy, sensitivity, and specificity table.**

Method	Probed				Validated				Symmetric			
	T	Acc	Sn	Sp	T	Acc	Sn	Sp	T	Acc	Sn	Sp
Riprap	3.90	0.86	0.91	0.82	4.47	0.63	0.73	0.54	6.11	0.58	0.53	0.63
Simple	2.13	0.73	0.73	0.73	2.24	0.60	0.59	0.62	2.63	0.57	0.49	0.65
RNA <span>sn</span> p-value	0.39	0.77	0.82	0.73	0.30	0.61	0.48	0.75	0.47	0.57	0.56	0.59
RNA <span>sn</span> p-value 0.05	0.05	0.50	0.00	1.00	0.05	0.52	0.13	0.92	0.05	0.52	0.09	0.95
RNA <span>sn</span> dmax	0.07	0.73	0.73	0.73	0.07	0.60	0.52	0.67	0.06	0.57	0.53	0.61
classSNitch **	0.35	0.50	0.45	0.55	0.32	0.47	0.48	0.59	0.35	0.46	0.58	0.50
remuRNA*	2.82	#	0.64	0.82	1.65	#	0.56	0.56	2.38	#	0.45	0.64
SNPfold*	0.92	#	0.73	0.73	0.93	#	0.60	0.60	0.93	#	0.54	0.61
Riprap <sup>^</sup>	-	-	-	-	3.90	0.61	0.75	0.48	3.90	0.56	0.63	0.49
Simple <sup>^</sup>	-	-	-	-	2.13	0.60	0.60	0.59	2.13	0.55	0.54	0.57
RNA <span>sn</span> p-value <sup>^</sup>	-	-	-	-	0.39	0.57	0.49	0.65	0.39	0.57	0.49	0.64
RNA <span>sn</span> dmax <sup>^</sup>	-	-	-	-	0.07	0.59	0.51	0.67	0.07	0.57	0.48	0.65
classSNitch <sup>^</sup> **	-	-	-	-	0.37	0.49	0.33	0.65	0.37	0.48	0.38	0.58

T: Threshold; Acc: Accuracy; Sn: Sensitivity; Sp: Specificity.

\*The values for the methods were directly extracted from Corley *et al.* (2)

\*\*The lower performance may result from the fact that classSNitch was trained on the experimental SHAPE data instead of analysing BPP profiles.

# The values were not listed in Corley *et al.* (2)

<sup>^</sup> The threshold is determined based on the “Probed” riboSNitches, and the performance is evaluated on “Validated” and “Symmetric” riboSNitches.

**Supplementary Table 2. Commands to run the methods.**

Method	command
RNAfold	RNAfold -p < input.fa > seq.out
RNAshp	RNAshp -f input.fa -s snp_input.txt -w 100 -l 10 -c 0.0
classSNitch	R: getFeatures(sample=mut,base=wt,trim=0,outfile = "out") data("mutmap") clf = classifyRNA(mutmap) f = read.delim("out", header=F) pred = predict(clf,f)

**Supplementary Table 3. Annotation datasets of riboSNitchDB.**

Annotation category	Data sources	Data description	The entire annotation dataset size	Genomic annotation version
RBP binding sites	POSTAR2 (4) ( <a href="http://lulab.life.tsinghua.edu.cn/postar/">http://lulab.life.tsinghua.edu.cn/postar/</a> )	The RBP binding sites overlapping with the flanking region of riboSNitches.	31,186,943 RBP binding sites	hg19
miRNA targeting sites	UCSC Table Browser (5) ( <a href="https://genome.ucsc.edu/cgi-bin/hgTables">https://genome.ucsc.edu/cgi-bin/hgTables</a> )	The miRNA targeting sites around the positions of riboSNitches	1,329,071 miRNA targeting sites	hg19
N6-methyladenosine (m <sup>6</sup> A) modification sites	m6AVar (6) ( <a href="http://m6avar.renlab.org/">http://m6avar.renlab.org/</a> )	The N6-methyladenosine (m <sup>6</sup> A) modification sites around the riboSNitches	301,529 m <sup>6</sup> A modification sites	hg19
Clinical significance	ClinVar (7) ( <a href="https://www.ncbi.nlm.nih.gov/clinvar/">https://www.ncbi.nlm.nih.gov/clinvar/</a> )	Clinical significance of the target SNVs	395,755 clinical significance records of the SNVs	hg19

## Reference:

1. Nackley, A.G., Shabalina, S.A., Tchivileva, I.E., Satterfield, K., Korchynskyi, O., Makarov, S.S., Maixner, W. and Diatchenko, L. (2006) Human catechol-O-methyltransferase haplotypes modulate protein expression by altering mRNA secondary structure. *Science*, **314**, 1930-1933.
2. Corley, M., Solem, A., Qu, K., Chang, H.Y. and Laederach, A. (2015) Detecting riboSNitches with RNA folding algorithms: a genome-wide benchmark. *Nucleic Acids Res*, **43**, 1859-1868.
3. Sabarinathan, R., Tafer, H., Seemann, S.E., Hofacker, I.L., Stadler, P.F. and Gorodkin, J. (2013) The RNAsnp web server: predicting SNP effects on local RNA secondary structure. *Nucleic Acids Res*, **41**, W475-479.
4. Zhu, Y., Xu, G., Yang, Y.T., Xu, Z., Chen, X., Shi, B., Xie, D., Lu, Z.J. and Wang, P. (2019) POSTAR2: deciphering the post-transcriptional regulatory logics. *Nucleic Acids Res*, **47**, D203-D211.
5. Karolchik, D., Hinrichs, A.S., Furey, T.S., Roskin, K.M., Sugnet, C.W., Haussler, D. and Kent, W.J. (2004) The UCSC Table Browser data retrieval tool. *Nucleic Acids Res*, **32**, D493-496.
6. Zheng, Y., Nie, P., Peng, D., He, Z., Liu, M., Xie, Y., Miao, Y., Zuo, Z. and Ren, J. (2018) m6AVar: a database of functional variants involved in m6A modification. *Nucleic Acids Res*, **46**, D139-D145.
7. Landrum, M.J., Lee, J.M., Riley, G.R., Jang, W., Rubinstein, W.S., Church, D.M. and Maglott, D.R. (2014) ClinVar: public archive of relationships among sequence variation and human phenotype. *Nucleic Acids Res*, **42**, D980-985.