ID	Name	Description	Note
		downstream 100bp of transcription	
1	TSS	start site.	
2		Downstream 100bp of transcription	
	TSS_A	start site with sequence A.	
3	alternative_exon	Alternatively spliced exons.	Dummy variables
4	constitutive_exon	Constitutively spliced exons.	site is overlapped to
5	internal_exon	Internal exons.	the topological region
6	long_exon	Long exons (length>400bp).	on the major RNA
7	last_exon	The last exon.	transcript.
8	last_exon_400bp	5' start 400 bp of the last exon.	
9	intron	Intron.	
10	pos_exons	relative positioning on exon.	
44	diat ai E #0000	distance to the upstream (5' end)	Nucleotide distances
11	dist_sj_5_p2000	splicing junction.	toward the splicing
		distance to the downstream (3' end)	junctions or the
12	dist_sj_3_p2000	splicing junction	nearest neighboring
			sites.
13	length_gene_ex	gene length-exons (z-score).	The region length in
14	length_gene_full	gene length full transcript (z-score).	bp.
15	AAACA	motif AAACA.	
16	AGACA	motif AGACA.	
17	AAACT	motif AAACT.	
18	AGACT	motif AGACT.	
19	AAACC	motif AAACC.	
20	AGACC	motif AGACC.	
21	GAACA	motif GAACA.	
22	GGACA	motif GGACA.	
23	GAACT	motif GAACT.	
24	GGACT	motif GGACT.	belong to the motif
25	GAACC	motif GAACC.	
26	GGACC	motif GGACC.	
27	TAACA	motif TAACA.	
28	TGACA	motif TGACA.	
29	TAACT	motif TAACT.	
30	TGACT	motif TGACT.	
31	TAACC	motif TAACC.	
32	TGACC	motif TGACC.	
22	clust_DRACH_f10	number of DRACH motif neighboors	clustering indicators
33	00	within 1000bp flanking regions.	and motif clustering

Table S1 Genomic features considered in LITHOPHONE

3/	clust_DRACH_f10	number of DRACH motif neighboors	
34	0	within 100bp flanking regions.	
35	dist_DRACH_p20	distance to the nearest DRACH motif	
	00	(peaked at 2000bp).	-
36	dist_DRACH_p20	distance to the nearest DRACH motif	
	0	(peaked at 200bp).	
37	PC_1bp	phast cons scores 1bp.	
38	PC_101bp	phast cons scores 101bp.	Scores related to
30	FC 1bp	fitness consequences scores 1bp z	evolutionary - conservation
39	FC 101bp	score.	
40		fitness consequences scores 101bp z	
40	10_10100	score.	
41	struct_hybridize	predicted hybridized region.	RNA secondary
40	atruat loop	inferred loop structures between	structures
42	struct_loop	hybridized region.	Siluciales
13		eCLIP data of HNRNPC RNA binding	
40		sites.	
11	VTHDC1 TREW	TREW data of YTHDC1 RNA binding	
		sites.	
45	YTHDE1 TREW	TREW data of YTHDF1 RNA binding	
-10		sites.	
46	YTHDF2 TREW	TREW data of YTHDF2 RNA binding	
		sites.	-
47	METTL3 TREW	TREW data of METTL3 RNA binding	
		sites	-
48	METTL14 TREW	TREW data of METTL14 RNA binding	
	_		
49	WTAP_TREW	TREW data of WTAP RNA binding	RNA annotations
		sites.	related to m ⁶ A biology.
50	METTL16_CLIP	CLIP data of METTL16 RNA binding	
51	ALKBH5_PARCLI	PARCLIP data of ALKBH5 RNA	
	۲	binding sites.	-
52	FTO_CLIP	CLIP data of FTO RNA binding sites.	
			-
53	FTO_eCLIP	eCLIP data of FTO RNA binding sites.	
	TargetScan	Predicted miRNA targeted sites by	-
54		TargetScan	
	Verified_miRtarge	miRNA targeted sites verified by	-
55		experiment	
	miR targeted ge		a dummy variable
56	nes	miRNA targeted genes.	indicating wheather
			J

			the IncRNA is a
			miRNA target
57	isoform_num	isoform number z score.	Isoform and exon
58	exon_num	exon number z score.	number z score
59	GC_cont_genes	gene level GC content z score is	
		generated.	GC content z score
60	GC_cont_101bp	101bp GC content z score.	

Table S2 R packages corresponding to five methods

Method	R package	
RF	randomforest	
SVM	e1071	
KNN	class	
LR	e1071	
XGBoost	xgboost	