

**Table S1. Summary of existing m<sup>6</sup>A predictors for different species.**

Classifier	Species	Prediction engine	Features used in the classifier	PMID
RFAthM6A	<i>A.thaliana</i>	RF	[1] Position-specific nucleotide sequence profile; [2] Position-specific dinucleotide sequence profile; [3] K-nucleotide frequencies; [4] K-spaced nucleotide pair frequencies	29340952
RAM-NPPS	<i>S.cerevisiae</i> ; <i>H.sapiens</i> ; <i>A.thaliana</i>	SVM	[1] Nucleotide pair position specificity	28440291
RAM-ESVM	<i>S.cerevisiae</i>	SVM	[1] Pseudo nucleotide composition; [2] Motif features; [3] K-mer features	28079126
M6ATH	<i>A.thaliana</i>	SVM	[1] Nucleotide chemical property; [2] Nucleotide composition	27590733
AthMethpre	<i>A.thaliana</i>	SVM	[1] Positional binary encoding; [2] Position-independent Kmer frequency	27550167
RNAMethPre	<i>H.sapiens</i> ; <i>M.musculus</i>	SVM	[1] Positional binary encoding; [2] Nucleotide k-mer frequency; [3] Relative position in mRNA; [4] Stability of the local structure	27723837
TargetM6A	<i>S.cerevisiae</i>	SVM	[1] Position-specific nucleotide sequence profile; [2] Position-specific dinucleotide sequence profile	27552763
RNA-MethylPred	<i>S.cerevisiae</i>	SVM	[1] Bi-profile Bayes; [2] Dinucleotide composition; [3] K-nearest neighbor encoding	27338301
M6A-HPCS	<i>S.cerevisiae</i>	SVM	[1] Physical-chemical property; [2] Pseudo dinucleotide composition; [3] auto-covariance; [4] cross-covariance transformations	27193216
MethyRNA	<i>H.sapiens</i> ; <i>M.musculus</i>	SVM	[1] Chemical property; [2] Nucleotide frequency	26912125
SRAMP	<i>Mammalian</i>	RF	[1] Positional binary encoding; [2] K-nearest	26896799

pRNAm-PC	<i>S.cerevisiae</i>	SVM	neighbor encoding; [3]K-spaced nucleotide pair frequencies	26748145
M6Apred	<i>S.cerevisiae</i>	SVM	[1] Physical-chemical property; [2] Pseudo dinucleotide composition; [3] auto-covariance; [4] cross-covariance transformations	26343792
iRNA-Methyl	<i>S.cerevisiae</i>	SVM	[1] Chemical property; [2] Accumulated nucleotide frequency	26314792
			[1] pseudo dinucleotide composition	26314792

Note: RF: random forest; SVM: support vector machine; *A.thaliana*: *Arabidopsis thaliana*; *S.cerevisiae*: *Saccharomyces cerevisiae*; *H.sapiens*: *Homo sapiens*; *M.musculus*: *Mus musculus*.

**Table S2. Optimized sliding window sizes for the *Saccharomyces cerevisiae* dataset.**

Sliding window size( <i>S</i> )	1	<b>2</b>	4	6	8	10	12	14
AUC Value	0.7602	<b>0.7639</b>	0.7358	0.6999	0.6884	0.6575	0.6555	0.6461
AUC01 Value	0.0249	<b>0.0264</b>	0.0220	0.0200	0.0173	0.0163	0.0141	0.0131

Note: For the *Saccharomyces cerevisiae* dataset, the sequence window size (*W*) was fixed as 51 [16]. The sliding window size (*S*) of 2, corresponding to the largest AUC and AUC01 values, was selected for the development of the algorithms.

**Table S3. Optimized window sizes for the Arabidopsis thaliana dataset.**

Sliding window size( $S$ )	2	4	6	<b>8</b>	10	12	14
AUC Value	0.8865	0.8914	0.8958	<b>0.8981</b>	0.8973	0.8953	0.8946
AUC01 Value	0.0503	0.0489	0.0491	<b>0.0511</b>	0.0488	0.0487	0.0477

*Note:* For *Arabidopsis thaliana* dataset, the sequence window size ( $W$ ) was fixed as 101 [17].

The sliding window size ( $S$ ) of 8, corresponding to the largest AUC and AUC01 values, was selected for the development of the algorithms.

**Table S4. Optimized window sizes for the Mammalian full transcript dataset.**

AUC/AUC01	Sequence window size ( $W$ )					
	51	61	71	<b>81</b>	91	101
1	0.8031/0.0323	0.8035/0.0323	0.8036/0.0330	0.8041/0.0336	0.8037/0.0326	0.8031/0.0305
<b>2</b>	0.8034/0.0332	0.8041/0.0332	0.8043/0.0335	<b>0.8056/0.0340</b>	0.8044/0.0324	0.8033/0.0319
4	0.7998/0.0319	0.8014/0.0310	0.8014//0.0326	0.8023/0.0323	0.8022/0.0313	0.8024/0.0319
6	0.7893/0.0318	0.7902/0.0303	0.7908/0.0317	0.7914/0.0305	0.7913/0.0306	0.7910/0.0307
8	0.7693/0.0285	0.7694/0.0273	0.7700/0.0294	0.7696/0.0282	0.7694/0.0276	0.7686/0.0293
10	0.7543/0.0256	0.7532/0.0257	0.7525/0.0250	0.7526/0.0266	0.7518/0.0265	0.7511/0.0259
12	0.7436/0.0250	0.7425/0.0255	0.7411/0.0256	0.7406/0.0251	0.7402/0.0248	0.7388/0.0242
14	0.7382/0.0236	0.7381/0.0232	0.7363/0.0241	0.7347/0.0240	0.7342/0.0240	0.7328/0.0234

*Note:* The sliding window size ( $S$ ) of 2 and the sequence window size ( $W$ ) of 81, corresponding to the largest AUC and AUC01 values, was selected for the development of the algorithms.

**Table S5. Optimized window sizes for the Mammalian mature mRNA dataset.**

AUC/AUC01		Sequence window size ( $W$ )					
		51	<b>61</b>	71	81	91	101
Sliding window size ( $S$ )	1	0.7530/0.0251	0.7560/0.0254	0.7532/0.0252	0.7547/0.0249	0.7537/0.0256	0.7506/0.0255
	<b>2</b>	0.7584/0.0255	<b>0.7606/0.0262</b>	0.7586/0.0255	0.7573/0.0253	0.7552/0.0252	0.7547/0.0251
	4	0.7546/0.0244	0.7549/0.0251	0.7543/0.0257	0.7543/0.0242	0.7540/0.0241	0.7540/0.0242
	6	0.7428/0.0250	0.7416/0.0233	0.7410/0.0247	0.7408/0.0235	0.7403/0.0235	0.7391/0.0237
	8	0.7169/0.0209	0.7147/0.0212	0.7128/0.0209	0.7125/0.0218	0.7114/0.0208	0.7100/0.0201
	10	0.6980/0.0191	0.6944/0.0188	0.6922/0.0201	0.6904/0.0192	0.6889/0.0190	0.6880/0.0183
	12	0.6844/0.0180	0.6795/0.0180	0.6756/0.0173	0.6728/0.0170	0.6710/0.0163	0.6698/0.0179
	14	0.6762/0.0180	0.6709/0.0172	0.6671/0.0173	0.6626/0.0153	0.6600/0.0165	0.6580/0.0163

*Note:* The sliding window size ( $S$ ) of 2 and the sequence window size ( $W$ ) of 61, corresponding to the largest AUC and AUC01 values, was selected for the development of the algorithms.

**Table S6. Optimized window sizes of the BGRU-based classifier for mammalian dataset.**

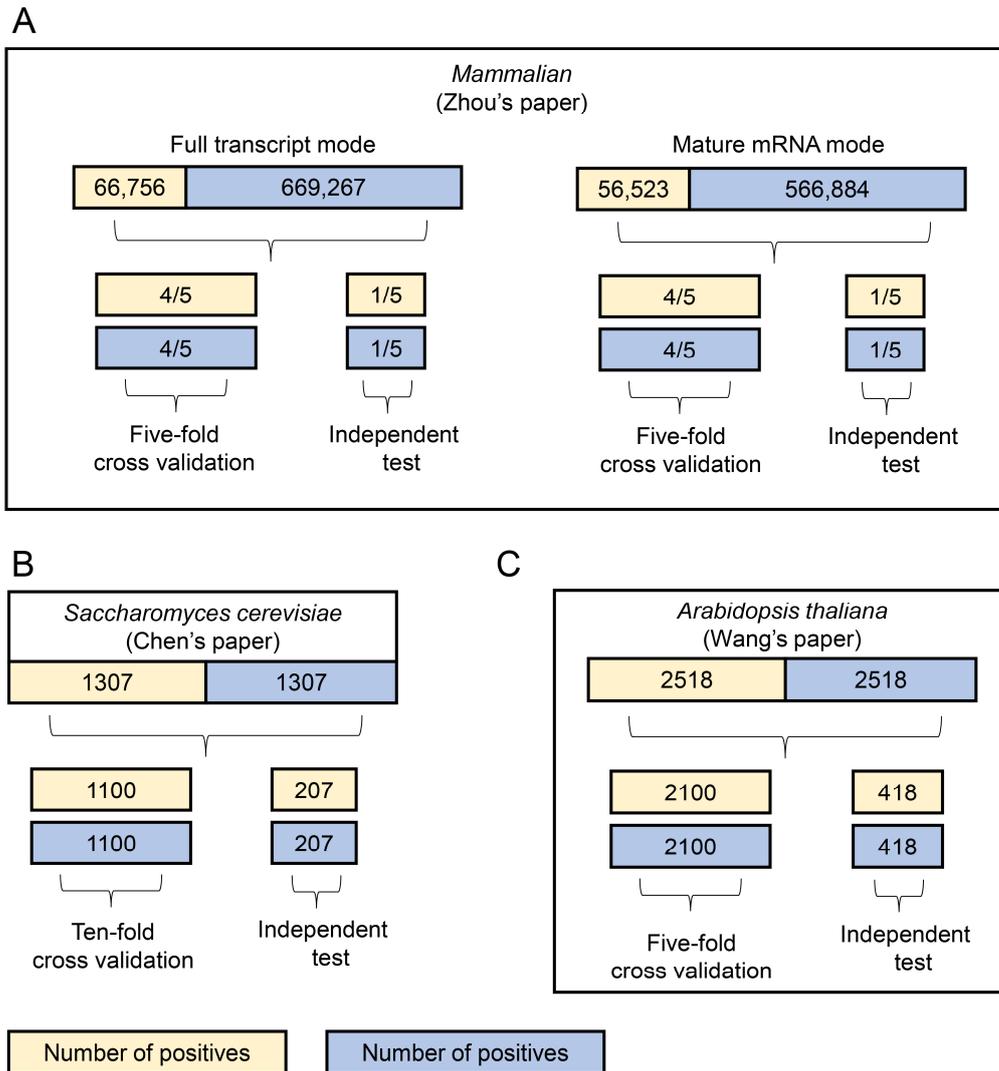
Prediction mode	Window size
Full transcript mode	301
Mature mRNA mode	191

**Table S7. Prediction results of different classifiers via independent tests.**

Species <sup>1</sup>	Classifiers <sup>2</sup>	Acc <sup>3</sup>	Sn <sup>3</sup>	Sp <sup>3</sup>	MCC <sup>3</sup>	AUC <sup>3</sup>	AUC01 <sup>3</sup>
<i>Mammalia</i> Full transcript	RF <sub>ENAC</sub>	86.74	51.89	90.21	0.353	0.828	0.0368
	RF <sub>Kmer</sub>	85.75	41.52	90.17	0.273	0.790	0.0284
	RF <sub>KSNPF</sub>	85.71	40.71	90.20	0.268	0.793	0.0264
	RF <sub>PseDNC</sub>	85.14	35.40	90.10	0.224	0.751	0.0242
	UGRU	87.97	66.61	90.10	0.455	0.900	0.0446
	BGRU	88.11	68.16	90.10	0.466	0.904	0.0459
	BERMP	88.34	70.72	90.10	0.484	0.907	0.0509
<i>Mammalia</i> Mature mRNA	RF <sub>ENAC</sub>	85.71	41.49	90.12	0.272	0.773	0.0268
	RF <sub>Kmer</sub>	84.24	25.56	90.15	0.143	0.678	0.0154
	RF <sub>KSNPF</sub>	84.03	22.59	90.16	0.117	0.642	0.0135
	RF <sub>PseDNC</sub>	83.90	21.91	90.09	0.110	0.633	0.0130
	UGRU	86.18	47.71	90.00	0.319	0.826	0.0293
	BGRU	86.20	47.66	90.05	0.318	0.827	0.0298
	BERMP	86.32	49.42	90.00	0.331	0.829	0.0317
<i>Saccharomyces cerevisiae</i> Mature mRNA	RF <sub>ENAC</sub>	69.32	48.31	90.34	0.426	0.801	0.0348
	RF <sub>Kmer</sub>	60.39	30.43	90.33	0.259	0.710	0.0182
	RF <sub>KSNPF</sub>	57.25	24.15	90.34	0.193	0.686	0.0146
	RF <sub>PseDNC</sub>	57.73	24.64	90.82	0.206	0.685	0.0152
	UGRU	53.86	17.39	90.34	0.113	0.622	0.0108
	BGRU	58.21	26.09	90.34	0.214	0.655	0.0149
	BERMP	68.84	47.34	90.34	0.417	0.800	0.0347
<i>Arabidopsis thaliana</i> Mature mRNA	RF <sub>ENAC</sub>	80.86	71.53	90.19	0.628	0.900	0.0479

RF <sub>Kmer</sub>	87.08	84.18	90.10	0.743	0.939	0.0657
RF <sub>KSNPF</sub>	83.97	77.75	90.19	0.685	0.925	0.0586
RF <sub>PseDNC</sub>	83.85	77.51	90.19	0.683	0.923	0.0568
UGRU	85.77	81.10	90.43	0.718	0.930	0.0593
BGRU	87.32	84.21	90.43	0.748	0.936	0.0613
BERMP	87.20	84.21	90.19	0.745	0.934	0.0607

*Note:* <sup>1</sup>The training and independent datasets were depicted in Figure S1. <sup>2</sup>RF<sub>ENAC</sub>=RF classifier with the ENAC encoding, RF<sub>KSNPF</sub>= RF classifier with the encoding of K-spaced nucleotide pair frequencies, RF<sub>PseDNC</sub>=RF classifier with the encoding of Pseudo dinucleotide composition, UGRU= the unidirectional GRU-based RNN classifier with word embedding, BGRU= the bidirectional GRU-based RNN classifier with word embedding, BERMP= BGRU-based Ensemble RNA Methylation site Predictor that integrating BGRU and RF<sub>ENAC</sub>. <sup>3</sup>Acc=accuracy, Sn=sensitivity, Sp=specificity, MCC=Matthew's Correlation Coefficient, AUC=area under the receiver operating characteristic, AUC01 = AUC with a <10% false positive rate (*i.e.*, specificity>90%).



1100
1100

Ten-fold cross validation

207
207

Independent test

2100
2100

Five-fold cross validation

418
418

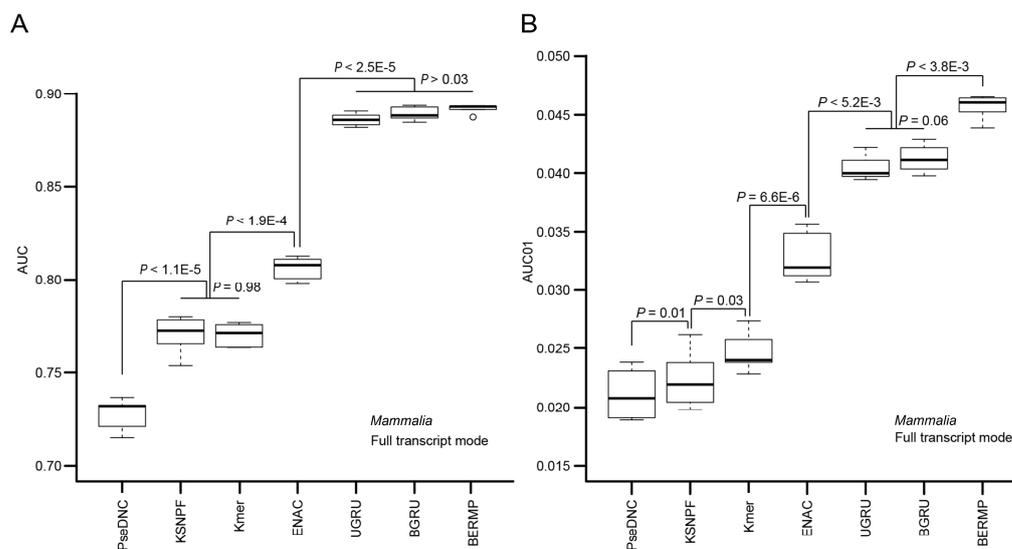
Independent test

Number of positives

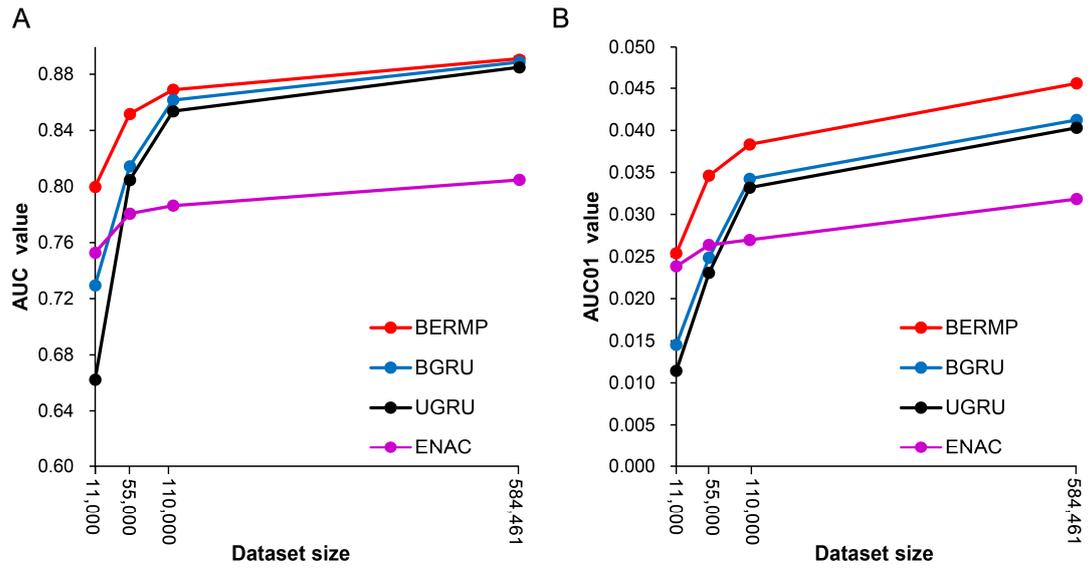
Number of positives

**Figure S1.** The flowchart of dataset process for *Mammalia* (A), *Saccharomyces cerevisiae* (B) and *Arabidopsis thaliana* (C). All the datasets were derived from previous publications as well as the data processing flow for *Mammalia* (A) and *Arabidopsis thaliana* (C).

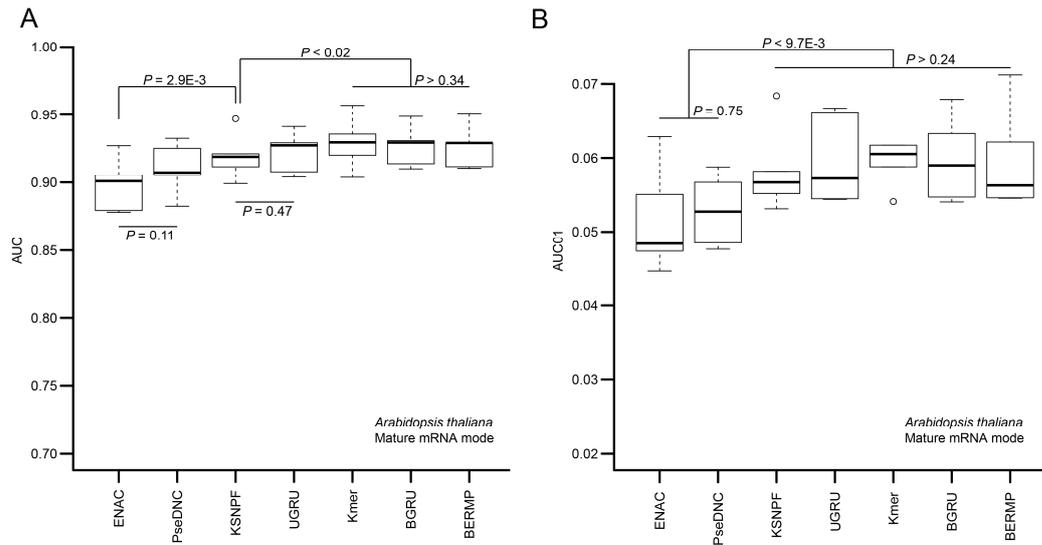




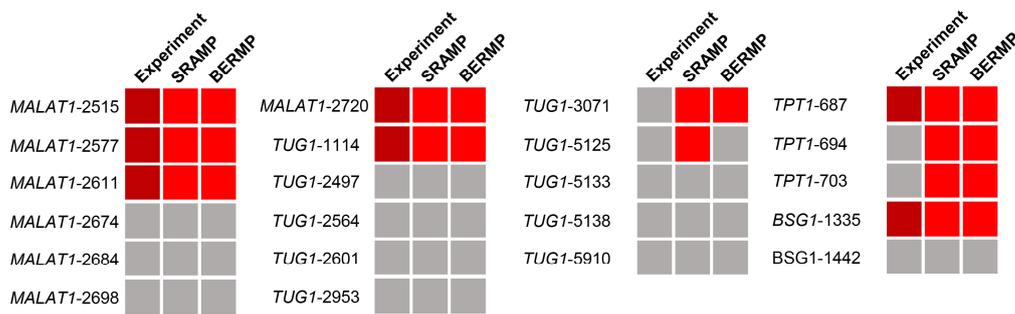
**Figure S3. Performance comparison of the seven m<sup>6</sup>A predictors using *Mammalia* full transcript dataset.** The AUC (A) and AUC01 values (B) were calculated via five-fold cross validation (Figure S1A). For each algorithm, the AUC or AUC01 values between the adjacent data sets were statistically compared and the horizontal line represented no statistical difference ( $P > 0.05$ ). The  $P$  value was calculated by a paired student's t-test.



**Figure S4. Relationship between data size and prediction performance of classifiers using the *Mammalia* full transcript dataset.** The AUC values (A) and AUC01 values (B) were calculated using four different data sizes (all, one-fifth, one-tenth and one-fiftieth) via five-fold cross validation (Figure S1A).



**Figure S5. Performance comparison of the seven m<sup>6</sup>A predictors using the *Arabidopsis thaliana* dataset.** The AUC (A) and AUC01 values (B) were calculated via five-fold cross validation (Figure S1C). For each algorithm, the AUC or AUC01 values between the adjacent data sets were statistically compared and the horizontal line represented no statistical difference ( $P > 0.05$ ). The  $P$  value was calculated by a paired student's t-test.



**Figure S6. The performance comparison of SRAMP and BERMP on the golden standard dataset.** The gene identifiers and site positions were in lines with the original publication by Liu *et al.* [25]. Experimental reference sites, the SRAMP prediction results and BERMP prediction results were denoted in the Experiment, SRAMP and BERMP columns, respectively. Experimentally identified m<sup>6</sup>A sites and non-m<sup>6</sup>A sites were indicated by deep red and grey boxes, respectively. Predicted above high confidence m<sup>6</sup>A sites and non-m<sup>6</sup>A sites were indicated by red and grey boxes, respectively.