

Supplemental Table 2: Ensemble predictor has no preference for coding to noncoding sequences.

CREMA prediction	CPAT prediction:			
	Coding		Noncoding	
	lncRNA positive	lncRNA negative	lncRNA positive	lncRNA negative
Species				
<i>A. thaliana</i>	243	37279	1067	2314
<i>O. sativa</i>	133	49325	723	1762
<i>E. salsugineum</i>	44	28079	154	1161

Transcripts were identified as either protein coding or non-coding using the CPAT software. Coding probability cutoffs were calculated by intersect of sensitivity and specificity via 10-fold cross validation. A coding probability cut off of 0.38 was used for *A. thaliana* and 0.52 for both *O. sativa* and *E. salsugineum*.