TF	Bound	FN (Both Models)	FN (Mouse Only)	Unbound	FP (Both Models)	FP (Mouse Only)
			DNA			
CTCF	10.1%	11.4%	7.3%	11.4%	8.9%	**9.0%**
CEBPA	12.3%	10.4%	8.3%	11.3%	13.0%	**9.2%**
HNF4A	10.7%	12.0%	11.3%	11.4%	9.5%	**9.0%**
RXRA	10.1%	11.7%	8.8%	11.4%	10.0%	**9.4%**
			LINE			
CTCF	18.3%	22.5%	21.3%	37.6%	17.8%	**31.6%**
CEBPA	25.6%	26.3%	25.0%	37.6%	29.0%	**32.3%**
HNF4A	21.0%	25.3%	26.3%	37.6%	21.5%	**30.5%**
RXRA	21.0%	27.9%	22.1%	37.8%	22.1%	**33.2%**
			Low Complexity			
CTCF	2.5%	1.0%	2.6%	1.9%	4.0%	**1.5%**
CEBPA	1.5%	1.4%	0.0%	1.9%	1.6%	**1.4%**
HNF4A	2.0%	1.1%	1.5%	1.9%	2.4%	**1.3%**
RXRA	2.1%	1.3%	2.0%	1.9%	2.2%	**1.5%**
			LTR			
CTCF	8.6%	12.6%	8.2%	17.6%	16.4%	**15.5%**
CEBPA	12.8%	12.5%	41.7%	17.6%	19.4%	**14.0%**
HNF4A	13.3%	15.0%	16.0%	17.6%	19.4%	**12.6%**
RXRA	12.2%	14.8%	9.8%	17.6%	18.2%	**11.5%**
			Simple Repeat			
CTCF	13.4%	10.1%	10.9%	11.5%	15.8%	**12.0%**
CEBPA	9.4%	8.6%	25.0%	11.6%	9.8%	**11.9%**
HNF4A	12.3%	11.4%	9.1%	11.5%	11.6%	**12.8%**
RXRA	11.5%	9.4%	13.3%	11.5%	11.1%	**13.4%**
			SINE			
CTCF	23.9%	23.2%	20.6%	31.2%	18.4%	**81.1%**
CEBPA	30.9%	22.6%	41.7%	31.2%	35.2%	**86.9%**
HNF4A	27.1%	23.6%	22.6%	31.2%	28.9%	**95.5%**
RXRA	27.1%	25.3%	18.9%	31.3%	32.2%	**97.6%**
			Unknown			
CTCF	0.2%	0.1%	0.0%	0.2%	0.1%	**0.0%**
CEBPA	0.3%	0.5%	0.0%	0.1%	0.2%	**0.0%**
HNF4A	0.2%	0.1%	0.0%	0.2%	0.2%	**0.0%**
RXRA	0.2%	0.3%	0.0%	0.1%	0.2%	**0.0%**

Table S1: Percent of windows overlapping various RepeatMasker-defined repeat elements, for different categories of genomic windows from the held-out test set. Only RepeatMasker repeat classes with at least 500 distinct annotations within the testset are shown. FPs: false positives. FNs: false negatives. Mouse Only: specific to mouse-trained models. See Methods for more details on site categorization.