

## General Motif Characteristics:

### Top Similar Looking Motifs:

SP1.p2	0.000		
ENCODE.SP1	0.041		
HOMER.Sp1(Zf).Promoter	0.050		
HCMC.SP1_f1.wm	0.062		
PATZ1.p2	0.070		
HCMC.SP1_f2.wm	0.099		
UNIPROBE.Klf7_primary.wm		0.099	
HCMC.SP3_f1.wm	0.108		
HTSELEX.SP1.C2H2.DBD.monomeric.wm1			0.117
ENCODE.ZNF281	0.117		

### Statistics:

- Motif name: SP1.p2
- Enrichment: 3.41120934365
- log-Likelihood Ratio: 613.533437327
- Area under precision recall curve: 0.2132

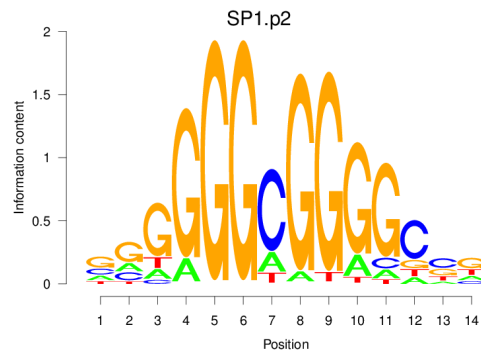


Figure 1: Logo of Weight Matrix

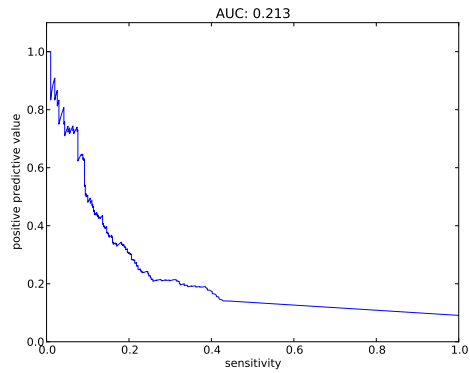


Figure 2: Precision and Recall curve

## Statistics for TFBS - Peak Centering and Peak Posterior - Peak Height Correlation:

### 0.1

Overall statistics:

- Number of true peaks out of total number of peaks: 2680/12571
- 21.32 percent are true.
- Cut-off: minimum summed posterior of 0.20
- Peak plots contain TFBS of posterior  $\geq 0.20$

Statistic for centering of TFBSs at peak centers:

- Enrichment at binding sites: 0.9024

Correlation between peak Z-score and number of binding sites at peak:  $-0.0399$

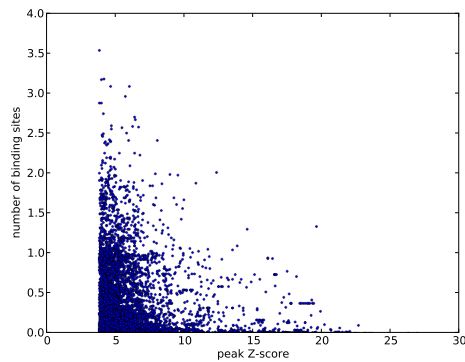


Figure 3: Scatter plot of peak Z-score versus number of binding sites at peak

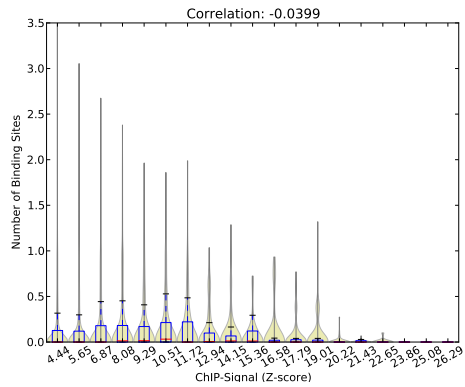


Figure 4: Violin plot of peak Z-score versus number of binding sites at peak

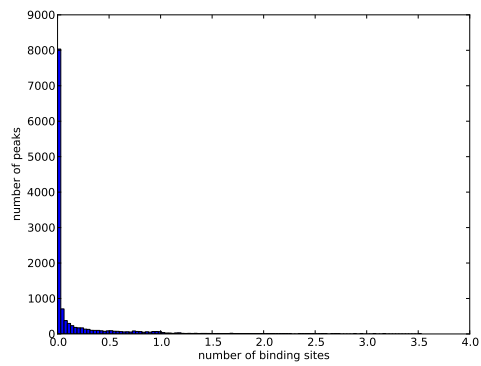


Figure 5: Histogram of number of binding sites at peaks

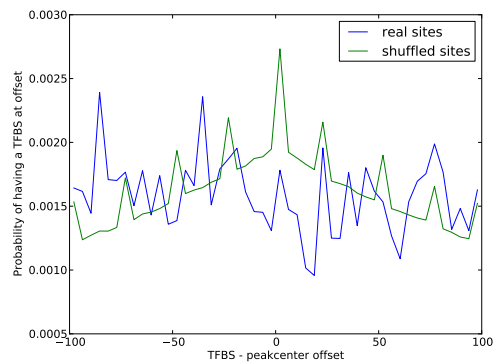


Figure 6: Histogram of distances of TFBS to its nearest peak center

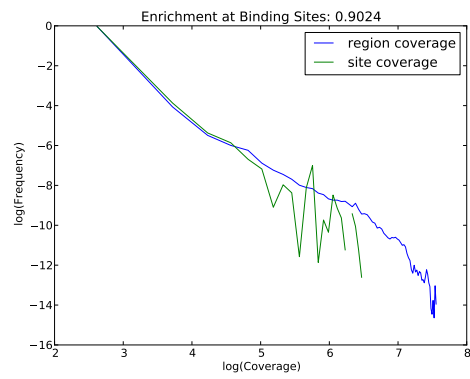


Figure 7: Histograms of coverage at sites and in total regions