

## General Motif Characteristics:

### Top Similar Looking Motifs:

HOMER.GFX(?) .Promoter	0.044	
HOMER.ZBTB33.GM12878-ZBTB33-ChIP-Seq	0.097	
ENCODE.UA1	0.101	
HTSELEX.ZBED1.znf_BED.DBD.dimeric.wml	0.208	
HCMC.KAISO_f1.wm	0.239	
HCMC.E2F3_si.wm	0.278	
HCMC.E2F5_do.wm	0.310	
HCMC.AHR_si.wm	0.322	
HTSELEX.TFE3.bHLH.DBD.dimeric.wml	0.329	
E2F1..5.p2	0.333	

### Statistics:

- Motif name: denovo\_WM\_13
- Enrichment: 2.55936603865
- log-Likelihood Ratio: 469.879793335
- Area under precision recall curve: 0.4793

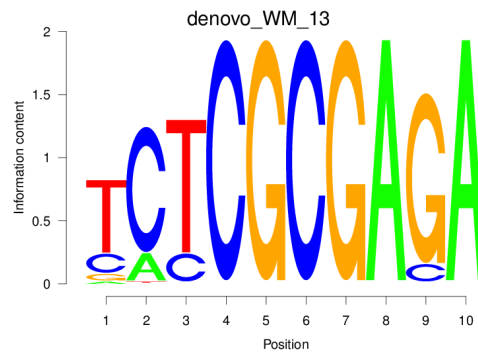


Figure 1: Logo of Weight Matrix

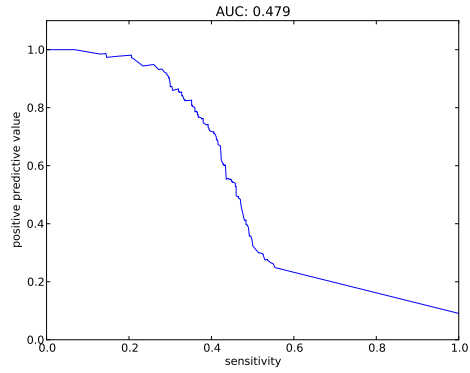


Figure 2: Precision and Recall curve

## 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: 831/37435
- 2.22 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: 22.7271

Correlation between peak Z-score and number of binding sites at peak: 0.4054

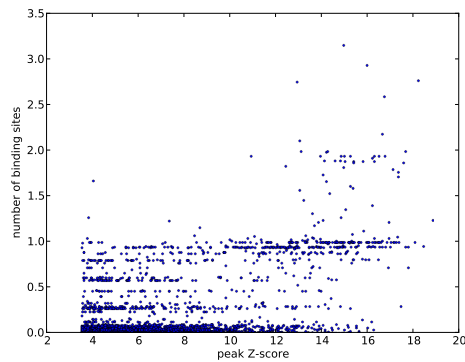


Figure 3: Scatter plot of peak Z-score versus number of binding sites at peak

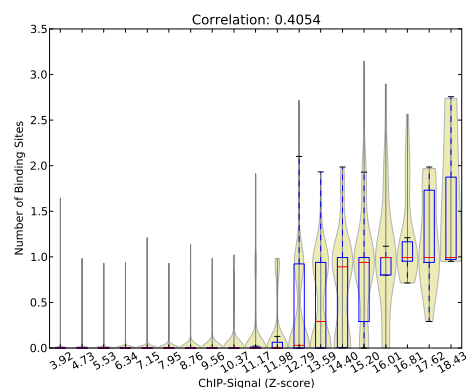


Figure 4: Violin plot of peak Z-score versus number of binding sites at peak

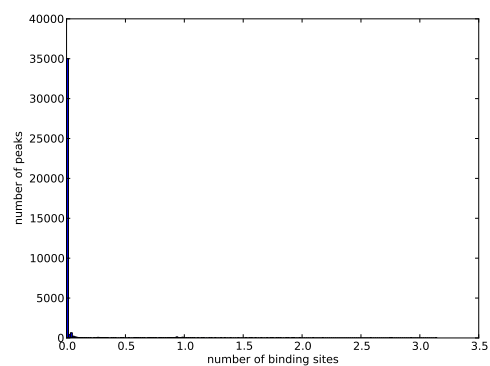


Figure 5: Histogram of number of binding sites at peaks

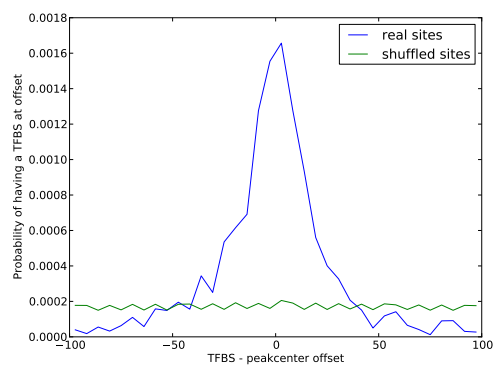


Figure 6: Histogram of distances of TFBS to its nearest peak center

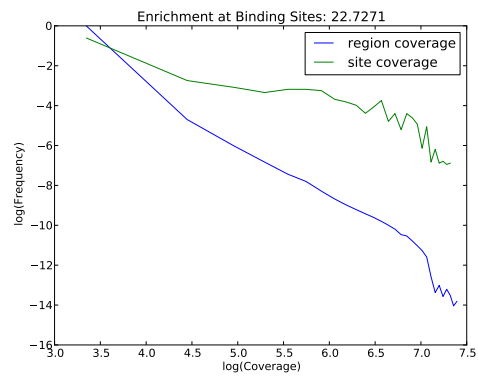


Figure 7: Histograms of coverage at sites and in total regions