

## General Motif Characteristics:

### Top Similar Looking Motifs:

HTSELEX.SOX8.HMG.full.dimeric.wm2	0.263	
HTSELEX.SOX8.HMG.DBD.dimeric.wm2	0.275	
HOMER.GFY(?) .Promoter	0.307	
ENCODE.ZNF143-ext	0.309	
RUNX1..3.p2	0.311	
HOMER.Sox6(HMG) .Myotubes-Sox6-ChIP-Seq (GSE32627)		0.315
HCMC.RUNX1_f1.wm	0.327	
UNIPROBE.IRC900814_primary.wm	0.332	
UNIPROBE.Foxj1_secondary.wm	0.339	
HTSELEX.SOX21.HMG.DBD.dimeric.wm4	0.339	

### Statistics:

- Motif name: denovo\_WM\_12
- Enrichment: 56.6265704025
- log-Likelihood Ratio: 2018.2391584
- Area under precision recall curve: 0.8717

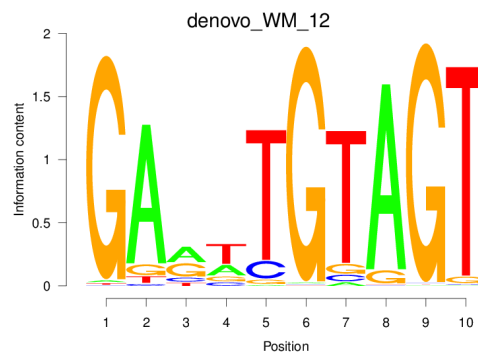


Figure 1: Logo of Weight Matrix

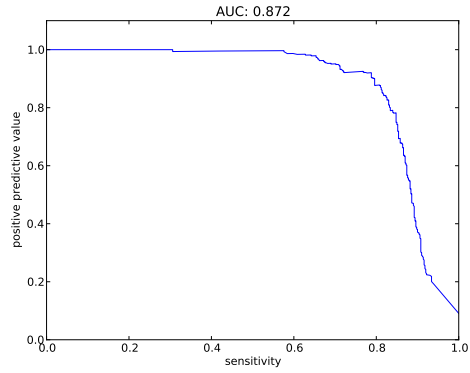


Figure 2: Precision and Recall curve

## 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: 3761/36955
- 10.18 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: 12.412

Correlation between peak Z-score and number of binding sites at peak: 0.5292

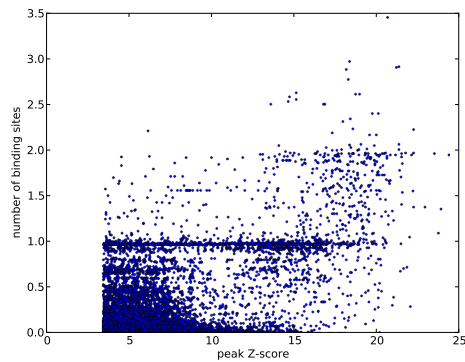


Figure 3: Scatter plot of peak Z-score versus number of binding sites at peak

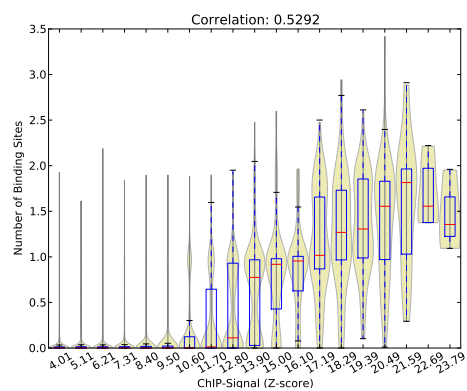


Figure 4: Violin plot of peak Z-score versus number of binding sites at peak

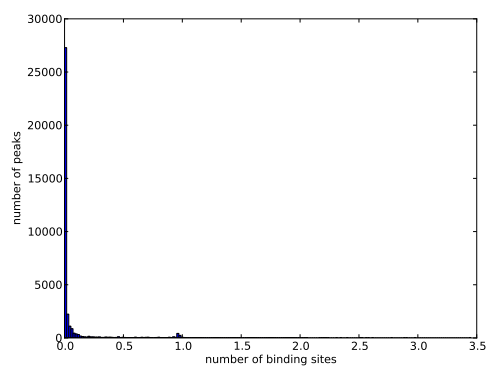


Figure 5: Histogram of number of binding sites at peaks

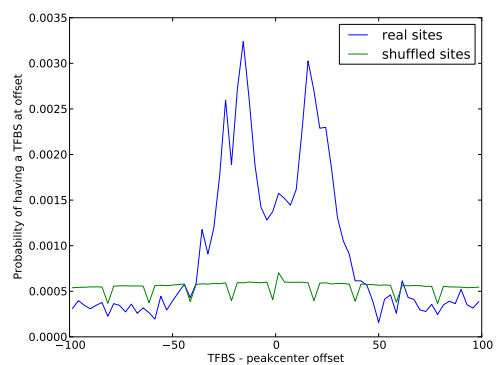


Figure 6: Histogram of distances of TFBS to its nearest peak center

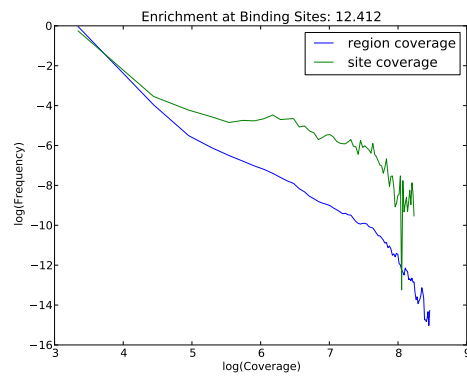


Figure 7: Histograms of coverage at sites and in total regions