

General Motif Characteristics:

Top Similar Looking Motifs:

| | | |
|---|-------|--|
| ENCODE.CTCF | 0.148 | |
| HCMC.CTCF_f2.wm | 0.166 | |
| JASPAR.CTCF.wm | 0.168 | |
| HOMER.BORIS(Zf).K562-CTCF-Seq | 0.176 | |
| CTCF.p2 | 0.177 | |
| HOMER.CTCF(Zf).CD4-CTCF-Seq | 0.182 | |
| HTSELEX.CTCF.C2H2.full.monomeric.wm1 | 0.193 | |
| HOMER.NeuroD1(bHLH).Islet-NeuroD1-Seq(GSE30298) | 0.202 | |
| UNIPROBE.UW.Motif.0004.wm | 0.218 | |
| HCMC.INSM1_f1.wm | 0.219 | |

Statistics:

- Motif name: denovo_WM_5
- Enrichment: 10.9950503896
- log-Likelihood Ratio: 1198.72260348
- Area under precision recall curve: 0.8173

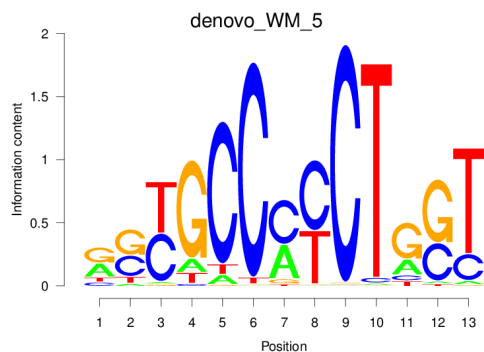


Figure 1: Logo of Weight Matrix

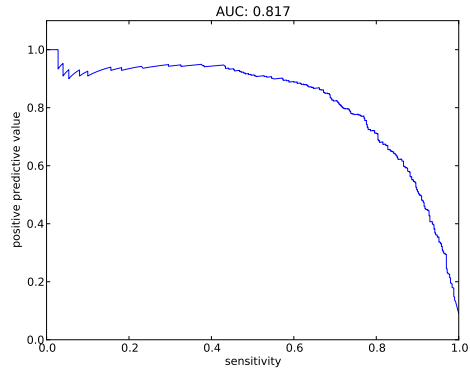


Figure 2: Precision and Recall curve

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: 27102/44017
- 61.57 percent are true.
- Cut-off: minimum summed posterior of 0.20
- Peak plots contain TFBS of posterior ≥ 0.20

Statistic for centering of TFBSs at peak centers:

- Enrichment at binding sites: 3.7859

Correlation between peak Z-score and number of binding sites at peak: 0.4123

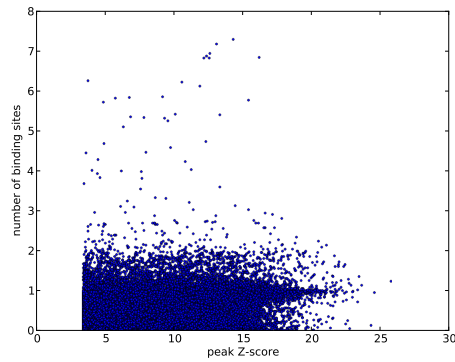


Figure 3: Scatter plot of peak Z-score versus number of binding sites at peak

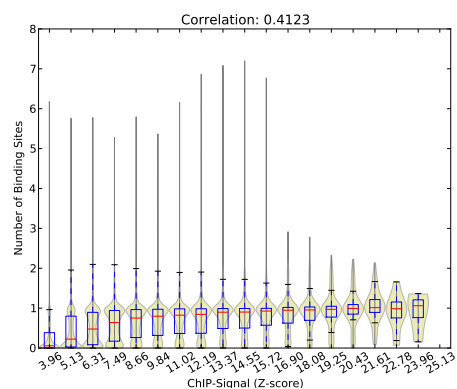


Figure 4: Violin plot of peak Z-score versus number of binding sites at peak

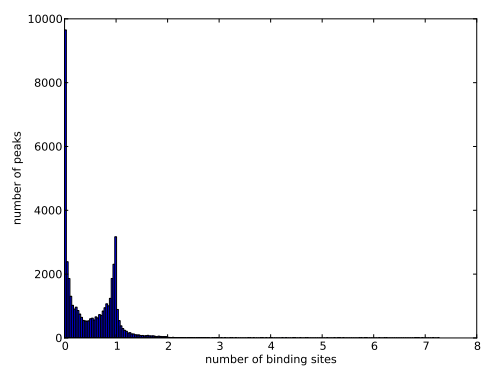


Figure 5: Histogram of number of binding sites at peaks

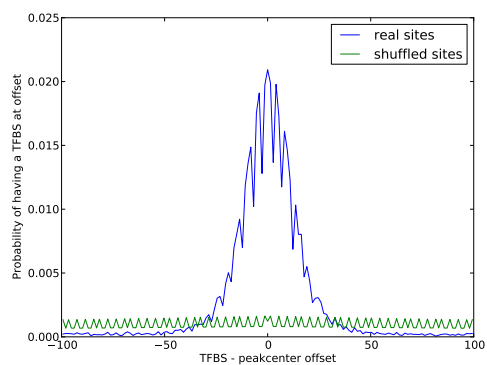


Figure 6: Histogram of distances of TFBS to its nearest peak center

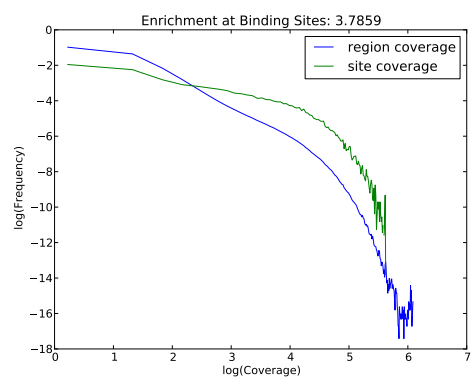


Figure 7: Histograms of coverage at sites and in total regions