

We further studied the differences in TF binding at promoters and enhancers (Figure 5). (The ENCODE consortium has ChIP-Seq data for 60 transcription related factors in H1-hESC cell line, including a few chromatin remodelers and histone modification enzymes. Collectively we call all these transcription related factors “TF”s for simplicity.) The majority of promoters and enhancers contain multiple TF-binding sites. However, the TF-binding of enhancers is more heterogeneous than promoters: more than 70% of the promoters bind to the same set of 2-3 TFs, which is not observed for enhancers. The majority of the promoters also contain peaks for several TATA-associated factors (TAF1, TAF7, and TBP). In addition, the ChIP-seq peaks of a few TFs are mostly within the predicted enhancers and promoters. Overall, the high heterogeneity associated with enhancer TF-binding is consistent with the absence of a sequence code (or grammar) which can be utilized to easily identify active enhancers on a genome-wide fashion.