Distance distribution of randomly selected 7 mers (pages 2-101). Each graph is a histogram that represents the number of occurrences of each motif (every 200 nt stretch) in the first 50 kb upstream region of protein-coding genes. In stark contrast to the top-predicted motifs, the control motifs do not have a strong preference to occur in the neighborhood of the transcription start sites of protein-coding genes.
Number of occurrences

Location of AAAACAG motifs (nt)
Number of occurrences vs. Location of AGTGAGT motifs (nt)
Number of occurrences

Location of CGGGATT motifs (nt)
Number of occurrences vs. Location of CTATCTC motifs (nt)
Number of occurrences of GACTGGC motifs vs. location (in nt)
Number of occurrences

Location of GAGCCAT motifs (nt)
Number of occurrences vs Location of GGCACTG motifs (nt)
Number of occurrences

Location of TTACTTA motifs (nt)