Distance distribution of randomly selected 8 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 AACTAAGA motifs (nt)
Number of occurrences vs. Location of AAGTAATA motifs (nt)
Number of occurrences vs Location of AAATATTA motifs (nt)
Number of occurrences

Location of ATAGGTAT motifs (nt)
Number of occurrences

Location of AGGATCCG motifs (nt)
Number of occurrences

Location of ATGCTGCA motifs (nt)
Number of occurrences

Location of ATTCTAAC motifs (nt)
Number of occurrences

Location of CGACAGCC motifs (nt)
Number of occurrences vs. Location of CTCTTTAG motifs (nt)
Number of occurrences vs. Location of GGCACGGC motifs (nt)
Number of occurrences vs. Location of GTGAGGAC motifs (nt)
Number of occurrences

Location of GTGATATT motifs (nt)
Number of occurrences vs Location of GTTGGGAG motifs (nt)
The graph shows the number of occurrences of TATCTGTC motifs against their location in nucleotides (nt).

- The x-axis represents the location of TATCTGTC motifs in nucleotides, ranging from -50,000 to 0.
- The y-axis represents the number of occurrences, ranging from 0 to 80.
- The data points are distributed across the graph, indicating the frequency of occurrence at various locations.
Number of occurrences of TCTATCCC motifs as a function of their location (nt).
Number of occurrences

Location of TGAACCTAA motifs (nt)
Number of occurrences

Location of TGCTACAC motifs (nt)
Number of occurrences of TGTCACTA motifs vs. location (nt)
Number of occurrences

Location of TGTTCACA motifs (nt)