

The binding sites of transcription factors (TFs) in upstream DNA regions are called transcription factor binding sites (TFBSs). TFBSs are important elements for regulating gene expression. To date, there have been few studies on the profiles of TFBSs in plants. In total, 4,873 sequences with 5' upstream regions from 8530 wheat fl-cDNA sequences were used to predict TFBSs. We found 4572 TFBSs for the MADS TF family, which was twice as many as for bHLH (1951), B3 (1951), HB superfamily (1914), ERF (1820), and AP2/ERF (1725) TFs, and was approximately four times higher than the remaining TFBS types. The percentage of TFBSs and TF members showed a distinct distribution in different tissues. Overall, the distribution of TFBSs in the upstream regions of wheat fl-cDNA sequences had significant difference. Meanwhile, high frequencies of some types of TFBSs were found in specific regions in the upstream sequences. Both TFs and fl-cDNA with TFBSs predicted in the same tissues exhibited specific distribution preferences for regulating gene expression. The tissue-specific analysis of TFs and fl-cDNA with TFBSs provides useful information for functional research, and can be used to identify relationships between tissue-specific TFs and fl-cDNA with TFBSs. Moreover, the positional distribution of TFBSs indicates that some types of wheat TFBS have different positional distribution preferences in the upstream regions of genes.

Identification and positional distribution analysis of transcription factor binding sites for genes from the wheat fl-cDNA sequences

