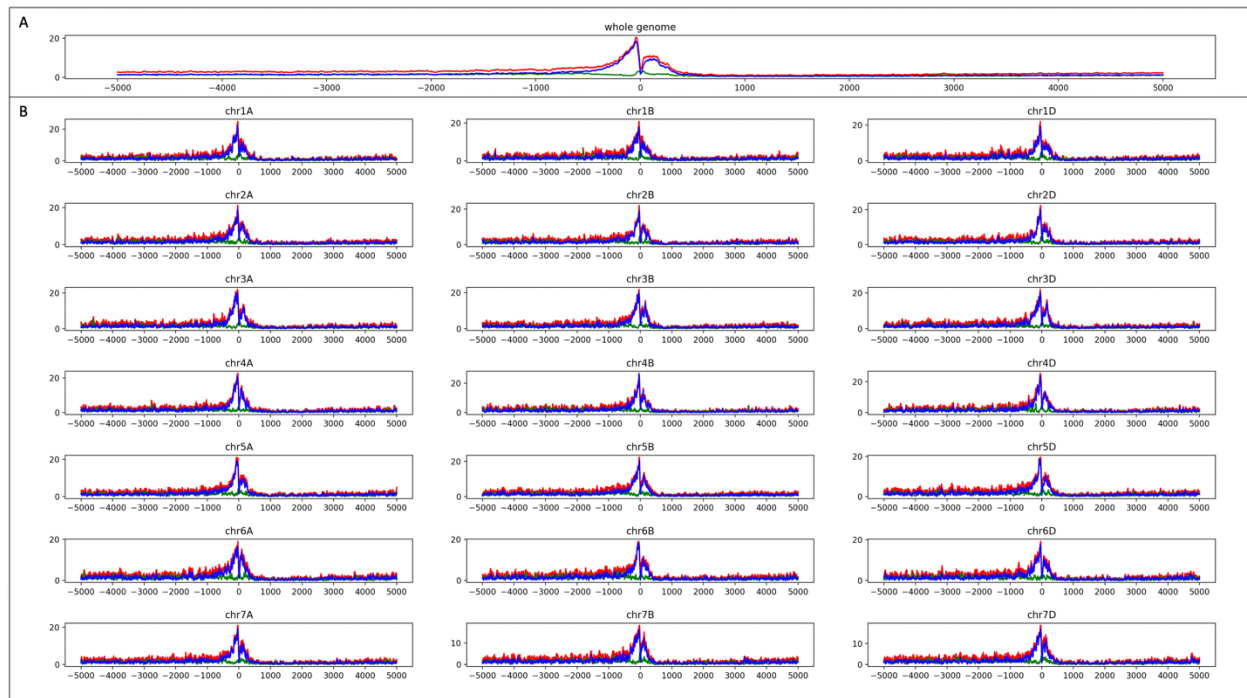


## Supplemental Material for Cagirici and Sen, 2020

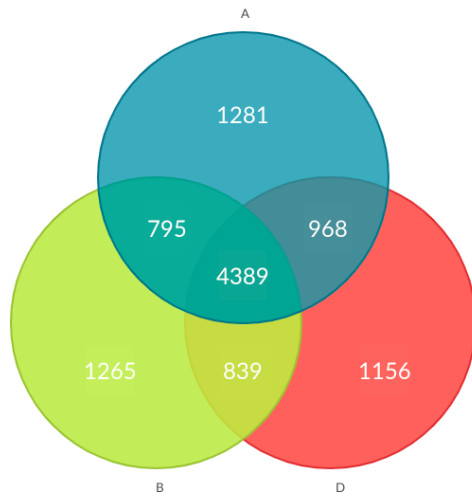
**File S1** BED file for the G4 motifs identified in wheat reference genome (v1.0). (Supplementary\_data\_s1.bed)

**File S2** Annotation of the transcripts containing G4 motifs within gene body and 1,000 bp upstream.

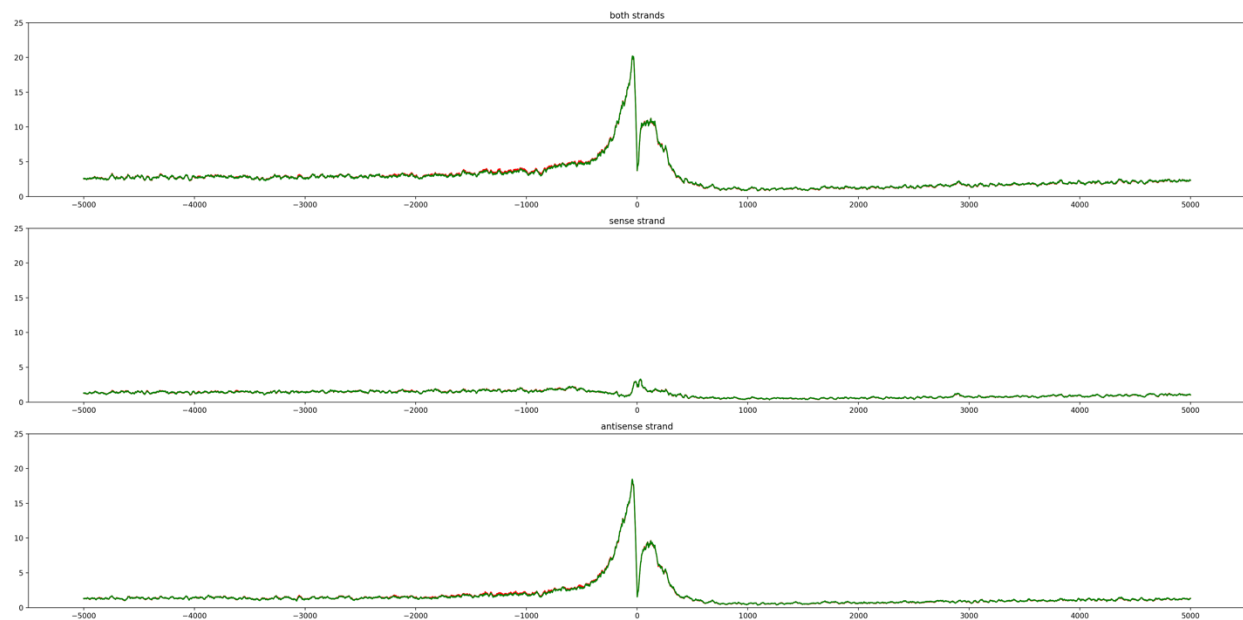
**File S3** Annotation of the transcripts containing G4 motifs within the peak regions.



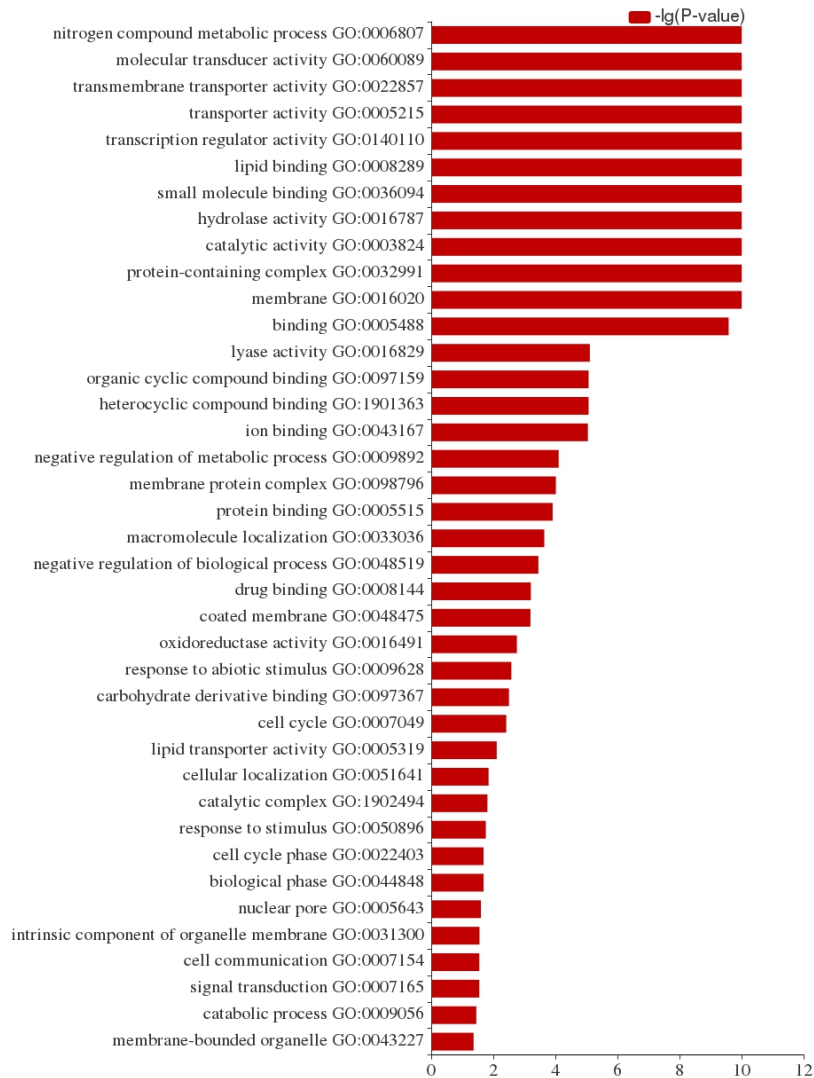
**Figure S1** Distribution of G4s over the 10 kb regions centered at the start codon. The x axis shows the distance from the start codon, and the y axis the G4 frequency at a given position per gene, per Kb. The G4 distributions on the sense (green), antisense (blue), and both strands (red) of the genes were shown **(A)** in the whole wheat genome and **(B)** in each chromosome.



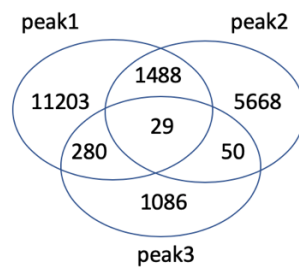
**Figure S2. Number of wheat homoeologs containing G4 motifs were shown.** Gene homoeolog triads only with 1:1:1 correspondence was included. 7,714 triads which does not contain any G4s is not shown.



**Figure S3 G4 motif distribution within 10kb from the start codon of high confidence genes (green) and transcripts (red) plotted on top of each other.** Total number of G4 motifs on both strands was shown on top, G4 motifs on the sense strand only was shown in the middle, and the number of G4 motifs on the antisense strand only was shown at the bottom.



**Figure S4. GO terms that are significantly different between the enrichment in high confidence genes and transcripts.** Y-axis shows the log of the p-values from chi-square test between GO term enrichment at the gene and transcript levels. GO terms with significant relationship (p-value < 0.05) up to level 3 are shown.



**Figure S5. Venn diagram showing the number of genes containing G4 motifs within three peak regions.**