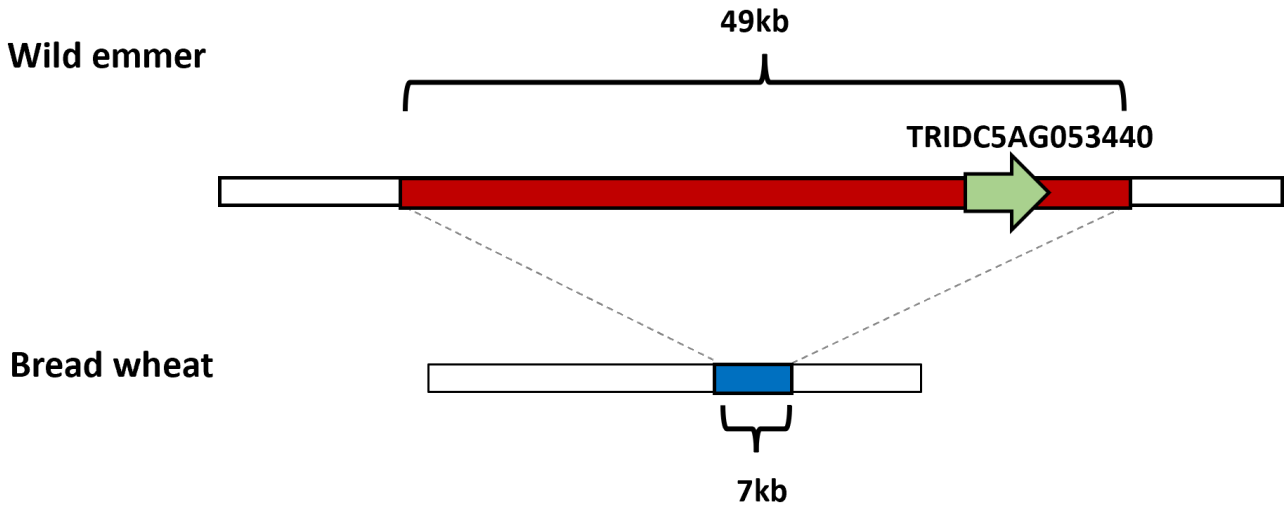

Fig S2a. Schematic representation of locus 5A11 in wild emmer (top) and bread wheat (bottom). An introgression of a new sequence into locus 5A11 in the wheat genome resulted in a 49kb segment in wild emmer compared to a 7kb segment in bread wheat. Sequence length is unscaled. A gene (accession number: TRIDC5AG053440) in wild emmer, represented by a green arrow, shows high sequence identity to a wall-associated receptor kinase 5-like protein. Brown and blue lines represent the wild emmer and bread wheat-specific sequences, respectively. Dashed lines connect orthologous sequence segments at the borders of the indel and the ends of the represented sequences.

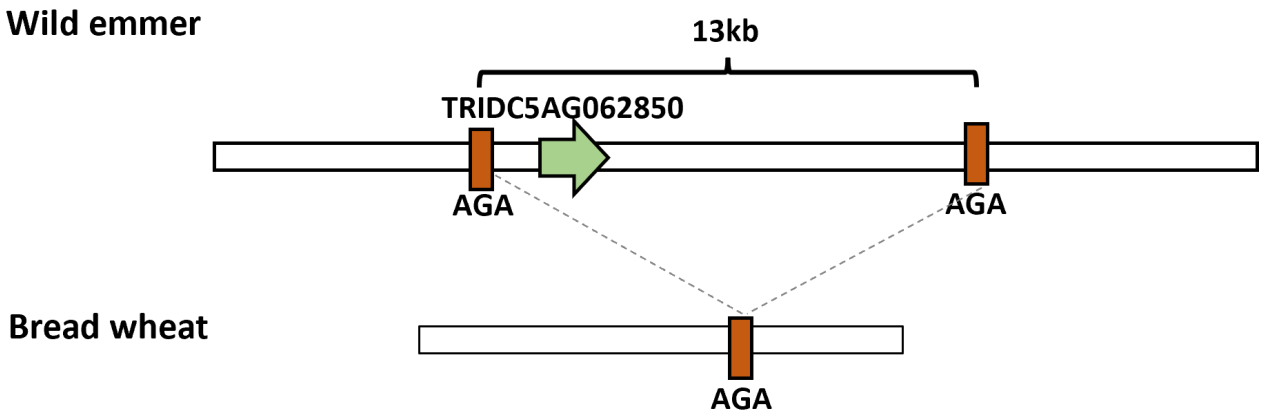

Fig. S2b. Schematic representation of the locus containing 5A13 in the wild emmer (top) and bread wheat genomes (bottom). A deletion of a 13.5kb segment in bread wheat compared to wild emmer resulted in a large-scale structural difference between the two genomes. Sequence lengths are not drawn to scale. A gene (TRIDC5AG062850) in wild emmer, represented by a green arrow, shows sequence similarity to the predicted protein. Indels result in sequence signatures (AGA) represented by brown boxes, characterizing double-strand break (DSB) repair via microhomology-mediated end joining (MMEJ). No genes were identified in the orthologous genomic locus in bread wheat. Dashed lines indicate the alignment between orthologous sequence segments at the indel borders and the terminal ends of the represented sequences.

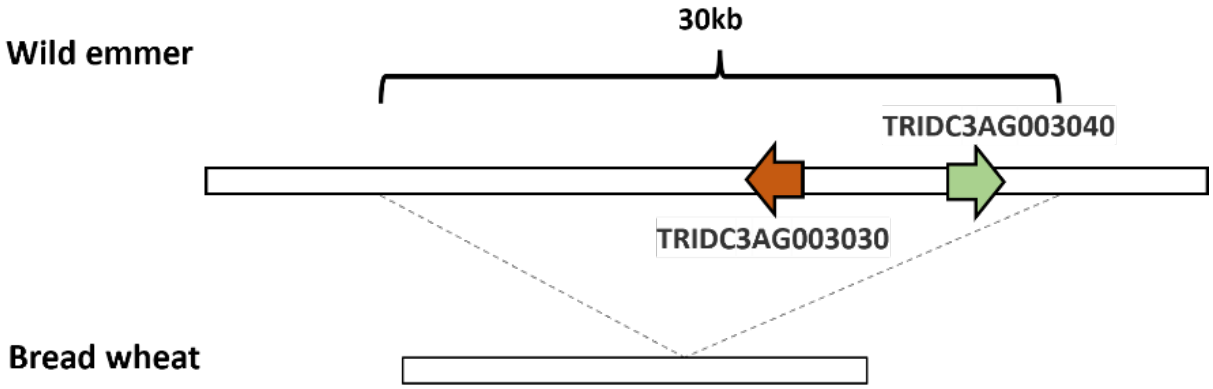

Fig S2c. Schematic representation of the locus containing 3A14 in the wild emmer (top) and bread wheat genomes (bottom). A deletion of a 30kbp segment in bread wheat compared to wild emmer resulted in a large-scale structural difference between the two genomes. Sequence lengths are not drawn to scale. A gene (TRIDC3AG003040) in wild emmer, represented by a green arrow, shows sequence similarity to the predicted protein EXECUTER 2, chloroplastic-like (LOC119266574). Another gene (TRIDC3AG003030), represented by a brown arrow, shows sequence similarity to homeobox-leucine zipper protein HOX15-like. No genes were identified in the orthologous genomic locus in bread wheat. Dashed lines indicate the alignment between orthologous sequence segments at the indel borders and the terminal ends of the represented sequences.

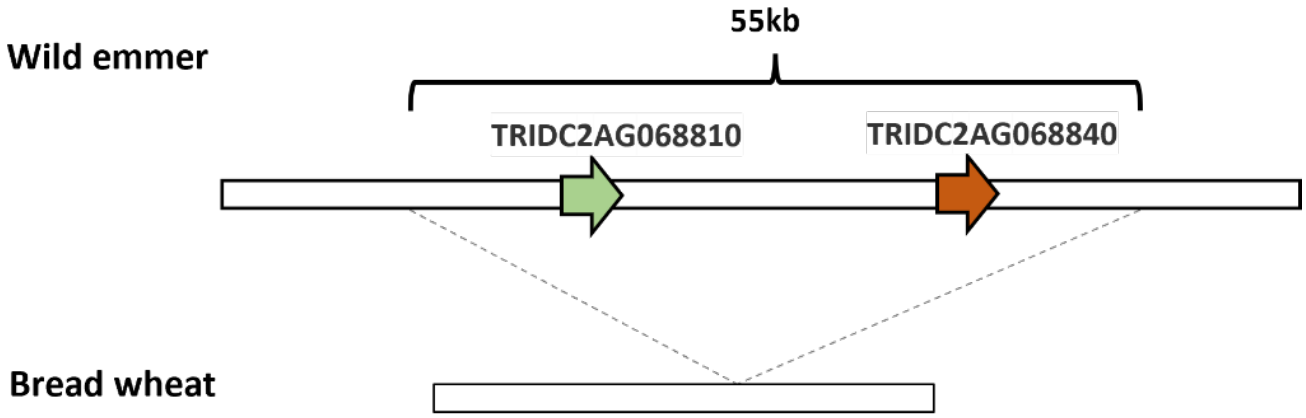

Fig S2d. Schematic representation of the locus containing 2A23 in the wild emmer (top) and bread wheat genomes (bottom). A deletion of a 55kbp segment in bread wheat compared to wild emmer resulted in a large-scale structural difference between the two genomes. Sequence lengths are not drawn to scale. A gene (TRIDC2AG068810) in wild emmer, represented by a green arrow, shows sequence similarity to the predicted protein mitochondrial transcription factor 1. Another gene (TRIDC2AG068840), represented by a brown arrow, shows sequence similarity to senescence-specific cysteine protease SAG39-like. No genes were identified in the orthologous genomic locus in bread wheat. Dashed lines indicate the alignment between orthologous sequence segments at the indel borders and the terminal ends of the represented sequences.

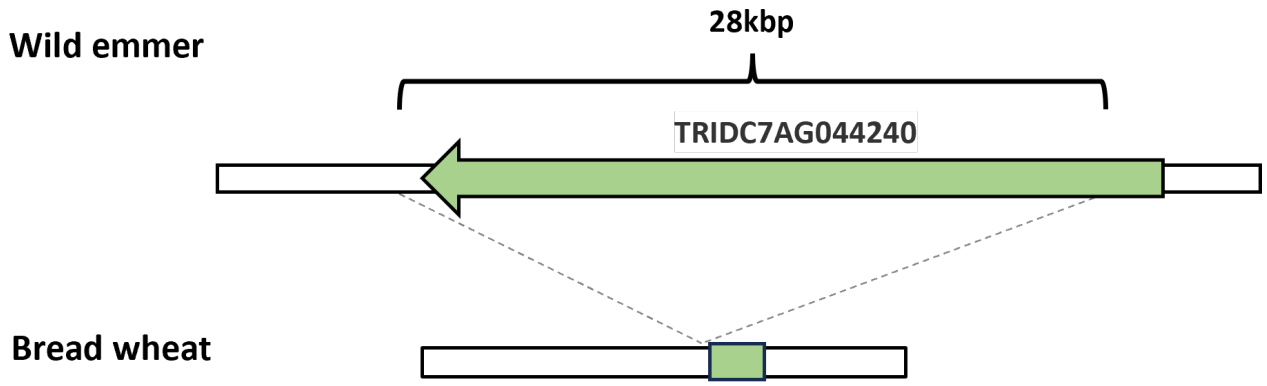

Fig. S2e. Schematic representation of the locus containing 7A30 in the wild emmer and bread wheat genomes. A deletion of a 28kbp segment in bread wheat compared to wild emmer resulted in a large-scale structural difference between the two genomes. Sequence lengths are not drawn to scale. A gene (TRIDC7AG044240) in wild emmer, represented by a green arrow, shows sequence similarity to four isoforms of vacuolar protein sorting-associated protein 13b-like. No genes were identified in the orthologous genomic locus in bread wheat; however, a green box represents the remaining part of the gene found in the 5' region of the indel in bread wheat. Dashed lines indicate the alignment between orthologous sequence segments at the indel borders and the terminal ends of the represented sequences.

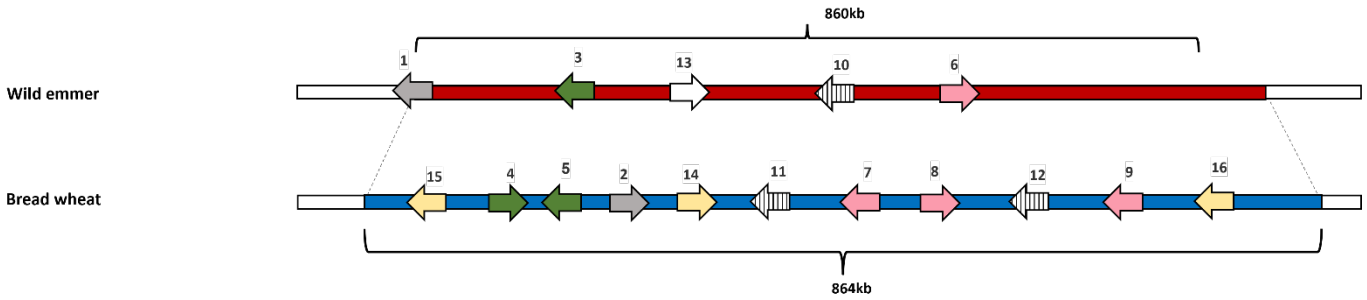

Fig. S2f. Schematic representation of locus 5B19 in wild emmer (top) and bread wheat (bottom). The sequence length is unscaled. Genes are represented by arrows, with each color representing similar genes. Gray arrows represent genes similar to disease resistance protein RGA4-like: (1) TRIDC5BG054370 and (2) TraesCS5B02G344300. Green arrows denote genes with similarity to disease resistance protein RGA5-like or its isoform: (3) TRIDC5BG054400, (4) TraesCS5B02G344100, and (5) TraesCS5B02G344200. Pink arrows represent genes with similarity to auxin-responsive protein SAUR36-like: (6) TRIDC5BG054470, (7) TraesCS5B02G344700, (8) TraesCS5B02G344800, and (9) TraesCS5B02G345000. White arrows indicate uncharacterized proteins, with striped, white arrows representing regions with sequence similarities (greater than 70% identity) between wild emmer and bread wheat: (10) TRIDC5BG054460, (11) TraesCS5B02G344600, (12) TraesCS5B02G344900, and (13) TRIDC5BG054430. Yellow arrows denote hypothetical proteins: (14) TraesCS5B02G344500 and (15) TraesCS5B02G344000. Red lines represent wild emmer-specific sequences, while blue lines represent bread wheat-specific sequences. Dashed lines indicate the alignment between orthologous sequence segments at the indel borders and the terminal ends of the represented sequences.