A global alfalfa diversity panel reveals genomic selection signatures in Chinese varieties and genomic associations with root development

Lin Chen†, Fei He†, Ruicai Long, Fan Zhang, Mingna Li, Zhen Wang*, Junmei Kang*, Qingchuan Yang*

Institute of Animal Science, Chinese Academy of Agricultural Sciences, Beijing 100193, China

†These authors contributed equally to this work.

*Correspondence: Zhen Wang (wangzhen@caas.cn); Junmei Kang, (kangjmei@126.com); Qingchuan Yang (qchyang66@163.com, Dr. Yang is fully responsible for the distribution of all materials associated with this article)

SUPPORTING INFORMATION

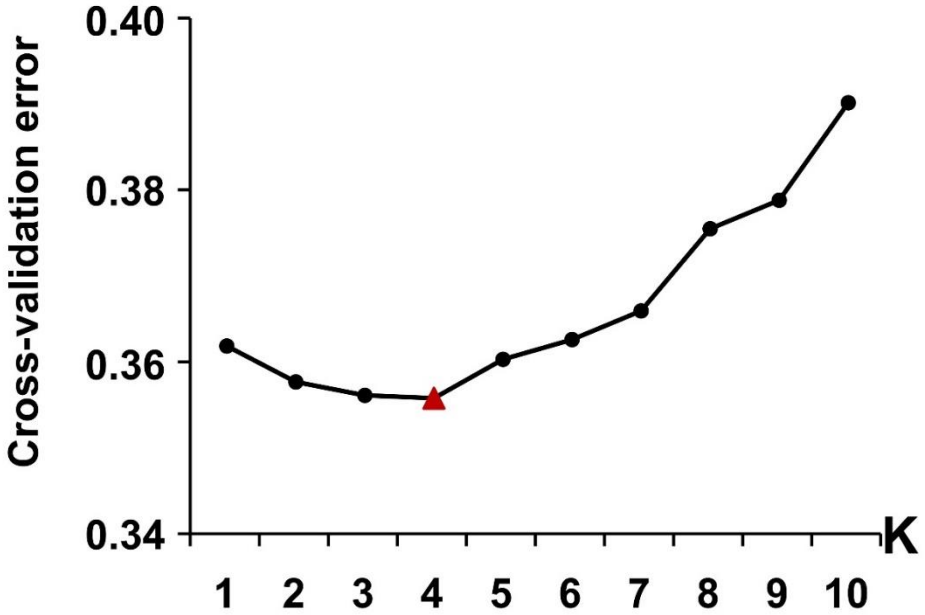


Figure S1. Cross-validation error value for each genetic subpopulation (K from 1 to 10)

The lowest value was achieved when K = 4 (indicated by a triangle), suggesting that the alfalfa diversity panel used in this study consists of four subgroups.

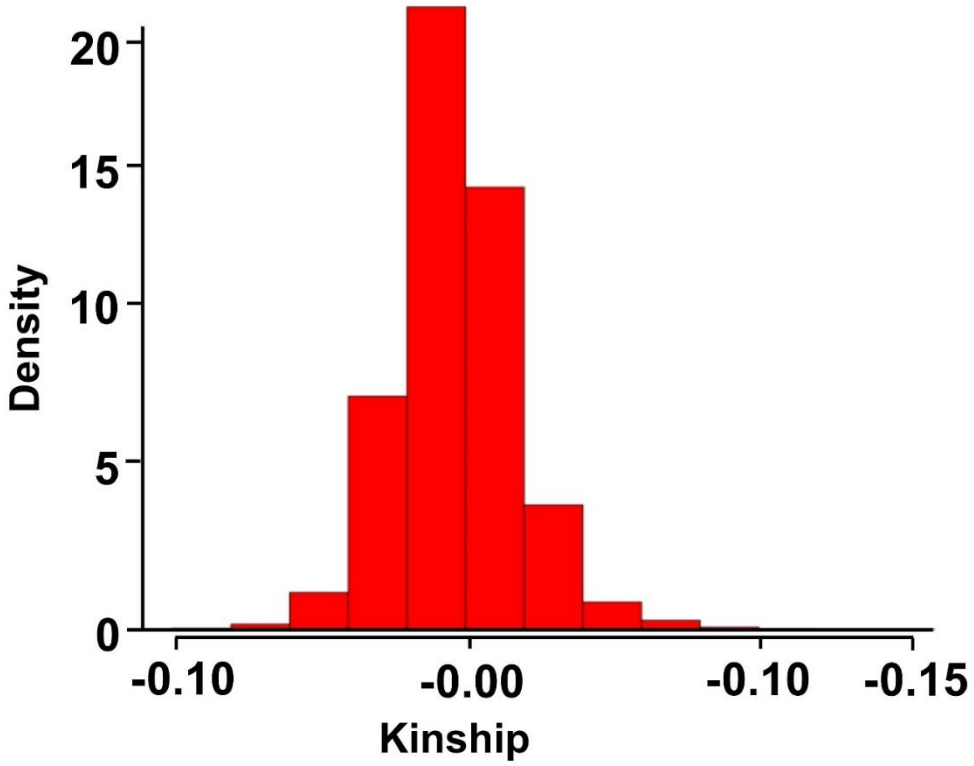


Figure S2. Frequency distribution of the kinship coefficient in the alfalfa diversity panel

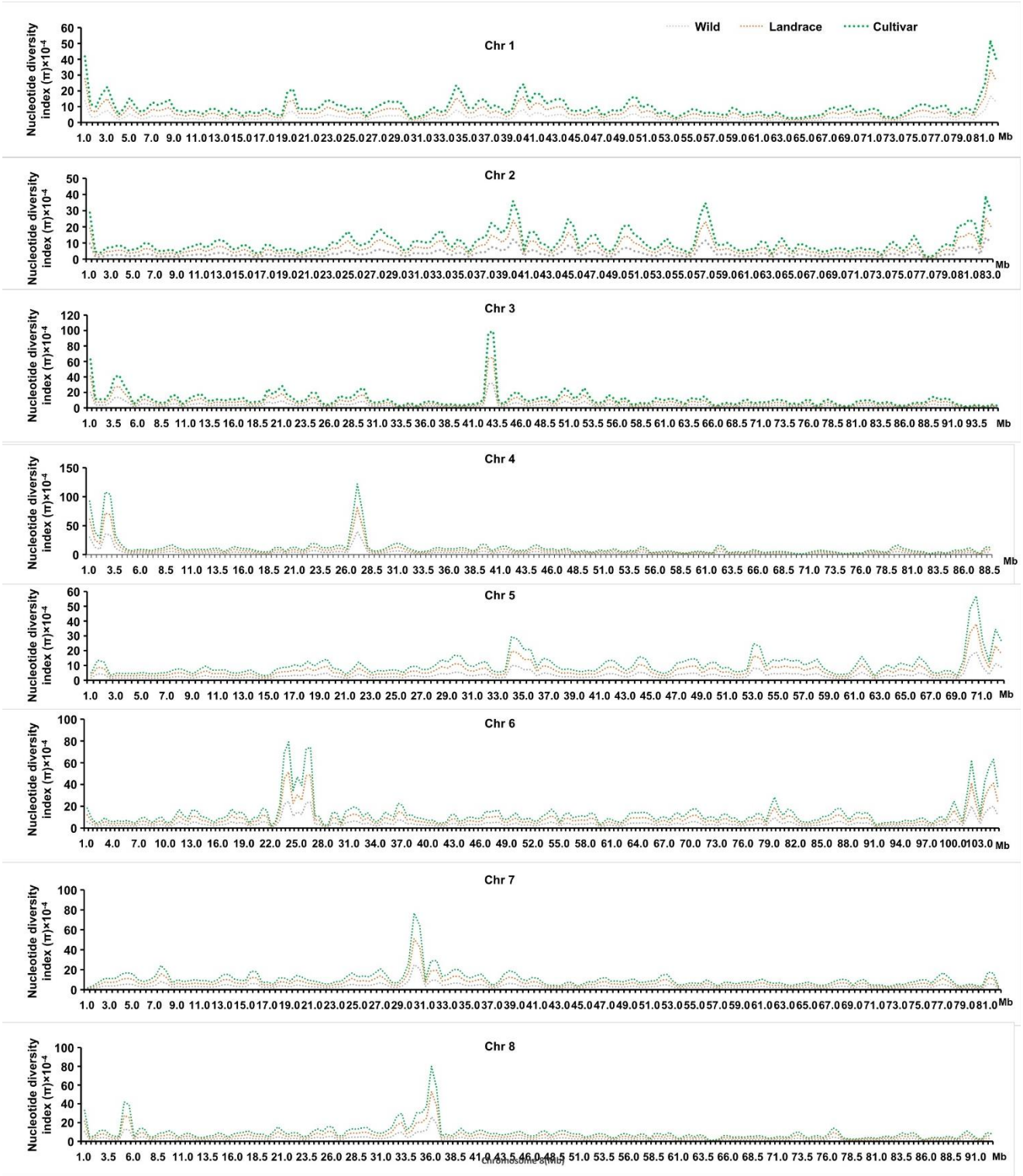


Figure S3. Genome-wide distribution of nucleotide diversity (π) in wild, landrace, and cultivated alfalfa varieties. Statistics are based on 10-Mb windows

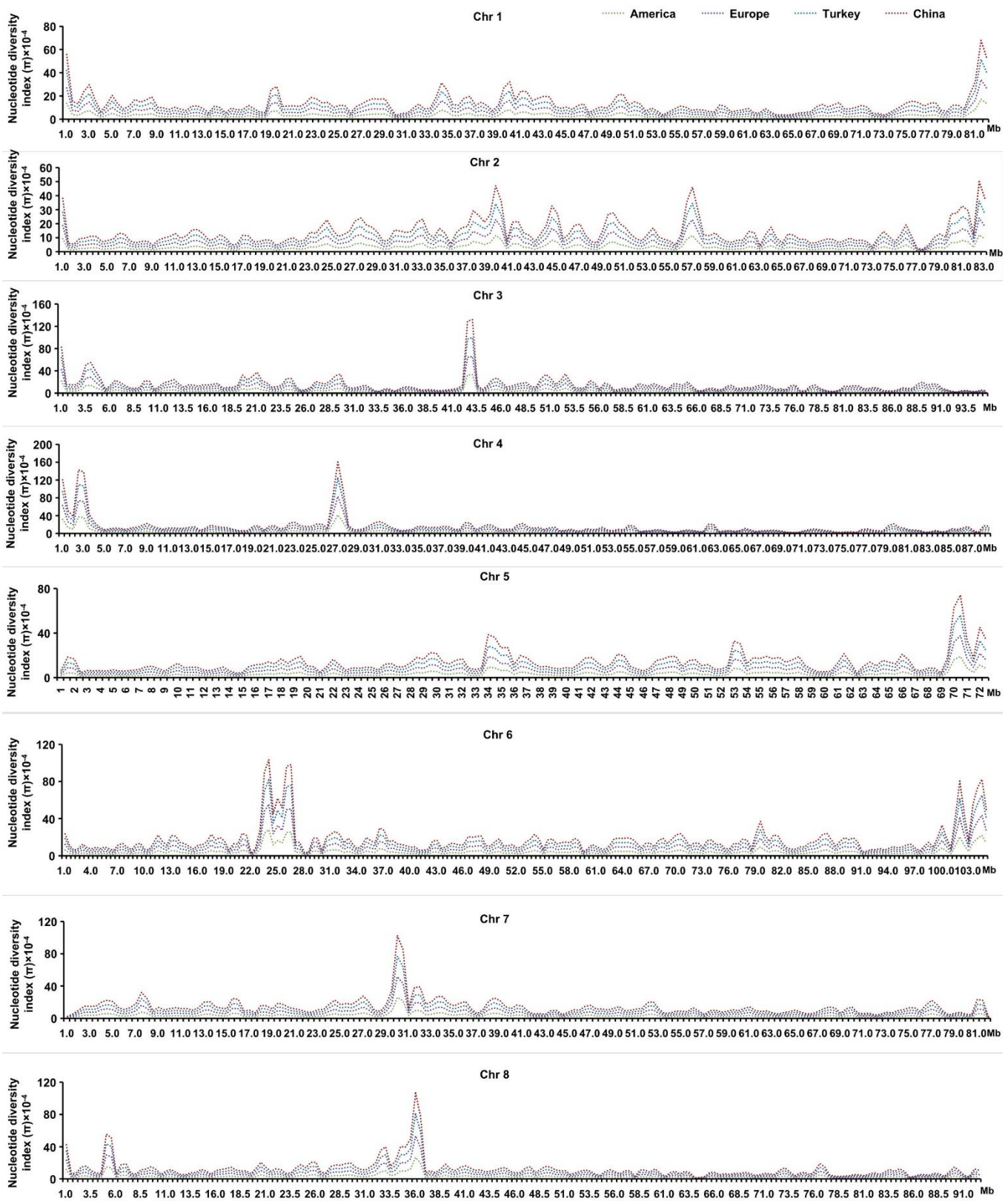


Figure S4. Genome-wide distribution of nucleotide diversity (π) in alfalfa

originating from America, Europe, Turkey, and China. Statistics are based on 10-Mb windows

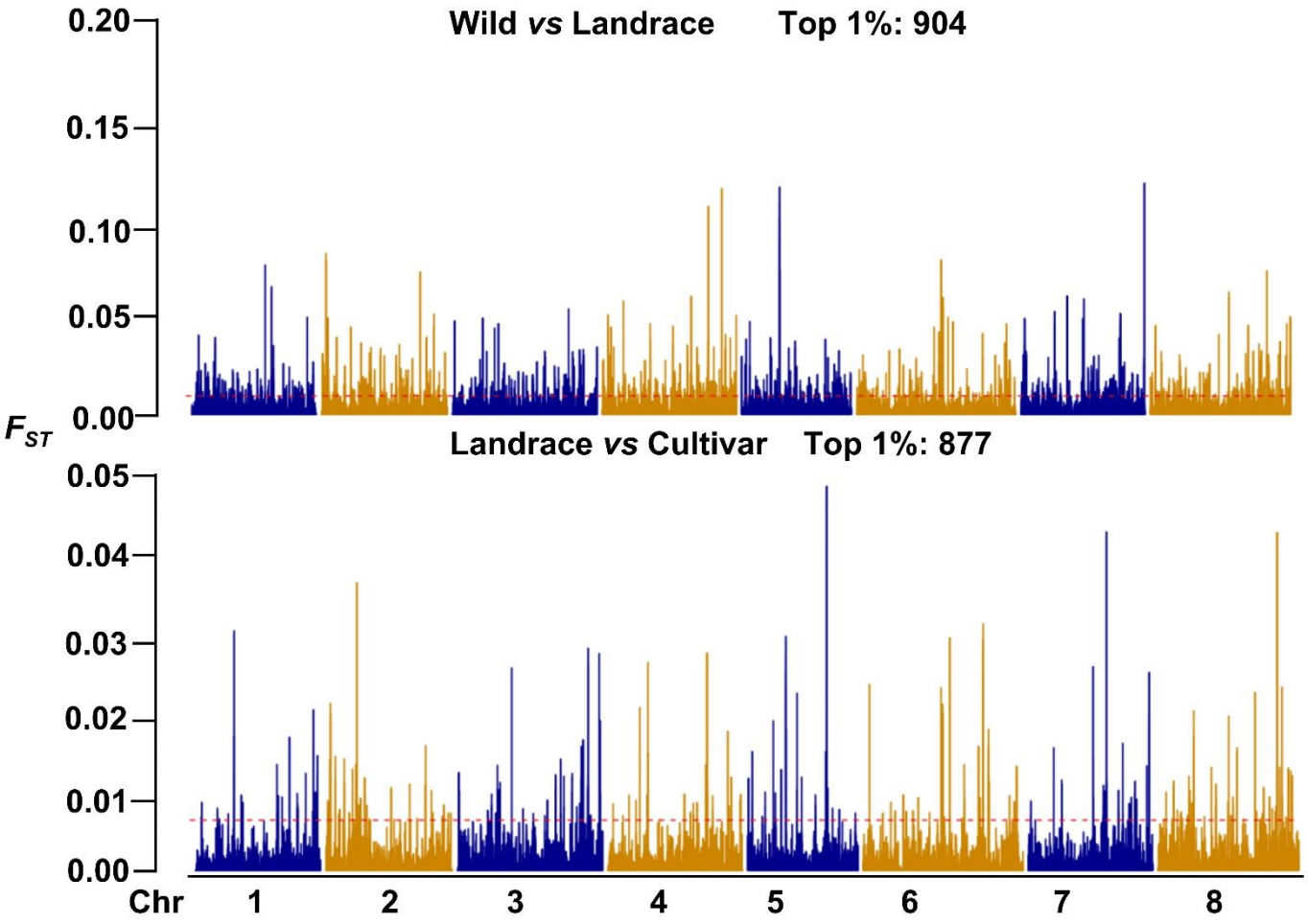


Figure S5. Genome-wide distribution of Fixation Index (F_{ST}) values between subgroups of different breeding status (wild, landrace, and cultivar) of the alfalfa diversity panel. The dashed line in red indicates the 99th percentile

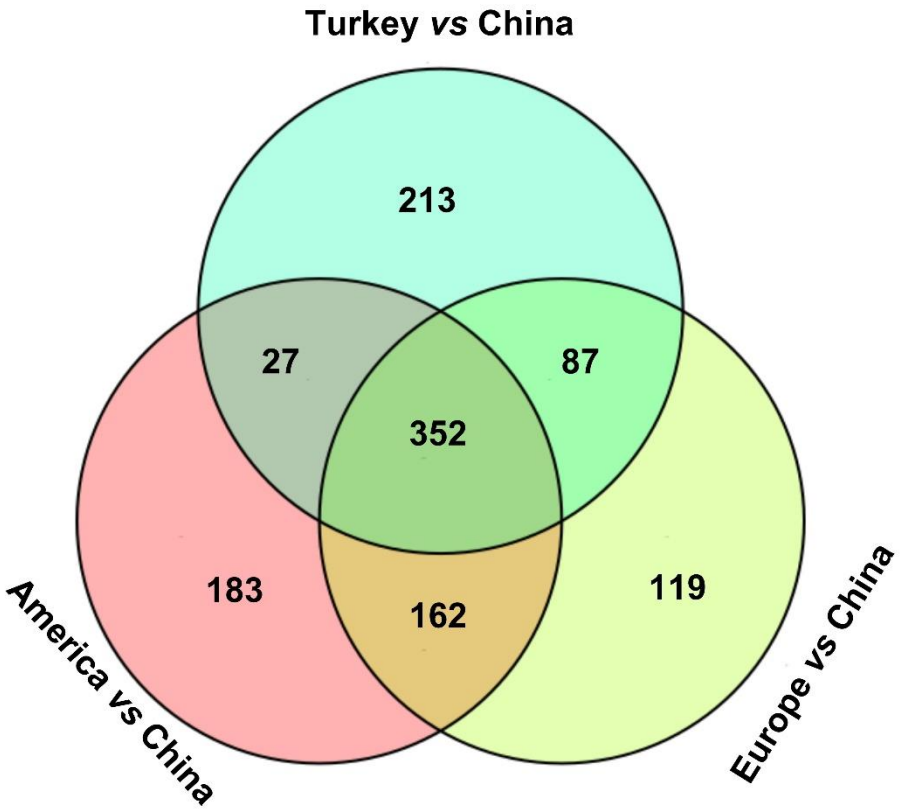


Figure S6. Venn diagram showing the extent of overlap between the number of selected regions in alfalfa varieties from America, Europe, Turkey, and China

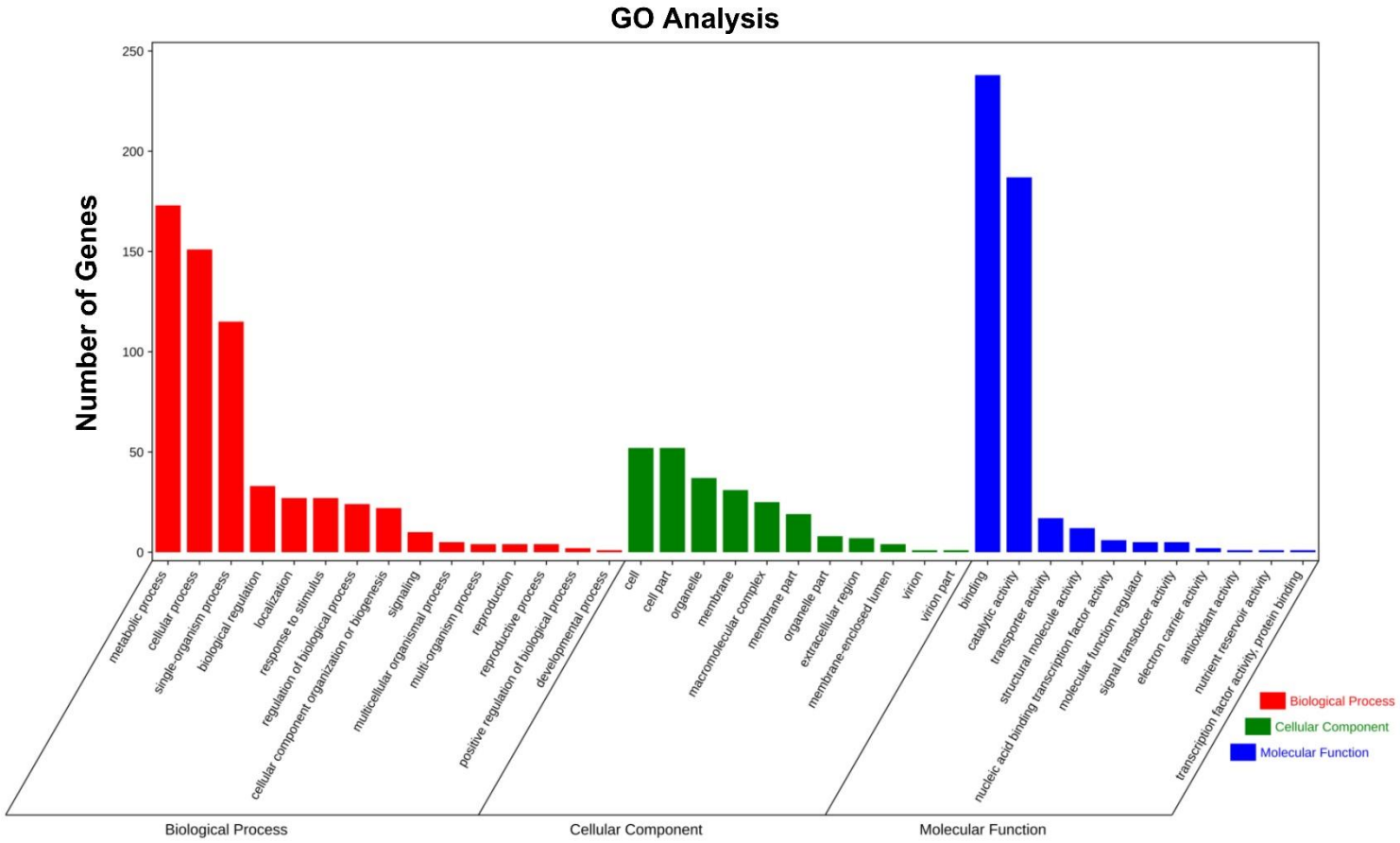


Figure S7. GO term enrichment analysis for the genes in the shared selected regions across subgroups

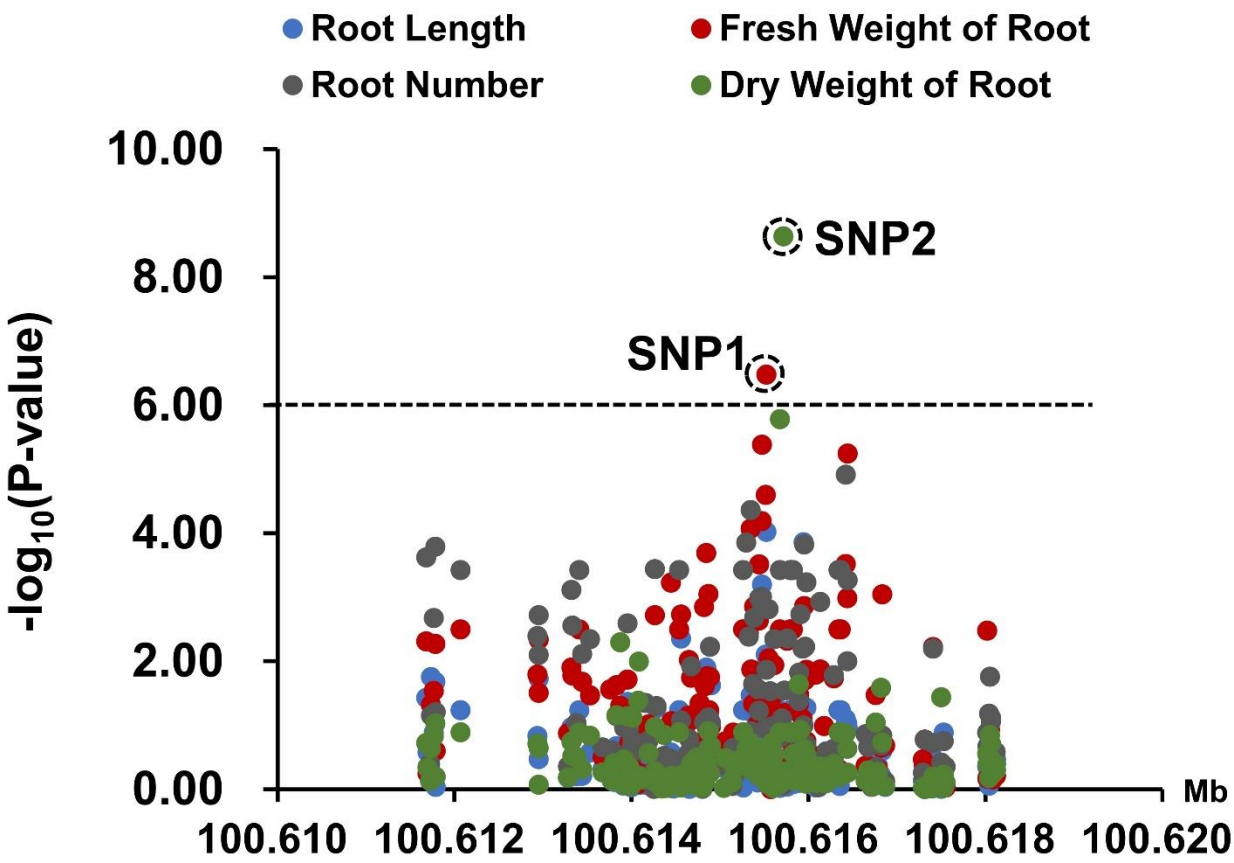
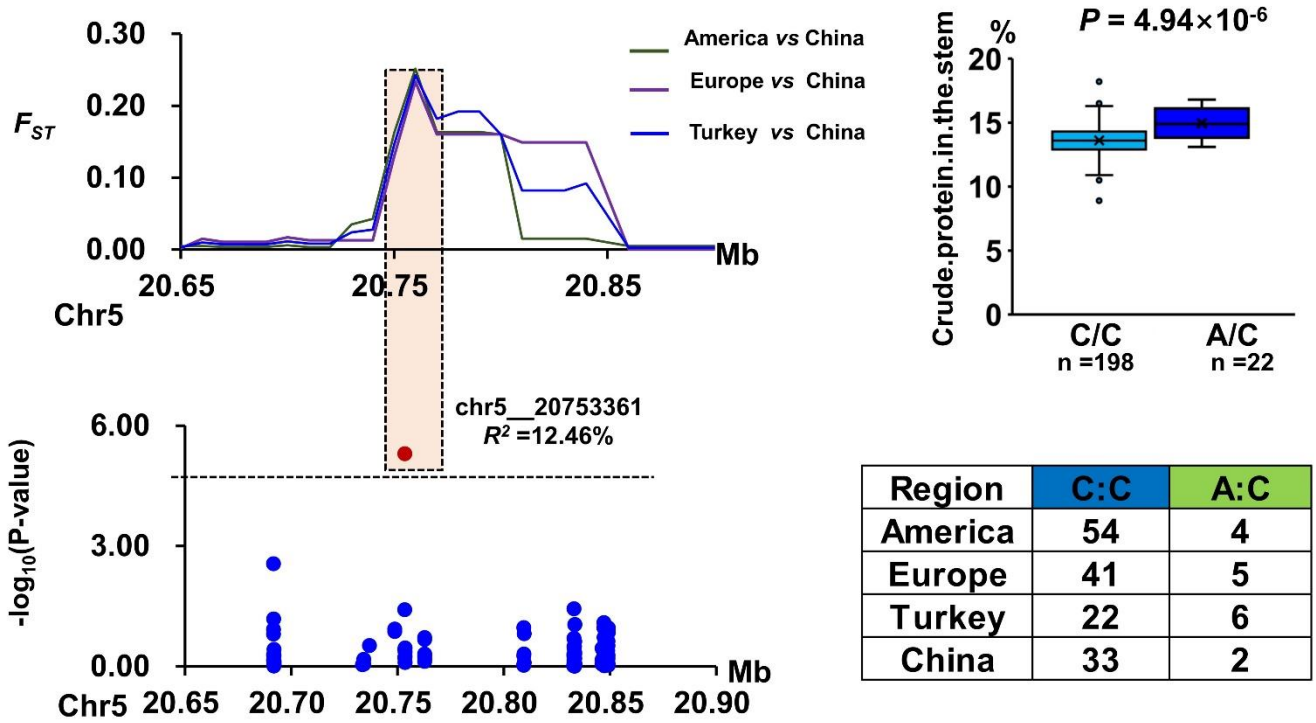


Figure S8. Association analysis for the root weight candidate gene *MsTIR*. Four root-related traits (length, number, fresh weight, and dry weight) were used. Two SNPs (SNP1 and SNP2) located in the fifth exon and the sixth intron of *MsTIR* were significantly associated with fresh weight and dry weight, respectively



Region	C:C	A:C
America	54	4
Europe	41	5
Turkey	22	6
China	33	2

Figure S9. The selected region on chromosome 5 overlaps with a GWAS region identified for crude protein content in stems

Table S1. Information of the alfalfa varieties used in this study

Table S2. Regions under genetic selection detected between the subgroups of wild and landraces, or landraces and cultivars

Table S3. Regions under genetic selection detected among the subgroups of America, Europe, Turkey, and China

Table S4. Common selected regions detected among the alfalfa subgroups of America, Europe, Turkey, and China

Table S5. Information on the genes within the 350 common selected regions

Table S6. Allele frequency for the candidate genes within the selected genomic regions under selection from different geographical origins

Table S7. Results of GO analysis for the 686 genes

Table S8. Phenotypic variation source for the 24 traits used in this study

Table S9. Results of GWAS analysis for the 24 phenotypic traits

Table S10. List of candidate genes for the 24 phenotypic traits