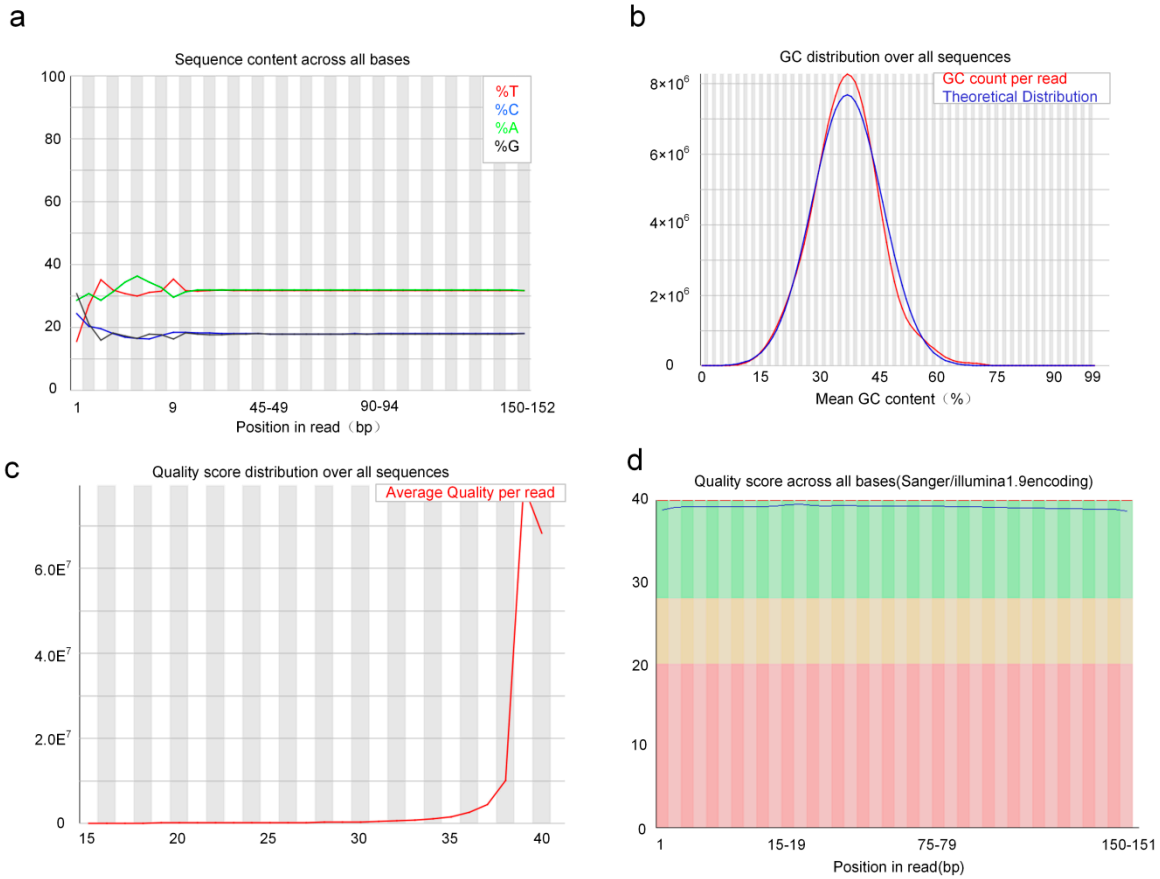


Figure S3. Quality assessment of the raw sequencing data for bulked segregant analysis (BSA).

(a) Base content distribution plot: The x-axis represents the base position within the reads (5' to 3'), and the y-axis represents the proportion of each base (A, T, C, G) at that corresponding position. (b) GC content distribution plot: The x-axis represents the GC content (%), and the y-axis represents the number of reads. The red curve indicates the observed distribution, while the blue curve represents the theoretical distribution (the peak of which is species-specific). (c) Sequence base quality distribution plot: The x-axis represents the mean quality score per read, and the y-axis represents the number of reads corresponding to each mean quality value. (d) Per-base sequence quality plot: The x-axis represents the base position within the reads (5' to 3'), and the y-axis represents the quality score (Q-score). The red line shows the median value, the blue line represents the mean, and the whiskers indicate the 10th to 90th percentile range.