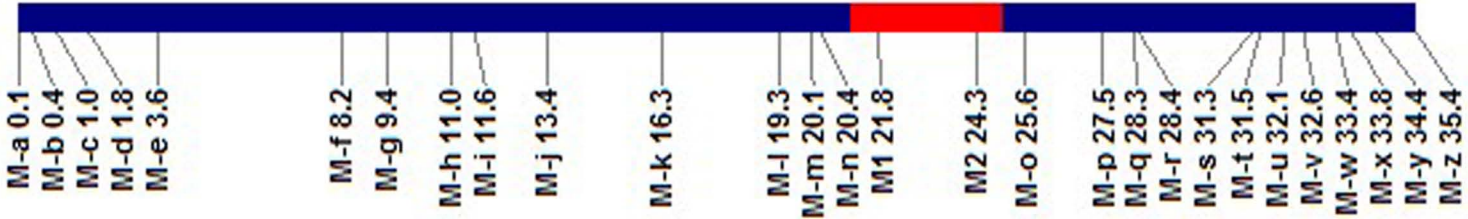

Figure S1. The size of W1943 segment in NIL8. Blue bars indicate Guangluai 4 genomic background. Red bar indicates the W1943 genomic segment. The underside show the order of the markers used in this study and the upside numbers show the physical distance (Mb) according to Nipponbare genome chromosome 4 sequence.

Figure S2. Plant and panicle architecture of NIL8 and Guangluai 4. A, Plant architecture of NIL8 (left) and Guangluai 4 (right). B, Panicles of NIL8 with black hull (left) and Guangluai 4 with straw-white hull (right).

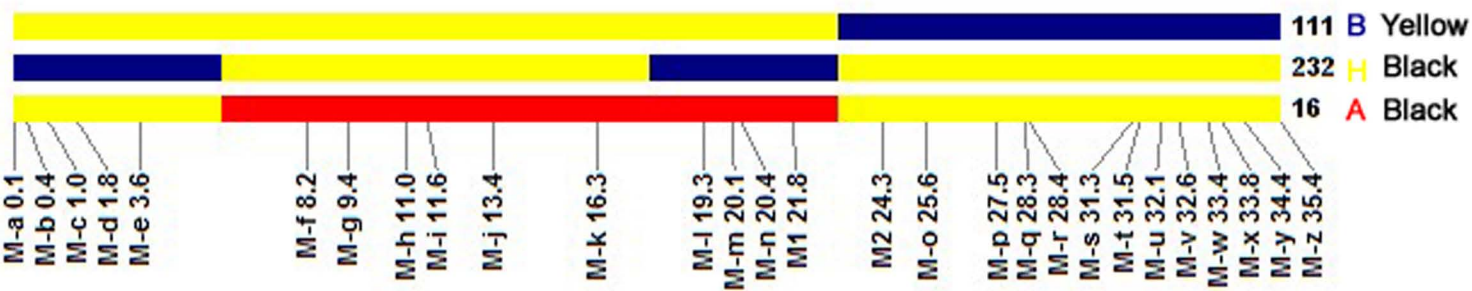

Figure S3. Some of F2 individuals used for *Bh4* primary mapping. The numbers and letters at right represent the serial number and *Bh4* genotypes of each F2 individual respectively. A: indicates SL4 (W1943) homozygous, B: indicates Guangluai4 homozygous, H: heterozygous. The genotype of *Bh4* locus of each F2 individual was judged from the corresponding F3 population. From these three individuals, the *Bh4* was localized in the region between the M1 and M2 markers.

Figure S4. Multiple sequence alignment of the rice BH4 protein and the amino acid transporter proteins from Arabidopsis and tomato.Black and gray shading indicate 100% and 80% conserved amino acid residues, respectively. The names of proteins are indicated on left side.

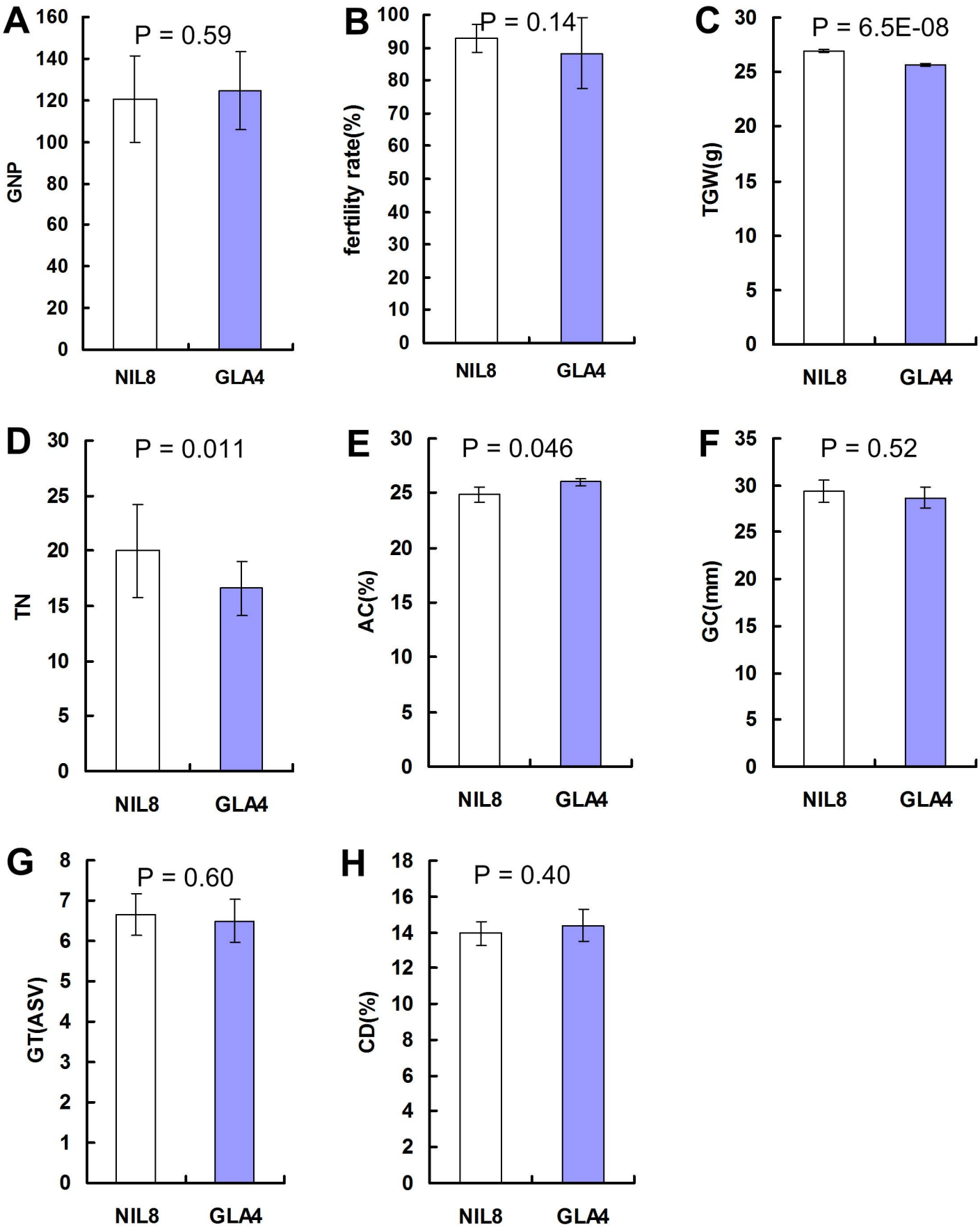

Figure S5. Comparison of grain number per panicle(GNP), fertility rate, thousand grain weight(TGW), tiller number per plant (TN), amylose content(AC), gel consistency(GC), gelatinization temperature(GT) and chalkness degree(CD) between NIL8 and Guangluai4(GLA4). For GNP, fertility rate, TN analysis, n=15; and TGW, AC, GC, GT, CT, n=3. All data are given as means ± s.e. A Student's t-test was used to generate the P values.

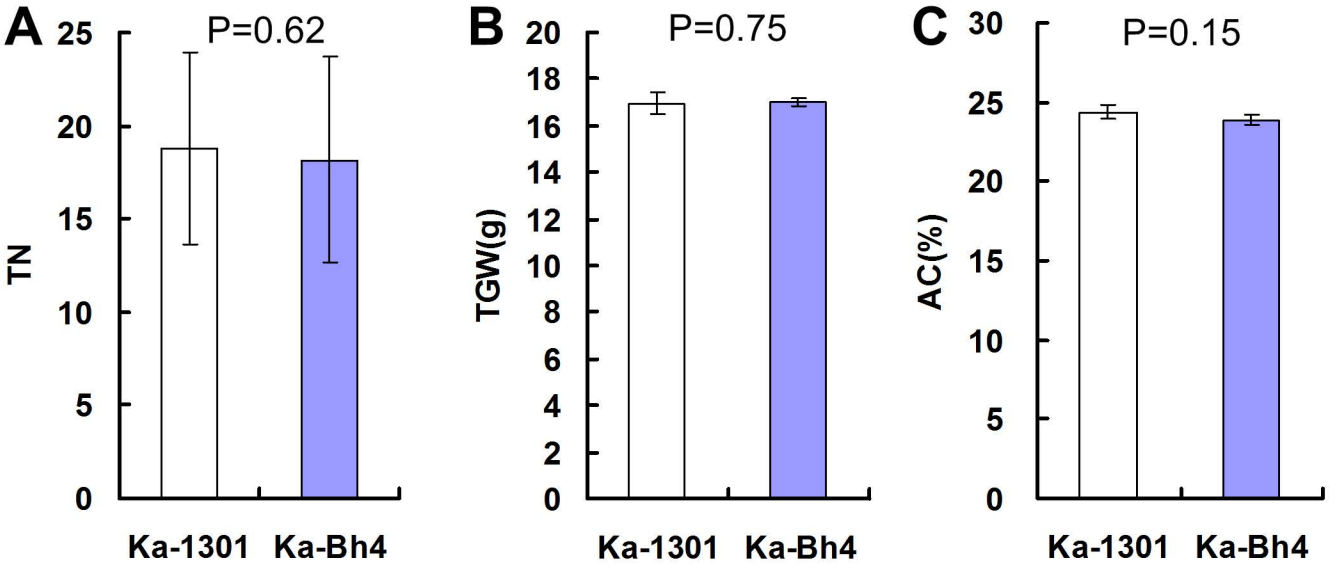

Figure S6. Comparison of TN (tiller number, A), TGW (Tousand grain weight, B) and AC (amylose content, C) between control (Ka-1301) and transgenic line(Ka-Bh4). There is no significant difference of the tree traits analyzed between control and transgenic line. This indicates the difference detected between NIL8 and GLA4 is caused by other genes. All data are given as means ± s.e. A Student's t-test was used to generate the P values.