File S2: SI TEXT

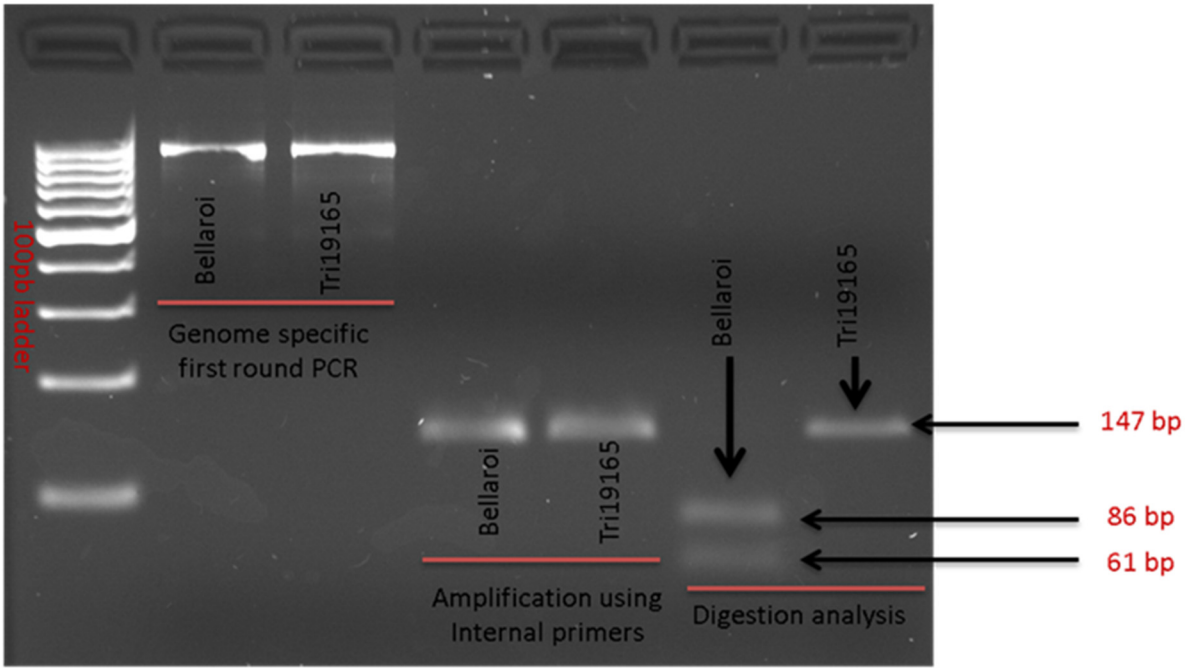
Candidate gene specific marker development in barley and tetraploid wheat

The orthologs of maize *BD1* and rice *FZP/BFL1* (CHUCK *et al.* 2002; KOMATSU *et al.* 2003; ZHU *et al.* 2003) were considered as candidate genes underlying *bh*^{t} and *com2.g* phenotypes in wheat and barley, respectively. To map the barley ortholog, the *COM2* gene, in barley the corresponding specific gene-based markers were developed. The two CAPS markers (table below) are both on the basis of SNP A873G found between the two parents of the population (Haruna Nijo x BW 192 *com2.g*). Barley marker information is provided in the table below. Each of the two markers could be utilized for mapping.

primer name	orientation	primer seq	Restriction Enzyme	Product size	Digested Product size_wild	Digested Product size_mutant	Digested Product size_heterozygote	Tm (°C)
HN_SNP2for	Forward	AACTCCGGGTACCTGAGCA	BtgZI	327 bp	219 bp; 108 bp	327 bp	327 bp; 215 bp; 112 bp	60,66
HN_SNP2rev	Reverse	CAGATCGGCCATTAAGTGAG						58,33
GW_3UTR_For3	Forward	GAGGACGTGGACGACCTG	Hpy99I	207 bp	10bp; 31bp; 168 bp	10bp; 31bp; 59 bp; 108 bp	10bp; 31bp; 59 bp; 108 bp; 168 bp	57,7
HN_SNP2rev	Reverse	CAGATCGGCCATTAAGTGAG						58,33

Wheat marker information for the candidate gene underlying *bh*^{t} is presented in the following table and exemplary image. This CAPS marker is developed on the basis of SNP T287C found between parents of the corresponding mapping populations. Two-step PCR reactions were followed. In the first round, any of the A genome specific primers 1 or 2 could be used. Second, the internal primer pair *(*TdFZP2A_in_F and TdFZP2A_in_R*)* was used to amplify a short fragment using the first round PCR product as template. This short fragment was used for CAPS marker development. At the image below, amplification of the candidate gene using genome-specific primers (left), amplification of region of interest using internal primers (middle), and digestion analysis (right) are depicted.

primer name	Product ID	primer seq	Restriction Enzyme	Product size	Digested Product size _ wild	Digested Product size _ mutant	Digested Product size _ heterozygote
Tafzp_2A_Forward 1	A geneome specific 1	AGCCAACCTCACTTCACTTC	-	946 bp	946 bp	946 bp	946 bp
Tafzp_2A_Reverse 1		GAGCAATGCCAGCGCGTCCGT					
Tafzp_2A_Foward 2	A geneome specific 2	CTAGGCGGGAGCAGTAGTA	-	963 bp	963 bp	963 bp	963 bp
Tafzp_2A_Reverse 2		AGCGCGTCCGTTTCAGTGG					
TdFZP2A_in_F	Internal for A.G. Specific 1 and 2	GACCCGACCACCAAGGAG	BstNI	147 bp	61 bp; 86 bp	147 bp	61 bp; 86 bp; 147 bp
TdFZP2A_in_R		GTAGTTGTTGTAGGCGGCGT					



Genetic mapping in barley

To newly map the phenotype in barley, we developed an F_2 mapping population comprising 286 individuals between the parental mutant BW 192 *com2.g* and barley cv. Haruna Nijo. The parental mutant BW 192 *com2.g* is the Bowman Near Isogenic Line of *com2.g* and was previously developed until BC7≤ generation (DRUKA *et al.* 2010). The segregation pattern of the phenotype among the corresponding F2 barley plants fitted well with a 3:1 ratio typical for a monogenic recessive trait. The syntenic information reported in the form of barley chromosome 2H Genome Zipper (MAYER *et al.* 2011) was explored to genetically localize the phenotype and to develop further markers surrounding the locus (Figure S2). This low resolution genetic mapping localized the *com2.g* phenotype in an interval of 5.5 cM along the barley chromosome 2H short arm. The interval was flanked by two barley gene based markers orthologous of rice genes Os07g0673700 (barley CAPS marker com2_p11/com2_p12) and Os07g0668900 (barley CAPS marker com2_p19/com2_p20), respectively. The two aforementioned flanking markers were used for screening a larger population consisting of 1750 F_2 plants from which 52 recombinant plants were identified.

While fine-mapping *com2.g*, we discovered that the observed genotypic alleles for the *COM2*-specific marker failed to fully match the *com2.g* phenotypic score since six out of the 52 F_2 recombinants showed a discrepant phenotype. While genotypically heterozygous for *COM2*, they showed the supposedly recessive branched phenotype. F_3 offsprings (17 to 28

progenies per F_2 totaling 128 F_3 plants) segregated for the respective spike-branching phenotype at both phenotypic and genotypic levels thus confirming the heterozygosity of the parental F_2 plants. However, also F_3 plants showed inconsistencies between genotypic and phenotypic data. While all 36 genotypically homozygous mutant plants showed clear spike-branching, 33 of the 70 heterozygous and three out of 22 homozygous wild type plants showed a faintly branched spike. The deviant F_3 plants were progenies from three out of six F_2 plants which suggest that additional genetic and/or environmental factors regulate spike-branching. This is supported by observations in maize that found an effect of genotypic background on the expressivity of the inflorescence branching (COLOMBO *et al.* 1998).

Genetic mapping in tetraploid wheat

A linkage map for *bh*^{t} in tetraploid wheat was established in parallel to genetic mapping of *com2.g* in barley. Two different mapping populations consisting of 279 and 159 F_2 plants were developed using a set of published microsatellite markers (RODER *et al.* 1998) (Figure S2), the phenotype could be genetically mapped to a region spanning ~20 cM on wheat chromosome 2A short arm (2AS) (Figure S2) that confirmed the previous finding of *bh*^{t} genetic mapping (KLINDWORTH *et al.* 1997). Both mutant parents TRI 27966 and TRI 19165 carry the same recessive allele of *bh*^{t} showing only a single amino acid substitution of leucine to proline at position 96 (L96P) as compared to the wild type. However, further genetic mapping for positional cloning of the gene was not followed in wheat.

SI MATERIALS AND METHODS

Marker development in barley

Barley chromosome 2H genome zipper (GZ) (MAYER *et al.* 2011) was utilized for initial marker development. Barley sequence information, the homologs of the rice genes ordered along the 2H-GZ was used for primer design. Barley sequence information was obtained from IPK Barley Blast Server (http://webblast.ipk-gatersleben.de/barley/viroblast.php). Corresponding rice genes were extracted from respective genome browser server (http://rice.plantbiology.msu.edu/cgi-bin/gbrowse/rice/). Rice gene sequences were blasted against the IPK barley blast server (http://webblast.ipk-gatersleben.de/barley/viroblast.php) to obtain barley sequences. Sequences were amplified in parental lines of the mapping population using primers designed to detect SNP polymorphisms. Identified SNPs were converted to restriction enzyme based CAPS (http://nc2.neb.com/NEBcutter2/). The barley ortholog of rice *FZP/BFL1* gene sequence (*Os07g0669500)* was used for candidate gene marker development. In case of wheat, publicly available SSR markers (RODER *et al.* 1998) were used for genetic mapping.

TILLING analysis

Barley

For identifying further mutant alleles of *COM2* in barley, a TILLING population consisting of 10,279 EMS (Ethyl methanesulfonate) treated plants of cv. Barke was screened. A primer combination (Table S6) was used to amplify the single exon of the *COM2* gene. The product was subjected to standard procedure of AdvanCE™ TILLING kit as described in (GAWRONSKI *et al.* 2014). Amplified products were digested with dsDNA cleavage kit followed by analysis via mutation discovery kit and gel - dsDNA reagent kit. These were performed on the AdvanCE™ FS96 system according to manufacturer's guidelines (advanced analytical, IA, USA). Amplified ORF was also re-sequenced by Sanger sequencing on all accessions carrying polymorphism identified to confirm SNPs. In addition to the Barke population, a different TILLING population, TillMore (in the cv. Morex) (TALAME *et al.* 2008) was also screened for additional mutant alleles of *COM2.* A set of TillMore lines with branched phenotype was screened. From these lines, the *COM2* single exon was re-sequenced for detection of causal SNPs.

Wheat

To identify *TtBH-1* mutants in wheat, a tetraploid TILLING population consisting of 1,139 EMS plants of cv. Kronos was screened (Uauy et al 2009). Homoeologue-specific primers were designed (Table S6) and tested for specificity; betaine was added at 1M final concentration to the PCR reactions. A 1,011 bp fragment of *TtBH-A1* and a 1,218 bp fragment of *TtBH-B1* were screened. Mutant detection was performed with *Cel*1 digestion followed by analysis on a capillary ABI3730 sequencer (Applied Biosystems, Foster City, California, USA) using published protocols (LE SIGNOR *et al.* 2009). Individual DNAs from positive pools were Sanger sequenced to identify the nature of the mutations.

Haplotype analysis

Genomic DNA was extracted using mixed alkyl trimethyl ammonium bromide method (SALLAUD *et al.* 2003) from a diverse set of barley accessions (Table S3) using standard protocols and the full coding sequence of the barley *COM2* gene was PCR-amplified using primers **Ptp #56:** GCATGCATGTCACTCGAACT (upstream of ATG) and **Ptp #67:** CTAGGGCACCGAAACAAGCC (downstream of stop codon). PCR reactions (15 μL) contained 40-50 ng genomic DNA, 0.3 μM of each primer, dNPT mix 0.5 mM, DMSO 5%, herculase buffer 1x, Herculase Hotstart DNA polymerase 0.74 U (Stratagene #600310). Thermal cycling protocol consisted of: 1 cycle (95°/5'), 4 cycles (95°/20", 65°/10" [with 0.5°reduction per cycle], 72°/45"), 2 cycles (95°/20", 62°/20",72°/45"), followed by 34 cycles (95°/20", 60°/30",72°/1') and the last step of 1 cycle (72°/10'). PCR products were purified using the ExoSAP-IT (Exo-nucleases) PCR clean-up protocol (Applied Biosystems®) following the manufacturer's instructions and sequenced using the Sanger method with BigDye™ Terminator v3.1 Matrix Standard Sequencing Kit (Applied Biosystems®, ABI PRISM® 3700 DNA Analyzer) using primers **Ptp** #56 and **Ptp** #23 ACCAACTTCGTCTACGCGCA.

Sequence annotation

Unpublished sequence information for the two BAC contigs (44575 and 47813; spanning the interval between M1 and M2) was made available from the international barley sequencing consortium (through Dr. Nils Stein). First, barley BAC contig overlap detection was performed by an all against-all alignment with megablast (ZHANG *et al.* 2000). This was to confirm the overlap between the two BAC contigs initially identified. The criteria were only BLAST hits longer than 2 kb and 99.5% sequence identity. Sequence annotation was performed as described by (MASCHER *et al.* 2014). Sequences were first subjected to k-mer-based repeat masking using Kmasker algorithm (SCHMUTZER *et al.* 2014). *Augustus* was implemented for structural gene annotation of repeat-masked contigs using the maize model. Finally, predicted protein sequences were functionally annotated with the AHRD pipeline. This included parses the description of BLASTP hits against the TAIR, Uniprot/trEMBL, and Uniprot/SwissProt databases as utilized by (ZHANG *et al.* 2000). Genes annotated as unknown proteins or transposable elements were excluded from further analysis.