Article

Title:

COMPOSITUM 1 (COM1) contributes to the architectural simplification of barley inflorescence via cell wall-mediated and meristem identity signals

Authors:

N. Poursarebani[1], C. Trautewig[1], M. Melzer[1], T. Nussbaumer[3,4], U. Lundqvist[5], T. Rutten[1], T. Schmutzer[1,2], R. Brandt[1], A. Himmelbach[1], L. Altschmied[1], R. Koppolu[1], H. M. Youssef[1,6,7], N. Stein[1], Z. Xin[8], T. Schnurbusch[1,2]

Affiliations:

[1] Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Corrensstr. 3 OT Gatersleben, D-06466 Seeland, Germany

[2] Martin Luther University Halle-Wittenberg, Faculty of Natural Sciences III, Institute of Agricultural and Nutritional Sciences, 06120 Halle, Germany

[3] Technical University of Munich and Helmholtz Center Munich, Institute of Environmental Medicine, UNIKA-T, Neusäßer Str. 47, 86156 Augsburg, Germany

[4] Helmholtz Zentrum München (HMGU), German Research Center for Environmental Health, Institute of Network Biology (INET), 85764 Neuherberg, Germany

[5] Nordic Genetic Resource Center (NordGen), Smedjevägen 3, Box P.O. 41, SE-230 53 Alnarp, Sweden

[6] Molecular Cell Biology, Department of Biology, Lund University, Sölvegatan 35, 223 62 – Lund, Sweden,

[7]Faculty of Agriculture, Cairo University, Giza, Egypt

[8] USDA-ARS, Plant Stress and Germplasm Development Unit, Cropping Systems Research Laboratory, Lubbock, TX 79415, USA

*Corresponding authors. Email: Poursarebani@ipk-gatersleben.de (NP); Schnurbusch@ipk-gatersleben.de (TS)

Abstract:

Grasses have varying inflorescence shapes; however, little is known about the genetic mechanisms specifying such shapes among tribes. We identified the grass-specific TCP transcription factor COMPOSITUM 1 (COM1) expressed in inflorescence meristematic boundaries of different grasses. COM1 specifies branch-inhibition in *Triticeae* (barley) versus branch-formation in non-*Triticeae* grasses. Analyses of cell size, cell walls and transcripts revealed barley COM1 regulates cell growth, affecting cell wall properties and signaling specifically in meristematic boundaries to establish identity of adjacent meristems. *COM1* acts upstream of the boundary gene *Liguleless1* and confers meristem identity independent of the *COM2* pathway. Furthermore, COM1 is subject to purifying natural selection, thereby contributing to specification of the spike inflorescence shape. This meristem identity module has conceptual implications for both inflorescence evolution and molecular breeding in *Triticeae*.

Main Text:

The grass family (Poaceae), one of the largest angiosperm families, has evolved a striking diversity of inflorescence morphologies bearing complex structures such as branches and specialized spikelets [1]. These structural features are key for sorting the grass family into tribes [1]. Current grass inflorescences are proposed to originate from a primitive ancestral shape exhibiting "a relatively small panicle-like branching system made up of primary and secondary paracladia (branches), each one standing single at the nodes" [2] (**Fig. 1A**). This ancestral panicle-like inflorescence is also known as a compound spike [3-5]. Several independent or combined diversification processes throughout the evolutionary history of the grass family have resulted in the broad diversity of today's grass inflorescences [2,3,6]. Some tribes, e.g. *Oryzeae* (rice) and *Andropogoneae* (maize and sorghum), still display ancestral and complex compound shapes, keeping true-lateral long primary and secondary branches. Other grasses, such as *Brachypodium distachyon*, show lower inflorescence complexity with branch length and number reduced to lateral, small pedicels ending in only one multi-floretted spikelet (**Fig. 1A–C**). Inflorescences within the tribe *Triticeae*, e.g. barley (*Hordeum vulgare* L.), probably evolved from the ancestral compound spike into the typical unbranched spike (**Fig. 1D)**. The spike displays the least-complex inflorescence shape due to the sessile nature of spikelets and reduction in rachis internodes [2,7]. Architectural variation is often manifested through subtle modifications of transcriptional programs during critical transitional windows of inflorescence meristem (IM) maturation [7,8] or functional divergence of key transcriptional regulators and/or other genes [9,10]. Identification of key genetic determinants is crucial for better understanding and explaining both the origin of grass inflorescence diversity and grass developmental gene evolution. Inflorescence

developmental patterning controls pollination, grain set and grain number, and is thus highly relevant to agronomy as a target of natural and human selection.

Results

Atypical for *Triticeae*—barley *com1.a* mutant forms a branched inflorescence

To provide insight into the inflorescence architecture of *Triticeae*, we conducted a detailed phenotypic inspection of an induced barley mutant displaying non-canonical, i.e. branched, spike morphology. Barley (and other *Triticeae*) wild-type (Wt) unbranched spikes are typically composed of sessile, single-flowered spikelets arranged in a regular distichous fashion of two opposite rows directly attached to the main axis (**Fig. 1E**). The Compositum-Barley *com1.a* (*compositum 1.a*) is an induced mutant with a branched spike introgressed into the two-rowed *cv.* Bowman (BW) (**Supplementary Fig. 1**). The BW near isogenic line (NIL) of the *com1.a* allele, BW-NIL(*com1.a*), is a backcross (BC6)-derived EMS/neutron-induced mutant from *cv.* Foma [11]. The inflorescence in this mutant develops a ramified or branched architecture, resembling an ancestral compound spike (**Fig. 1E–I**), but lacks an organ called pulvinus. The pulvinus is present at the axil of lateral long branches in panicles and compound spikes of non-*Triticeae* grass species, defining branch angle extent. We observed differences in spike shape between BW and *com1.a* during early spike differentiation at the triple mound (TM) to early glume primordium (GP; in which the mutant central spikelet meristem (SM) is elongated; **Fig. 1J**) stage, becoming more apparent during later reproductive stages of late glume primordium onwards (**Fig. 1K–N**). At GP, predominantly in the basal part of the spike, meristems of the central spikelet positions undergo apparent floral reversion, displaying branch- or IM-like meristems (**Fig. 1N**). Instead of generating florets, the meristem potentially grows indefinitely and functions as an indeterminate spikelet multimer in the form of a primary branch-like structure (**Fig. 1M–N**). Such branch-like

structures occasionally replace other spikelet-related organs, such as the rachilla primordium (RP, the spikelet axis) or glumes (**Fig. 1M–N**). The *com1* branching phenotype resembles that of the previously described *compositum 2* mutant, *com2*, in which formation of branch-like structures results from lack of SM identity (in Supplementary Fig. 4 of [12] .

COM1 restricts palea cell size by thickening their cell walls

Besides the branch phenotype, *com1.a* exhibits a deviation in adaxial palea morphology, having a flat plane (**Fig. 1O**) versus the conventional distinct infolding observed in BW (**Fig. 1P),** *cv.* Foma, and wild barley (*H. vulgare* subsp. *spontaneum*). This deviation was visible in all paleae independent of their position along the spike. Histological analyses using cross sections of paleae middle-areas (**Fig. 1O**) revealed distinct features of *com1.a* in which sclerenchymtous cells, in particular, were expanded in size and numbers (**Fig. 1Q–R**). Cell expansion is thought to be tightly linked to cell wall extensibility [13,14]. We used transmission electron microscopy (TEM) to verify whether *com1.a* palea cells had altered cell wall features. Notably, mutant palea cells had clearly thinner cell wall structures, thus fewer mechanical obstructions for cell expansion, indicating that COM1 functions as a regulator of cell growth via cell wall modifications (**Fig. 1S–Z and Supplementary Fig. 2**). Moreover, mutant paleae generally formed three vascular bundles (VB) (**Fig. 1Q**) compared with two VBs in BW (**Fig. 1R**). By analogy to changes in palea cell walls, such alterations might also explain the rescission of SM identity, providing that COM1 similarly affects cell wall integrity in meristematic cells, e.g. SM cells or boundary cells (cells separating IM from SMs) (see below).

COM1 encodes a TCP transcription factor that inhibits inflorescence branching independent of *COM2*

To investigate the genetic basis of the *com1.a* phenotype, we constructed a genetic map by screening ~6,000 gametes for recombination events in an F_2 population (Bowman × *com1.a*) followed by further analysis of F_3 families (**Supplementary Fig. 1C–F, Supplementary Table 1, 2 and 3**) (**Supplementary Note**). This delimited a ~1.4 Mb interval carrying eight genes, one of which is a predicted transcription factor (HORVU 5Hr1G061270) entirely absent in *com1.a*, likely due to an induced deletion (**Fig. 2A**). The remaining seven genes either were not expressed or not differentially regulated between Wt and *com1.a* mutant (see below, the transcriptome analysis) (Fig1. A). To validate our candidate gene, we sequenced it in a barley TILLING population and in a set of 20 induced barley spike-branching mutants. We identified five branched mutants (*Mut.3906*, *int-h.42*, *int-h.43*, *int-h.44* and *com1.j*) missing the same transcription factor as *com1.a* (**Supplementary Fig. 3, Supplementary Table 1 and 4**) and six TILLING mutants with non-synonymous amino acid substitutions (**Fig. 2B; Supplementary Figs. 4–5; Supplementary Table 4**). *Mut.3906* was used for confirming allelism with *com1.a* (**Supplementary Note**). Together, these data confirmed unambiguously that the transcriptional regulator was responsible for the spike-branching phenotype in *com1.a*. Annotation analysis of the COM1 protein showed that it belongs to the plant-specific TCP (Teosinte branched1 (TB1)/Cycloidea/Proliferating Cell Factor) transcription factor family; COM1 contains 273 amino acids and features one basic helix-loop-helix TCP domain (**Fig. 2B**). Proteins of the TCP family fall into two classes, with COM1 belonging to class II, subclass CYC/TB1 [15,16].

We detected a higher phenotypic penetrance for spike-branching in *com1.a* (96.3%) compared to *com2.g* (78%). Double mutant (DM) plants outperformed either single mutant in grain number per spike, and supernumerary spikelet and branch production. Moreover, DM plants showed additional floral reversion in lateral spikelet positions (**Supplementary Fig. 6; Supplementary**

Note), further indicating that the two loci act independently/additively in branch inhibition in barley.

Barley COM1 function evolved under purifying natural selection and affects boundary signaling

We next asked whether COM1 has experienced functional conservation or divergence within the grasses and whether its sequence composition supports possible functional alteration. We used the comprehensive phylogenetic analyses available for grass TCPs [16] and the references therein) as a starting point for our own COM1-specific phylogenic analyses. We identified first-and second-best (closest) homologs (FBH and SBH) of COM1 from sequenced grass genomes, including rice, maize, sorghum, hexaploid wheat and *Brachypodium distachyon*, as well as *Arabidopsis thaliana* (**Fig. 2C**). The homolog of maize TB1, obtained from the aforementioned grasses, was added as an out-group to the phylogeny. Our phylogenetic analysis confirmed that COM1 is restricted to grasses (**Fig. 2C**) [17-19]. The FBHs of COM1 in maize and rice were reported previously as *ZmBAD1/WAB1* and *OsREP1/DBOP* (60.3% and 65.5% sequence similarity to COM1), respectively [17-20]. Except for maize, none of the COM1 FBHs showed a duplicated copy (no in-paralogs resulting from within-genome duplication after a speciation, [21] (**Supplementary Fig. 7A–B**)). Instead, COM1 seems to be an out-paralog [21] of SBHs including the sorghum gene *SbMSD1* (44.1% sequence similarity to COM1) [22]**.** Functional characterization of COM1 homologs is only available for maize and rice (Table 1) [17-19].

Maize *BAD1/WAB1* transcripts are mainly detected at the IM-to-BM (branch meristem) boundary region as well as between pulvinus and lateral branches (in Fig. 3J of [18]). Consequently, loss-of-function *bad1/wab1* mutants display organ fusion (a known boundary formation defect) resulting in reduced branch number (from 5.8 in Wt to 1.3 in mutant siblings) and angle size, and more

upright tassel branches [17,18]. This gene was dubbed a boundary formation gene promoting lateral meristem (e.g. branch) and axillary organ (e.g. pulvinus) formation in Wt maize [17,18].

Our TILLING analysis for the *BAD1/WAB1* ortholog in sorghum revealed one mutant (ARS180 line; A144T) with both upright tassel branches (10.95° in Wt vs. 5.2° in mutant, $P \leq 0.001$; **Fig. 3, Supplementary Table 5**) and reduced primary branch number per node (5.4 in Wt vs. 4.2 in mutant, $P \leq 0.05$; **Supplementary Table 5**). These data suggest a similar positive role of sorghum *BAD1/WAB1* in pulvinus development and branch initiation/formation, revealing functional conservation of the protein between sorghum and maize. Moreover, we detected no obvious change in sorghum palea morphology except one additional vascular bundle, similar to maize (**Table 1**).

The rice homolog of *COM1*, *OsREP1/DBOP*, shows a major effect in promoting palea identity, growth and development, with no effect on branch angle or branch initiation [19,20]. Loss-of-function mutants display smaller paleae due to less differentiation and severely reduced size of palea cells; a clear contrast to palea defects in barley (**Table 1**). Our TILLING analysis of COM1 homologs in *Brachypodium distachyon* (*Bd*) identified several mutants. Phenotypic investigation of two lines (5446: Q116* and 8373: S146N) (**Supplementary Table 4, Supplementary Note**) revealed similar phenotypes to the aforementioned non-*Triticeae* species (**Table 1**) (**Fig. 3F–P**). Similarly, we observed a palea defect (**Fig. 3G**) but histological analyses revealed no changes in cell expansion, except the formation of one additional vascular bundle in each mutant (**Fig. 3L–M**). We also observed a reduction in branch angle because of smaller or absent pulvini; however, the number of lateral branches was not altered in either *Brachypodium* mutants (**Fig. 3O–Q**). In conclusion, COM1 homologs within non-*Triticeae* grasses primarily promote boundary formation and cell differentiation (as in rice palea)/proliferation (as seen for pulvinus) (**Table 1**)

but also contribute to formation of lateral axillary organs, e.g. branch or pulvinus, creating more complex inflorescence structures.

To better understand the contrasting COM1 function of branch-inhibition in barley versus branch-formation in non-*Triticeae* grasses, we analyzed barley *COM1* expression using qRT-PCR and semi-qPCR (Fig. 4A-C) followed by mRNA *in-situ* hybridization (**Fig. 4D-G**). Barley *COM1* transcripts were detected in paleae (**Fig. 4C, F-G**), VB of the rachis (Fig. 4E), and importantly at the base of forming SMs throughout the boundary region separating SMs from IM (IM-to-SM boundary) and between lateral and central SMs (**Fig. 4E-F**), similar to non-*Triticeae* grass species, e.g. maize. This expression pattern suggests involvement of barley *COM1* in specification of the spikelet meristematic boundary. However, since central and lateral spikelets do not fuse into each other or to the IM (as long branches do in maize or sorghum), barley COM1 may not be involved in boundary formation *per se*. In combination with our cell wall analysis in palea cells, this implies that barley COM1 may be involved in formation of meristem identity signals released from the boundary region through thickened cell walls encompassing boundary cells; thus, COM1 affects boundary signaling via cell wall modifications [23]. Recently acquired protein motifs specific to *Triticeae* COM1 may support this functional modification (**Fig. 2D** Motifs 7, 13, 15 and 17 **and Supplementary Fig. 8**).

We checked whether natural selection has acted upon barley COM1 sequence composition and function, and consequently formation of unbranched spikes in barley. Re-sequencing of the barley *COM1* coding sequence in a panel of 146 diverse barley landraces and 90 wild barleys [24,25] (**Supplementary Table 6**) revealed very little natural sequence variation (site diversity of pi = 0.0006). Eleven SNPs resulted in a simple 12-haplotype network (**Supplementary Fig. 9**) comprising only two main haplotypes, neither of the 12 showed mutant spike or palea phenotypes

(**Supplementary Fig. 9**). This suggests that barley COM1 underwent purifying natural selection most likely for maintaining inflorescence shape.

Putative transcriptional regulation for barley spike

To further examine the molecular basis of COM1 branch inhibition within the barley spike, we performed qRT-PCR to locate *COM1* relative to other previously known spike architecture genes (**Fig. 5A**, black arrows). We localized *COM1* downstream of *VRS4 (HvRA2*; orthologous to maize *RAMOSA2*), the main regulator of row type and branch inhibition [7,12] (**Supplementary Fig. 10**). *COM2* transcript levels in immature spikes of *com1.a* were slightly lower only during later stages of development (**Supplementary Fig. 10F**).

We performed comparative RNA-seq using mRNAs from immature spikes of BW and *com1.a* as well as the mutant progenitor, *cv.* Foma, when spike patterning begins to differ between genotypes, plus two subsequent stages (**Figs. 1 and 5B; Online Materials**). Differentially expressed (DE) genes were identified in comparisons of *com1.a* versus BW and mutant versus *cv.* Foma. We found 83 genes (Log2 FoldChanges; LFC $|\geq 0.5$; adjusted $P < 0.05$) DE in at least one stage in both comparisons (**Fig. 5; Supplementary Figs. 11–12; Supplementary Source Data 1**): 18 and 65 genes up- and downregulated in BW-NIL(*com1.a*), respectively.

Among significantly downregulated genes across all three stages (**Fig. 5B**), we detected one *SQUAMOSA PROMOTER-BINDING-LIKE 8* gene (*SPL8*, HORVU2Hr1G111620) homologous to the boundary gene *LIGULELESS 1* in maize (*LG1*; *ZmSPL4*), rice *OsLG1* (*OsSPL8*) and hexaploid wheat *TaLG1* (*TaSPL8*) [26]. Similar to the known maize module (*RA2*→*WAB1*/*BAD1*→*LG1*; [17,18], we found that *VRS4*/*HvRA2*→*COM1*→*HvLG1* regulation appears to be maintained in barley. Transcriptome analysis of leaf tissues in a wheat *liguleless1*

mutant revealed *TaSPL8* as a cell wall-related gene [26]. Notably, no spike-branching phenotype was reported for this erected-leaf *liguleless* mutant, most likely due to genetic redundancy.

Among other significantly downregulated genes in *com1.a*, we found important genes associated with cell wall properties and integrity (**Fig. 5D**). These include HORVU5Hr1G006430, a leucine-rich repeat receptor kinase (LRR-RLK), and HORVU3Hr1G030260 belonging to the cytochrome P450 superfamily. LRR-RLKs and CYP450s are involved in lignin deposition to cell walls upon cellulose biosynthesis inhibition and during lignin biosynthesis *per se*, respectively [27,28]. Other cell wall-related genes include two genes encoding xyloglucan endotransglucosylase/hydrolase (XTH) 25 (HORVU7Hr1G098280 and HORVU7Hr1G098260) and barley *Low Silicon Influx 1* (*HvLSI1*; *HORVU6Hr1G075850*) [29], both downregulated in the mutant. These cell wall-related genes may support *COM1* involvement in regulation of cell wall mechanics of palea cells and the IM-to-SM boundary, and indirectly, putative signaling required for acquiring SM identity.

Discussion

Here we report that barley COM1 affects cell growth through regulation of cell wall properties specifically in palea and IM-to-SM boundary cells; the latter provide identity signals to barley SMs [30]. Signaling to the SM to establish its identity is a key genetic switch by which barley inflorescences acquire spike architecture, not seen in non-*Triticeae* grasses.

COM1 is present only in grasses, with no true *Arabidopsis* ortholog; intriguingly, we observed functional modification of COM1 between barley and non-*Triticeae* grass species. The modification in COM1 function was clear by comparing mutant versus wild type inflorescence phenotypes across grass species, and was further elucidated by our analysis at the cellular/molecular level. At the phenotypic level, barley COM1 inhibits spike-branching to simplify floral architecture; whereas in non-*Triticeae* COM1 homologs promote formation of

lateral branches (e.g. up to 60% more branches in maize when compared to mutants [18]) to sustain the ancestral inflorescence complexity.

At the cellular level in non-*Triticeae* grasses, COM1 has evolved as a boundary formation factor, its putative ancestral role (**Fig. 5C-D**). Consequently, loss-of-function of COM1 homologs result in lack of boundaries and subsequent organ fusion, e.g. BM into IM as demonstrated by a low number of lateral branches in maize mutants. Notably, this loss of function did not change the overall inflorescence architecture in non-*Triticeae* grasses. Barley *com1* loss-of-function, however, increases branch formation/extension mostly from SMs, a clear deviation from the canonical spike form. As barley *COM1* displayed a similar boundary mRNA expression as seen in maize, we presume that barley COM1 functions through boundary signaling [30], thereby affecting the identity of adjacent SMs. The formation of boundary regions in barley *com1* mutants (no organ fusion) via pathway(s) independent of *COM1* (**Fig. 5E-F**), and thus separation of meristematic zones in this mutant, implies that barley IM-to-SM boundary cells fail to deliver proper identity-defining signals to SMs. This signaling failure may perturb transcriptional programs required to establish identity in barley SMs; such meristems eventually revert back to IM-like meristems forming a branch-like structure (**Fig. 5F)**. The function of the boundary, and boundary-expressed genes (e.g., maize *RAMOSA1–3*), as a signaling center for adjacent meristems, e.g. SMs, has been proposed in grasses, yet features of these signals remain unknown [30]. Signals associated with COM1 might include micromechanical forces derived from formation of rigid cell walls enclosing boundary cells. Involvement of COM1 in printing such mechanical regulation is supported by our anatomical analysis of palea cell walls and further confirmed by our transcriptome analysis of immature barley spike samples. *HvLG1*, *HvLSI* and genes encoding one LRR-RLK, one CYP450 and two XTHs were among the most downregulated in the mutant and involved in defining cell wall

properties [26-28,31]. The contribution of boundary cell wall mechanics in guiding organogenesis within reproductive tissues has been well described in dicot species [32] .

Such functional modification usually includes constraints on expression patterns, protein sequence/structure or participation in molecular networks, often assumed to be associated with gene duplication [21]. Notably, *COM1* shows no sign of duplication within the barley genome and as mentioned above displays a similar expression pattern to maize [17,18]. Thus, COM1's functional modification and implication in boundary-derived signaling seem to be associated with its protein sequence (**Fig. 2D**) and the respective downstream molecular networks. Furthermore, COM1's role in regulating floral complexity-levels in grasses fits well with the view that TCP transcription factors are growth regulators and evolutionary architects of plant forms that create diversity [33]. They influence the final architecture of plants in response to endogenous and/or external conditions. Thus, the barley floral reductionism (from compound spike to spike form; **Fig. 1A-D**) contributed by COM1, might be a response to the ecological expansion of the *Triticeae* into more temperate climates [3].

In summary, our findings enabled identification of a barley SM identity module, *VRS4* (*HvRA2*) → *COM1* → *HvLG1*, which works independently of *COM2* and inhibits spike-branching via boundary-defined signals (**Fig. 5A and Supplementary Fig. 12**). Our model of branch-inhibition in barley spikes opens a new window into grass inflorescence evolution and molecular crop breeding, and the elevated grain number per spike in *com1.a/com2.g* double mutants supports this notion.

Methods

Barley Plant material

The Nordic Genetic Resource center, the National Small Grains Collection (US Department of Agriculture), and the IPK gene bank were inquired to access 'Compositum-Barley' mutants (Supplementary Table 4). Bowman near isogenic line carrying *com1.a* allele ((i.e., BW-NIL(*com1.a*); syn. BW189 or CIho 11333)), its two-rowed progenitor Foma and Wt barley cv. Bowman were used for phenotypic descriptions, whole genome shotgun sequencing (WGS) (see below) as well as SEM analysis. Plant material used to generate mapping populations is reported in the corresponding section for genetic mapping. For haplotype analysis, a core collection including of 146 diverse barley landraces and 90 diverse wild barleys were sequenced [24,25] (Supplementary Table 6).

Plant phenotyping

Barley; For phenotyping the mapping population, barley BW-NIL(*com1.a*) , Bowman and the corresponding segregating populations (F_2 and F_3) were grown side by side under greenhouse conditions at the IPK. For a plant to be assigned as a branched spike mutant, spike shape at all tillers was visually inspected for presence of at least one extra spikelet at any rachis node. Grain related characters such as weight, number, etc. were also measured at harvest for the two parental lines of the mapping population. In case of phenotyping of the barley TILLING population (see below and the Supplementary Table 4), other induced mutants (Supplementary Table 4) as well as the BW-NIL(*com1.a*) / BW-NIL(*com2.g*) double mutants (see below), visual phenotyping for variation in palea structure was also applied in addition to the aforementioned phenotyping approach used for spike branching in F_2 and F_3 progenies. In case of TILLING, from the six mutants for which the spike-branching phenotype was observed at M4, only three (carrying mutation inside the protein domain; M4.15104, M4.4406, and M4. 2598) were subjected for

further study at M5 generation. For which, one M4 plant was selected from which 16 M5 plants were grown and phenotyped.

Brachypodium distachyon: An already published TILLING population and the corresponding Wt accession Bd21-3 were used for phenotyping [34]. That included measurement of branch angle, as proxy for pulvinus size, spikelet number per spike, floret number per spikelet and palea structure. Hence, per M4 plants, only homozygous M5 plants either with mutant genotype aa (3 to 4 plants) or wild type bb (3 to 4 plants) were selected. Per M5 plants, 10 M6 plants were grown under greenhouse conditions at the IPK and used for measurement. Thus, 30 to 40 plants per group and for each plant angles of basal spikelets in main tillers were considered for measurement. To this end, spikes were first imaged and then imported to the ImageJ tool (https://imagej.nih.gov/ij/index.html) for angle measurement. In case of original wild type Bd21-3, five plants were grown and measured. The same set of plants and the corresponding spike images were used to calculate number of spikelets per spike and number of floret per spikelet. In case of palea phenotyping: paleae were visually inspected across all spikes per plant. We detected plants with paleae being sensitive to exogenous finger-pressure, and thus such plants were scored as mutants. A gentle finger-pressure led the mutant paleae to crash from the middle longitude-line so that a scissors-like structure was formed (Fig. 3G). The crashing was not evident in Wt plants even with severe exogenous hand-pressure.

Sorghum: An already published TILLING population and the corresponding Wt accession BTx623 were used for phenotyping [35]. To measure primary branch number and angle, 5 to 8 plants, either M5 or M6 generations, per family including a Wt sorghum family cv. BTx623 were grown under greenhouse conditions at the IPK. Average branch number per panicle, e.g. per

plant, was calculated by counting all branches that originated per each rachis node (Supplementary Table 5). The Average branch number per family was then used to compare with the same value obtained from Wt family BTx623. To measure the branch angle, for each plant 3 to 4 basal nodes per panicle were separately photographed. Each node contained at least 1 and up to 5 lateral branches. To cover angles of each individual branch per node, each node was photographed multiple time. Images were then imported to ImageJ for angle measurement as described for *Brachypodium* (see above). Spikelet organs of palea and glume as well as overall grain set were also visually inspected for any visible alteration.

Marker development

Bowman near isogenic line BW-NIL(*com1.a*) and two-rowed progenitor of *com1.a,* cv. Foma, were survey sequenced using WGS approach (see below). These sequence information were compared against already available WGS of Bowmann [36], as present in **Supplementary Fig. 1**. Polymorphisms e.g. SNPs detected from this comparison (named as Next Generation Sequencing based markers (NGS-based markers)) between the two parental lines were converted to restriction enzyme based CAPS (http://nc2.neb.com/NEBcutter2/) markers to derive a restriction based genetic marker as previously described [12]. The developed genetic markers (**Supplementary Table 1**) were used to screen the corresponding mapping population.

Genetic mapping and map-based cloning of *com1.a*

com1.a was initially proposed to be located in chromosome 5HL with unknown genetic position [11]. A barley F_2 mapping population was developed by crossing Bowman introgression

line BW-NIL(*com1.a*) and barley cv. Bowman. For initial mapping 180 individuals were analyzed and genotyped using the aforementioned NGS based markers. The pattern of segregation between mutant and Wt F_2 plants fitted a 3:1 ratio typical for a monogenic recessive gene. Linkage analysis of segregation data was carried out using maximum likelihood algorithm of Joinmap 4.0. Kosambi mapping function was used to convert recombination fractions into map distances. The linkage mapping was further followed by a high-resolution genetic mapping in which almost 6,000 gametes were screened with the flanking markers NGS045 and NGS049. For narrowing down the *com1.a* genetic interval; the identified recombinants (a set of 109) were used. From 109, a set 15 F_2 were labeled (**Supplementary Table 2-3**) to be critical recombinants for precisely defining the *com1.a* genetic interval. From each of the 15 critical plants, 16 F_3 progenies were evaluated for their phenotypes and marker genotypes at the *com1.a* candidate gene. (**Supplementary Table 2 and S3**). Based on F_2 high-resolution mapping and F_3 genetic analysis described, two tightly linked markers, NGS084 and NGS094, were taken to harvest the available barley genome BAC sequence data (data not shown). A single BAC contig spanning 1.4 Mb of the minimal tiling path (MTP) was identified. Genes in this region were utilized for marker development and further genetic mapping that resulted in identification of a ~380 kb region deleted in the mutant BW-NIL(*com1.a*). The deleted fragment contains a single gene, i.e., *com1.a.*

Allelism test of *com1* mutants.

Mut.3906 mutant (**Supplementary Table 4**) was crossed with BW-NIL(*com1.a*) to test for allelism. The resultant F_1 plants showed a mutant spike phenotype confirming its allelism with

com1. All alleles showed phenotypic similarities with *com1* and mutations in the *COM1* gene sequences.

Double-mutant analysis

Double mutants (DM) were generated by crossing mutant BW-NIL(*com1.a*) to BW-NIL(*com2.g*), followed by selfing of the F_1 progeny. All obtained 183 F_2 plants were subsequently genotyped (**Supplementary Table 1**). In case of *com2.g* mutation detection, a primer pair (Com2-Bw_SfiI_FR; **Supplementary Table 1**) spanning the A300C haplotype (that differentiate the Wt Bowman allele A from *com2*.g mutant C allele at position 300bp [12] were used for sequencing and to classify F_2 genotypes for the *com2* locus. Thus, genotypic classes include C300C allele as homozygous mutant, AA as Wt and CA as heterozygous. In case of *com1.a*, a presence/absence marker was used (**Supplementary Table 1**), where absence of the *COM1* gene was considered as homozygous *com1.a* mutant. A total number of five plants were recovered as homozygous double mutants (from 183 F_2 plants) (Supplementary Note) that were used for generating F_3 plants used in subsequent DM phenotypic analysis (**Supplementary Fig. 6**). Two DM F_3 families, each consisting of 20 plants along with 20 plants from each of the single mutants and 20 wild type cv. Bowman plants, were grown and used for phenotyping (**Supplementary Fig. 5 and 6**).

TILLING analysis

Barley: For identifying further mutant alleles of *COM1* in barley TILLING populations including EMS (Ethyl methanesulfonate) treated population of cv. Barke consisting 10279

individuals were screened [37]. A primer combination (**Supplementary Table 1**) was used to amplify the coding region of the *COM1* gene. The amplicon was subjected to standard procedures using the AdvanCETM TILLING kit as described in [12]. 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 AdvanCETM FS96 system according to manufacturer's guidelines (advancedanalytical, IA, USA). The amplified ORF was also re-sequenced by Sanger sequencing using primers listed in **Supplementary Table 1**.

Brachypodium distachyon: Mutation detection screenings were performed in the TILLING collection of chemically induced *Brachypodium* mutants, described in [34]. TILLING by NGS consists to sequence 500 bp PCR fragments libraries prepared from 2600 individual genomic DNA pooled in two dimensions. A dual indexing system, one placed on the 5'adaptater, and the second one on the 3'adaptater, added by a two-step PCR (for primer sequence; see **Supplementary Table 1**) allow a direct identification of the sequence identities. The first PCR amplification is a standard PCR with target-specific primers carrying Illumina's tail (**Supplementary Table 1**) and 10 ng of *Brachypodium* genomic DNA. Two microliter of the first PCR product served as a template for the second PCR amplification, with a combination of Illumina indexed primers (**Supplementary Table 1**). The sequencing step of PCR fragments was done on an Illumina Miseq personal sequencer using the MiSeq Reagent Kit v3 (Illumina®) followed by quality control processes for libraries using the PippinHT system from SAGE Sciences for libraries purification, and the Bioanalyzer™ system from Agilent®. To identify induced mutations, a bioinformatic pipeline, called "Sentinel" was used to analyze the data sequences (IDDN.FR.001.240004.000.R.P.2016.000.10000). Prediction of the impact of each mutation (**Supplementary Table 4**) was made with SIFT software as described in in [34]. The

amplified ORF was also re-sequenced by Sanger sequencing using primers listed in **Supplementary Table 1**.

Sorghum: A pedigreed sorghum mutant library was established in the inbred line BTx623, which was used to produce the sorghum reference genome. This mutant library consists of 6400 M4 seed pools derived from EMS-treated sorghum seeds by single seed descent. Whole genome sequencing of a set of 256 lines uncovered 1.8 million canonical EMS-induced mutations [34]. We searched the sorghum ortholog of the barley *COM1* in the aforementioned sequence database to identity plants carrying mutation. To confirm the mutations, the amplified ORF was also re-sequenced by Sanger sequencing using primers listed in Supplementary Table 1.

Haplotype and network analysis

Genomic DNA from a core collection including 146 landrace and intermedium barley accessions as well as 90 wild barley (Supplementary Table 6) was PCR-amplified using specific primers to amplify full coding sequence of the barley *COM1* gene. Amplified fragments were used for direct PCR sequencing (Sanger method; BigDye Terminator v3.1 cycle sequencing kit; Applied Biosystems). A capillary-based ABI3730xl sequencing system (Applied Biosystems) at the sequencing facility of IPK was used to separate the fluorescently terminated extension products. Sequence assembly was performed using Sequencher 5.2.2.3. Visual inspection of sequence chromatograms was carried out using Sequencher to detect the corresponding SNPs. Network analysis of the nucleotide haplotypes was carried out using TCS v1.21 software (http://darwin.uvigo.es/software/tcs.html) [38].

RNA extraction, sequencing and data analysis

RNA Extraction; For the RNA-seq study, immature spike tissues were collected from BW-NIL(*com1.a*) and WT progenitor Bowman and the donor cultivar Foma. Plants were grown under phytochamber conditions of 12h light (12 ºC) and 12h dark (8 ºC). Tissues were always collected at the same time slot (14:00 to 17:00) during the day at three different developmental stages including TM and GP, and pooled stages of LP+SP. Three biological replicated were applied that resulted in 27 individual tissue samples. The TRIzol method (Invitrogen) was applied to extract total RNA from immature spike tissues followed by removal of genomic DNA contamination using RNAse-free DNAse (Invitrogen). RNA integrity and quantities were analyzed via Agilent 2100 Bioanalyzer (Agilent Technologies) and Qubit (Invitrogen), respectively.

Preparation and sequencing of mRNA-Seq libraries: SENSE mRNA-Seq libraries (27 = 3 reps/3 stages /3 genotype) were prepared from 2 µg total RNA according to the protocol provided by the manufacturer (Lexogen GmbH, Vienna, Austria). Libraries were pooled in an equimolar manner and analysed electrophoretially using the Agilent 4200 TapeStation System (Agilent Technologies, Inc., Santa Clara, CA, USA). Quantification of libraries and sequencing (rapid run, paired-end sequencing, 2 x 100 cycles, on-board clustering) using the Illumina HiSeq2500 device (Illumina, San Diego, California, USA) were as described previously [39].

Analysis of the RNAseq data:

The reads from all three biological replicates were pooled per stage and each pool was independently mapped to barley pseudomolecules [36], (160404_barley_pseudomolecules_masked.fasta) using TopHAT2 [40]. Gene expression was estimated as read counts for each gene locus with the help of featureCounts [41] using the gene

annotation file Hv_IBSC_PGSB_r1_HighConf.gtf and fragment per million (FPM) values were extracted from the BWA-aligned reads using Salmon [42]. Genes that showed FPM of 0 across all 45 samples were excluded from expression levels calculations. Expression levels were normalized by TMM method and *p*-values were calculated by an exact negative binomial test along with the gene-specific variations estimated by empirical Bayes method in edgeR [43]. The Benjamini-Hochberg method was applied on the *p*-values to calculate *q*-values and to control the false discovery rate (FDR). Differentially expressed genes (DEGs) were defined as *q*-value < 0.05, log2 fold change > 1 or < -1.

Quantitative RT-PCR

Tissue sampling, RNA extraction, qualification and quantification was performed as described above. Reverse transcription and cDNA synthesis were carried out using SuperScript III Reverse Transcriptase kit (Invitrogen). Real-time PCR was performed using QuantiTect SYBR green PCR kit (Qiagen) and the ABI prism 7900HT sequence detection system (Applied Biosystems). Each qRT-PCR comprised at least four technical replicates, and each sample was represented by three biological replicates. The *Actin* gene based primers (**Supplementary Table 1**) were used as the reference sequence. qRT-PCR results were analyzed using SDS2.2 tool (Applied Biosystems) in which the presence of a unique PCR product was verified by dissociation analysis. Significance values were calculated using Student's *t*-test (two-tailed). The relevant primer sequences per species are detailed in **Supplementary Table 1**.

Phylogenetic analysis

A comprehensive analysis of TCP proteins in grasses was already available we therefore focused only on constructing a detailed phylogeny of the COM1 protein among grasses and the barley TCP genes. Thus, barley COM1 was then queried against Ensembl Plants database to retrieve its orthologs or homologs from other grasses. The same database was also used to extract all barley TCP proteins. In case of COM1, protein and DNA sequence of the first and the second best hit (FBHs and SBHs) to each of the grass species were retrieved. To re-check their homology with barley COM1, the retrieved sequences were blasted back against the barley genome. For phylogenetic analysis, protein sequences were initially aligned using the algorithm implemented in CLC sequence viewer V7.8.1 (https://www.qiagenbioinformatics.com). UPGMA tree construction method and the distance measure of Jukes-Cantor were implemented for constructing the phylogenetic tree using CLC sequence viewer. The bootstrap consensus tree inferred from 1000 replicates was taken to represent the evolutionary relationship of the sequences analyzed.

mRNA in situ hybridization

Three separated segments (excluding the TCP domain) from the *COM1* gene each containing 300-360 bp were synthesized (probe 1 and 2, GenScript Biotech , Netherlands) or amplified (probe 3) using cDNAs isolated from immature spikes of cv. Bonus and specific primers (Supplementary Table 1). The resulting products were cloned into pBluescript II KS (+) vector (Stratagene, La Jolla, CA, USA and GenScript Biotech , Netherlands). Linearized clones by HindIII or NotI were used as templates to generate antisense (HindIII) and sense (NotI) probes using T3 or T7 RNA polymerase. *In situ* hybridization was conducted with a single pool of the three aforementioned probes as described previously [44].

Scanning electron microscopy

Scanning electron microscopy (SEM) was performed on immature spike tissues at five stages including triple mound, glume, lemma, stamen, and awn primordium from greenhouse-grown plants. SEM was conducted as described elsewhere [45].

DNA preparation

DNA was extracted from leaf samples at the seedling as described before [12]. Plants for which the DNA was prepared included barley, *Sorghum* and *Barchypodium*. That included either mapping population, TILLING mutants or both.

Palea anatomical and transmission electron microscopy analyses.

Plant material consisting of intact spikes shortly before anthesis was collected. Spikelets containing no grains were used for dissecting paleae that were subsequently stored in fixative (4% FA, 1% GA in 50 mM phosphate buffer). Central spikelets (in case of barley) were isolated and placed in a 15 ml test tube containing 10 ml fixative, followed by extensive degassing until all probes had settled. Material was stored in a fridge until use. After three washes with A.D., lemma and palea were isolated by cutting away a small part at the base of the spikelet. Isolated paleae were placed in a flat bottomed mould filled with 4% liquid agarose (~60°C). After setting, agarose blocks were removed from the mold and the encapsuled Palea was cut into 1-2mm wide sections using fresh razor blades. The embedding in agarose facilitated the cuttings while preventing unnecessary damage to the probes. After embedding in Spurr resin (see next page) semithin sections of 2 µm were cut on an Leica Ultracut. Sections were allowed to be baked in on a droplet of 0,02% Methylene blue/Azur blue on a heating plate set at 90°C. Recordings were

made using a Keyence VHX-5000 digital microscope (Keyence Germany GmbH, Neu-Isenburg, Germany).

Sequence information and analysis

Unpublished sequence information for the BAC contigs 44150 spanning the interval between NGS084 and NGS094) was made available from the international barley sequencing consortium (through Nils Stein). This sequence information was used for marker development during high resolution mapping, map-based cloning and *COM1* gene identification. Later on, the initial contigs 44150 sequence information was re- checked and confirmed with the higher-quality barley genome assembly and annotation data [23].

Whole genome shotgun sequencing of BW-NIL (*com1.a)*

A whole-genome shotgun library was constructed using standard procedures (TruSeq DNA; Illumina) and quantified using real-time PCR. Cluster formation using the cBot device and paired-end sequencing (HiSeq2000, 2 x 101 cycles) were performed according to the manufacturer's instructions (Illumina).

References

1. Zhang, D.B. & Yuan, Z. Molecular Control of Grass Inflorescence Development. *Annual Review of Plant Biology, Vol 65* **65**, 553-78 (2014).

2. Vegetti, A. & Anton, A.M. Some Evolution Trends in the Inflorescence of Poaceae. in *Flora* Vol. 190 225-228 (1995).

3. Kellogg, E.A. *et al.* Early inflorescence development in the grasses (Poaceae). *Frontiers in Plant Science* **4** (2013).

4. Endress, P.K. Disentangling confusions in inflorescence morphology: Patterns and diversity of reproductive shoot ramification in angiosperms. *Journal of Systematics and Evolution* **48**, 225-239 (2010).

5. Remizowa, M.V., Rudall, P.J., Choob, V.V. & Sokoloff, D.D. Racemose inflorescences of monocots: structural and morphogenetic interaction at the flower/inflorescence level. *Annals of Botany* **112**, 1553-1566 (2013).

6. Kellogg, E.A. Evolutionary history of the grasses. *Plant Physiology* **125**, 1198-1205 (2001).

618 7. Koppolu, R. & Schnurbusch, T. Developmental pathways for shaping spike inflorescence
619 architecture in barley and wheat. *J Integr Plant Biol* **61**, 278-295 (2019).
620
621 8. Lemmon, Z.H. *et al.* The evolution of inflorescence diversity in the nightshades and
622 heterochrony during meristem maturation. *Genome Research* **26**, 1676-1686 (2016).
623
624 9. Malcomber, S.T., Preston, J.C., Reinheimer, R., Kossuth, J. & Kellogg, E.A. Developmental
625 gene evolution and the origin of grass inflorescence diversity. *Advances in Botanical*
626 *Research: Incorporating Advances in Plant Pathology, Vol 44* **44**, 425-481 (2006).
627
628 10. Vollbrecht, E., Springer, P.S., Goh, L., Buckler, E.S. & Martienssen, R. Architecture of
629 floral branch systems in maize and related grasses. *Nature* **436**, 1119-1126 (2005).
630
631 11. Druka, A. *et al.* Genetic dissection of barley morphology and development. *Plant Physiol*
632 **155**, 617-27 (2011).
633
634 12. Poursarebani, N. *et al.* The Genetic Basis of Composite Spike Form in Barley and 'Miracle-
635 Wheat'. *Genetics* **201**, 155-165 (2015).
636
637 13. Cosgrove, D. Biophysical Control of Plant-Cell Growth. *Annual Review of Plant*
638 *Physiology and Plant Molecular Biology* **37**, 377-405 (1986).
639
640 14. Cosgrove, D.J. Plant cell wall extensibility: connecting plant cell growth with cell wall
641 structure, mechanics, and the action of wall-modifying enzymes. *Journal of Experimental*
642 *Botany* **67**, 463-476 (2016).
643 15. Martin-Trillo, M. & Cubas, P. TCP genes: a family snapshot ten years later. *Trends Plant*
644 *Sci* **15**, 31-9 (2010).
645
646 16. Zhao, J. *et al.* Genome-Wide Identification and Expression Profiling of the TCP Family
647 Genes in Spike and Grain Development of Wheat (Triticum aestivum L.). *Frontiers in*
648 *Plant Science* **9** (2018).
649
650 17. Lewis, M.W. *et al.* Gene regulatory interactions at lateral organ boundaries in maize.
651 *Development* **141**, 4590-4597 (2014).
652
653 18. Bai, F., Reinheimer, R., Durantini, D., Kellogg, E.A. & Schmidt, R.J. TCP transcription
654 factor, BRANCH ANGLE DEFECTIVE 1 (BAD1), is required for normal tassel branch
655 angle formation in maize. *Proceedings of the National Academy of Sciences of the United*
656 *States of America* **109**, 12225-12230 (2012).
657

658 19. Zeng, D.-D. *et al.* DBOP specifies palea development by suppressing the expansion of the
659 margin of palea in rice. *Genes & Genomics* **38**, 1095-1103 (2016).
660
661 20. Yuan, Z. *et al.* RETARDED PALEA1 Controls Palea Development and Floral
662 Zygomorphy in Rice. *Plant Physiology* **149**, 235-244 (2009).
663
664 21. Studer, R.A. & Robinson-Rechavi, M. How confident can we be that orthologs are similar,
665 but paralogs differ? *Trends in Genetics* **25**, 210-216 (2009).
666
667 22. Jiao, Y.P. *et al.* MSD1 regulates pedicellate spikelet fertility in sorghum through the
668 jasmonic acid pathway. *Nature Communications* **9** (2018).
669
670 23. Aguilar-Martinez, J.A., Poza-Carrion, C. & Cubas, P. Arabidopsis BRANCHED1 acts as
671 an integrator of branching signals within axillary buds. *Plant Cell* **19**, 458-472 (2007).
672
673 24. Koppolu, R. *et al. Six-rowed spike4* (*Vrs4*) controls spikelet determinacy and row-type in
674 barley. *Proceedings of the National Academy of Sciences of the United States of America*
675 **110**, 13198-13203 (2013).
676
677 25. Russell, J. *et al.* Exome sequencing of geographically diverse barley landraces and wild
678 relatives gives insights into environmental adaptation. *Nature Genetics* **48**, 1024-1030
679 (2016).
680
681 26. Liu, K. *et al.* Wheat TaSPL8 Modulates Leaf Angle Through Auxin and Brassinosteroid
682 Signaling. *Plant Physiology* **181**, 179-194 (2019).
683
684 27. Van der Does, D. *et al.* The Arabidopsis leucine-rich repeat receptor kinase MIK2/LRR-
685 KISS connects cell wall integrity sensing, root growth and response to abiotic and biotic
686 stresses. *Plos Genetics* **13** (2017).
687
688 28. Gou, M.Y., Ran, X.Z., Martin, D.W. & Liu, C.J. The scaffold proteins of lignin biosynthetic
689 cytochrome P450 enzymes. *Nature Plants* **4**, 299-310 (2018).
690
691 29. Chiba, Y., Mitani, N., Yamaji, N. & Ma, J.F. HvLsi1 is a silicon influx transporter in barley.
692 *Plant Journal* **57**, 810-818 (2009).
693
694 30. Whipple, C.J. Grass inflorescence architecture and evolution: the origin of novel signaling
695 centers. *New Phytol* **216**, 367-372 (2017).
696

697 31. Hara, Y., Yokoyama, R., Osakabe, K., Toki, S. & Nishitani, K. Function of xyloglucan
698 endotransglucosylase/hydrolases in rice. *Annals of Botany* **114**, 1309-1318 (2014).
699
700 32. Landrein, B. & Ingram, G. Connected through the force: mechanical signals in plant
701 development. *Journal of Experimental Botany* **70**, 3507-3519 (2019).
702
703 33. Manassero, N.G., Viola, I.L., Welchen, E. & Gonzalez, D.H. TCP transcription factors:
704 architectures of plant form. *Biomol Concepts* **4**, 111-27 (2013).
705
706 34. Dalmais, M. *et al.* A TILLING Platform for Functional Genomics in Brachypodium
707 distachyon. *PLoS One* **8**, e65503 (2013).
708
709 35. Jiao, Y.P. *et al.* A Sorghum Mutant Resource as an Efficient Platform for Gene Discovery
710 in Grasses. *Plant Cell* **28**, 1551-1562 (2016).
711
712 36. Mascher, M. *et al.* A chromosome conformation capture ordered sequence of the barley
713 genome. *Nature* **544**, 427 (2017).
714
715 37. Gottwald, S., Bauer, P., Komatsuda, T., Lundqvist, U. & Stein, N. TILLING in the two-
716 rowed barley cultivar 'Barke' reveals preferred sites of functional diversity in the gene
717 HvHox1. *BMC Res Notes* **2**, 258 (2009).
718
719 38. Clement, M., Posada, D. & Crandall, K.A. TCS: a computer program to estimate gene
720 genealogies. *Mol Ecol* **9**, 1657-9 (2000).
721
722 39. Beier, S. *et al.* Construction of a map-based reference genome sequence for barley,
723 Hordeum vulgare L. *Scientific Data* **4**, 170044 (2017).
724
725 40. Kim, D. *et al.* TopHat2: accurate alignment of transcriptomes in the presence of
726 insertions, deletions and gene fusions. *Genome Biology* **14**, R36 (2013).
727
728 41. Liao, Y., Smyth, G.K. & Shi, W. featureCounts: an efficient general purpose program for
729 assigning sequence reads to genomic features. *Bioinformatics* **30**, 923-930 (2014).
730
731 42. Patro, R., Duggal, G., Love, M.I., Irizarry, R.A. & Kingsford, C. Salmon provides fast and
732 bias-aware quantification of transcript expression. *Nature Methods* **14**, 417-419 (2017).
733
734 43. Robinson, M.D., McCarthy, D.J. & Smyth, G.K. edgeR: a Bioconductor package for
735 differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140
736 (2010).

44. Komatsuda, T. *et al.* Six-rowed barley originated from a mutation in a homeodomain-leucine zipper I-class homeobox gene. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 1424-1429 (2007).

45. Lolas, I.B. *et al.* The transcript elongation factor FACT affects Arabidopsis vegetative and reproductive development and genetically interacts with HUB1/2. *Plant J* **61**, 686-97 (2010).

Acknowledgments: We thank the US Department of Agriculture–Agricultural Research Service (USDA-ARS), the National Small Grains Collection, Aberdeen (ID); the Nordic Genetic Resources Center (NordGen), Alnarp, Sweden; and the IPK Genebank, Germany, for providing the mutants and germplasm for haplotype analysis. The authors would like to thank Dr. Shun Sakuma for fruitful discussions and help in conducting mRNA *in-situ* hybridizations. We are thankful to Dr. Marion Dalmais, INRA Evry, France, for providing access to and screening of the *Brachypodium distachyon* TILLING resource, Anne Fiebig for help with data submission to ENA and Sandra Driesslein, Jenny Knibbiche, Mechthild Pürschel, Ines Walde, Kerstin Wolf, Marion Benecke and Kirsten Hoffie for excellent technical assistance.

Funding: During this study, research in the Schnurbusch laboratory received financial support from the Federal Ministry of Education and Research (BMBF) FKZ 0315954A and 031B0201A, HEISENBERG Program of the German Research Foundation (DFG), grant No. SCHN 768/8-1 and SCHN 768/15-1, as well as IPK core budget.

Author contributions: T.S. conceived the idea for the study, designed and monitored experiments, and analyzed data. N.P. expanded the idea for the study, designed and performed

experiments and analyzed data; C.T. executed the mRNA *in-situ* hybridizations. M.M. conducted microscopic analyses of cellular structures in paleae. T.N. analyzed RNA-seq data. U.L. provided irregular and intermedium barley spike mutants. T.R. executed SEM analyses. T.Schm. conducted sequence read mapping to unpublished barley genomic sequences for SNP calling. R.B., A.H. and L.A. performed the initial whole-genome shotgun sequencing of the parental genotypes for mapping. R.K. was involved in the phenotypic analysis of *com1.a* and RT-qPCR analyses of *COM1* in barley *vrs4* mutant. H.M.Y. provided sequences from a barley diversity panel for haplotype analysis and was involved in the RT-qPCR analysis. N.S. provided the barley TILLING resource; Z.X. provided the sorghum TILLING resource. N.P. and T.S. wrote the manuscript including contributions from co-authors. All authors have seen and agreed upon the final version of the manuscript.

Competing interests:

The authors declare no conflict of interest.

Data and materials availability:

Barley mutants are available from TS under a material transfer agreement (MTA) with IPK-Gatersleben. All data are available in the main text or online materials. The RNA-seq data and the whole genome shotgun (WGS) sequences of *com1.a* mutant have been submitted to the European Nucleotide Archive under accession number PRJEB35746 and PRJEB35761, respectively. COM1 sequences are available with the corresponding ID mentioned in the current

study in the public databases https://plants.ensembl.org/ & https://apex.ipk-gatersleben.de/apex/f?p=284:10 and are in the process of submission to NCBI as well.

Figure Legends:

Main Figures

Fig. 1 Proposed evolutionary pattern of grass inflorescences, and the spike/palea morphology of wild-type and *com1.a* mutant barley. (**A–D**) Model for grass inflorescence evolution from ancestral compound form to spike in *Triticeae*; re-drawn from [2]. (**E**) Spike morphology of wild-type (Wt) barley *cv.* Bowman. **(F–G**) Branched (mutant) BW-NIL(*com1.a*). (**H–I**) Floral reversion of spikelet to small spike-like branch structure from severe (**H**) to weak appearance as an extended rachilla (**I**). (**J–M**) Consecutive developmental stages of immature BW-NIL(*com1.a*) mutant spike from early glume primordium (**J**), advanced glume primordium (**K**), to advanced lemma (**L**) and early stamen primordium (**M**). (**N**) Dorsal view of immature BW-NIL(*com1.a*) mutant spike at early stamen primordia. (**O)** Longitudinal adaxial palea view of the BW-NIL(*com1.a*); white rectangle corresponds to the area used to take sections for histological analysis and to the lower image depicting the flat-plane surface of a cross section. (P) Longitudinal adaxial view of Wt palea; the lower image corresponds to the infolding surface of a cross section. (**Q–R**) Histological analyses of transverse sections (from O and P; white rectangles) of palea in mutant (**Q**) and Wt **(R**). Paleae are from spikelets shortly before anthesis. (**S–Z**) Walls of palea cells in mutant (**S-V**) versus Wt (**W-Z**).

Fig. 2 Map-based cloning of the gene underlying *com1.a*, phylogenetic analysis and protein structural variation of COM1. (**A**) Physical and genetic maps of *com1.a* from 100 recombinant plants or ~6,000 gametes. A single gene (red; HORVU5Hr1G061270, a single-exon TCP transcription factor) was the strongest candidate and deleted in the mutant parent BW-NIL(*com1.a*). (**B**) *COM1* gene model containing one TCP DNA binding domain. Six TILLING alleles are shown with prefix M3 (**Supplementary Note**). **(C)** UPGMA phylogenetic tree, using 1000 bootstrap replications, of COM1 first best (highlighted in gray) and second best homologs. (**D**) Evolutionarily conserved motifs among proteins present in the phylogenetic tree using SALAD. (see also **Supplementary Fig. 8**).

Fig. 3 Inflorescence morphology and gene expression patterns in sorghum (A–F) and *Brachypodium* (G–R). (**A**) Inflorescence shape in wild type (Wt) *cv.* BT623. (**B**). Compact inflorescence in TILLING mutant ARS180. (**C**) Expanded branch angle of Wt. (**D**) Acute branch angle in mutant ARS180. (**E**) Dissected dorsal view of the pulvinus (red circles). (**F**) RT-qPCR of *SbBad1/Wab1* in organs of Wt plants. 1_1, 1_2 and 1_3 represent first, second and third branch meristem stages, respectively. **(G)** Mutant palea scissor-like structure collapses easily due to external mechanical pressure; **(H)** normal/solid palea structure in Wt plants. **(I)** Acute branch angle in mutant **(J)**; expanded branch angle in Wt. (**K**) SEM view of transverse section of Wt palea; mutant has extra VB in middle **(L)** lacking in Wt **(M)**. (**N**) RT-qPCR of *BdBad1/Wab1* gene expression in Wt. **(O)** Branch angle measurement, as proxy for pulvinus size, among contrasting M6 homozygous TILLING lines; aa and bb refer to mutant and Wt homozygous, respectively (**Supplementary Note**). Values above x-axis indicate number of angles measured. **(P–Q)** were used for measuring number of spikelets per individual inflorescence (panicle) (**P**)

and number of florets per spikelet (**Q**) in the same M6 TILLING plant material (illustrated in **O**); mutants (gray boxes) showed no significant difference to Wt. P values were determined using *Student's t* test. Values above x-axis indicate number of plants**.** Genotype IDs below x-axis refer to the parental line of the respective M6 family.

Fig. 4. Transcript analyses of *COM1* in two-rowed barley. . **(A)** Relative *COM1* expression at different stages of immature spike and (**B)** organs along with spike sections (basal, central and apical at AP) of *cv.* Bowman.. Despite expression in tiller buds, no difference in tiller number was observed (**Supplementary Note, Supplementary Fig. 5**). y-axis value: expression relative to *HvActin*. (**C)** Semi-qPCR of *COM1* mRNA. **(D)** Control hybridization using sense probe. (**E–G**) mRNA in-situ hybridization of *COM1* in two-rowed Wt barley *cv.* Bonus. Tissues represent cross-section through a spikelet triplet at TM **(E)** and AP stages (**F–G**).

Fig. 5. Model of *COM1* regulation based on transcriptome analysis in barley (A–B) and schematic representation of *COM1* functional modification from non-*Triticeae* (C–D). Wt *COM1* transcriptional regulation model based on downregulation of the Wt allele. **(B)** RNA-seq-based heat map of selected differentially expressed (DE) genes; for the remaining DE genes see **Supplementary Fig. 11**. **(C–D)** IM-to-BM boundary formation due to Wt gene function (**C**) and lack of boundary due to the loss-of-function allele (**D**) in non-*Triticeae* grasses. Of note, *COM1* involvement in thickening the boundary cell walls within *non-Triticeae* species cannot be excluded. (**E**) IM-to-SM boundary formation in Wt barley; restriction of *COM1* function to cell wall thickening (the blue program), due to the evolutionary functional modification. (**F**). SM-to-

IM floral reversion observed in barley *com1.a* due to lack of the corresponding wall-amplified micromechanical signals needed to confer SM identify.

Table1. Functional variation of *COM1* homologs observed among grass species

		Effect on the corresponding organ/meristem				
Species	Gene function in boundary	On branch formation	On pulvinus size/formation	Growth of palea cells	Number of VB [1] in palea	Pollen fertility[2]
Barley	signaling	Inhibition	- [3]	Inhibition[4]	Promotion	Normal
Brachypodium	formation[5]	No effect	Promotion	No effect [6]	Promotion	Normal
Rice	formation	Not known[7]	Promotion	Promotion	Promotion	Reduced
Sorghum	formation	Promotion	Promotion	No effect	Promotion	Reduced
Maize	formation	Promotion	Promotion	No effect	Not known	Not known

1 stands for Vasculature Bundles

2 revealed by grain setting measurements as a proxy

3 pulvinus is typically absent in Wt spike of *Triticeae* including barley as well as in the branched mutant spikes.

4 apparent at the longitudinally-middle palea part resulting in formation of the infloded area

5 Refers to the formation of boundary between pulvinus and the lateral branch without which fusion of the two happened. Reflects intermediate evolutionary phylogenetic positon of *Brachypodium* among grasses.

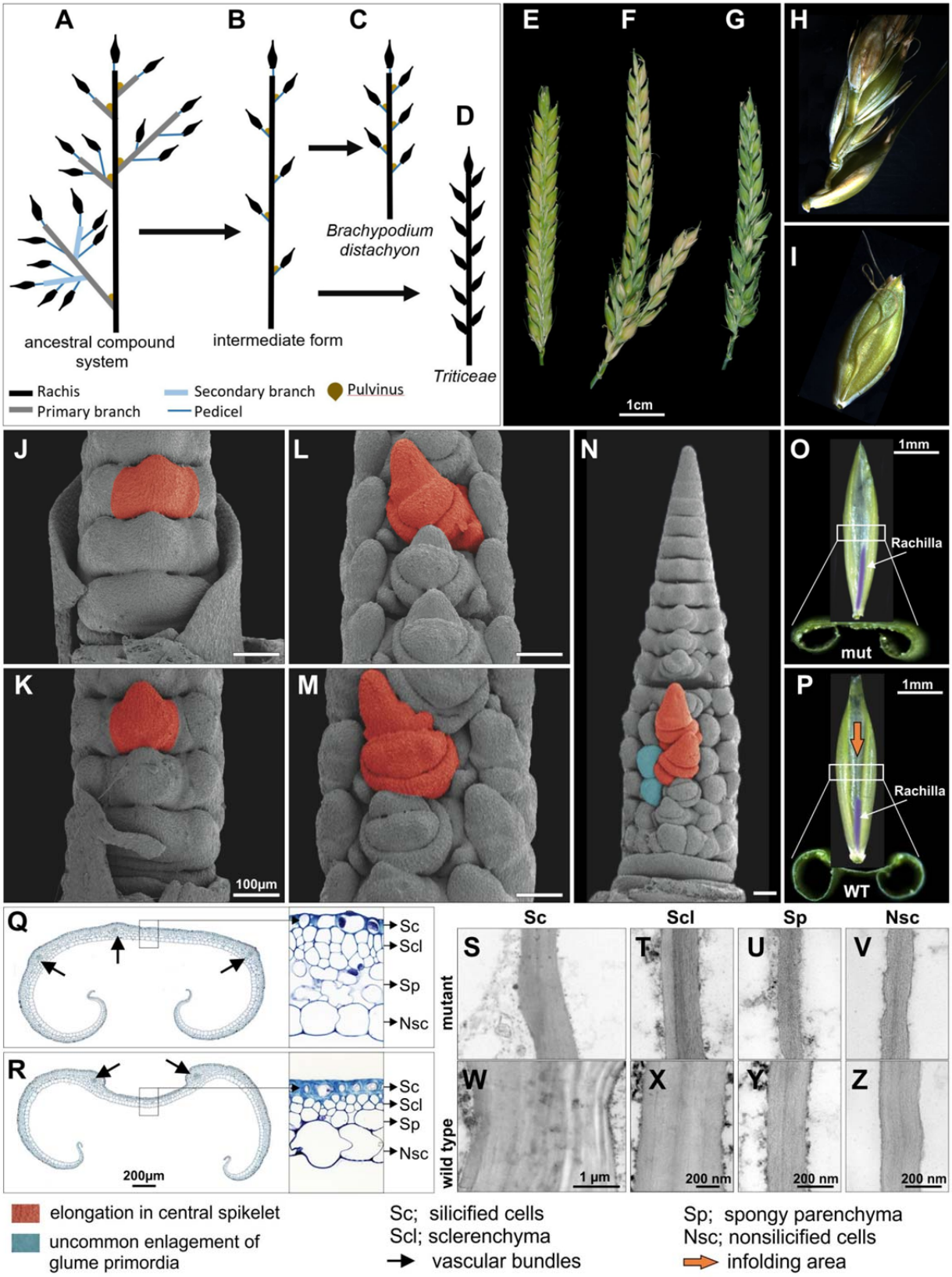
6 not visible at the microscopic level

7 perhaps because the rice cultivars used in the corresponding studies (*cv.* Nipponbare and *cv.* 9522) are known to exhibit panicles with acute lateral branches.

852

853

854 Figure 1



855

Figure 2

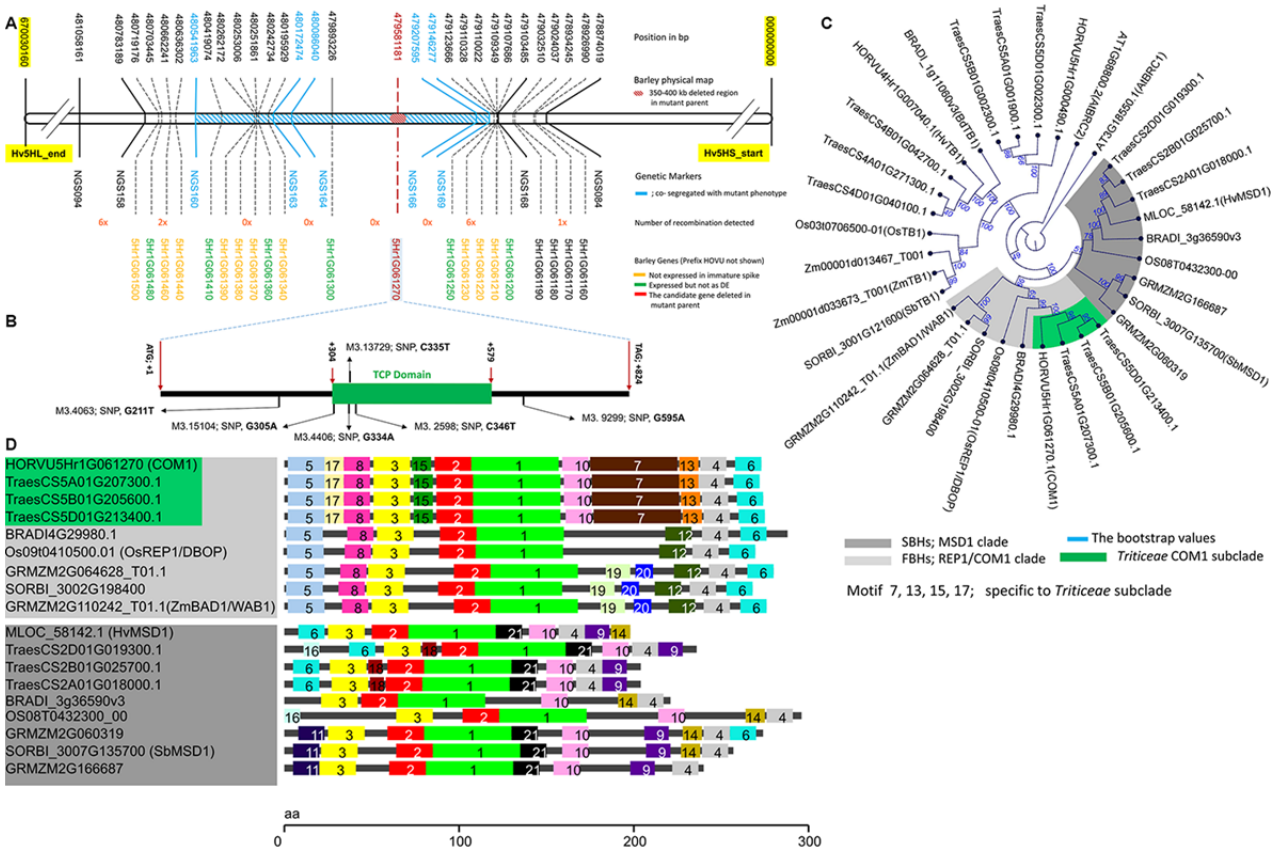

A
Position in bp
Barley physical map
350-400 kb deleted region in mutant parent
Hv5HL_end
Hv5HS_start
Genetic Markers
Number of recombination detected
Barley Genes (Prefix HOVU not shown)
Not expressed in immature spike
Expressed but not as DE
The candidate gene deleted in mutant parent
B
TCP Domain
M3.4063; SNP, G211T
M3.15104; SNP, G305A
M3.4406; SNP, G334A
M3. 2598; SNP, C346T
M3. 9299; SNP, G595A
M3.13729; SNP, C335T
C
SBHs; MSD1 clade
FBHs; REP1/COM1 clade
The bootstrap values
Triticeae COM1 subclade
Motif 7, 13, 15, 17; specific to Triticeae subclade
D
HORVU5Hr1G061270 (COM1)
TraesCS5A01G207300.1
TraesCS5B01G205600.1
TraesCS5D01G213400.1
BRADI4G29980.1
Os09t0410500.01 (OsREP1/DBOP)
GRMZM2G064628_T01.1
SORBI_3002G198400
GRMZM2G110242_T01.1(ZmBAD1/WAB1)
MLOC_58142.1 (HvMSD1)
TraesCS2D01G019300.1
TraesCS2B01G025700.1
TraesCS2A01G018000.1
BRADI_3g36590v3
OS08T0432300_00
GRMZM2G060319
SORBI_3007G135700 (SbMSD1)
GRMZM2G166687
aa
0
100
200
300

Figure 3

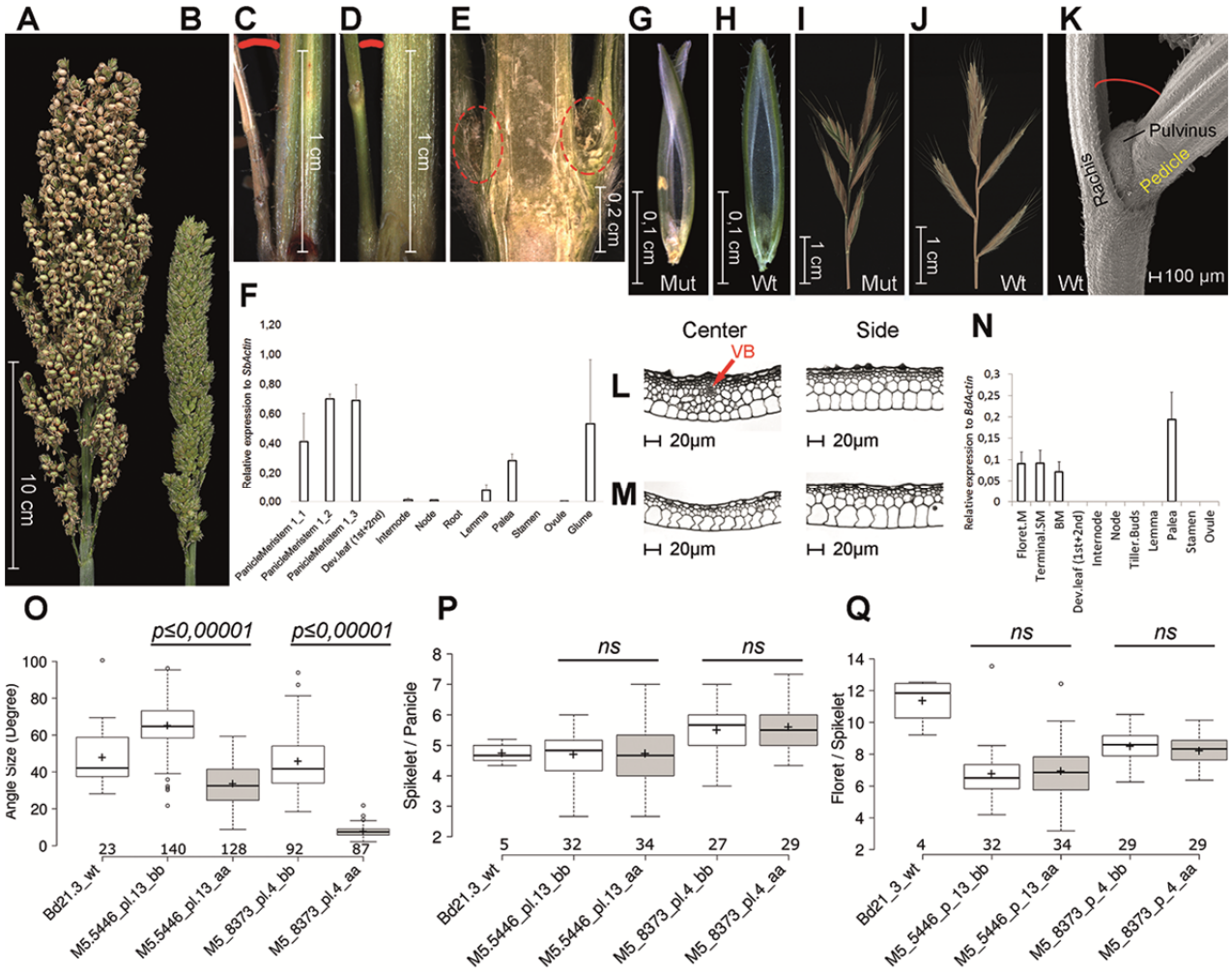
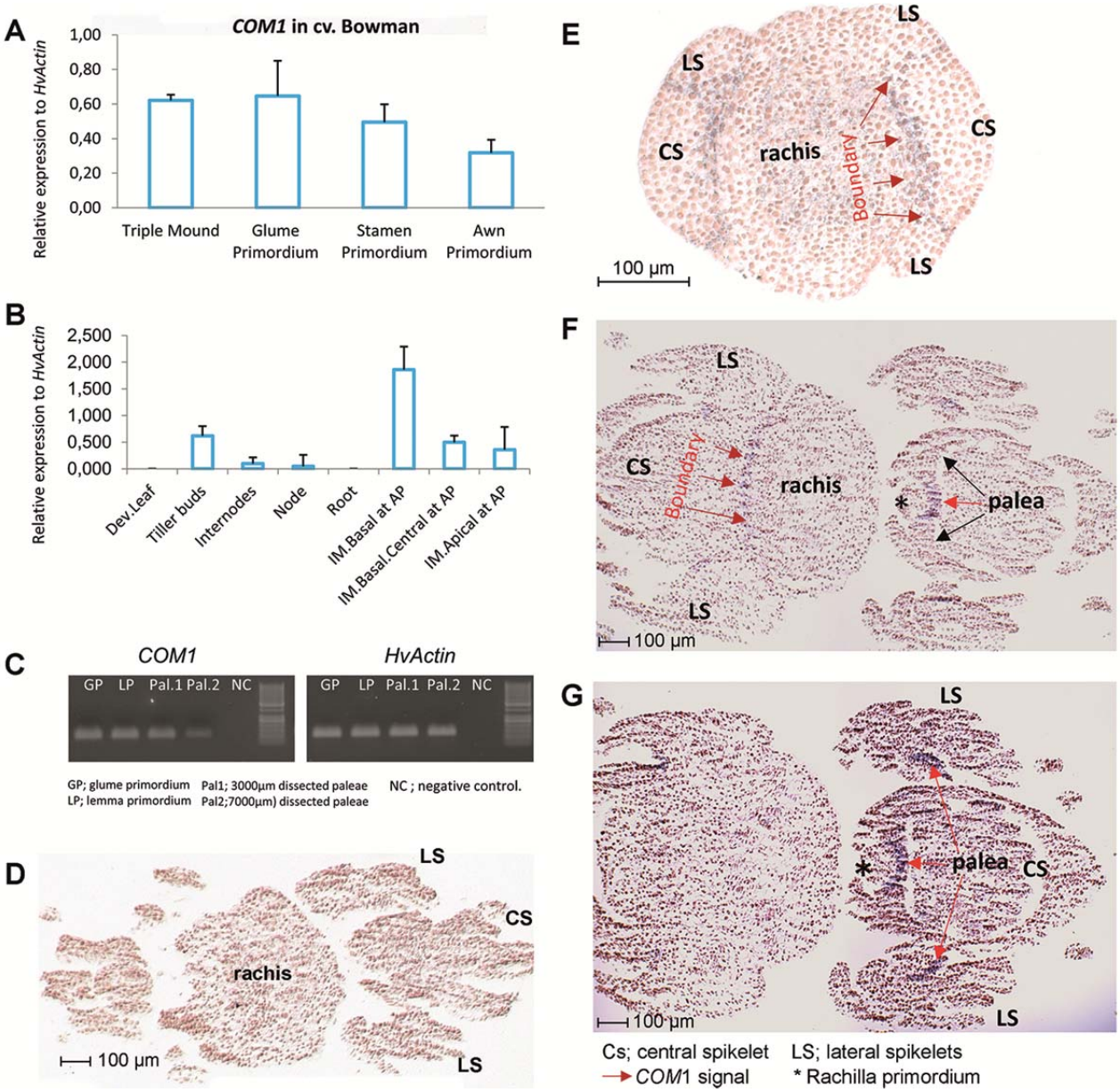
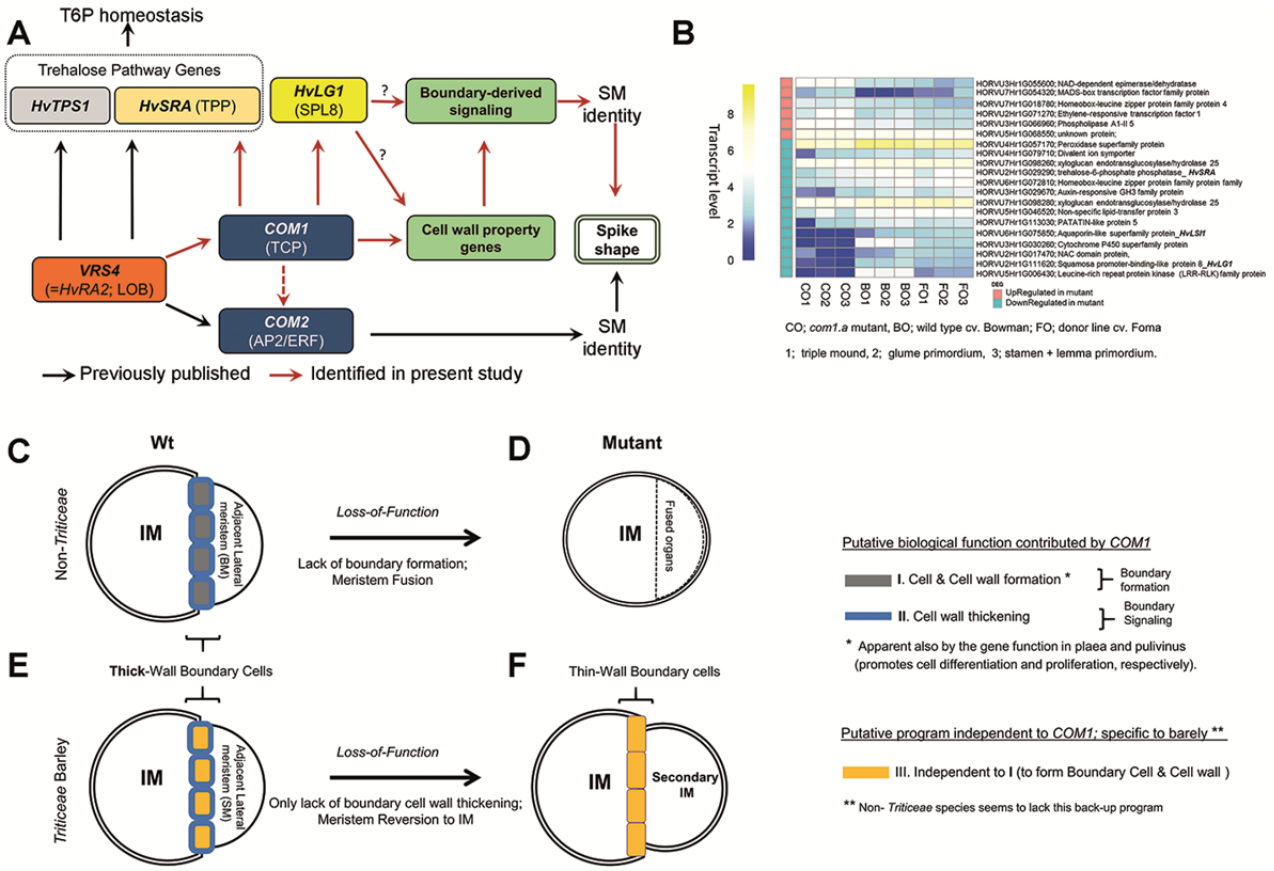

A
B
C
D
E
G
H
I
J
K
10 cm
1 cm
1 cm
0.2 cm
0.1 cm
Mut
0.1 cm
Wt
1 cm
Mut
1 cm
Wt
Wt
100 µm
Pulvinus
Pedicle
Rachis
F
Relative expression to SbActin
PanicleMeristem 1_1
PanicleMeristem 1_2
PanicleMeristem 1_3
Dev.leaf (1st+2nd)
Internode
Node
Root
Lemma
Palea
Stamen
Ovule
Glume
Center
Side
VB
L
M
20µm
N
Relative expression to BdActin
Floret.M
Terminal.SM
BM
Dev.leaf (1st+2nd)
Internode
Node
Tiller.Buds
Lemma
Palea
Stamen
Ovule
O
p≤0,00001
p≤0,00001
Angle Size (Degree)
23
140
128
92
87
Bd21.3_wt
M5.5446_pl.13_bb
M5.5446_pl.13_aa
M5_8373_pl.4_bb
M5_8373_pl.4_aa
P
ns
ns
Spikelet / Panicle
5
32
34
27
29
Bd21.3_wt
M5.5446_pl.13_bb
M5.5446_pl.13_aa
M5_8373_pl.4_bb
M5_8373_pl.4_aa
Q
ns
ns
Floret / Spikelet
4
32
34
29
29
Bd21_3_wt
M5_5446_p_13_bb
M5_5446_p_13_aa
M5_8373_p_4_bb
M5_8373_p_4_aa

Figure 4


A
COM1 in cv. Bowman
Relative expression to HvActin
1,00
0,80
0,60
0,40
0,20
0,00
Triple Mound
Glume Primordium
Stamen Primordium
Awn Primordium
B
Relative expression to HvActin
2,500
2,000
1,500
1,000
0,500
0,000
Dev.Leaf
Tiller buds
Internodes
Node
Root
IM.Basal at AP
IM.Basal.Central at AP
IM.Apical at AP
C
COM1
HvActin
GP LP Pal.1 Pal.2 NC
GP LP Pal.1 Pal.2 NC
GP; glume primordium
LP; lemma primordium
Pal1; 3000µm dissected paleae
Pal2;7000µm) dissected paleae
NC ; negative control.
D
LS
CS
rachis
100 µm
LS
E
LS
LS
CS
CS
rachis
Boundary
100 µm
LS
F
LS
CS
Boundary
rachis
*
palea
LS
100 µm
G
LS
*
palea
CS
100 µm
LS
Cs; central spikelet
LS; lateral spikelets
COM1 signal
* Rachilla primordium

866 Figure 5



867