Spatial and single-cell expression analyses reveal complex expression domains in early wheat spike development

Xiaosa Xu[1*‡], Huiqiong Lin[2, 3*], Junli Zhang[2], German Burguener[2, 3], Francine Paraiso[2, 3], Connor Tumelty[2], Chengxia Li [2, 3], Yuchen Liu[1] and Jorge Dubcovsky[2, 3‡].

[1] Department of Plant Biology, University of California, Davis, CA 95616, USA

[2] Department of Plant Sciences, University of California, Davis, CA 95616, USA

[3] Howard Hughes Medical Institute, Chevy Chase, MD 20815, USA

[*] The first two authors contributed equally to this paper.

[‡] Co-corresponding authors

Co-corresponding authors

Jorge Dubcovsky, Email: jdubcovsky@ucdavis.edu

Xiaosa Xu, Email: xjkxu@ucdavis.edu

Emails from other authors:

Huiqiong Lin: hqlin@ucdavis.edu

Junli Zhang: psjzhang@ucdavis.edu

German Burguener: gfburguener@ucdavis.edu

Francine Paraiso: fjparaiso@ucdavis.edu

Connor Tumelty: cwtumelty@ucdavis.edu

Chengxia Li: chxli@ucdavis.edu

Yuchen Liu: yvcliu@ucdavis.edu

ABSTRACT

Wheat is an important crop for global food security and a better understanding of the molecular mechanisms driving spike and spikelet development can inform the engineering of more productive varieties. In this study, we integrated single-molecule fluorescence *in situ* hybridization (smFISH) and single-cell RNA sequencing (scRNA-seq) to generate an atlas of cell types and expression domains during the early stages of wheat spike development. We characterized spatiotemporal expression of 99 genes by smFISH in 48,225 cells at the early transition (W1.5), late double ridge (W2.5) and lemma primordia (W3.25) stages. These cells were clustered into 21 different expression domains, including four beneath the developing spikelets and three different meristematic regions, which were consistent across spikelets and sections. We also identified genes differentially expressed during the transition of the inflorescence meristem into a terminal spikelet. The smFISH study was complemented by the scRNA-seq profiling of 26,009 cells from W2.5 and W3.25 that were clustered into 23 distinct cell types. We annotated the single-cell clusters as meristem, boundary domains, cortex, central spike, epidermis, and vasculature by using known markers, co-expression analysis of genes selected from the smFISH study, and correlation matrices of clusters identified in both studies. The scRNA-seq data uncovered genes associated with boundary genes *TCP24* and *FZP*, as well as the meristematic genes *AGL6* and *ULT1.* The smFISH and scRNA-seq studies provided complementary tools for characterizing the gene networks that regulate spike development, supporting the functional characterization of key regulatory genes, and informing the engineering of more productive wheat spikes.

Keywords

Wheat, spike development, single-molecule fluorescence *in situ* hybridization, single-cell RNA sequencing, spatial transcriptomics, co-expression analysis, gene expression, transcription factors.

BACKGROUND

In 2023, the world produced approximately 799 million tons of wheat grain [1], but further increases are necessary to feed a growing human population. Wheat grains develop in inflorescences called spikes, so improving spike productivity is an important breeding objective. Traditional and molecular breeding methods have made small incremental advances in spike productivity, but a deeper understanding of the regulatory processes governing spike development has the potential to accelerate the engineering of higher-yielding wheat varieties.

Wheat spike development starts with the transition of a shoot apical meristem (SAM) that produces leaves to an inflorescence meristem (IM) that produces spikelets. During the vegetative phase, the SAM produces lateral meristems that develop into leaves, whereas the subtended axillary meristems generally fail to grow out immediately [2]. By contrast, the wheat IM produces lateral meristems with two distinct ridges (double ridge stage). The lower ridges (also known as leaf ridges) are suppressed, whereas the upper ridges rapidly develop into spikelet meristems (SMs) that generate sessile spikelets attached directly to the inflorescence axis (rachis). Therefore, the transition from the vegetative to the reproductive phase involves both a drastic change in the bract and axillary meristems developmental rates and a change in the identity of the axillary meristems.

Wheat MADS-box genes from the *SQUAMOSA* clade (*VRN1*, *FUL2*, and *FUL3*) play an essential role in the repression of the lower ridge. Wheat plants with simultaneous loss-of-function mutations in these three genes show a normal transition of the SAM to the double ridge stage and stem elongation, but the lower ridges are derepressed and develop into leaves or bracts [3]. Bract outgrowth has also been reported in wheat plants with loss-of-function mutations in the *SQUAMOSA PROMOTER BINDING PROTEIN LIKE17* (*SPL17*) and *SPL14*, two regulators of the *SQUAMOSA* genes [4]. The maize orthologs of these genes (*TSH4*, *UB2*, and *UB3*) play a similar role [5, 6], suggesting a conserved function across grass species. Additional grass genes involved in the bract-suppression pathway include *NECK LEAF1* (*NL1*) [7] and *LG2* [8].

The wheat *SQUAMOSA* genes also play a critical role in the determination of the meristem identity of the upper ridge. In the absence of these genes, the upper ridges develop into vegetative meristems rather than SMs, resulting in a “spike of tillers” [3]. Taken together, these results indicate that the *SQUAMOSA* genes are essential for conferring SM identity to the

86 axillary meristem and suppressing bract development, but not for the initial transition of the
87 SAM to the double ridge stage and the subsequent stem elongation [3].

88 The MADS-box genes from the *SVP* clade (*VRT2, SVP1*, SVP3) play similar roles as the
89 *SQUAMOSA* genes during the initial stages of spike development. Individual loss-of-function
90 mutations in *VRN1* or *FUL2* [3], or simultaneous mutations in *VRT2* and *SVP1* [9], both result in
91 delayed heading, shorter stems and reduced spikelet number per spike (SNS). However, at later
92 stages of spike development, the SQUAMOSA proteins transcriptionally repress the *SVP*-clade
93 genes to facilitate interactions with SEPALLATA1 proteins, which are required for normal
94 spikelet development [9, 10].

95 The normal development of the wheat lateral spikelets begins with the formation of two sterile
96 bracts called glumes, followed by slightly different bracts called lemmas that subtend the floral
97 organs. Differences between glumes and lemmas are regulated by the *APETALA2-like* genes
98 *AP2L5* (*Q* gene) and *AP2L2*, which are cleaved by the microRNA miR172. Mutations in the
99 miR172 binding sites of these genes result in increased transcript levels and conversions of
100 sterile glumes into fertile lemmas [11, 12]. Conversely, loss-of-function mutation in *AP2L5* and
101 *AP2L2* result in conversions of fertile lemmas into sterile glumes [13]. These results indicate an
102 important role of *AP2L5* and *AP2L2* in the regulation of glume and lemma organ identities and
103 the development of the axillary floret meristem (FM).

104 In grasses, the first whorl of floral organs generated by the FM is the palea, which is regulated by
105 the MADS-box gene *AGL6* [14]. Development of the internal floral whorls (lodicules, stamens,
106 and pistils) is severely disrupted in wheat plants with loss-of-function mutations in *LEAFY* (*LFY*)
107 or *WHEAT ORTHOLOGUE of APO1* (*WAPO1*), two genes expressed in partially overlapping
108 domains at the FM and that encode interacting proteins [15, 16]. In addition to the disruptions in
109 floral development, loss-of-function mutations in either or both genes significantly reduce the
110 rate at which the IM generates lateral meristems, leading to drastic reductions in SNS [15].
111 Natural alleles with increased SNS have been identified for both *WAPO1* [16-19] and *LFY* [20],
112 and allele combinations that maximize SNS have been also reported [20].

113 Despite progress in the genetic characterization of wheat spike development, the diversity of cell
114 types or expression domains within the wheat spike and their associated gene networks remain
115 poorly understood. In this study, we characterized the spatial expression of 99 wheat genes

during the initial SAM-IM transition, the late double ridge and lemma primordia stages using single-molecule fluorescence *in situ* hybridization (smFISH). To expand the list of characterized genes, we integrated the smFISH data with single-cell RNA sequencing (scRNA-seq) from protoplasts isolated during the late double ridge and lemma primordia stages. The smFISH and scRNA-seq data provided complementary resources to study the gene networks regulating spike development and revealed opportunities to engineer more productive wheat spikes.

RESULTAS AND DISCUSSION

1. Spatial expression and cell segmentation

To identify the distinct cell types and expression domains involved in early wheat spike development, we analyzed the expression of 99 genes simultaneously (Data S1) by smFISH at three early stages of spike development (W1.5, W2.5, and W3.25 as defined by the Waddington scale [21]). W1.5 marks the initial transition of the shoot apical meristem (SAM) to the reproductive stage (Fig. 1A). W2.5 represents the late double ridge stage before glume primordia initiation (Fig. 1B), and W3.25 marks the lemma primordia stage (Fig. 1A-C). At W3.25, the inflorescence meristem (IM) begins its transition to a terminal spikelet (IM→TS), and the most advanced spikelet meristems (SMs) at the center of the spike initiate floret meristems (FMs, Fig. 1C).

We analyzed two sections from each developmental stage (Figs. 1, S1) using the cell segmentation program QuPath [22], identifying a total of 48,225 cells. In the same sections, cell walls were visualized with calcofluor-white (Figs. S2, S3). Based on the expression of 99 genes (Data S1), these cells were grouped into 21 clusters, represented by consistent colors across all three developmental stages and replications (Figs. 1, S1). To facilitate visualization, the different cell clusters are presented separately in Fig. S4.

On average, we characterized 3,807 cells at W1.5, 8,200 at W2.5, and 12,105 at W3.25 (Fig. 1 and S1), showing a clear increase in cell number during spike development. These cells also exhibited increasing differentiation, with 14 clusters at W1.5, 19 at W2.5, and 21 at W3.25 (Fig. 1 and S1). At the two later stages, cell clusters formed at the SMs and beneath them were

consistent across spikelets within the same spike and between sections in Figs. 1 and S1, demonstrating good reproducibility.

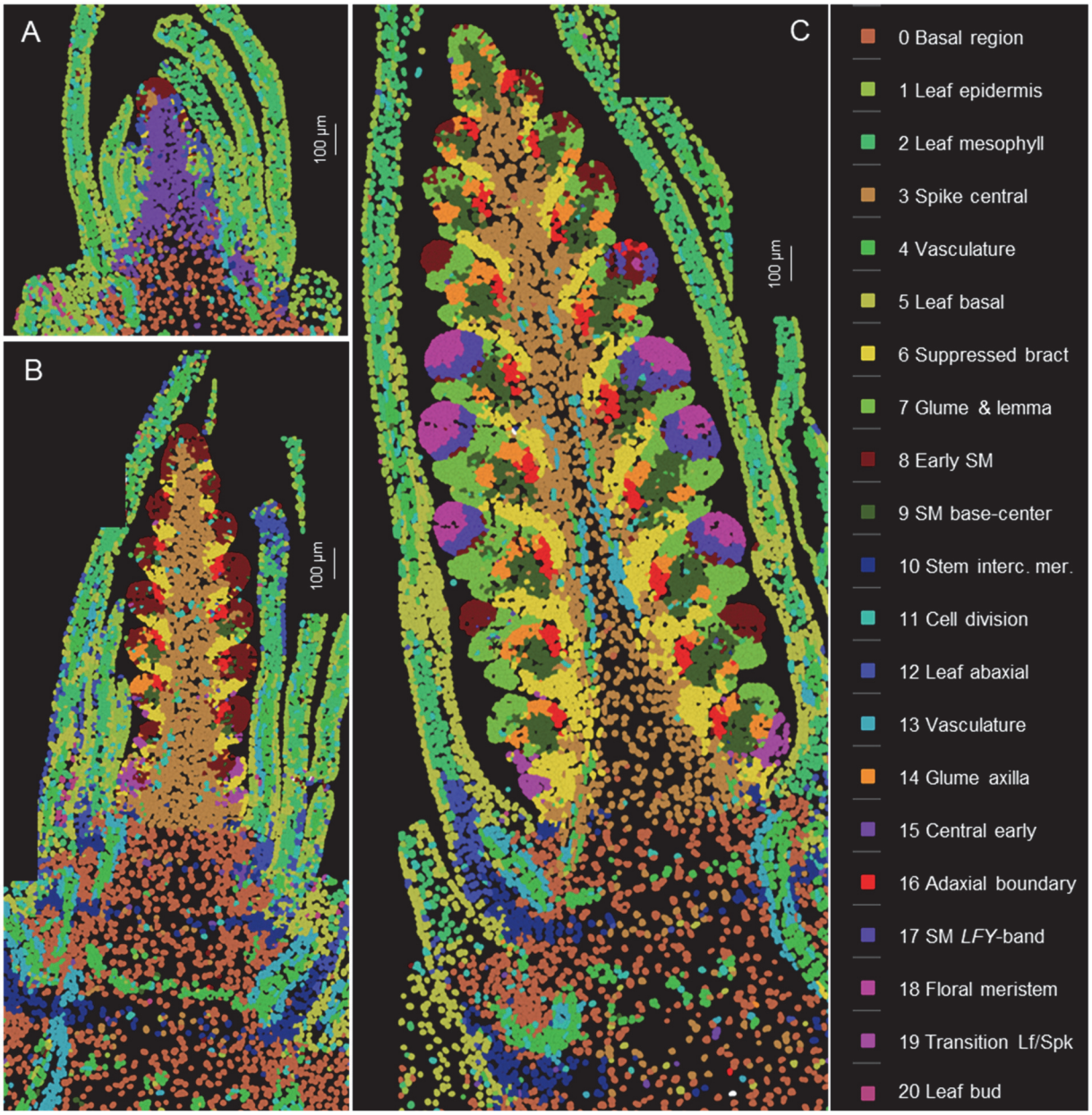


Fig. 1. Cell clusters based on the expression of 99 genes analyzed by smFISH at three stages of wheat spike development. Cell clusters were generated from the three sections shown in this figure and replicate sections presented in Fig. S1. Identical colors represent the same clusters across the six sections. **A** W1.5: initial transition from the vegetative to the reproductive stage (section C1-1). **B** W2.5: late double ridge stage (section B1-3). **C** W3.25: lemma primordia stage (section A1-2). Scale bars = 100 µm. A detailed description of the cell clusters, cell counts, and preferentially expressed genes in each cluster is available in Data S2.

Annotation of the cell clusters and their preferentially expressed genes are summarized in Data S2, whereas the correlations among cell clusters are presented in Data S3. The five cell clusters detected in the leaves (c1, c2, c5, c12, and c20) were highly correlated ($R \geq 0.78$, Data S3) and are not discussed further in this study. To organize the description of these clusters and their changes during spike development, we separated the developing spike into five regions: the basal region below the spike, the spike central region, the region below the SMs, the SMs, and IM.

1.1. The basal region below the spike

The basal region below the spike consists of cells that will later differentiate into nodes and internodes, forming the stem and contributing to important agronomic traits such as plant height and stem strength. This region is primarily composed of ground tissue cells grouped in cluster c0. At W2.5 and W3.25 this basal region is interrupted by alternating bands of cells from clusters c4 (vasculature) and c10 (stem intercalary meristem, Figs. 1, S1, S4).

In longitudinal sections stained with calcofluor-white at W2.5 and W3.25 (Figs. S2, S3), two alternating zones are visible at the basal region: one with shorter cells likely corresponding to the nodes, and one with longer cells likely corresponding to the internodes of the future stem. Enriched expression of the procambial marker *LAX4* [23] and the phloem marker *APL* [24] was observed in the smaller cells, supporting their association with future nodes (Fig. 2A). These cells, together with the cells expressing these vascular genes in the leaf and spike vascular tissue were all included in cluster c4 (Fig. 2B).

Cells classified as cluster c10 (Fig. 2B) appeared as incomplete bands in the future internode regions at W2.5 and W3.25 (Figs. 1, S4). In the three-dimensional space, c10 cells may form a transversal disc across the stem located where the abaxial base of the leaves joins the stem. These c10 cells were characterized by the differential expression of *LFY*, *PLATZ1* (Fig. 2A), *NAL1*, and *SVP1* (Fig. S5, Data S2). Our previous study showed that loss-of-function mutations in wheat *PLATZ1* [25] or *SVP1* [9] significantly reduced plant height, suggesting that cluster c10 may correspond to the stem intercalary meristem region (IcM), a region at the base of grass internodes that contributes new cells to the elongating internode [26].

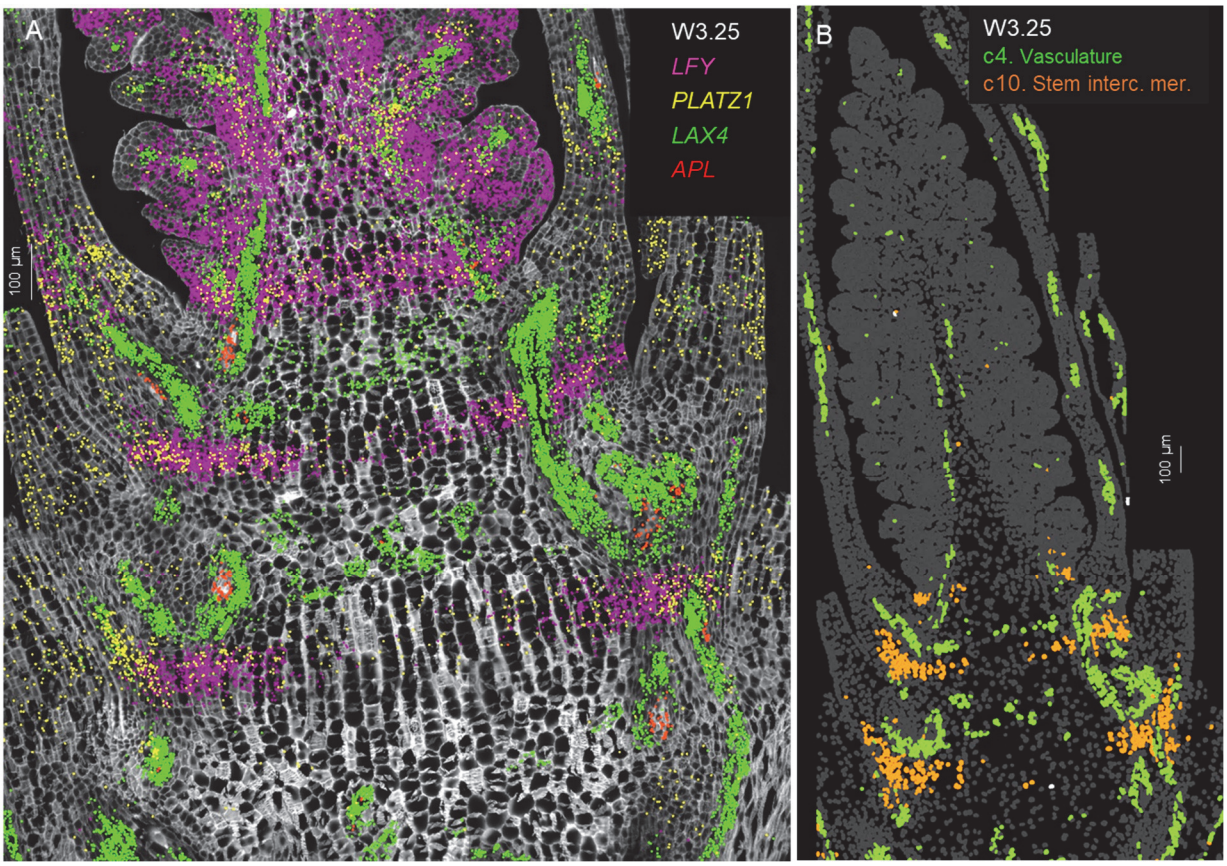


Fig. 2. Single-molecule fluorescence *in-situ* hybridization (smFISH) characterization of the basal region below the spike at W3.25. Cell walls were stained with calcofluor-white. **A** Vasculature patterning genes *LAX4* (green) and *APL* (red) are concentrated in the region with shorter cells (likely future nodes). *LFY* (purple) and *PLATZ1* (yellow) mark a band in the region with longer cells (likely future internodes). **B** Cell segmentation with highlighted clusters c4 in green (vascular tissue) and c10 in orange (stem intercalary meristem). Scale bar= 100 µm.

The role of *LFY* in stem development is complex. Loss-of-function mutations in the rice *LFY* ortholog *APO2* had limited effect on plant height [27]. However, the rice natural mutant *Short and Solid Culm* (*ssc*), which carries a *LFY* allele with a D16N amino acid polymorphism and exhibits 3-fold higher transcript levels than the wildtype, showed reduced stem length and diameter. Overexpression of the wildtype *LFY* allele led to reduced stem length and diameter [28]. In this study, we measured the stems and analyzed node and internode sections of Kronos *lfy*-null mutants developed in a previous study [15]. We found no significant differences in stem length, but the *lfy* mutant showed significantly wider diameter for nodes (22.3% ± 1.8%) and internodes (11.5% ± 2.0%). Moreover, the *lfy* mutants had more solid stems (62% ± 17% reduction in the diameter of the central hole) than the WT (Data S4). These results confirmed the conserved role of *LFY* on stem development in wheat and rice.

199 Since the gibberellic acid (GA) hormone plays a critical role in stem elongation, we examined
200 the spatial distribution of two genes involved in GA biosynthesis (*GA20ox1* and *GA20ox2*) and
201 one involved in GA degradation (*GA2ox1*). At W1.5, the higher expression of *GA2ox1* relative to
202 the two biosynthetic genes, was a distinguishing feature of the early central region (cluster c15)
203 relative to the same region at later stages (cluster c3) (Fig. S6A, Data S2). At W2.5 and W3.25,
204 the two GA biosynthetic genes were enriched in the basal region below the spike, outside the c10
205 cluster (Fig. S6B-C). At these two stages, *GA2ox1* expression was restricted to the periphery of
206 the region enriched in the GA biosynthetic genes, possibly forming a ring of reduced GA
207 abundance around the elongating cells. We hypothesize that this spatial distribution of GA
208 metabolic and catabolic genes may be important for normal wheat stem elongation.

209 **1.2. The transition zone between leaf and spikelet meristems**

210 At W2.5 and W3.25, a group of cells within cluster c19 was detected in the transition region
211 between the leaves and the spike (Fig. S7A-B). Proximal to this boundary, lateral meristems
212 exhibit rapid leaf development and axillary meristems undergo a temporary arrest. Conversely, in
213 the region distal to this boundary lower ridges are suppressed (suppressed bracts) and upper
214 ridges rapidly develop into spikelet meristems. We hypothesize that cells in cluster c19
215 contribute to these drastic functional changes.

216 The genes preferentially expressed in cluster c19 include *SPL* transcription factors *SPL2*, *SPL14*,
217 and *SPL17* (Fig. S7C-D), the *TEOSINTE BRANCHED1* (*TB1*), *CYCLOIDEA*, *PCF1* (*TCP*)
218 transcription factors *TB1* and *TB2* (Fig. S7E-F), and the early inflorescence MADS-box genes
219 *FUL2* and *SOC1-1* (Data S2). Several of these genes have previously been associated with the
220 regulation of spike architecture. In tetraploid wheat, combined *vrn1 ful2* mutations result in de-
221 repression of the lower ridge, leading to bract outgrowth, and development of the upper ridge
222 into tiller-like structures with residual floral organs [3]. The preferential expression of *FUL2* in
223 cluster c19 may contribute to enhancing these functions.

224 *SQUAMOSA* genes are direct targets of *SPL* genes [29, 30], and mutations in the *SPL14* and
225 *SPL17* genes result in the formation of ectopic bracts at the base of the inflorescence in rice [31],
226 maize [5] and wheat [4]. Based on the basal location of these bracts, we hypothesize that the
227 preferential expression of *SPL14* and *SPL17* in cluster c19 may contribute to bract repression in
228 the transition zone. *TSH4,* the maize ortholog of *SPL17* and its close paralogues *UB2* and UB3

229 (*SPL14*) work together to target genes responsible for axillary meristem suppression such as *TB1*
230 and auxin response regulators [32].

231 In rice, *SPL14* is a direct regulator of *TB1* [33], which in our study was preferentially expressed
232 in cluster c19. *TB1* reduces lateral branching, and mutants of this gene in maize [34] and rice
233 [35] show increased tillering. *TB1* mutants in barley and wheat also show increased tillering and
234 produce additional lateral spikelets in barley [36] and paired spikelets in wheat [37, 38]. These
235 findings suggest that the preferential expression of *TB1* and its close paralog *TB2* in cluster c19
236 may contribute to the changes in the regulation of both the lower and upper ridges in the
237 transition region.

238 **1.3. The central region of the spike**

239 The ground tissue cells in the central region of developing spikes were primarily assigned to
240 cluster c15 at W1.5 and cluster c3 at both W2.5 and W3.25 (Figs. 1, S4). Cells from cluster c15
241 showed stronger correlations with the leaf mesophyll (R= 0.74) than cluster c3 (R= 0.46, Data
242 S3), suggesting progressive differentiation of the ground tissue at the center of the spike. Gene
243 expression profiles in clusters c15 and c3 were highly correlated with each other (R= 0.70).
244 These two clusters also showed strong correlations with cluster c0 (Data S3), which represents
245 ground tissue from the adjacent basal region beneath the spike. Several genes were preferentially
246 expressed at the spike central region across all three developmental stages, including *AP3*,
247 *bZIPC4* (Fig. S8A-C), *IDD4* and *OSH1* (synonymous with *KNOX3*, Fig. S8D-F). The homeobox
248 gene *OSH1*was expressed at the base and center of the spikelets, but not in developing glumes
249 and lemmas. This expression pattern resembles that of its orthologs in maize and rice [39].

250 Another important function of the central region of the spike is to support the development of an
251 efficient vascular system for transporting water and nutrients to the developing spike. We
252 characterized four vascular markers: the auxin influx carrier *LAX4* [40], the phloem marker *APL*
253 [24], and the duplicated auxin efflux carriers *PIN1a* and *PIN1b* [23, 41] (Fig. 3). In
254 *Brachypodium*, *PIN1b* plays a role in sink-finding and connecting new primordia to existing
255 vasculature, whereas *PIN1a* forms narrower canals than *PIN1b* and is likely involved in vein
256 patterning [41]. A similar pattern is observed in wheat, where *PIN1b* is expressed in a broader
257 area than *PIN1a* (Fig. 3).

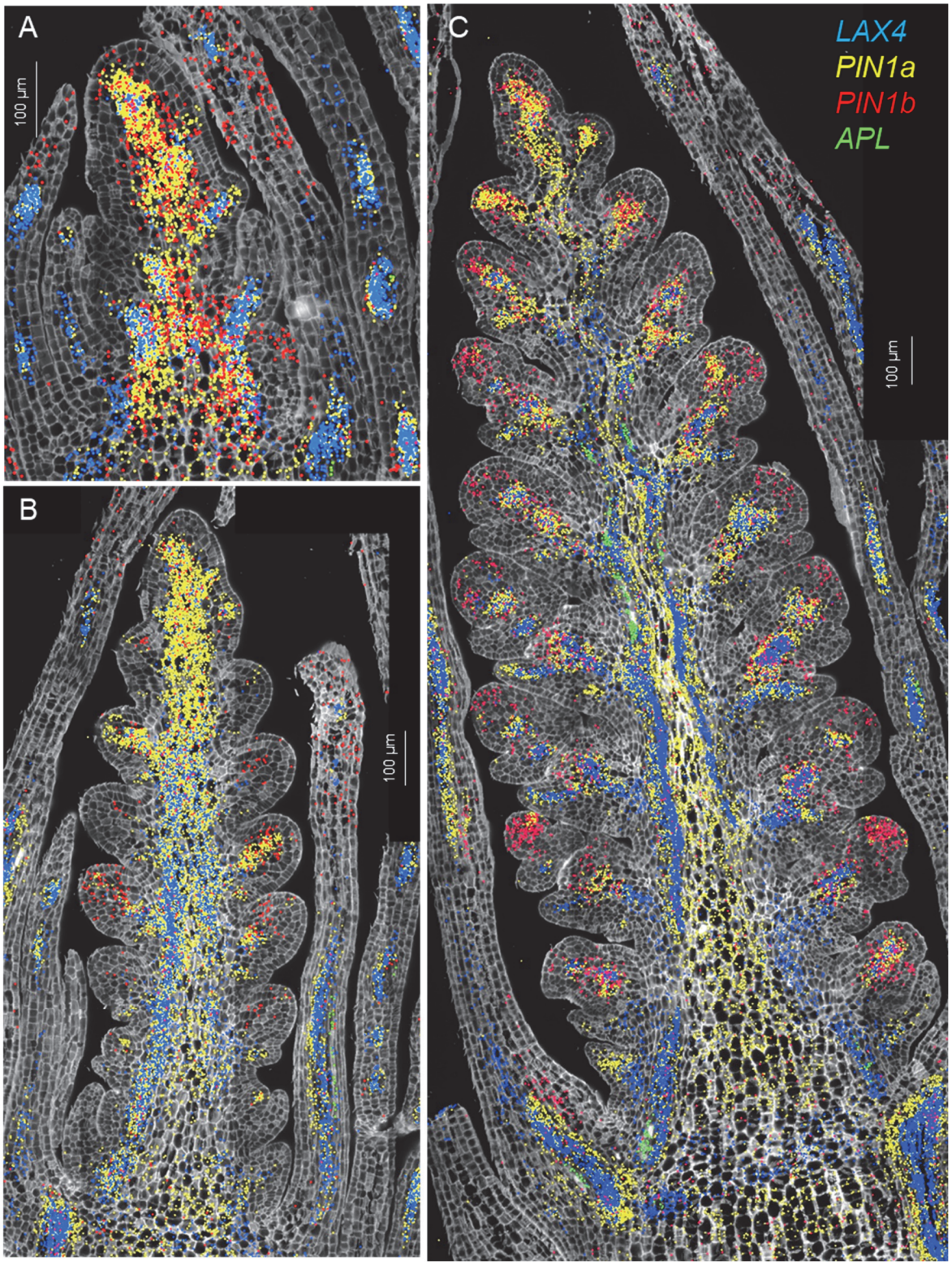


Fig. 3. Single-molecule fluorescence *in-situ* hybridization characterization of the spike central region (cluster c3). Spike central region (cluster c3). **A** W1.5, **B** W2.5, **C** W3.25. In all three panels, *LAX4* hybridization signal is shown in blue, *APL* in green, *PIN1a* in yellow, and *PIN1b* in red. Cell walls are stained with calcofluor-white.

262 In maize, *LAX4* and *PIN1a* have antagonistic functions and are frequently co-expressed in pro-
263 vascular cells [23, 40], an overlap that was also apparent in wheat spikes at all three
264 developmental stages (Fig. 3). The phloem marker *APL* was detected at developing vascular
265 tissue of leaves at W1.5 and W2.5 (Fig. 3A-B), and in the central region of the spike at W3.25
266 (Fig. 3A). Taken together, the expression profiles of these vascular genes provide an initial
267 framework for understanding the complex regulatory networks that govern vascular development
268 within the spike rachis.

269 **The region proximal to the spikelet meristem (SM)**

270 At W2.5 (Fig. 4A), cortex cells proximal to the spikelet meristem were primarily classified into
271 clusters c6 (suppressed bract), c9 (central SM basal region), and c16 (adaxial boundary). These
272 clusters remained evident at W3.25, with the addition of clusters c7 (glume and lemma) and c14
273 (glume axilla). A few cells from clusters c7 and c14 were also detected at W2.5 (Fig. 4A).

274 Cluster c6 corresponds to the suppressed bract region (Fig. S9). This cluster showed a relatively
275 high correlation with cluster c19 in the transition zone (R= 0.71, Fig. S9 and Data S3), which
276 was also associated with the repression of the lower ridge in the lateral meristems. Both clusters
277 showed preferential expression of *SPL14, FUL2*, and *LFY* (Data S2), but cluster c6 differed from
278 cluster c19 by its preferential expression of *SPL13* and lower expression of *TB1* and *TB2*.

279 Cluster c9 is located at the center of the region proximal to the SM (Fig. 4) and showed high
280 correlations (R= 0.75-0.80) with the surrounding clusters c3, c7, c8, c14, and c16 (Data S3). This
281 central cluster was enriched in *PIN1a* (Fig. 4A-D), *PIN1b, SOC1-1*, and *OSH1* (Data S2),
282 suggesting roles in transport of water and nutrients from the rachis and in SM development.

283 Cluster c16 is located adaxial to c9 and was defined by the differential expression of the
284 transcription factor *TCP24* (Figs. 4A-D, S10). *TCP24* was detected in 93% of the cells of cluster
285 c16 and only 2% of the cells in other clusters (Fig. S10, Data S2). *TCP24* and its closest paralogs
286 *TCP22*, *TB1*, and *TB2* all belong to the CYC/TB1 subclass of TCP transcription factors (Fig.
287 S7E-F). In yeast-two-hybrid assays, *TCP24* can form both homodimers and heterodimers with all
288 three other members of the CYC/TB1 subclass [42]. Loss-of-function mutations in barley *TCP24*
289 (also known as *COM1*) were associated with branched spikes, a phenotype not observed in *tb1*
290 mutants, suggesting functional differentiation.

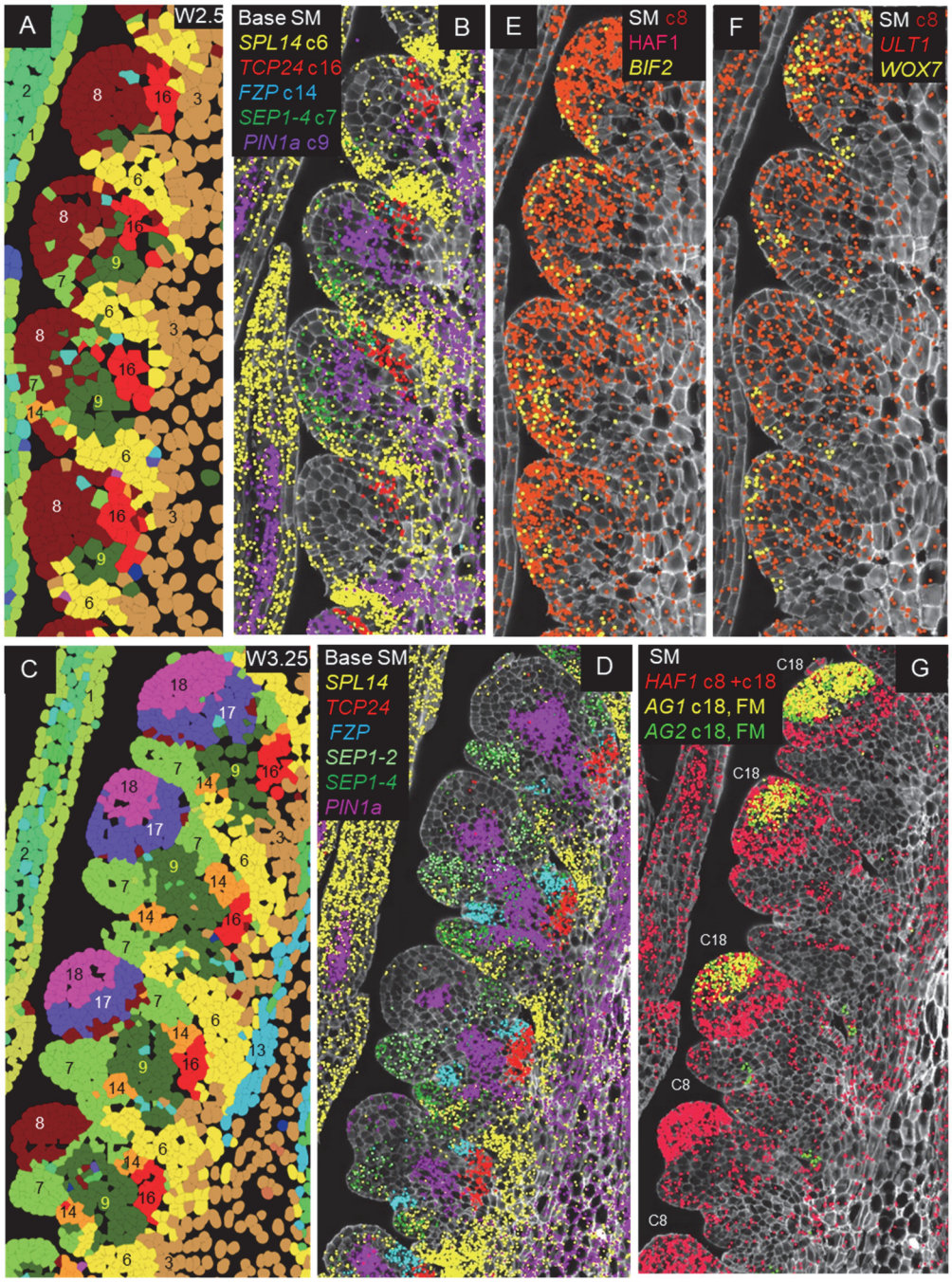


Fig. 4. Expression profiles of genes differentially expressed during spikelet development. **A-B, E-F** W2.5 (late double ridge stage). (**C-D, G**) W3.25 (lemma primordia stage). **A, C** Cell segmentation and clustering. **B, D** Genes differentially expressed at the base of the SM. **E-G** Genes differentially expressed in the SMs. In **G**, the more basal SMs are at an earlier stage (cluster c8) than the central spikelets, which have already transitioned to the floret development stage (FM, cluster c18), and are already expressing floral homeotic genes *AG1* and *AG2*. *SEP1-4* marks the developing glume, *SEP1-2* the developing lemma, *PIN1a* the developing vasculature, *SPL14* the suppressed bract, *FZP* the glume axilla and *TCP24* the adaxial boundary. Complete spike images showing expression of these genes are presented in Figs. S10-S12.

300 The glume and lemma primordia, which became visible at W3.25, were included in cluster c7.
301 The similar expression profiles observed in this study, together with the gradual interconversions
302 between glumes and lemmas observed in the wheat *AP2L* mutants and overexpression lines in
303 previous studies [11, 13], suggest a similar origin. The MADS-box gene *SEP1-4* was
304 preferentially expressed in glumes and young meristems, whereas *SEP1-2* was preferentially
305 expressed in lemmas (Figs. 4C-D, S10B). The latter result is consistent with previous reports in
306 rice [43]. Glume and lemma primordia also showed preferential expression of *FUL2* (Fig. S9F)
307 and reduced expression of *OSH1* (Fig. S8E). The *SPL14* and *SPL13* genes preferentially
308 expressed in the suppressed bract region (c6), were also detected in the glume (Fig. S9B, D, Data
309 S2). Cluster c7 showed the highest correlation with the early meristem cluster c8 (R= 0.81), from
310 which glumes and lemmas originate. Relatively high correlations were also observed between c7
311 and the cell clusters identified in leaves (R= 0.63-0.75, Data S3), supporting the hypothesis that
312 glumes and lemmas are modified leaves [44].

313 Cluster c14, which is also correlated with c7 (R= 0.78) and c9 (R= 0.80, Data S3), is located at
314 the axillae of the developing glumes (Fig. 4C-D). In spike longitudinal sections, c14 cells
315 frequently appear as two small cell clusters flanking the central cluster c9, with the adaxial
316 region of c14 adjacent to c16 (Figs. 4C-D, S10, S11). This restricted localization is consistent
317 with previous *FRIZZY PANICLE* (*FZP*) *in situ* hybridization studies in rice and maize [45, 46].
318 Cluster c14 is defined by the expression of *FZP*, which is present in 65% of the c14 cells but
319 only in 1% of cells from other clusters (Fig. S10B, Data S2). Cells from cluster c14 also exhibit
320 preferential expression of the homeobox domain gene *WOX7* (Fig. S11A) and the PINOID-like
321 serine/threonine kinase *BIF2* (Fig. S11B), although both genes are also expressed outside of c14.
322 Loss-of-function mutations in *FZP* result in branch formation at the base of the spikes in both
323 wheat (*bh1* mutant) and barley (*com2* mutant) [47, 48], as well as conversions of determinate
324 SMs to indeterminate branch meristems in rice (*fzp*) [45] and maize (*bd1*) [46]. These results
325 suggest that *FZP* is required to prevent branch development from the glume's axillary meristems
326 [45].

327 In summary, the region proximal to the wheat SM encompasses several expression domains and
328 cell types, consistent with recent models proposing the basal region of the grass spikelets as an
329 important signaling center for spikelet development [49].

330 **1.4. Spikelet meristem (SM) and floret meristem (FM) regions**

331 At all three developmental stages, cells in the early SMs were assigned to cluster c8. However, in
332 the more developed central spikelets at the lemma primordia stage, meristems were assigned to
333 cluster c18 (FM, Figs. 1, S1). In the central spikelets, a group of cells forming a distinct band
334 between the FM and the lemma primordia was assigned to cluster c17 (Figs. 1, 4C).

335 At W2.5, early SMs exhibited preferential expression of *HAF1*, *BIF2*, (Figs. 4E, S12A), *ULT1*
336 and *WOX7* (Figs. 4F, S12C-D). *HAF1*, a C3HC4 RING domain-containing E3 ubiquitin ligase,
337 has no reported role in inflorescence development [50, 51]. By contrast, the other three genes are
338 well-established regulators of meristem function: *i*) *BIF2* regulates auxin transport and is
339 required for axillary meristem initiation and lateral primordia development in maize and rice
340 [52]; *ii*) *ULT1*, which encodes a SAND domain transcription factor, controls shoot and FM
341 activity in Arabidopsis [53, 54]; and *iii*) *WOX7* is a *WUSCHEL*-related homeobox gene known to
342 shapes inflorescence architecture in rice [55].

343 At W3.25, the gradients between the more developed central spikelets and the less developed
344 ones at the tip or base of the spike helped us organize gene expression within the SM over time.
345 *HAF1* and *ULT1* were expressed in both the early and more advanced meristems, whereas *BIF2*
346 and *WOX7* were restricted to the youngest SMs adjacent to the IM (Fig. S12B, S12D). At W3.25
347 these two genes were expressed in c14 and other clusters (Fig. S11). The MADS-box gene *AGL6*
348 was initially detected in the SMs of spikelets more developed than those expressing *BIF2* and
349 *WOX7*, but at an earlier stage than spikelets at the lemma differentiation stage at the central
350 region of the spike, which express floral homeotic genes (Fig. S12).

351 In the central spikelets at W3.25, *AGL6* was detected in a region between the lemma primordia
352 and the FM, consistent with its reported role in palea and lodicule development [14]. The
353 meristematic cells of these central spikelets were included in cluster c18, defined by the
354 preferential expression of floral homeotic genes *SEP3-1*, *SEP3-2* (Fig. S12E), *PI1*, *AP3* (Fig.
355 S12F), *AG1* and *AG2* (Figs. 4G, S12G, Data S2). *WAPO1* was also expressed in cluster c18 (Fig.
356 S12H) and partially overlapped with the adjacent cluster c17. Cluster c17 also showed
357 preferential expression of *LFY* (Fig. S12H, Data S2), a protein that physically interacts with
358 *WAPO1* [15]. Loss-of-function mutations in either of these genes led to the downregulation of
359 most floral homeotic genes, and severe floral defects [15, 16]. Taken together, these findings

360 suggest that the LFY-WAPO1 protein complex plays a crucial role in the spatial regulation of
361 proper floral development in wheat [15].

362 **1.5. Changes within the inflorescence meristem (IM) during the transition to a terminal**
363 **spikelet (IM→TS)**

364 The IM plays a central role in regulating spikelet number, an important grain yield component.
365 In wheat, the rate of lateral meristem formation and/or the timing of IM transition to a terminal
366 spikelet (IM→TS) determine the final SNS. Our previous studies have demonstrated that *LFY* or
367 *WAPO1* loss-of-function mutations reduce the rate of SM formation, leading to reduced SNS
368 [15]. The reductions in SNS were comparable in the single and combined *lfy wapo1* mutants,
369 suggesting that the LFY-WAPO1 protein complex is essential for establishing and/or
370 maintaining a normal rate of SM generation. However, *LFY* and *WAPO1* are co-expressed in the
371 IM only during the earliest stages of spike development (W1.5), implying that the early presence
372 of the LFY-WAPO1 protein complex may induce a persistent change in IM function [15].

373 Since changes in the timing of the IM→TS transition can also influence SNS, we focused on
374 genes involved in the early stages of this transition. No specific cell cluster was identified for the
375 IM, so we lacked a predefined list of genes preferentially expressed in this region. To address
376 this limitation, we manually delineated the IM region and compared the hybridization density of
377 the 99 genes in this region before the IM→TS (W2.5) and at the early stages of the transition
378 (W3.25). These comparisons were conducted within genes and, therefore, were not influenced by
379 probe hybridization efficiency. For each stage, we analyzed four sections. To verify that the
380 selected W3.25 sections had already initiated the IM→TS transition, we confirmed the
381 expression of *FZP* in the youngest lateral meristems (LM, Fig. S13). The selected region
382 encompassed both the IM and the two youngest LMs (IM+2L), which together will form the
383 terminal spikelet. Using the same approach, we previously observed a significant decrease in
384 *LFY* expression and a significant increase in *FUL2* expression between W2.5 and W3.25 in the
385 IM+2L region [15].

386 In this study, we identified 11 additional genes with significantly altered hybridization densities
387 (spots/100 μm^2) in the IM+2L region between W2.5 and W3.25 (FDR<0.05, Fig. S14, Data S5).
388 Among them, *SPL14* and *HAF1* exhibited a decrease in hybridization density during the

389 transition, with *SPL14* having the most pronounced effect (6.7-fold reduction, FDR *P*= 0.0059,
390 Figs. S13, S14, Data S5). Since SPL proteins directly regulate the expression of *SQUAMOSA*
391 MADS-box genes, we hypothesize that the reduced *SPL14* expression may be linked to the
392 increase in *FUL2* and *VRN1* expression in the same region observed during the IM→TS
393 transition [15]. Supporting this hypothesis, a mutation in a predicted *SPL*-binding site in the
394 *VRN1* promoter was previously associated with a significant increase in *VRN1* expression during
395 early spike development [56].

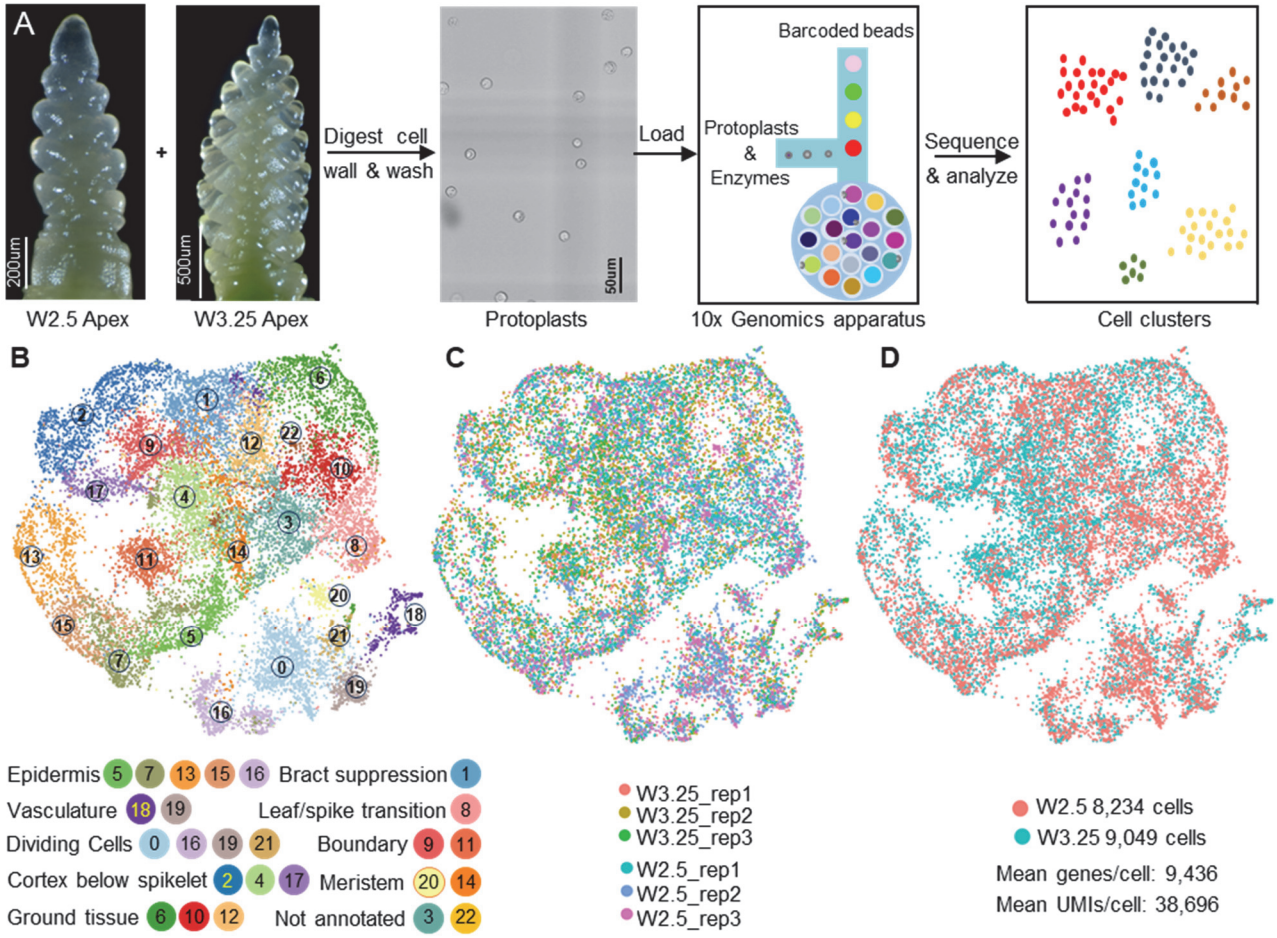
396 Among the genes upregulated at W3.25 in the IM+2L region, we identified two additional
397 MADS-box genes (*SOC1-2* and *SEP1-4*), four bZIP transcription factors (*bZIPC1*, *bZIPC3*,
398 *FDL2*, and *FDL6*), one bHLH transcription factor (*BC1*), one auxin efflux carrier (*PIN1b*), and
399 one receptor-like kinase (*CRN*) involved in the CLV-WUS negative feedback loop (Fig. S14).
400 Several of these genes have been associated with inflorescence or floral phenotypes. For
401 example, mutations in *bZIPC1* drastically reduce SNS in wheat [57]. In rice, mutations in the
402 ortholog of *SEP1-4* (*OsMADS5*) impact inflorescence branching [58], whereas mutations in *CRN*
403 affect flower development in Arabidopsis. In summary, this analysis identified multiple wheat
404 genes involved in the early IM→TS transition that will be prioritized for functional
405 characterization.

406 **2. Single-cell RNA sequencing (scRNA-seq) of the developing wheat spike**

407 To identify additional genes regulating wheat spike development, we conducted single-cell
408 RNA-seq (scRNA-seq) using protoplasts isolated from *Kronos* spikes at the W2.5 and W3.25
409 stages (Fig. 5A). We optimized the cell wall digestion method to enhance meristem cell
410 recovery, following our previous single-cell study on maize inflorescence stem cell [59] (see
411 Methods). Raw scRNA-seq data are available in the NCBI Sequence Read Archive (SRA) under
412 BioProject accession number PRJNA1199502. Reads were mapped to the Kronos reference
413 genome v1.1 (annotation v1.0) [60].

414 In total, we profiled 12,304 and 13,705 individual cells from three independent replicates at
415 W2.5 and W3.25 stages, respectively (Fig. S15, Data S6). Biological replicates exhibited high
416 reproducibility at both stages (Fig. S15A). We detected expression of 67,629 genes at W2.5 and
417 70,047 genes at W3.25 (Data S5), comparable to the number of genes detected in bulk RNA-seq
418 of Kronos spikes at the same developmental stages [61]. The mean and median unique molecular

419 indices (UMIs), the number of detected genes per cell, and other statistics were reported in Fig.
420 S15A and Data S6. Cell clustering using Seurat v5 [62] and integration of 26,009 cells from both
421 stages identified 25 distinct clusters (Fig. S15A).



422

423 **Fig. 5. Single-cell RNA-seq (scRNA-seq) analysis of wheat inflorescence spikes at the W2.5 and W3.25 stages.**
424 **A** Experiment workflow, including wheat protoplast isolation and subsequent scRNA-seq. **B-D** UMAP plots
425 showing cells clustering by **B** cell type annotations, **C** sample identity, and **D** developmental stage.
426

427 Nine clusters were identified with cells at either the G2/M or S phase, marked by *CYCB2* and
428 *HIS2A1* expression, respectively (Fig. S15B-C). Another cell cycle marker gene *CDC20* was
429 enriched in G2/M clusters, and its spatial expression profile highlighted the positions of dividing
430 cells (Fig. S15D-E). To improve the resolution of biologically relevant clusters, cells from the
431 nine cell-cycle clusters were removed, and the remaining 17,283 cells were re-clustered into 23
432 groups (sc0 to sc22). These clusters were visualized using two-dimensional (2D) uniform
433 manifold approximation and projection (UMAP) plots (Fig. 5B). To facilitate data visualization,

434 we generated a 3D-UMAP (https://dubcovskylab.ucdavis.edu/content/kronos-scrna-3d-clusters)
435 and a Shiny App (https://dubcovskylab.ucdavis.edu/content/kronos-integrated-clusters). We also
436 quantified relationships among the single-cell clusters by calculating Pearson correlations for all
437 cluster pairs, using combined normalized gene expression profiles across all cells in each cluster
438 (Data S7).

439 **2.1. Characterization of scRNA-seq cell clusters and comparison with spatial clusters**

440 The scRNA-seq clusters were annotated using multiple methods, as detailed in the Materials and
441 Methods section. Briefly, annotation methods included known marker genes (Data S1),
442 identification of genes preferentially expressed in each cluster (Data S8), KEGG pathway
443 annotation (Data S9), correlations among single-cell clusters (Data S7), and correlations with
444 spatial expression clusters from this study (Data S10) and from a published study in hexaploid
445 wheat (Data S11-13) [63]. Additional spatial information for single-cell cluster annotation was
446 obtained from an RNA-seq study of the base, middle, and top of the Kronos developing spike at
447 W3.0 (Fig. S16 and Data S14). These data, along with the final annotation of the single-cell
448 clusters, are summarized in Data S15. The use of similar names in the annotation of spatial and
449 single-cell clusters (sc) does not imply a perfect one-to-one correspondence, but a higher overlap
450 than with other clusters.

451 Spike samples used in the single-cell study lacked certain tissues and developmental stages
452 present in the smFISH analysis, including leaves (c1, c2, c5, c12, c20), the basal region below
453 the spike (c0), and the W1.5 stage (c15, Data S10). The low correlation values observed between
454 these seven smFISH clusters and all the single-cell clusters (Data S10) served as good negative
455 controls.

456 We identified two single-cell clusters, sc3 and sc22, with a high risk of being affected by the
457 protoplasting process based on two criteria. First, these clusters showed the highest proportion of
458 preferentially expressed genes not detected in RNA-seq studies of Kronos developing spikes
459 (this study and [61]). Approximately 19% of the preferentially expressed genes in sc3 and 28%
460 in sc22, were not detected at the same stages in Kronos spikes under normal conditions (Data
461 S15). The 235 genes detected in the scRNA-seq and not in the RNA-seq were enriched in stress
462 related GO terms (Data S16), further supporting the induction by the protoplasting process. The
463 preferentially expressed genes in sc3 and sc22 also showed a relatively high overlap with a list of

464 genes previously annotated as potentially affected by the protoplasting process [64] (sc3= 11%
465 and sc22= 13%, Data S15). The percentage of genes identified by these two criteria across the 22
466 single-cell cluster showed a high correlation (*R*= 0.81), pointing to a similar underlying process.
467 Cluster sc3 showed a high correlation with sc10 (*R*= 0.85, Data S7) and sc22 with cluster sc14
468 (*R*= 0.88, Data S7), suggesting that these clusters may contain cells similar to those in sc10 and
469 sc14 but potentially affected by the protoplasting process. Based on these results, we annotated
470 sc3 and sc22 as "unknown".

471 **2.2. Cell cycle enriched clusters**

472 Although cells from the initial cell-cycle clusters were removed in the subsequent scRNA-seq re-
473 analysis, four of the 22 final clusters still exhibited enrichment in cell-division marker genes
474 combined with markers for other tissues. The S-phase markers *HIS2A1-1* and *HIS2A1-2* were
475 enriched in clusters sc0, sc16, sc19, and sc21 (Fig. S17). Cluster sc16 was also enriched in G2/M
476 phase cyclin markers (*CYC7*), cyclin-dependent kinases (*CDKB2*), and the *GROWTH*
477 *REGULATING FACTORs GRF2*, *GRF4*, and their cofactor *GIF1* (Fig. S17). Gene IDs and
478 references for all marker genes are provided in Data S1.

479 Among the genes preferentially expressed in cluster sc21 (Data S8), those that exhibited
480 significant differences in the RNA-seq study among the bottom, middle, and tip of the spike
481 were enriched in genes expressed at higher levels at the basal region (Data S14-S15). This basal
482 region showed a green coloration that extended upward through the central part of the spike (Fig.
483 5A), and that was indicative of the presence of chlorophyll. This result was also supported by the
484 KEGG analysis, which indicated that sc21 was enriched in genes within the photosynthetic
485 pathway (Data S9). Based on the previous observations, cluster sc21 was annotated as "cell cycle
486 / photosynthetic ground tissue / basal region" (see section 2.5 for additional support).

487 The preferentially expressed genes in sc0, sc16, and sc19 (Data S8) were enriched in genes
488 expressed at higher levels in the distal third of the developing spikes (Data S15). This enrichment
489 was strongest in sc0 (100% of the genes), intermediate in sc16 (78%), and weak but significant
490 in sc19 (47%). Based on its distal position and enrichment in meristematic markers (section
491 2.10), sc0 was annotated as "cell cycle / meristematic / distal region". Cluster sc19 was also
492 enriched in vascular marker (see section 2.4) and was annotated as "cell cycle / vasculature /
493 enriched in spike tip". Finally, sc16 showed multiple epidermal markers (section 2.3) and KEGG

494 enrichment in cutin, suberin, and wax biosynthesis (Data S9), so it was annotated as "cell-cycle /
495 epidermis / enriched in spike tip".

496 **2.3. Epidermal clusters**

497 A similar enrichment in cutin, suberin, and wax biosynthesis was detected in the KEGG analysis
498 for cluster sc15, supporting its annotation as an epidermal cluster. In addition to sc15 and sc16,
499 clusters sc5, sc7, and sc13 were also annotated as epidermal clusters based on the preferential
500 expression of known epidermal markers, including *ONI* (synonymous with *KCS10*), *PEL3*, *SCR*,
501 *ROC2*, *ROC3*, *ROC7* (Data S1, Fig. S18). A list of additional genes preferentially expressed in
502 each epidermal cluster is provided in Data S8.

503 Our spatial expression study did not include epidermal markers, so we compared the single-cell
504 epidermal clusters to a recent hexaploid spatial transcriptomics study that identified two
505 expression domains, ED1 and ED4, associated with epidermal cells [63]. These two epidermal
506 expression domains had the highest correlations with sc7, sc13, sc15, and sc16 (Data S11-13),
507 supporting the epidermal annotation of these clusters. However, their correlations with sc5 were
508 intermediate (Data S11-13).

509 Epidermal clusters sc16 and sc7 were highly correlated with each other (R= 0.79) and were both
510 enriched in genes expressed at higher levels at the spike tip based on the RNA-seq data (Data
511 S14-S15). Genes encoding ribosomal proteins were expressed at high levels in both clusters
512 (Data S15), suggesting rapid cell growth and division. The correlations between cluster sc16 and
513 cell-cycle clusters sc0, sc19 and sc21 were consistently higher than those for sc7 (Data S7),
514 justifying the previous annotation of sc16 in section 2.2 as "cell cycle / epidermis / enriched in
515 spike tip".

516 Epidermal cluster sc7 was annotated as "epidermis / meristem / enriched in spike tip" based on
517 the above observations and its high correlations (Data S7) with the meristematic clusters sc14
518 and sc20 (section 2.10). The highest correlations observed for the spatial meristematic clusters c8
519 in tetraploid wheat (Data S10) and ED0 in hexaploid wheat (Data S11-S13) were with epidermal
520 clusters sc7 and sc16, and meristematic clusters sc14 and sc20. However, sc7 also showed high
521 correlations with expression domains adjacent to the SM (Data S10-S13). These results suggest
522 that sc7 may include epidermal cells surrounding the meristems and adjacent tissues.

523 Epidermal cluster sc15 was highly correlated with sc7 (*R*= 0.81) but differed from the latter in its
524 lower expression of genes encoding ribosomal proteins (Data S15). Cluster sc15 also differed
525 from sc7 in its higher correlation with suppressed bract cluster sc1 (*R*= 0.63) than with
526 meristematic clusters sc14 and sc20 (Data S7). The high sc15 and sc1 correlation was also
527 supported by the co-expression of *SPL13* and *SPL14* in these two clusters (Fig. S18G). When
528 sc15 was compared with the non-epidermal spatial clusters, it showed relatively high correlations
529 with suppressed bract clusters c6 in tetraploid wheat (*R*= 0.66, Data S10) and ED12 in hexaploid
530 wheat early stages of spike development (Data S11-S12). Based on these results, sc15 was
531 annotated as "epidermis / suppressed bract".

532 Epidermal clusters sc13 and sc5 showed higher correlations with sc15 (*R*= 0.77) and sc7 (*R*=
533 0.63, Data S7), respectively. These two epidermal clusters showed no strong preferential
534 expression region in the RNA-seq data and exhibited low levels of expression of genes encoding
535 ribosomal proteins. Since no consistent results were detected across the different correlation
536 matrices, these two clusters were annotated simply as "epidermis".

537 **2.4. Vascular clusters**

538 The vascular genes *APL*, *LAX4*, *CEN5*, *PIN1a*, and *PIN1b* characterized by smFISH (Fig. 3)
539 were all preferentially expressed in sc18 and sc19, along with additional known vascular
540 markers, including *JUL1*, *SHR*, *WAT1*, *XCP1*, *TOM5L3*, and *TMAAT* (Fig. S19, Data S1),
541 supporting their annotation as vascular clusters. This conclusion agreed with the relatively high
542 correlation of these two single-cell clusters with smFISH vascular cluster c13 in tetraploid wheat
543 (Data S10) and the spike vasculature expression-domain ED6 in hexaploid wheat [63] (Data
544 S11-13).

545 Both vascular clusters shared relatively high expression levels of *OSH1* and genes encoding
546 ribosomal proteins (Data S15). However, the genes preferentially expressed in sc19 (DataS8)
547 were expressed at higher levels at the tip of the spike (Data S15), whereas those preferentially
548 expressed in sc18 (Data S8) were expressed at higher levels at the base and center of the
549 developing spikes (Data S15). In addition, sc19 was enriched in *HIS2A* expression, whereas sc18
550 was not (Fig. S17). Based on these results, sc18 was annotated as "vasculature / enriched in spike
551 base and middle", and sc19 as "cell cycle / vasculature / enriched in spike tip"

552 Cluster sc18 showed two distinct regions: one preferentially expressing phloem markers (Figs.
553 S19B and D) and the other expressing xylem markers (Figs. S19C, E-F). Interestingly, the floral
554 homeotic genes *AP3* and *PI1* (Fig. S12F) were also detected in cluster sc18 (Figs. S20A-B). A
555 closer examination of the smFISH profiles revealed that these two MADS-box genes were co-
556 expressed with the vascular marker *APL* at all three developmental stages (Figs. S20C-E).
557 Expression of class-B MADS-box genes in the developing vasculature has been reported in a few
558 eudicot plant species [65]. Our findings extend this conserved function to monocots.

559 **2.4. Central spike clusters**

560 In addition to vascular and pro-vascular cells, the central region of the developing spike contains
561 ground tissue cells, which correspond to cluster c3 (Fig. 1 central spike) in our smFISH study,
562 and expression domains ED3 (developing rachis) and ED5 (rachilla) in the hexaploid wheat
563 study [63]. These ground tissue cells showed a green coloration indicative of the presence of
564 chlorophyll, which was more abundant at the base of the spike (Fig. 5A). The KEGG analysis
565 indicated that genes preferentially expressed in single-cell clusters sc6 and sc10 were enriched in
566 photosynthesis pathways (Data S9). Eighty two percent of the top 50 genes preferentially
567 expressed in sc6 were annotated with photosynthetic-related functions (Data S8), including
568 *CAB2R* and *LHCB1.1* (Fig. S21). The expression levels of these photosynthetic genes across all
569 clusters are presented in a heat-map in Fig. S22.

570 The highest average enrichment in photosynthetic genes was detected for sc6 followed by the
571 highly correlated (Data S7) cell-cycle cluster sc21 (Fig. S22, Data S15). The latter result
572 supported the previous annotation of sc21 in section 2.2 as "cell cycle / photosynthetic ground
573 tissue / basal region". Single-cell clusters sc10 and sc12 (at lower level) also showed enrichment
574 in photosynthetic genes (Data S15 and Fig. S22). A large proportion of the genes preferentially
575 expressed in sc10 (44%) and sc12 (36%) were also preferentially expressed in sc6 (Fig. S23),
576 supporting the connections among these clusters.

577 The genes preferentially expressed in sc6, sc10, and sc12 showed higher levels of expression at
578 the basal region of the developing spikes based on the RNA-seq data (Data S14-S15). This
579 location was further supported by the enrichment of *FWL3* and *SVP1* in both sc6 and sc10,
580 which are expressed in the central and basal parts of the spike (Figs. S23A-C and H). In addition,
581 both sc10 and sc12 were enriched in *OSH1*, a gene highly expressed in the central spike region

582 (Fig. S8). Cluster sc6 was enriched in additional genes expressed in the central spike region,
583 including *AP3*, *bZIPC3*, *FDL15*, *FT2*, *OSH1*, *SPL3*, and *VRT2* in tetraploid wheat (Figs. S8 and
584 S21), and both *CAB2R* and *LHCB1.1* in hexaploid wheat [63] (Fig. S21A). The central location
585 of sc6, sc10, and sc12 was further supported by higher correlations with the spatial cluster c3
586 (spike central, Data S10) and with the hexaploid wheat expression domains ED3 (developing
587 rachis) and ED5 (rachilla, Data S11-13) [63] than with most other spatial clusters or expression
588 domains. Based on the above information, clusters sc6 and sc10 were annotated as "central spike
589 / photosynthetic ground tissue / spike base", and sc12 as "tentative ground tissue / enriched spike
590 base".

591 **2.5. Transition zone between leaves and spike**

592 At the base of the developing spikes, a drastic change occurs in the relative dominance of the leaf
593 and axillary meristems, and in the identity of the latter (see section 1.2.). This region is
594 represented in the smFISH study by cluster c19 (Figs. 1, S4), which showed high correlations
595 with the suppressed bract cluster c6 and the central spike clusters c3 (Data S3). Similarly, high
596 correlations were observed between single-cell clusters sc8 (transition leaf-spike) and both sc1
597 (suppressed bract) and sc10 (ground tissue / spike base, Data S7). Cluster sc8 also exhibited high
598 correlations with spatial clusters c6, c3 and c19 (Data S10). The strong correlations between the
599 transition zone and repressed bract clusters in both studies align well with the strong repression
600 of the leaf primordia in the transition zone. The RNA-seq data indicated that single-cell cluster
601 sc8 was enriched in genes expressed at higher levels at the base of the spike (Data S15).

602 In addition to their similar spatial localization, spatial cluster c19 and single-cell cluster sc8
603 shared preferential expressed genes. Those genes include *TB1*, *TB2*, *SPL14*, *GNI2*, *IDD10*, and
604 *GA20ox1* from our smFISH study, along with *LEC1* from the hexaploid study [63] (Fig. S24,
605 Data S2 and S8). Based on these results, sc8 was annotated as "transition leaf-spike / spike
606 base".

607 **2.6. Suppressed bract cluster**

608 The suppressed bract cluster c6 was consistently detected in the smFISH study at an abaxial
609 position at the base of each developing spikelet (Fig. 1). This spatial cluster showed the highest
610 correlation with the single-cell cluster sc1 (R= 0.74, Data S10), which was uniformly distributed

611 across the spike based on the RNA-seq data (Data S15). *SPL13* and *SPL14* are strong differential
612 markers for spatial cluster c6, (Data S2), and both were enriched in sc1 (Fig. S25). The role of
613 *SPL14* in bract suppression has been functionally validated in hexaploid wheat mutants [4].
614 Based on these results, the sc1 cluster was annotated as "suppressed bract".

615 In spite of their similar annotation, the correspondence between single-cell cluster sc1 and spatial
616 cluster c6 is likely not perfect. The bract suppression marker *NECK LEAF1* (*NL1*) and several
617 *YABBY* markers (*YAB1*, *YAB3*, *YAB4*, and *YAB7*), which were recently reported to hybridize to
618 the suppressed bract region in hexaploid wheat [63], were not preferentially expressed in sc1. In
619 our scRNA-seq study, these genes were preferentially expressed in the epidermal clusters sc7
620 and sc16, while *NL1* was also enriched in vascular clusters sc18 and sc19. In addition to its
621 highest correlation with spatial cluster c6, sc1 exhibited high correlations with neighboring
622 spatial clusters c3 (central spike) and c16 (adaxial boundary, Fig. 4 and Data S10). These results
623 suggest that sc1 may include cells from the suppressed bract and adjacent expression domains,
624 and that some bract suppression-related cells may also be present in clusters sc7 and sc16.

625 **2.7. Boundary clusters**

626 Among the genes analyzed in the smFISH experiment, *FZP* (c14) and *TCP24* (c16) showed the
627 most specific expression domains (Figs. 4, S10, Data S2). *FZP* was expressed at the glume
628 axilla, while *TCP24* was expressed at the adaxial boundary below each spikelet (Fig. 1). These
629 precise expression domains enabled us to identify co-expressed genes using Pearson correlations
630 between each of these genes and all other genes across all cells included in the scRNA-seq study.
631 As expected, most genes had correlations with *FZP* and *TCP24* near 0, so our analysis focused
632 on genes with *R* values exceeding four standard deviations above the mean. The highest
633 correlations for *FZP* and *TCP24* were observed with their respective homeologs, which served as
634 a positive control (Data S17-S18).

635 We found 608 genes significantly co-expressed with *FZP* and 606 with *TCP24*, including 226
636 shared by both. Among the shared genes, we identified the well-studied *CUP-SHAPED*
637 *COTYLEDON 3* (*CUC3*, Data S17-S18), which encodes a NAC-domain transcription factor
638 involved in meristem / organ boundary specification in rice [66] and Arabidopsis [67]. *In situ*
639 and scRNA-seq experiments in maize demonstrated that *CUC3* is expressed in both the *FZP* and
640 *TCP24* domains [64, 68, 69]. The *MULTI-FLORET SPIKELET 1* (*MFS1*, synonymous with

641 *DUO* and *ERF130*) is also significantly co-expressed with both *FZP* and *TCP24* (Data S17-S18),
642 a result validated by spatial expression data (Fig. S26A-B). Mutations in *MFS-B1* in wheat are
643 associated with mild supernumerary spikelets and increased grain number per spike,
644 documenting its important role on spike development [70].

645 *FZP* was also co-expressed with *WOX-9C* (Data S17), a close paralog of *WOX7* (synonymous
646 with *WOX9A*) [71, 72] (Fig. S11). In our spatial data, *WOX7* expression overlapped with *FZP* at
647 W3.25, but was also detected in other regions where *FZP* was absent (Fig. S11), likely reducing
648 the correlation between these two genes. Analysis of hexaploid wheat spatial transcriptomics
649 data [63] showed that *WOX9C* (TraesCS1B02G427400) was also expressed in the glume axilla
650 during the early spikelet development and in additional tissues at later stages. In Arabidopsis,
651 RNA interference of *WOX9* in reproductive meristems resulted in the conversion of flowers into
652 inflorescence-like structures [73], resembling the conversion of spikelets to branches observed in
653 *fzp* mutants in wheat and other grass species [45-48].

654 Three lines of evidence supported the annotation of single-cell cluster sc11 as "glume axilla".
655 First, the correlation between sc11 and the spatial clusters c14 (glume axilla, R= 0.41) was the
656 highest among the sc11 correlations (Data S10). Second, *FZP* and *WOX9C* were both
657 preferentially expressed in sc11 (Fig. S27). Finally, RNA-seq results showed that sc11 was
658 significantly enriched in genes with higher expression levels at the center of the spike (Data
659 S15), a pattern also observed for *FZP* and *WOX9C* (Data S17). However, we noticed that only
660 37% of the sc11 cells expressed *WOX9C* and 6% expressed *FZP*, suggesting that this cluster may
661 include cells from other spatial domains. A potential candidate is c9 (SM base center, Fig. 4),
662 which is adjacent to c14 and showed the second-highest correlation with sc11 (R= 0.34, Data
663 S10).

664 *TCP24* showed high co-expression with its close paralog *TCP22* (Data S18), which in hexaploid
665 wheat [63] is expressed in a similar domain as the one reported for *TCP24* in this study. In the
666 maize male inflorescence, the orthologue of *TCP24, BRANCH ANGLE DEFECTIVE 1* (*BAD1*),
667 is preferentially expressed in the branch axilla (the pulvinus) and influences branch angle [74].
668 The barley ortholog *HvTB2* (also known as *COMPOSITUM*) is also expressed at the adaxial
669 boundary of the spikelet meristem. Loss-of-function mutations in this gene result in branched
670 spikes [42] providing functional validation.

671 *TCP24* exhibited high co-expression with *BARREN STALK FASTIGIATE1* (*BAF1*, Data S18),
672 which is expressed in a narrow domain adaxial to developing axillary meristems in maize [69].
673 Mutations in this gene resulted in the absence of ears or their fusion to the main stalk, supporting
674 its role in demarcating the boundary of the developing axillary meristem. *TCP24* was also co-
675 expressed with *LAX PANICLE1* (*LAX1*) (Data S18), a gene required for the initiation of aerial
676 lateral meristems. The rice *LAX1* gene and its maize ortholog *BARREN STALK1* (*ZmBA1*) were
677 both detected in the adaxial boundary region of the spikelet meristems [75, 76]. A strong
678 correlation was also observed between *TCP24* and *RAMOSA2* (*RA2*, *R*= 0.13-0.21), which is
679 expressed in a group of cells predicting the position of the axillary meristem formation [77]. In
680 summary, the *TCP24* co-expression analysis identified several genes that share a similar
681 expression domain and that are involved in defining boundaries and regulating lateral meristem
682 formation in other grass species.

683 *TCP24* was preferentially expressed in single-cell clusters sc9 (~24% of cells) and sc13 (~12%
684 of cells, Fig. S28). These clusters were highly correlated with spatial cluster c16, which
685 represents the adaxial boundary expression domain, as well as with adjacent clusters c6, c9, and
686 c14 (Data S10). UMAPS and bubble plots of *TCP24* and its highly co-expressed genes
687 confirmed their preferential expression in sc9 and sc13 (Fig. S28). The significant co-expression
688 of *TCP24* and *NAL1* was validated using spatial expression data (Fig. S26C-E), which showed
689 that most cells expressing *TCP24* also expressed *NAL1.* This figure also revealed *TCP24*
690 expression in some epidermal cells, potentially explaining its preferential expression in
691 epidermal cluster sc13.

692 Additional genes co-expressed with *TCP24* are listed in Data S18, and markers preferentially
693 expressed in sc9 and sc13 are provided in Data S8. These two methods complemented each other
694 in the identification of genes previously identified in the adaxial boundary domain in other grass
695 species and predicted genes of unknown function that have the potential to expand our
696 understanding of these intriguing boundary domains.

697 **2.8. Cortex cells below developing spikelets**

698 In the wheat spike, the cortex consists of a layer of cells situated between the vascular tissues and
699 developing spikelets (Fig. 1). The parenchymatic cells in this region are expected to resemble
700 those in the central ground tissue. As shown in Fig. 4C, the cell layer between the vasculature

701 and the SM base includes ground tissue cells from the central cluster c3 as well as the innermost
702 cells from clusters c6 (suppressed bract), c9 (SM base-center), c16 (adaxial boundary) and c14
703 (glume axilla). Since cortical cell show no greening and are not meristematic, clusters including
704 these cells are expected to show low enrichment in photosynthetic genes and reduced expression
705 of *OSH1* or ribosomal protein-encoding genes.

706 Clusters sc2, sc4, and sc17 exhibited these characteristics. They showed high correlations with
707 the spatial clusters located proximal to the developing spikelets in tetraploid wheat (c3, c6, c9,
708 c14, and c16, Data S10) and with the corresponding spatial clusters in hexaploid wheat (ED0,
709 ED3, ED5, and ED11, Data S11-S13) [63]. They also exhibited low expression levels of
710 photosynthetic and meristematic genes (Data S15) and showed uniform expression along the
711 spike based on the RNA-seq data. Taken together, these results suggested that sc2, sc4, and sc17
712 may correspond to the cortical cells below the spikelets. Cluster sc17 showed high correlations
713 with both sc2 and sc4 (R= 0.71-0.72, Data S7).

714 The annotation of sc2 as a cortex cluster was further supported by the preferential expression of
715 *LAX1* and *DEPRESSED PALEA1* (*DP1*) (Fig. S29), two genes previously shown to be
716 preferentially expressed in the adaxial boundary below the spikelets [68, 75, 76, 78]. Two
717 additional genes, *KAN1* and *RL9*, expressed in the adaxial boundary region below the spikelet in
718 hexaploid wheat [63], were also preferentially expressed in sc2 (Fig. S29). Although these genes
719 were associated with the adaxial boundary they exhibited stronger expression in sc2 than in the
720 adaxial boundary cluster sc9 suggesting a differentiated function (Fig. S29). Additional markers
721 enriched in the correlated sc2 and sc13 clusters (Data S7) include *ABCG11* [79] and *SOT17* [80]
722 and are listed in Fig. S29 and Data S8.

723 Single-cell cluster sc4 showed higher correlations with spatial clusters c9 (the central cortex
724 region below the spikelet) and c14 (glume axilla) than with other spatial clusters (Data S10). A
725 more central location of sc4, relative to sc2 and sc17, was supported by the 2.4 to 3.4-fold higher
726 expression of *OSH1* in sc4 than in sc17 and sc2, respectively (Figs. S8, S30, Data S15). In
727 addition, the cortex marker gene *NPF3.1* (also known as *NPF6.4-LIKE2* in maize [64]) was
728 enriched in sc4 (Fig. S30F-G), suggesting that this cluster corresponds to cortical cells.
729 Additional genes preferentially expressed in sc4 are presented in Fig. S30 and Data S8.

730 Overall, the correlations between single-cell cluster sc17 and the spatial clusters were lower than
731 those observed with sc2, but they were still higher for the contiguous cortex clusters than for the
732 rest of the spatial clusters (Data S10). We annotated sc17 as "tentative cortex region below
733 spikelet lateral regions" due to the lack of validation with marker genes of known spatial
734 distribution. Additional genes preferentially expressed in sc17 and neighboring clusters were
735 listed in Fig. S30 and Data S8.

736 **2.9.Meristem clusters**

737 The spatial cluster c8 (early meristem) was characterized by the preferential expression of *ULT1*,
738 *HAF1, WOX7* and *BIF2* (Fig. S12, Data S2). In the UMAP graphs, these four genes were
739 expressed in multiple single-cell clusters, but all shared expression in cluster sc20 (Fig. S31).
740 The *WOX7* and *BIF2* genes were also expressed in single-cell clusters sc11 and sc1 (Fig. S31D-
741 E), mirroring their expression in the early meristems at W2.5 (Fig. S12A, C) and their later
742 colocalization with *FZP* at W3.25 (Fig. S11, S12D).

743 The limited overlap between the cells expressing *ULT-A1* (red) and *ULT-B1* (blue) in the UMAP
744 (Fig. S31B) was reflected in a low correlation between the two *ULT1* homeologs in the co-
745 expression analysis (Data S19). These results indicate a functional differentiation between the
746 two homeologs. This was also evident in the bubble plot in Fig. S11G, where *ULT-A1* was
747 enriched in cluster sc20 and *ULT-B1* was not. Interestingly, among the 100 genes with highest
748 correlations with *ULT-A1*, 82 encoded ribosomal proteins (Data S19), suggesting a high rate of
749 protein production and rapid cell growth and division. This enrichment was not observed for
750 *ULT-B1*, indicating that *ULT-A1* is the homeolog predominantly expressed in meristematic cells.
751 Among the 22 single-cell clusters, sc20 showed the highest expression levels of ribosomal
752 protein genes in the heatmap (Fig. S32, Data S15). Taken together these results indicate that the
753 sc20 cluster includes meristematic cells.

754 The presence of meristematic cells in sc20 was also supported by the correlations between the
755 meristematic spatial clusters and sc20. The correlations between spatial cluster c8 (early
756 meristem) in tetraploid wheat were higher for sc20 and the associated epidermal-meristem-cell-
757 cycle clusters sc7 and sc16 (Data S10) than for other single-cell clusters. In hexaploid wheat
758 [63], the highest correlations between the meristematic expression domain ED0 (Data S11-S13)
759 were observed for sc20 at the earliest stage (Data S11) and for sc20 and sc5 (rachilla) at the two

later stages (Data S12-S13). However, when correlations were analyzed from the other direction, sc20 also showed high correlations with spatial clusters located proximally to the SMs in tetraploid wheat (c3, c6, c9, c14, and c16, Data S10) and hexaploid wheat (ED3, ED5, and ED11, Data S11-S13). These results suggest that, in addition to meristematic cells, sc20 may include cells from the region between the SM and its adjacent proximal clusters.

Based on the previous results, and the enrichment of sc20 in genes expressed in the distal one-third of the spike (Data S15), sc20 was annotated as "meristem / enriched in spike tip". Among the additional genes preferentially expressed in sc20, we have prioritized chromatin remodeling genes *DDM1a*, *BTBN16*, and *CHR11*, and auxin response genes *AFB5*, *SAUR25*, and *SHI1* for future functional characterization (Fig. S31G and Data S1).

In addition to sc20, the unannotated cluster sc14 also exhibited some meristematic characteristics, including intermediate expression levels of *OSH1* (Fig. S31G) and of the ribosomal protein genes (Fig. S32, Data S15). When compared with the spatial clusters, the correlations for sc14 were higher with the five clusters detected below the SMs than with other spatial clusters in tetraploid wheat (Data S10), and with ED0 (meristem/meristem base), ED11 (adaxial boundary), ED3, and ED5 (developing rachis/rachilla) in hexaploid wheat (Data S11-13). These results, along with the enrichment of sc14 in the meristematic MADS-box gene *AGL6* (palea development, Fig. S33B), supported the annotation of sc14 as a "region proximal to SM". Additional genes preferentially expressed in sc14 are listed in Data S8.

Co-expression analysis of *AGL6* revealed significant co-expression of several floral homeotic genes, including *SEP3-1, SEP3-2*, *WAPO1*, *AG1*, and *PI1* (Data S20). These floral meristem genes were detected in clusters sc20 in the UMAP graphs, albeit at low levels (Fig. S33). Among the 100 genes with highest co-expression values with *AGL6*, 43 were annotated as genes encoding ribosomal proteins (Data S20). A similar enrichment of ribosomal protein encoding genes has been reported in a single-cell study of maize stem cells [59]. These results provide additional support for the presence of meristematic cells in sc14 and sc20.

3. CONCLUSIONS

Our smFISH and scRNA-seq data complemented each other, providing valuable tools and rich datasets for studying wheat spike development. The spatial study identified several genes

789 expressed in the intercalary stem meristem that influence stem length and width, which are both
790 important agronomic traits. The smFISH results also revealed consistent and previously
791 unknown expression domains in the region proximal to the developing wheat spikelets. Some of
792 these expression domains likely correspond to the proposed interacting "signaling centers"
793 adjacent to the meristems, which have been hypothesized in maize to regulate SM development
794 and determinacy [81].

795 The smFISH experiments provide both quantitative and precise spatial information but their cost
796 limits the number of genes that can be analyzed. This limitation can be overcome by
797 complementing spatial transcriptomics with scRNA-seq of the same tissues and developmental
798 stages. The spatial transcriptomics informs the annotation of scRNA-seq clusters, and once these
799 connections are established, the scRNA expands the limited list of genes that can be included in
800 smFISH studies. Using this approach, we identified multiple genes co-expressed with *FZP* (sc11)
801 and *TCP24* (sc9) that were previously detected in maize in similar domains, as well as previously
802 uncharacterized genes. The uncharacterized genes from these boundary clusters, together with
803 those from meristematic cluster sc20, provide the opportunity to expand our understanding of the
804 meristems and signaling centers located beneath the developing spikelets and to enhance our
805 ability to engineer more productive wheat spikes.

806

807 **4. MATERIAL AND METHODS**

808 **4.1. Spatial Expression**.

809 ***4.1.1. Sample preparation***

810 Developing wheat spikes from the tetraploid wheat Kronos were collected at three
811 developmental stages: early transition from the vegetative to the reproductive stage (W1.5), late
812 double ridge stage before the development of glume primordia (W2.5), and lemma primordia
813 stage, when the central spikelets initiated the transition to the floral meristem (W3.25). After
814 collection, samples were fixed in 4% paraformaldehyde (PFA), dehydrated, and embedded in
815 paraffin. Longitudinal sections (10 µm-thick) from the central plane of the developing spikes
816 were positioned within the capture areas of the slides. The slides were dried overnight at 37°C,
817 followed by a 10-minute bake at 50°C to enhance section adhesion. The sections were then

818 deparaffinized, permeabilized, and refixed. After complete dehydration, they were mounted
819 using SlowFade-Gold Antifade reagent, covered with a thin glass coverslip, and shipped on dry
820 ice to Resolve BioSciences (Monheim am Rhein, Germany) for analysis.

821 Spatial transcriptomics analyses were performed using Molecular Cartography™ and a 100-plex
822 combinatorial single-molecule fluorescence in-situ hybridization. Upon arrival, tissue sections
823 were washed twice in 1x PBS for two minutes, followed by one-minute washing in 50% and
824 70% ethanol at room temperature. Ethanol was removed by aspiration, and DST1 buffer was
825 added followed by tissue priming and hybridization. Tissues were primed for 30 minutes at
826 37°C, followed by hybridization for 48 h using probes specific to the target genes. The following
827 day, samples were washed to remove excess probes and fluorescently tagged in a two-step color
828 development process. Regions of interest were imaged as described below, and fluorescent
829 signals were removed during decolorization. Color development, imaging, and decolorization
830 were repeated across multiple cycles to generate a unique combinatorial code for each target
831 gene, derived from raw images as described below.

832 ***4.1.2. Probe design***

833 For each of the selected tetraploid wheat genes, we selected the homeolog with higher expression
834 based on published RNA-seq data for Kronos spike development [61]. We also provided Resolve
835 Biosciences with the sequences of the lower-expressing homeologs to exclude them during probe
836 specificity evaluations conducted against all the coding sequences of the wheat genome (Ref Seq
837 v1.1, Data S1). Therefore, although probes were designed based on the sequence of the highest-
838 expressing homeolog, they are not genome-specific and can detect both homeologs for each
839 gene.

840 Probes for 100 genes were designed using Resolve Biosciences' proprietary design algorithm
841 and gene annotations from the Chinese Spring RefSeqv1.1 database. In brief, probe design was
842 conducted at the gene level. For each targeted gene, all full-length protein-coding transcript
843 sequences from the ENSEMBL database were utilized as design targets. To identify potential
844 off-target sites, searches were confined to the coding regions. Each target sequence underwent a
845 single scan for all k-mers, favoring regions with rare k-mers as seeds for full probe design. A
846 probe candidate was generated by extending a seed sequence until reaching a certain target
847 stability. Sequences known to cause problems experimentally were eliminated. Following these

848 initial screens, each retained probe candidate was aligned with the background transcriptome,
849 and probes with stable off-target hits were discarded. Specific probes were then scored based on
850 the number of on-target matches (isoforms), giving preference to principal isoforms over others.
851 A bonus was added if the binding site was within the protein-coding region. From the pool of
852 accepted probes, the final set was selected from the highest-scoring probes.

853 ***4.1.3. Imaging***

854 Samples were imaged by Resolve Biosciences using a Zeiss Celldiscoverer 7 (CD7) with a 50×
855 Plan Apochromat water immersion objective having an NA of 1.2 and a 0.5× magnification
856 changer, resulting in a final magnification of 25×. Standard CD7 LED excitation light sources,
857 filters, and dichroic mirrors were used in conjunction with customized emission filters optimized
858 for detecting specific signals. The excitation time per image was fixed at 1,000 ms for each
859 channel, 20 ms for DAPI, and 1 ms for Calcofluor White. A z-stack was captured at each region
860 with a z-slice spacing according to the Nyquist-Shannon sampling theorem. A custom CD7
861 CMOS camera (Zeiss Axiocam Mono 712, 3.45 µm pixel size) was employed. For each region, a
862 z-stack was acquired for each of two fluorescent colors in each imaging round. A total of 8
863 imaging rounds were performed for each position, resulting in 32 z-stacks per region. The
864 entirely automated imaging process per round, which included water immersion generation and
865 precise relocation of regions for imaging in all three dimensions, was implemented through a
866 custom Python script utilizing the scripting API of the Zeiss ZEN software (Open application
867 development). Imaging for the cell wall-specific stain, Calcofluor White, was conducted after the
868 primary 8-round imaging.

869 ***4.1.4. Spot segmentation and preprocessing***

870 The algorithms for spot segmentation were written in Java and are based on the ImageJ library
871 functionalities. Only the iterative closest point algorithm is written in C++ based on the
872 libpointmatcher library (https://github.com/ethz-asl/libpointmatcher).

873 In the initial image processing steps, all images underwent background fluorescence correction.
874 The allowed number of maxima, representing the brightest points in the images, was determined
875 based on the slice's area and an empirically optimized factor. Each image slice was processed
876 independently. Maxima without neighboring counterparts in adjacent slices (referred to as z-

877 groups) were discarded. An iterative filtering process adjusted thresholds for factors like absolute
878 brightness, local background, and background of the periphery, all based on 3D image volume.
879 Only maxima remaining in z-groups of at least 2 were retained. These maxima, with the highest
880 absolute brightness within their z-groups, were documented as 3D-point cloud features.

881 For signal segmentation and decoding, the raw data images were aligned using transformation
882 matrices derived from the feature point clouds. An iterative closest point cloud algorithm
883 minimized error between point clouds. Point clouds from each round were aligned to the round
884 one reference, and the corresponding point clouds were stored for downstream use. Aligned
885 images were used to create pixel profiles with 16 values, representing two color channels across
886 eight imaging rounds. Profiles were filtered for variance from zero, normalized by total profile
887 brightness.

888 Matched pixel profiles were assigned IDs, and neighboring pixels with the same ID formed
889 groups. These pixel groups underwent further filtering, considering group size, adjacent pixel
890 count, and dimensions with a size of two pixels. Local 3D-maxima in groups were identified as
891 potential transcript locations. Maxima were filtered based on expectations from raw data images.
892 The remaining maxima were evaluated by fitting to corresponding codes, and they were
893 documented as transcripts. Specificity, assessed by the ratio of experiment-matching signals to
894 non-experiment-matching signals, was used to estimate false positives. This comprehensive
895 process enabled precise image analysis and transcript identification. Final image analysis was
896 performed in ImageJ using the Polylux tool plugin from Resolve BioSciences to examine
897 specific Molecular Cartography^{TM} signals. Images, programs, and gene results are deposited in
898 Zenodo [82].

899 ***4.1.5. Cell segmentation and clustering***

900 We identified cell boundaries from the DAPI-stained images using the cell segmentation
901 program QuPath [22]. To account for different cell sizes and shapes, we defined different regions
902 in each image, applying region-specific parameter sets. Across all regions we used the
903 parameters pixel size = 0, background radius = 8 µm, intensity threshold = 5, and opening by
904 reconstruction with the "include cell nucleus" parameter disabled. For regular, well-defined cells,
905 "split by shape" was activated, and the cell expansion was set to 7 µm, while for irregular,
906 elongated cells, "split by shape" was turned off, and the cell expansion was set to 5 µm. The cell

907 walls, visualized with Calcofluor White in the same sections, served as controls. The six images
908 were merged into a single composite image and converted into a format suitable for clustering
909 using ImageJ.

910 Hybridization signals per cell and clustering were calculated using the Resolve Biosciences
911 "Recognize" software with the following parameters: Min cell = 5, Min Features = 1, Min
912 transcripts = 1, Dimensions = 20, Resolution = 0.75, and Selected genes = 99. "Recognize"
913 utilizes Seurat v4 [83] as the backend, with the following methods: NormalizeData
914 (normalization.method= "LogNormalize", scale.factor= 1), FindVariableFeatures, ScaleData,
915 RunPCA, FindNeighbors (dims= 1:20), FindClusters (resolution= 0.75), RunUMAP (dims=
916 1:20, n.components= 3, and FindAllMarkers. The R script to replicate the analysis using Seurat
917 v5 can be found in Zenodo [82]. The average log2 fold changes of marker genes were generated
918 locally using Seurat v5 [62], which implements a pseudocount at the group level instead of the
919 cell level as in Seurat v4.

920 Cells from the three developmental stages and two replications were clustered together. Clusters
921 are represented with the same color across stages and replicates. To facilitate visualization of the
922 individual cell clusters, we also present them separately in Fig. S4. The Seurat object was
923 deposited in Zenodo [82].

924 **4.2. RNA-seq of different parts of the developing wheat spike**

925 Kronos spikes at the glume primordia stage (W3.0) were sectioned into three parts (Fig S16) and
926 placed into 500 µl of DNA/RNA Shield solution in 1.5-ml Eppendorf tubes. Approximately 20
927 to 25 apices were used per replicate, and three replicates were collected for each section.
928 Samples in DNA/RNA Shield were stored at 4℃ prior to RNA extraction.

929 We followed the published protocol by Marina Millán Blánquez for RNA extraction
930 (dx.doi.org/10.17504/protocols.io.36wgq7kj3vk5/v1) with minor modifications. Briefly, 400 µl
931 of DNA/RNA Shield solution was removed, and the tissues were homogenized with plastic
932 pestles in the remaining 100 µl of solution. RNAs were then extracted using TRIzol reagent,
933 cleaned, and concentrated with the RNA Clean & Concentrator-5 kit from Zymo Research. RNA
934 samples were shipped to Novogene for quality control, library preparation, and sequencing on
935 the NovaSeq X Plus (PE150) platform.

936 Data have been deposited in bioproject number PRJNA1217025. The three replicates from each
937 region were clustered together in a principal component analysis, indicating high reproducibility
938 among biological replicates (Fig. S16). Of the 49,786 detected genes, 13,024 exhibited
939 significant differences among regions in Tukey tests (Data S14). For each region we calculated
940 the average TMM per region and performed Tukey tests. The genes showing significant
941 differences and the corresponding Tukey tests are listed in Data S14. Scripts for RNA-seq
942 mapping, principal component plot, and Tukey tests were deposited in Zenodo [82].

943 **4.3. Single-cell RNA sequencing**

944 ***4.3.1. Protoplast preparation and 10x Genomics library construction and sequencing***

945 We used a cell wall digestion method optimized to preserve meristem cells in maize
946 inflorescence [59]. Briefly, the inflorescence spike was finely dissected (Fig. 1A) under a
947 microscope and digested in a MOPS buffer-based enzyme mix (10 mM MOPS, pH 7.5; 10 mM
948 L-Arginine HCl, pH 7.5; 0.65 M mannitol; 1.5% Cellulase RS; 1.5% Cellulase R10; 1.0%
949 Macerozyme; 1.0% hemicellulose; 0.1% BSA; 5 mM β-mercaptoethanol; 1 mM $CaCl_2$) for 2
950 hours. Before adding $CaCl_2$, β-mercaptoethanol, and BSA, the solution was heated to 55°C for
951 10 minutes.

952 The digested cells and enzyme mix were passed through a 30 µm pluriStrainer (pluriSelect,
953 Catalog No. 43-50030-50) for filtration. The filtered protoplasts were pelleted by centrifugation
954 at 350g for 2 minutes at 4°C and washed with wash buffer (10 mM MOPS, pH 7.5; 10 mM L-
955 Arginine HCl, pH 7.5; 0.65 M mannitol). The protoplasts were pelleted again by centrifugation
956 at 350g for 2 minutes at 4°C. The clean protoplast pellets were gently resuspended in wash
957 buffer, and viability was checked using trypan blue. High-viability protoplasts (≥ 70%) were
958 immediately loaded into a 10x Genomics Chromium System using Next GEM Single Cell 3' Kit
959 v3.1 chemistry. The scRNA-seq libraries were sequenced on the Element Bio AVITI platform
960 using the 10x Genomics recommended sequencing format, with 122 cycles used for Read 2
961 (Data S6).

962 ***4.3.2. Single-cell RNA-seq analysis and clustering***

963 Single-cell RNA-seq reads were mapped to the Kronos genome assembly v1.1. (annotation v1.0)
964 [60] (https://doi.org/10.5281/zenodo.11106422) using STARsolo [84, 85]. Since scRNA-seq

965 reads were generated from the 3' end of the transcripts and not all Kronos genes have good
966 annotations of 3' UTR regions, we extended the 3′ end of all gene models by 500 bp.

967 We used the Seurat V5 pipeline [62] to cluster the cells. Briefly, genes expressed in at least three
968 cells were retained. All cells from both stages were filtered to include those with 1,500 to 25,000
969 expressed genes and 8,000 to 300,000 unique molecular identifiers (UMIs).

970 Reads from individual samples were log-normalized, and 2,000 highly variable genes were
971 selected for downstream analysis. The integrated matrix was scaled using the 'ScaleData'
972 function. Principal component analysis (PCA) was performed with the 'RunPCA' function.
973 Further integrative analysis of all W2.5 and W3.25 stage samples was conducted using the
974 'IntegrateLayers' function with the anchor-based RPCA integration method. The top 35 principal
975 components (PCs) were used. A resolution of 1 was applied with the 'FindClusters' function to
976 identify clusters. Cell cycle clusters were identified (Fig. S15), and the remaining clusters were
977 re-normalized, scaled, integrated using the RPCA method, and re-clustered. The top 35 PCs were
978 used to generate the final clusters, at a resolution of 1 (Fig. 5B-D). Scripts used for mapping
979 reads and clustering analysis and the final Seurat object files were deposited in Zenodo [82].

980 ***4.3.3. Correlations among cell clusters***

981 Normalized expression data for the 99 genes analyzed by smFISH in this study were extracted
982 from the Seurat object file created by the Resolve Biosciences' software "Recognize" (see M&M
983 section 1.5). These expression profiles were used to calculate Pearson correlation coefficients
984 between all possible pairs of spatial clusters (Data S3). Normalized expression data for the 200
985 genes used in the spatial transcriptomics study in hexaploid wheat were kindly provided by Dr.
986 Cristobal Uauy (John Innes Centre, UK) [63]. The hexaploid wheat data was provided separately
987 for the four different stages included in their study, but we only used the three early stages (late
988 double ridge, lemma primordia and terminal spikelet) that overlapped with the stages of spike
989 development used in our scRNAseq study.

990 The average normalized gene expression across all cells for each single-cell cluster were
991 obtained with the "AverageExpression" function in Seurat v5 in R after joining the six
992 replications (three for each stage) of the scRNA-seq data. For this analysis, we used the final
993 integrated data after removing the initial cell-cycle clusters and filtered the average normalized

994 data per cluster using two criteria. First, we selected genes with coefficients of variation higher
995 than 0.5 to eliminate genes with very similar expression levels across all clusters. We also
996 eliminated genes with a total sum across clusters <0.1 and >20 to minimize the effect of genes
997 with very high or very low expression levels (and multiple 0s). Using these criteria, we selected
998 14,903 informative genes that were used to calculate Pearson correlation coefficients between all
999 possible pairs of single-cell clusters using their average normalized expression across all
1000 common genes (Data S7).

1001 To calculate the correlations between the single-cell clusters and the spatial transcriptomics
1002 clusters from tetraploid and hexaploid wheat, we first identified the overlapping genes for each
1003 comparison. The coding regions of wheat A and B homeologs are on average 97.5% identical, so
1004 most of the probes used in the wheat spatial transcriptomics experiment are expected to hybridize
1005 with both homeologs. To make the single-cell data comparable, we added the normalized
1006 expression of the two Kronos homeologs before calculating the Pearson correlation coefficients
1007 between all possible pairs of clusters (Data S10). A similar procedure was followed for
1008 hexaploid wheat, but separate correlation matrices were generated for the late double ridge (Data
1009 S11), lemma primordia (Data S12) and the terminal spikelet stage (Data S13).

1010 Clustering of the correlation matrices was done by the function "heatmap" from R package [86]
1011 with default parameters (clustering_distance_rows = "euclidean", clustering_distance_cols =
1012 "euclidean", clustering_method = "complete"). Scripts used to extract cluster averages and to
1013 perform the correlation analyses are deposited in Zenodo [82].

1014 ***4.3.4. Annotation of scRNA-seq cell clusters***

1015 To annotate the scRNA-seq clusters, we used markers expressed in specific cell types and/or
1016 previously associated with known functions in other species (Data S1). We supplemented this
1017 information with correlations between the single-cell and spatial expression clusters identified in
1018 this study (Data S10) and in a published spatial transcriptomics study in hexaploid wheat [63]
1019 (Data S11-13). We calculated Pearson correlations for all possible cluster pairs across the
1020 compared studies using the average normalized expression profiles of shared genes across all the
1021 cells within each cluster. The sum of homeologs in single-cell clusters were used for each shared
1022 gene.

1023 Additional information for each cluster was extracted from the genes preferentially expressed in
1024 each cluster (Data S8). These genes were selected based on their low abundance in other clusters
1025 (pct.2<0.10) relative to their abundance in the studied cluster (pct.1>0.10), adjusted *P* values
1026 <0.001, and a greater than 2-fold difference in expression levels (log2Fold>1, Data S8). Clusters
1027 sc1 and sc22 had few enriched genes, so we relaxed the filtering parameters (average log2FC>1,
1028 pct.1 no filtering, pct.2<0.2, adjusted *P* value <0.01) (Data S8). We subsequently determined the
1029 pathways enriched in these preferentially expressed genes using KEGG analyses (Kyoto
1030 Encyclopedia of Genes and Genomes, Data S9) through the DAVID webtool
1031 (https://davidbioinformatics.nih.gov/tools.jsp) [87]. We converted the Kronos v1.0 gene IDs into
1032 Chinese Spring annotation v1.1 gene IDs, uploaded the CS gene IDs as "ENSEMBL_GENE_ID"
1033 (step 2: Select Identifier) and as "Gene List" (step 3: List Type), used the "*Triticum aestivum*" as
1034 background species, and started the analysis using "Functional Annotation Tool". In addition, we
1035 examined the average log2Fold expression across clusters for a subset of 41 photosynthetic
1036 genes to identify clusters associated with photosynthetic functions, and for 45 ribosomal protein-
1037 encoding genes to identify cells with high protein production, typically associated with elevated
1038 cell division rates. These results were visualized using heat maps.

1039 To incorporate additional spatial information into our cluster annotation, we used the RNA-seq
1040 data generated from the base, middle and top third of the spike described above (henceforth
1041 described as spike base, middle and tip). To determine if a cell cluster was enriched in any of
1042 these three spike regions, we identified genes with significant differences in Tukey tests among
1043 the preferentially expressed genes for each cluster and quantified the number of genes expressed
1044 at higher levels in each region. To determine if the distribution across the three regions was
1045 significantly different from a random uniform distribution, we performed chi-square tests for
1046 each cluster (Data S15).

1047 The protoplasting process may trigger a stress response in some cells, potentially influencing the
1048 clustering process. We employed two independent approaches to assess the potential effects of
1049 protoplasting on gene expression in the different clusters. First, we used a list of 713 maize genes
1050 previously identified as potentially affected by the protoplasting process [64], identified the
1051 closest Kronos homologs, and determined the percentage of preferentially expressed genes in
1052 each cluster overlapping with this list (totaling 215 unique genes). The second approach, which
1053 identified 235 unique genes, involved the identification of preferentially expressed genes in each

1054 single-cell cluster that were not present in two RNA-seq data from Kronos developing spikes
1055 (this study and [61]). A GO-term analysis of the unique genes identified in both analyses
1056 revealed enrichment in stress-related annotations, particularly among the genes not detected in
1057 the two RNA-seq datasets (Data S16). These results suggest that many of these genes may have
1058 been induced by the protoplasting process and are typically not expressed under normal
1059 conditions in the developing wheat spike. The percentages of genes identified by these two
1060 independent methods across clusters were highly correlated (R= 0.813, P< 0.0001) suggesting a
1061 similar underlying phenomenon. Both methods identified clusters sc3 and sc22 as having the
1062 highest risk of being affected by the protoplasting process (Data S15).

1063 ***4.3.5.* scRNA-seq co-expression analysis**

1064 Using the cell x gene normalized count matrix, we calculated Pearson correlation coefficients (R)
1065 between genes of interest with known spatial distribution (*AGL6, FZP, TCP24*, *ULT-A1*) and all
1066 other detected genes in the scRNA-seq across all cells (26,009 cells. We focused on co-expressed
1067 genes with correlation coefficients higher than the average plus four standard deviations. The
1068 script used for co-expression analysis was deposited in Zenodo [82].

1069

1070 **ACKNOWLEDGEMENTS**

1071 The authors thank Kun Li, Chaozhong Zhang and Qiujie Liu(UC Davis) for their help with the
1072 selection of the genes included in the smFISH experiment, Samik Bhattacharya from Resolve
1073 Biosciences for his continuous help with the spatial transcriptomics project, Cristobal Uauy from
1074 the John Innes Centre UK for his advice with the analysis of the spatial transcriptomics data and
1075 early access to his spatial transcriptomics data, and Ksenia Krasileva from UC Berkeley for early
1076 access to the Kronos genome. JD acknowledges funding from the Howard Hughes Medical
1077 institute and the United States Department of Agriculture (USDA) – National Institute of Food
1078 and Agriculture (NIFA) grants 2022-68013-36439 (WheatCAP) and 2022-67013-36209. X.X
1079 acknowledges USDA-NIFA HATCH project (CA-D-PLB-2850-H) and UC Davis new faculty
1080 start-up funding.

1081

1082

AUTHORS CONTRIBUTIONS

XX, HL, and JD conceived the project. Experimental work was performed by HL, XX, FP, CT, and YL. JZ, GB, HL, CL, XX and JD performed data analyses. JD and XX obtained funding and directed the project. HL, XX and JD generated the first draft, and all authors contributed to the revision of the manuscript. XX, CL and JD provided supervision.

CONFLICT OF INTERESTS

The authors declare no conflict of interests.

DATA AVAILABILITY

Data and scripts used in the different analyses are available in Supplemental Data and in Zenodo [82]. The ImageJ version with the Polylux tool plugin, the high-resolution images for the six sections used in the paper (Calcofluor White Figs. S2, S3 and DAPI), the hybridization results for the 99 genes for each image, and instructions for download and installation are also available at our web site (https://dubcovskylab.ucdavis.edu/content/spatial-transcriptomic) and in Zenodo [82]. The raw RNA-seq data of the three regions of the Kronos spike at W3.0 were deposited in NCBI's Sequence Read Archive (SRA) under BioProject number PRJNA1217025. The processed data, scripts used for mapping and feature counts, and a Readme file with instructions are deposited in Zenodo [82].

The scRNA-seq data were deposited in NCBI's Sequence Read Archive (SRA) under BioProject accession number PRJNA1199502. The final single-cell clusters can be accessed via the Shiny App (https://dubcovskylab.ucdavis.edu/content/kronos-integrated-clusters). Additionally, a 3D-UMAP (https://dubcovskylab.ucdavis.edu/content/kronos-scrna-3d-clusters) was generated to visualize cluster distributions. The Kronos genome assembly v1.1 (annotation v1.0) is available in Zenodo https://doi.org/10.5281/zenodo.11106422. Processed single-cell data, Seurat objects and R scripts are also available in Zenodo [82].

Data and scripts for cluster correlations, the Seurat object produced by Resolve BioSciences "Recognize" online tool, and the R script to replicate the online analysis locally can be also found in Zenodo [82].

1112 **REFERENCES**

1113 1. **FAOSTAT: Food and Agriculture Organization (FAO) of the United Nations.**
1114 http://www.fao.org/faostat/en/#data. 2023.

1115 2. Kellogg EA: **Genetic control of branching patterns in grass inflorescences.** *Plant Cell*
1116 2022, **34:**2518-2533.

1117 3. Li C, Lin H, Chen A, Lau M, Jernstedt J, Dubcovsky J: **Wheat *VRN1*, *FUL2* and *FUL3***
1118 **play critical and redundant roles in spikelet development and spike determinacy.**
1119 *Development* 2019, **146:**dev175398.

1120 4. Chen H, Zhang X, Xu SH, Song CX, Mao HL: ***TaSPL17*s act redundantly with**
1121 ***TaSPL14*s to control spike development and their elite haplotypes may improve**
1122 **wheat grain yield.** *Frontiers in Plant Science* 2023, **14**.

1123 5. Chuck G, Whipple C, Jackson D, Hake S: **The maize SBP-box transcription factor**
1124 **encoded by *tasselsheath4* regulates bract development and the establishment of**
1125 **meristem boundaries.** *Development* 2010, **137:**1243-1250.

1126 6. Chuck GS, Brown PJ, Meeley R, Hake S: **Maize SBP-box transcription factors**
1127 ***unbranched2* and *unbranched3* affect yield traits by regulating the rate of lateral**
1128 **primordia initiation.** *Proc Natl Acad Sci U S A* 2014, **111:**18775-18780.

1129 7. Wang L, Yin H, Qian Q, Yang J, Huang C, Hu X, Luo D: ***NECK LEAF 1*, a GATA type**
1130 **transcription factor, modulates organogenesis by regulating the expression of**
1131 **multiple regulatory genes during reproductive development in rice.** *Cell Res* 2009,
1132 **19:**598-611.

1133 8. Walsh J, Freeling M: **The *liguleless2* gene of maize functions during the transition**
1134 **from the vegetative to the reproductive shoot apex.** *Plant Journal* 1999, **19:**489-495.

1135 9. Li K, Debernardi JM, Li C, Lin H, Zhang C, Jernstedt J, von Korff M, Zhong J,
1136 Dubcovsky J: **Interactions between SQUAMOSA and SVP MADS-box proteins**
1137 **regulate meristem transitions during wheat spike development.** *The Plant Cell* 2021,
1138 **33:**3621-3644.

1139 10. Adamski NM, Simmonds J, Brinton JF, Backhaus AE, Chen Y, Smedley M, Hayta S,
1140 Florio T, Crane P, Scott P, et al: **Ectopic expression of *Triticum polonicum VRT-A2***

1141 **underlies elongated glumes and grains in hexaploid wheat in a dosage-dependent**
1142 **manner.** *The Plant Cell* 2021, **33:**2296-2319.

1143 11. Debernardi JM, Lin H, Chuck G, Faris JD, Dubcovsky J: **microRNA172 plays a crucial**
1144 **role in wheat spike morphogenesis and grain threshability.** *Development* 2017,
1145 **144:**1966-1975.

1146 12. Greenwood JR, Finnegan EJ, Watanabe N, Trevaskis B, Swain SM: **New alleles of the**
1147 **wheat domestication gene *Q* reveal multiple roles in growth and reproductive**
1148 **development.** *Development* 2017, **144:**1959-1965.

1149 13. Debernardi JM, Greenwood JR, Jean Finnegan E, Jernstedt J, Dubcovsky J: ***APETALA***
1150 ***2*-like genes *AP2L2* and *Q* specify lemma identity and axillary floral meristem**
1151 **development in wheat.** *Plant Journal* 2020, **101:**171-187.

1152 14. Reinheimer R, Kellogg EA: **Evolution of *AGL6-like* MADS box genes in grasses**
1153 **(Poaceae): ovule expression is ancient and palea expression is new.** *Plant Cell* 2009,
1154 **21:**2591-2605.

1155 15. Paraiso F, Lin H, Li C, Woods DP, Lan T, Tumelty C, Debernardi JM, Joe A, Dubcovsky
1156 J: **LEAFY and WAPO1 jointly regulate spikelet number per spike and floret**
1157 **development in wheat.** *Development* 2024, **151:**dev202803.

1158 16. Kuzay S, Lin H, Li C, Chen S, Woods DP, Zhang J, Lan T, von Korff M, Dubcovsky J:
1159 ***WAPO-A1* is the causal gene of the 7AL QTL for spikelet number per spike in**
1160 **wheat.** *PLoS Genetics* 2022, **18:**e1009747.

1161 17. Kuzay S, Xu Y, Zhang J, Katz A, Pearce S, Su Z, Fraser M, Anderson JA, Brown-
1162 Guedira G, DeWitt N, et al: **Identification of a candidate gene for a QTL for spikelet**
1163 **number per spike on wheat chromosome arm 7AL by high-resolution genetic**
1164 **mapping.** *Theor Appl Genet* 2019, **132:**2689–2705.

1165 18. Muqaddasi QH, Brassac J, Koppolu R, Plieske J, Ganal MW, Roder MS: ***TaAPO-A1*, an**
1166 **ortholog of rice *ABERRANT PANICLE ORGANIZATION 1*, is associated with total**
1167 **spikelet number per spike in elite European hexaploid winter wheat (*Triticum***
1168 ***aestivum* L.) varieties.** *Scientific Reports* 2019, **9:**13853.

1169 19. Voss-Fels KP, Keeble-Gagnere G, Hickey LT, Tibbits J, Nagornyy S, Hayden MJ, Pasam
1170 RK, Kant S, Friedt W, Snowdon R, et al: **High-resolution mapping of rachis nodes per**

1171 **rachis, a critical determinant of grain yield components in wheat.** *Theoretical and*
1172 *Applied Genetics* 2019, **132:**2707-2719.

1173 20. Zhang J, Burguener GF, Paraiso F, Dubcovsky J: **Natural alleles of LEAFY and**
1174 **WAPO1 interact to regulate spikelet number per spike in wheat.** *Theor Appl Genet*
1175 2024, **137:**257.

1176 21. Waddington SR, Cartwright PM, Wall PC: **A quantitative scale of spike initial and**
1177 **pistil development in barley and wheat.** *Annals of Botany* 1983, **51:**119-130.

1178 22. Bankhead P, Loughrey MB, Fernandez JA, Dombrowski Y, McArt DG, Dunne PD,
1179 McQuaid S, Gray RT, Murray LJ, Coleman HG, et al: **QuPath: Open source software**
1180 **for digital pathology image analysis.** *Sci Rep* 2017, **7:**16878.

1181 23. Perico C, Zaidem M, Sedelnikova O, Bhattacharya S, Korfhage C, Langdale JA:
1182 **Multiplexed in situ hybridization reveals distinct lineage identities for major and**
1183 **minor vein initiation during maize leaf development.** *Proceedings of the National*
1184 *Academy of Sciences of the United States of America* 2024, **121**.

1185 24. Bonke M, Thitamadee S, Mähönen AP, Hauser MT, Helariutta Y: **APL regulates**
1186 **vascular tissue identity in *Arabidopsis*.** *Nature* 2003, **426:**181-186.

1187 25. Zhang J, Li C, Zhang W, Zhang X, Mo Y, Tranquilli GE, Vanzetti LS, Dubcovsky J:
1188 **Wheat plant height locus *RHT25* encodes a PLATZ transcription factor that**
1189 **interacts with DELLA (RHT1).** *Proceedings of the National Academy of Sciences of*
1190 *the United States of America* 2023, **120:**e2300203120.

1191 26. McKim SM: **How plants grow up.** *J Integr Plant Biol* 2019, **61:**257-277.

1192 27. Ikeda-Kawakatsu K, Maekawa M, Izawa T, Itoh JI, Nagato Y: ***ABERRANT PANICLE***
1193 ***ORGANIZATION 2/RFL*, the rice ortholog of Arabidopsis *LEAFY*, suppresses the**
1194 **transition from inflorescence meristem to floral meristem through interaction with**
1195 ***APO1*.** *Plant Journal* 2012, **69:**168-180.

1196 28. Wang J, Wang R, Wang Y, Zhang L, Zhang L, Xu Y, Yao S: ***Short and Solid***
1197 ***Culm/RFL/APO2* for culm development in rice.** *Plant J* 2017, **91:**85-96.

1198 29. Huijser P, Klein J, Lonnig WE, Meijer H, Saedler H, Sommer H: **Bracteomania, an**
1199 **inflorescence anomaly, is caused by the loss of function of the MADS-box gene**
1200 ***squamosa* in *Antirrhinum majus*.** *Embo Journal* 1992, **11:**1239-1249.

1201 30. Klein J, Saedler H, Huijser P: **A new family of DNA binding proteins includes**
1202 **putative transcriptional regulators of the *Antirrhinum majus* floral meristem**
1203 **identity gene *SQUAMOSA*.** *Molecular and General Genetics* 1996, **250:**7-16.

1204 31. Wang L, Ming LC, Liao KY, Xia CJ, Sun SY, Chang Y, Wang HK, Fu DB, Xu CH,
1205 Wang ZJ, et al: **Bract suppression regulated by the *SPLs-NL1-PLA1* module is**
1206 **required for the transition from vegetative to reproductive branching in rice.**
1207 *Molecular Plant* 2021, **14:**1168-1184.

1208 32. Dong Z, Hu G, Chen Q, Shemyakina EA, Chau G, Whipple CJ, Fletcher JC, Chuck G: **A**
1209 **regulatory network controlling developmental boundaries and meristem fates**
1210 **contributed to maize domestication.** *Nat Genet* 2024, **56:**2528-2537.

1211 33. Lu Z, Yu H, Xiong G, Wang J, Jiao Y, Liu G, Jing Y, Meng X, Hu X, Qian Q, et al:
1212 **Genome-wide binding analysis of the transcription activator IDEAL PLANT**
1213 **ARCHITECTURE1 reveals a complex network regulating rice plant architecture.**
1214 *Plant Cell* 2013, **25:**3743-3759.

1215 34. Doebley J, Stec A, Hubbard L: **The evolution of apical dominance in maize.** *Nature*
1216 1997, **386:**485-488.

1217 35. Takeda T, Suwa Y, Suzuki M, Kitano H, Ueguchi-Tanaka M, Ashikari M, Matsuoka M,
1218 Ueguchi C: **The *OsTB1* gene negatively regulates lateral branching in rice.** *Plant J*
1219 2003, **33:**513-520.

1220 36. Ramsay L, Comadran J, Druka A, Marshall DF, Thomas WT, Macaulay M, MacKenzie
1221 K, Simpson C, Fuller J, Bonar N, et al: ***INTERMEDIUM-C*, a modifier of lateral**
1222 **spikelet fertility in barley, is an ortholog of the maize domestication gene**
1223 ***TEOSINTE BRANCHED 1*.** *Nat Genet* 2011, **43:**169-172.

1224 37. Dixon LE, Greenwood JR, Bencivenga S, Zhang P, Cockram J, Mellers G, Ramm K,
1225 Cavanagh C, Swain SM, Boden SA: ***TEOSINTE BRANCHED1* regulates inflorescence**
1226 **architecture and development in bread wheat (*Triticum aestivum*).** *Plant Cell* 2018,
1227 **30:**563-581.

1228 38. Volkman MM, Martin JM, Hogg AC, Wright L, Hale C, Carr PM, Giroux MJ: **Durum**
1229 **wheat *Teosinte Branched1* null mutations increase tillering.** *Crop Science* 2022,
1230 **62:**1522-1530.

1231 39. Reiser L, Sánchez-Baracaldo P, Hake S: **Knots in the family tree: evolutionary**
1232 **relationships and functions of homeobox genes.** *Plant Molecular Biology* 2000,
1233 **42:**151-166.

1234 40. Leiboff S, Li XR, Hu HC, Todt N, Yang JL, Li X, Yu XQ, Muehlbauer GJ, Timmermans
1235 MCP, Yu JM, et al: **Genetic control of morphometric diversity in the maize shoot**
1236 **apical meristem.** *Nature Communications* 2015, **6:**8974.

1237 41. O'Connor DL, Runions A, Sluis A, Bragg J, Vogel JP, Prusinkiewicz P, Hake S: **A**
1238 **division in PIN-mediated auxin patterning during organ initiation in grasses.** *Plos*
1239 *Computational Biology* 2014, **10:**e1003447.

1240 42. Moraes TD, van Es SW, Hernández-Pinzón I, Kirschner GK, van der Wal F, da Silveira
1241 SR, Busscher-Lange J, Angenent GC, Moscou M, Immink RGH, van Esse GW: **The**
1242 **TCP transcription factor HvTB2 heterodimerizes with VRS5 and controls spike**
1243 **architecture in barley.** *Plant Reproduction* 2022, **35:**205-220.

1244 43. Zong J, Wang L, Zhu L, Bian L, Zhang B, Chen X, Huang G, Zhang X, Fan J, Cao L, et
1245 al: **A rice single cell transcriptomic atlas defines the developmental trajectories of**
1246 **rice floret and inflorescence meristems.** *New Phytol* 2022, **234:**494-512.

1247 44. Patterson EL, Richardson A, Bartlett M: **Pushing the boundaries of organ identity:**
1248 **Homology of the grass lemma.** *American Journal of Botany* 2023, **110:**e16161.

1249 45. Komatsu M, Chujo A, Nagato Y, Shimamoto K, Kyozuka J: ***FRIZZY PANICLE* is**
1250 **required to prevent the formation of axillary meristems and to establish floral**
1251 **meristem identity in rice spikelets.** *Development* 2003, **130:**3841-3850.

1252 46. Chuck G, Muszynski M, Kellogg E, Hake S, Schmidt RJ: **The control of spikelet**
1253 **meristem identity by the *branched silkless1* gene in maize.** *Science* 2002, **298:**1238-
1254 1241.

1255 47. Dobrovolskaya O, Pont C, Sibout R, Martinek P, Badaeva E, Murat F, Chosson A,
1256 Watanabe N, Prat E, Gautier N, et al: ***FRIZZY PANICLE* drives supernumerary**
1257 **spikelets in bread wheat.** *Plant Physiology* 2015, **167:**189-199.

1258 48. Poursarebani N, Seidensticker T, Koppolu R, Trautewig C, Gawronski P, Bini F, Govind
1259 G, Rutten T, Sakuma S, Tagiri A, et al: **The genetic basis of composite spike form in**
1260 **barley and 'Miracle-Wheat'.** *Genetics* 2015, **201:**155-165.

1261 49. Bommert P, Whipple C: **Grass inflorescence architecture and meristem determinacy.**
1262 *Seminars in Cell & Developmental Biology* 2018, **79:**37-47.

1263 50. Yang Y, Fu D, Zhu C, He Y, Zhang H, Liu T, Li X, Wu C: **The RING-finger ubiquitin**
1264 **ligase HAF1 mediates Heading date 1 degradation during photoperiodic flowering**
1265 **in rice.** *Plant Cell* 2015, **27:**2455-2468.

1266 51. Zhu C, Peng Q, Fu D, Zhuang D, Yu Y, Duan M, Xie W, Cai Y, Ouyang Y, Lian X, Wu
1267 C: **The E3 ubiquitin ligase HAF1 modulates circadian accumulation of EARLY**
1268 **FLOWERING3 to control heading date in rice under long-day conditions.** *Plant Cell*
1269 2018, **30:**2352-2367.

1270 52. McSteen P, Malcomber S, Skirpan A, Lunde C, Wu XT, Kellogg E, Hake S: ***barren***
1271 ***inflorescence2*** **encodes a co-ortholog of the** ***PINOID*** **serine/threonine kinase and is**
1272 **required for organogenesis during inflorescence and vegetative development in**
1273 **maize.** *Plant Physiology* 2007, **144:**1000-1011.

1274 53. Carles CC, Choffnes-Inada D, Reville K, Lertpiriyapong K, Fletcher JC:
1275 ***ULTRAPETALA1*** **encodes a SAND domain putative transcriptional regulator that**
1276 **controls shoot and floral meristem activity in Arabidopsis.** *Development* 2005,
1277 **132:**897-911.

1278 54. Monfared MM, Carles CC, Rossignol P, Pires HR, Fletcher JC: **The** ***ULT1*** **and** ***ULT2***
1279 **trxG genes play overlapping roles in** ***Arabidopsis*** **development and gene regulation.**
1280 *Molecular Plant* 2013, **6:**1564-1579.

1281 55. Wang WF, Li G, Zhao J, Chu HW, Lin WH, Zhang DB, Wang ZY, Liang WQ: **DWARF**
1282 **TILLER1, a WUSCHEL-related homeobox transcription factor, is required for**
1283 **tiller growth in rice.** *Plos Genetics* 2014, **10**.

1284 56. Zhang J, Xiong H, Burguener GF, Vasquez-Gross H, Liu Q, Debernardi JM, Akhunova
1285 A, Garland-Campbell K, Kianian SF, Brown-Guedira G, et al: **Sequencing 4.3 million**
1286 **mutations in wheat promoters to understand and modify gene expression.**
1287 *Proceedings of the National Academy of Sciences of the United States of America* 2023,
1288 **120:**e2306494120.

1289 57. Glenn P, Woods DP, Zhang J, Gabay G, Odle N, Dubcovsky J: **Wheat bZIPC1**
1290 **interacts with FT2 and contributes to the regulation of spikelet number per spike.**
1291 *Theoretical and Applied Genetics* 2023, **136:**237.

1292 58. Zhu W, Yang L, Wu D, Meng Q, Deng X, Huang G, Zhang J, Chen X, Ferrandiz C,
1293 Liang W, et al: **Rice *SEPALLATA* genes *OsMADS5* and *OsMADS34* cooperate to**
1294 **limit inflorescence branching by repressing the *TERMINAL FLOWER1-like* gene**
1295 ***RCN4*.** *New Phytol* 2022, **233:**1682-1700.

1296 59. Xu X, Passalacqua M, Rice B, Demesa-Arevalo E, Takebayashi MKY, Harris B,
1297 Sakakibara H, Gallavotti A, Gillis J, Jackson D: **Large-scale single-cell profiling of**
1298 **stem cells uncovers redundant regulators of shoot development and yield trait**
1299 **variation.** *bioRxiv* 2024, **doi.org/10.1101/2024.03.04.583414**.

1300 60. Seong K, Kumar R, Lunde C, Krasileva K: **Chromosome-level genome assembly of**
1301 ***Triticum turgidum* var 'Kronos' [Data set].** Zenodo.
1302 https://doi.org/10.5281/zenodo.11106422 2023.

1303 61. VanGessel C, Hamilton J, Tabbita F, Dubcovsky J, Pearce S: **Transcriptional**
1304 **signatures of wheat inflorescence development.** *Sci Rep* 2022, **12:**17224.

1305 62. Hao Y, Stuart T, Kowalski MH, Choudhary S, Hoffman P, Hartman A, Srivastava A,
1306 Molla G, Madad S, Fernandez-Granda C, Satija R: **Dictionary learning for integrative,**
1307 **multimodal and scalable single-cell analysis.** *Nat Biotechnol* 2024, **42:**293-304.

1308 63. Long KA, Lister A, Jones MRW, Adamski NM, Ellis RE, Chedid C, Carpenter SJ, Liu X,
1309 Backhaus AE, Goldson A, et al: **Spatial transcriptomics reveals expression gradients**
1310 **in developing wheat inflorescences at cellular resolution.** *bioRxiv* 2024,
1311 https://doi.org/10.1101/2024.12.19.629411.

1312 64. Xu XS, Crow M, Rice BR, Li F, Harris B, Liu L, Demesa-Arevalo E, Lu ZF, Wang LY,
1313 Fox N, et al: **Single-cell RNA sequencing of developing maize ears facilitates**
1314 **functional analysis and trait candidate gene discovery.** *Developmental Cell* 2021,
1315 **56:**557-568.

1316 65. Skipper M: **Genes from the *APETALA3* and *PISTILLATA* lineages are expressed in**
1317 **developing vascular bundles of the tuberous rhizome, flowering stem and flower**
1318 **primordia of *Eranthis hyemalis*.** *Ann Bot* 2002, **89:**83-88.

1319 66. Wang J, Bao JL, Zhou BB, Li M, Li XZ, Jin J: **The *osa-miR164* target *OsCUC1***
1320 **functions redundantly with *OsCUC3* in controlling rice meristem/organ boundary**
1321 **specification.** *New Phytologist* 2021, **229:**1566-1581.

1322 67. Hibara K, Karim MR, Takada S, Taoka KI, Furutani M, Aida M, Tasaka M: **Arabidopsis**
1323 ***CUP-SHAPED COTYLEDON3* regulates postembryonic shoot meristem and organ**
1324 **boundary formation.** *Plant Cell* 2006, **18:**2946-2957.

1325 68. Gallavotti A, Zhao Q, Kyozuka J, Meeley RB, Ritter MK, Doebley JF, Pe ME, Schmidt
1326 RJ: **The role of *barren stalk1* in the architecture of maize.** *Nature* 2004, **432:**630-635.

1327 69. Gallavotti A, Malcomber S, Gaines C, Stanfield S, Whipple C, Kellogg E, Schmidt RJ:
1328 **BARREN STALK FASTIGIATE1 is an AT-hook protein required for the**
1329 **formation of maize ears.** *Plant Cell* 2011, **23:**1756-1771.

1330 70. Wang YG, Du F, Wang J, Wang K, Tian CH, Qi XQ, Lu F, Liu XG, Ye XG, Jiao YL:
1331 **Improving bread wheat yield through modulating an unselected AP2/ERF gene.**
1332 *Nature Plants* 2022, **8:**930-939.

1333 71. Li Z, Liu D, Xia Y, Li Z, Jing D, Du J, Niu N, Ma S, Wang J, Song Y, et al:
1334 **Identification of the WUSCHEL-related homeobox (*WOX*) gene family, and**
1335 **interaction and functional analysis of TaWOX9 and TaWUS in wheat.** *Int J Mol Sci*
1336 2020, **21:**1581.

1337 72. Fang F, Ye S, Tang J, Bennett MJ, Liang W: **DWT1/DWL2 act together with**
1338 **OsPIP5K1 to regulate plant uniform growth in rice.** *New Phytol* 2020, **225:**1234-
1339 1246.

1340 73. Hendelman A, Zebell S, Rodriguez-Leal D, Dukler N, Robitaille G, Wu XL, Kostyun J,
1341 Tal L, Wang PP, Bartlett ME, et al: **Conserved pleiotropy of an ancient plant**
1342 **homeobox gene uncovered by *cis*-regulatory dissection.** *Cell* 2021, **184:**1724-1739.

1343 74. Bai F, Reinheimer R, Durantini D, Kellogg EA, Schmidt RJ: **TCP transcription factor,**
1344 **BRANCH ANGLE DEFECTIVE 1 (BAD1), is required for normal tassel branch**
1345 **angle formation in maize.** *Proceedings of the National Academy of Sciences of the*
1346 *United States of America* 2012, **109:**12225-12230.

1347 75. Oikawa T, Kyozuka J: **Two-step regulation of LAX PANICLE1 protein accumulation**
1348 **in axillary meristem formation in rice.** *Plant Cell* 2009, **21:**1095-1108.

1349 76. Galli M, Liu Q, Moss BL, Malcomber S, Li W, Gaines C, Federici S, Roshkovan J,
1350 Meeley R, Nemhauser JL, Gallavotti A: **Auxin signaling modules regulate maize**
1351 **inflorescence architecture.** *Proc Natl Acad Sci U S A* 2015, **112:**13372-13377.

1352 77. Bortiri E, Chuck G, Vollbrecht E, Rocheford T, Martienssen R, Hake S: ***ramosa2***
1353 **encodes a LATERAL ORGAN BOUNDARY domain protein that determines the**
1354 **fate of stem cells in branch meristems of maize.** *Plant Cell* 2006, **18:**574-585.

1355 78. Jin Y, Luo Q, Tong H, Wang A, Cheng Z, Tang J, Li D, Zhao X, Li X, Wan J, et al: **An**
1356 **AT-hook gene is required for palea formation and floral organ number control in**
1357 **rice.** *Dev Biol* 2011, **359:**277-288.

1358 79. Panikashvili D, Shi JX, Bocobza S, Franke RB, Schreiber L, Aharoni A: **The**
1359 ***Arabidopsis* DSO/ABCG11 transporter affects cutin metabolism in reproductive**
1360 **organs and suberin in roots.** *Mol Plant* 2010, **3:**563-575.

1361 80. Douglas SJ, Li B, Kliebenstein DJ, Nambara E, Riggs CD: **A novel *Filamentous Flower***
1362 **mutant suppresses *brevipedicellus* developmental defects and modulates**
1363 **glucosinolate and auxin levels.** *PLoS One* 2017, **12:**e0177045.

1364 81. Whipple CJ: **Grass inflorescence architecture and evolution: the origin of novel**
1365 **signaling centers.** *New Phytologist* 2017, **216:**367-372.

1366 82. Xu X, Lin H, Zhang J, Burguener G, Paraiso F, Tumelty C, Li C, Liu Y, Dubcovsky J:
1367 **Spatial and single-cell expression analyses reveal complex expression domains in**
1368 **early wheat spike development (v1.0) [Data set]** In *Zenodo*
1369 *https://doi.org/10.5281/zenodo.14872905*; 2025.

1370 83. Hao Y, Hao S, Andersen-Nissen E, Mauck WM, 3rd, Zheng S, Butler A, Lee MJ, Wilk
1371 AJ, Darby C, Zager M, et al: **Integrated analysis of multimodal single-cell data.** *Cell*
1372 2021, **184:**3573-3587 e3529.

1373 84. Kaminow B, Yunusov D, Dobin A: **STARsolo: accurate, fast and versatile**
1374 **mapping/quantification of single-cell and single-nucleus RNA-seq data.** *bioRxiv*
1375 2021**:**2021.2005.2005.442755.

1376 85. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M,
1377 Gingeras TR: **STAR: ultrafast universal RNA-seq aligner.** *Bioinformatics* 2013,
1378 **29:**15-21.

1379 86. Kolde R: **pheatmap: Pretty Heatmaps.** https://CRAN.R-project.org/package=pheatmap.
1380 2019.
1381 87. Huang DW, Sherman BT, Lempicki RA: **Systematic and integrative analysis of large**
1382 **gene lists using DAVID bioinformatics resources.** *Nature Protocols* 2009, **4:**44-57.
1383