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], Kun Li[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 three 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

Kun Li: kzli@ucdavis.edu

Connor Tumelty: cwtumelty@ucdavis.edu

Chengxia Li: chxli@ucdavis.edu

Yuchen Liu: yvcliu@ucdavis.edu

ABSTRACT

Wheat is important for global food security and understanding the molecular mechanisms driving spike and spikelet development can inform the development 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 clusters 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, late double ridge, and floret primordia stages. These cells were grouped into 21 different expression domains, including four in the basal region of the developing spikelets and three different meristematic regions, which were consistent across spikelets and sections. Using induced mutants, we revealed functional roles associated with the specific expression patterns of *LFY* in intercalary meristems, *SPL14* in inflorescence meristems, and *FZP* in glume axillae. Complementary scRNA-seq profiling of 26,009 cells from W2.5 and W3.5 stages identified 23 distinct cell clusters. We used the scRNA-seq information to impute the expression of 74,464 genes into the spatially anchored smFISH-labelled cells and generated a public website to visualize them. We used experimental and imputed expression profiles, together with co-expression studies and correlation matrices, to annotate the scRNA-seq clusters. From co-expression analyses, we 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 dissecting gene networks that regulate spike development and identifying new co-expressed genes for functional characterization.

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], yet further increases are necessary to feed a growing human population. Wheat grains develop within inflorescences called spikes, thus 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 mechanisms underlying spike development could 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 (Fig.S1A), whereas the subtended axillary meristems generally fail to grow out immediately [2]. By contrast, the wheat IM produces lateral meristems (Fig.S1B), that later differentiate into two distinct ridges (the double ridge stage, Fig. S1C). 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. Fig. S1D). Therefore, the transition from the vegetative to the reproductive phase involves both a drastic shift in the relative developmental rates of bract and axillary meristems 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.

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

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

89 The MADS-box genes from the *SVP* clade (*VRT2, SVP1*, *SVP3*) play similar roles as the
90 *SQUAMOSA* genes during the initial stages of spike development. Individual loss-of-function
91 mutations in *VRN1* or *FUL2* [3], or simultaneous mutations in *VRT2* and *SVP1* [7], both result in
92 delayed heading, shorter stems and increased spikelet number per spike (SNS). However, at later
93 stages of spike development, the SQUAMOSA proteins transcriptionally repress the *SVP*-clade
94 genes to facilitate interactions with SEPALLATA1 proteins, which are required for normal
95 spikelet development [7, 8].

96 The normal development of the wheat lateral spikelets begins with the formation of two sterile
97 bracts called glumes, followed by slightly different bracts called lemmas that subtend the floral
98 organs. Differences between glumes and lemmas are regulated by the *APETALA2-like* genes
99 *AP2L5* (*Q* gene) and *AP2L2*, which are cleaved by the microRNA miR172. Mutations in the
100 miR172 binding sites of these genes result in increased transcript levels and conversions of
101 sterile glumes into fertile lemmas [9, 10]. Conversely, loss-of-function mutation in *AP2L5* and
102 *AP2L2* result in conversions of fertile lemmas into sterile glumes [11]. These results indicate an
103 important role of *AP2L5* and *AP2L2* in the regulation of glume and lemma organ identities and
104 the development of the axillary floret meristem (FM).

105 In grasses, the first whorl of floral organs generated by the FM is the palea, which is regulated by
106 the MADS-box gene *AGL6* [12]. Development of the internal floral whorls (lodicules, stamens,
107 and pistils) is severely disrupted in wheat plants with loss-of-function mutations in *LEAFY* (*LFY*)
108 or *WHEAT ORTHOLOGUE of APO1* (*WAPO1*), two genes that are expressed in partially
109 overlapping domains at the FM and encode interacting proteins [13, 14]. In addition to the
110 disruptions in floral development, loss-of-function mutations in either or both genes significantly
111 reduce the rate at which the IM generates lateral meristems, leading to drastic reductions in SNS
112 [13]. Natural alleles with increased SNS have been identified for both *WAPO1* [14-17] and *LFY*
113 [18], and allele combinations that maximize SNS have been also reported [18].

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

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

123

124

125 **RESULTS AND DISCUSSION**

126 **1. Spatial expression and cell segmentation**

127 To identify the distinct cell clusters and expression domains involved in early wheat spike
128 development, we analyzed the expression of 99 genes simultaneously (Data S1) by smFISH at
129 three early stages of spike development (W1.5, W2.5, and W3.5 as defined by the Waddington
130 scale [19]). W1.5 marks the initial transition of the shoot apical meristem (SAM) to the
131 reproductive stage (Fig. 1A). W2.5 is defined as the late double ridge stage before glume
132 primordia initiation (Fig. 1B and S1A), and W3.5 as the floret primordia stage (Fig. 1C and
133 S1B). At W3.5, the inflorescence meristem (IM) begins its transition to a terminal spikelet
134 (IM→TS), and the most advanced spikelet meristems (SMs) at the center of the spike initiate
135 floret meristems (FMs, Figs. 1C and S1B).

136 We analyzed two sections from each developmental stage (Figs. 1, S2) using the cell
137 segmentation program QuPath [20] and identified a total of 48,225 cells. In the same sections,
138 cell walls were visualized with calcofluor-white (Figs. S3, S4). Based on the expression of 99
139 genes (Data S1), these cells were grouped into 21 clusters, represented by consistent colors
140 across all three developmental stages and replications (Figs. 1, S2). The individual cell clusters
141 are presented separately in Fig. S5 and can be also visualized, together with heatmaps of the 99
142 genes, at https://dubcovskylab.ucdavis.edu/JD99-wheat-spike-smFISH.

143 We characterized 7,624 cells at W1.5, 16,400 at W2.5, and 24,210 at W3.5 (Fig. 1 and S2),
144 showing a clear increase in cell number during spike development. These cells also exhibited
145 increasing differentiation, with 14 clusters at W1.5, 19 at W2.5, and 21 at W3.5 (Fig. 1 and S2,
146 Data S2). 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 S2, demonstrating good reproducibility. The gene expression profiles obtained with Molecular Cartography also showed good reproducibility with previous *in situ* hybridization results (Fig. S6) and recently published wheat spatial transcriptomics results generated using MERFISH [21] (Fig. S7).

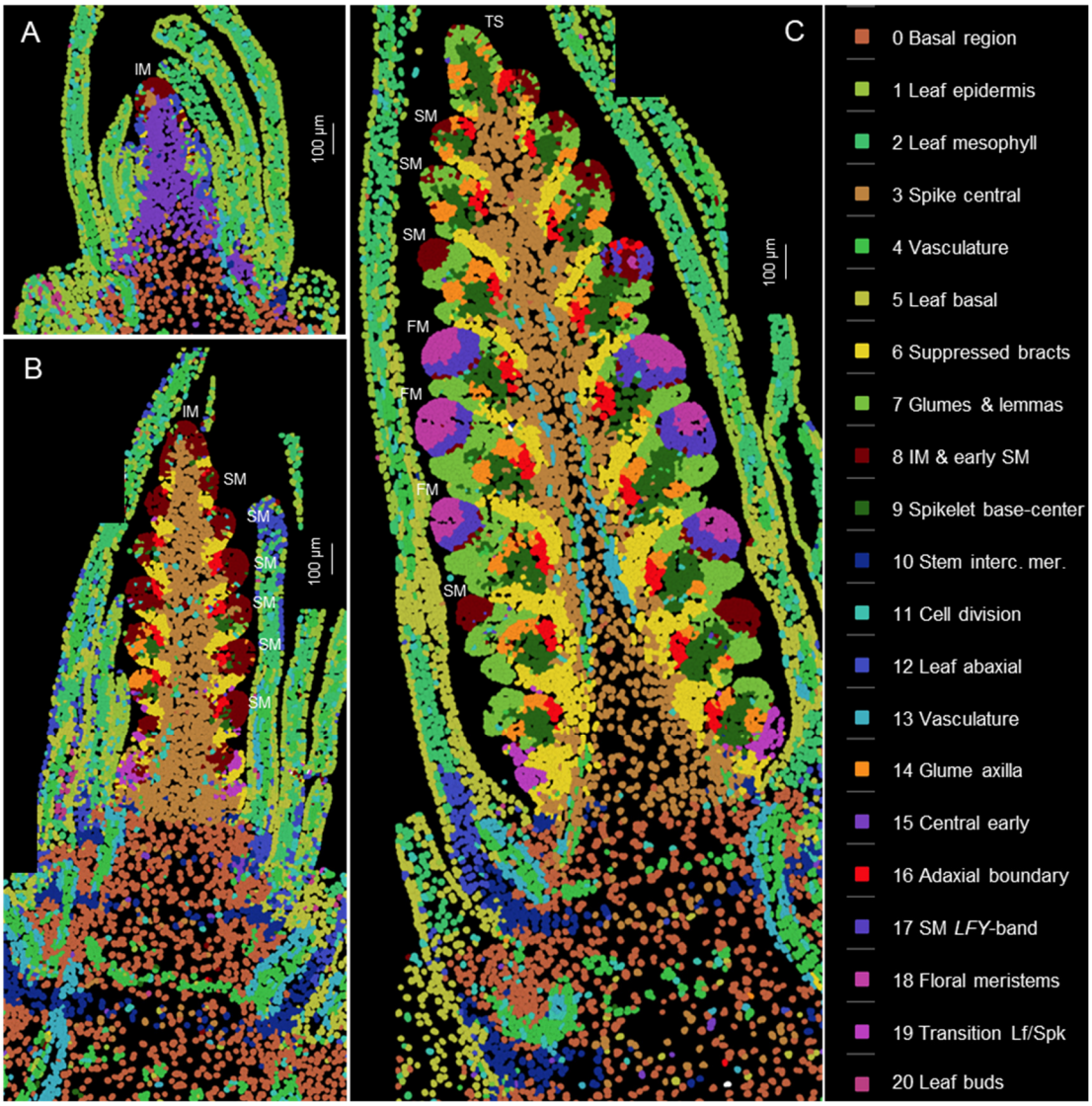

Fig. 1. Cell clusters based on the expression of 99 genes analyzed by smFISH at three stages of wheat spike development. Cell clusters were generated from the three sections shown in this figure and replicate sections presented in Fig. S2. Identical colors represent the same clusters across the six sections. **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** W3.5: floret 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. IM= inflorescence meristem, SM= spikelet meristem, FM= floral meristem, TS= terminal spikelet.

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 cells grouped in cluster c0, which are located below the epidermis and outside the vascular tissue and are referred hereafter as ground tissue. At W2.5 and W3.5 this basal region is interrupted by alternating bands of cells from clusters c4 (vasculature) and c10 (stem intercalary meristem, Figs. 1, S2, S5).

In longitudinal sections stained with calcofluor-white at W2.5 and W3.5 (Figs. S3, S4), two alternating zones are visible at the basal region: one with shorter cells likely corresponding to the nodes, and the other with longer cells likely corresponding to the internodes of the future stem. Enriched expression of the procambial marker *LAX4* [22] and the phloem marker *APL* [23] was observed in the smaller cells, supporting their association with future nodes (Fig. 2A). These cells, together with those expressing the vascular genes in the leaf and spike vascular tissues were all included in cluster c4 (Fig. 2B). Based on the location of c4 cells in the nodes and the orientation of the vascular bundles we hypothesize that these cells may interconnect the parallel vascular bundles of the internodes.

Cells classified as cluster c10 (Fig. 2B) appeared as incomplete bands in the future internode regions at W2.5 and W3.5 (Figs. 1, S5). In the three-dimensional space, c10 cells may form a transversal disc across the stem located at the point 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*, *ULT1*, and *SVP1* (Fig. S8, Data S2). Previous studies have demonstrated that loss-of-function mutations in wheat *PLATZ1* [24] or *SVP1* [7] significantly reduced plant height, and that overexpression of these genes partially restored normal plant height. The functional validation of the roles of *PLATZ1* and *SVP1* in stem development suggested that cluster c10 may

correspond to the stem intercalary meristem, a zone at the base of grass internodes that contributes new cells to the elongating internodes [25]. This hypothesis was further supported by our finding that *LFY* also plays an important role in wheat stem development.

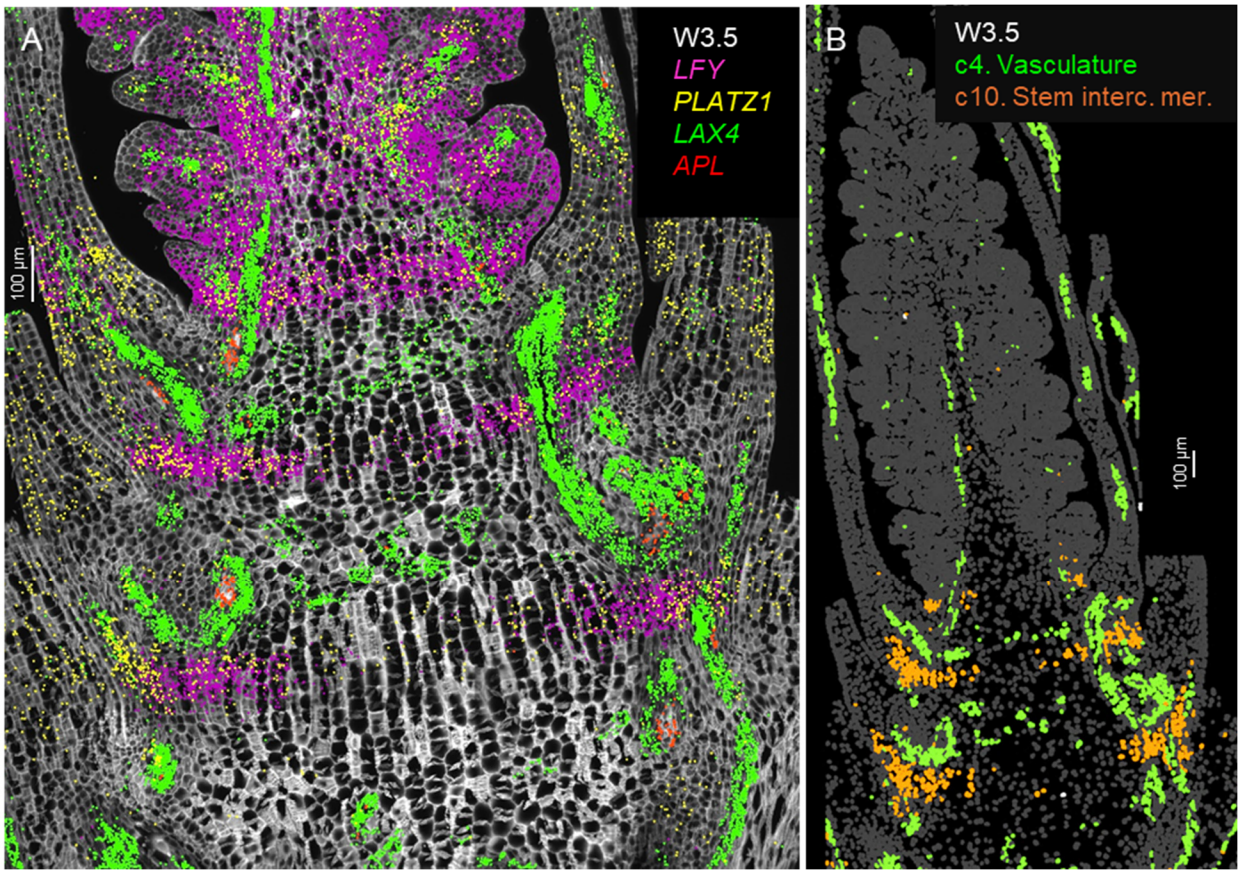

Fig. 2. Single-molecule fluorescence *in-situ* hybridization (smFISH) characterization of the basal region below the spike at W3.5. Cell walls were stained with calcofluor-white. **A** Vasculature patterning genes *LAX4* (green) and *APL* (red) are concentrated in the region with shorter cells (likely 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.

Using the *lfy*-null mutant developed in a previous study [13], we showed here that the mutant has no significant effect on stem length, but resulted in significantly wider nodes (22.3% ± 1.8%) and internodes (11.5% ± 2.0%). Moreover, the mutants produced more solid stems, showing a 62% ± 17% reduction of the diameter of the hollow central region relative to the WT (Fig. S9, Data S4). The *Short and Solid Culm* (*ssc*) rice mutant is a *LFY* allele with a D16N amino acid polymorphism and a 3-fold higher transcript levels than the wildtype, which has been associated with reduced stem length and diameter [26]. These results revealed a shared role of *LFY* on stem

development between wheat and rice and are a good example of the value of the smFISH data to generate functional hypothesis.

Since the gibberellic acid (GA) hormone plays a critical role in stem elongation, we examined the spatial distribution of two genes involved in GA biosynthesis (*GA20ox1* and *GA20ox2*) and one involved in GA degradation (*GA2ox1*). At W2.5 and W3.5, the two GA biosynthetic genes were enriched in c0 at the basal region below the spike (Fig. S10B-C, Data S2). Loss-of-function mutations in either *GA20ox1* or *GA20ox2* reduced the levels of bioactive GA_1 and GA_4, resulting in shorter plants than the wildtype [27]. These findings support a positive role of these genes in wheat stem elongation, consistent with their spatial expression adjacent to the stem intercalary meristem region (c10).

The catabolic gene *GA2ox1* showed higher levels of expression in the central region of the developing spikes at W1.5 (cluster c15) than at later stages (cluster c3, Fig. S10A, Data S2). We speculate that the elevated *GA2ox1* expression at early spike development prevents premature cell elongation and induction of *SOC1* and *LFY*, which are associated with increased GA levels [28]. At W2.5 and W3.5, *GA2ox1* expression was restricted to the periphery of the region enriched in GA biosynthetic genes, possibly forming a ring of reduced GA abundance around the elongating cells. The negative role of *GA2ox1* in stem elongation has been functionally validated in rice plants by both over-expression and ectopic expression of *GA2ox1* under the *OsGA3ox2* promoter [29]. We hypothesize that the spatial distribution of GA metabolic and catabolic genes described in this study may play an important role in normal wheat stem elongation.

1.2. The transition zone between leaf and spike

At W2.5 and W3.5, a group of cells within cluster c19 was detected in the transition region between leaves and spike (Fig. S11A-B). The genes preferentially expressed in cluster c19 include the *SPL* transcription factors *SPL2*, *SPL14*, and *SPL17* (Fig. S11C-D), the *TCP* transcription factor (*TEOSINTE BRANCHED1/CYCLOIDEA/PROLIFERATING CELL FACTOR*) *TB1* and *TB2* (Fig. S11E-F), and the MADS-box genes *FUL2* and *SOC1-1* (Data S2). Several of these genes have previously been associated with the regulation of spike architecture.

Mutations in the *SPL14* and *SPL17* genes result in the formation of ectopic bracts at the base of the inflorescence in rice [30], maize [5] and hexaploid wheat [4]. Based on the basal location of

236 these bracts, we hypothesize that the preferential expression of *SPL14* and *SPL17* in cluster c19
237 may contribute to bract repression in the transition zone. *TSH4,* the maize ortholog of *SPL17* and
238 its close paralogues *UB2* and UB3 (*SPL14*) work together to target genes responsible for axillary
239 meristem suppression such as *TB1* and auxin response regulators [31].

240 In rice, *SPL14* is a direct regulator of *TB1* [32], which in our study was also preferentially
241 expressed in cluster c19. *TB1* reduces lateral branching, and mutants of this gene in maize [33]
242 and rice [34] show increased tillering. *TB1* mutants in barley and wheat also show increased
243 tillering and produce additional lateral spikelets in barley [35] and paired spikelets in wheat [36,
244 37]. These findings indicate a conserved role of *TB1* in the regulation of branching.

245 *SQUAMOSA* genes are direct targets of the SPL proteins [38, 39] and combined loss-of-function
246 mutations in the wheat *SQUAMOSA* genes *VRN1* and *FUL2* result in de-repression of the lower
247 ridge, leading to bract outgrowth below the spikelets at the base of the spike [3]. Individual *vrn1*
248 or *ful2* mutants show no bract outgrowth, demonstrating that *FUL2* and *VRN1* play redundant
249 roles in bract repression.

250 Based on the physical location of the c19 cells and the function of the genes preferentially
251 expressed in this region (Data S2), we hypothesize that the C19 cells contribute to the transition
252 between the proximal region (where leaf meristems exhibit rapid development and axillary
253 meristems undergo a temporary arrest) and the distal region (where the subtending bracts are
254 suppressed and upper ridges rapidly develop into spikelet meristems).

255 **1.3. The central region of the spike**

256 The ground tissue cells in the central region of developing spikes were primarily assigned to
257 cluster c15 at W1.5 and cluster c3 at both W2.5 and W3.5 (Figs. 1, S5), which were highly
258 correlated with each other (R= 0.70). These two clusters also showed strong correlations with the
259 ground tissue at the adjacent basal region beneath the spike (cluster c0, Data S3). Cells from
260 cluster c15 showed stronger correlations with the leaf mesophyll (R= 0.74) than cluster c3 (R=
261 0.46, Data S3), suggesting progressive differentiation of the ground tissue at the center of the
262 spike. Several genes were preferentially expressed at the spike central region across all three
263 developmental stages, including *AP3*, *bZIPC4* (Fig. S12A-C), *IDD4* and *OSH1* (synonymous
264 with *KNOX3*, Fig. S12D-F). The homeobox gene *OSH1* was expressed at the base and center of

265 the spikelets, but not in developing glumes and lemmas. This expression pattern resembles that
266 of its orthologs in maize and rice [40].

267 In Fig. 3, we present the expression profiles of four vascular markers: the auxin influx carrier
268 *LAX4* [41], the phloem marker *APL* [23], and the duplicated auxin efflux carriers *PIN1a* and
269 *PIN1b* [22, 42] (Fig. 3). In *Brachypodium*, *PIN1b* plays a role in sink-finding and connecting
270 new primordia to existing vasculature, whereas *PIN1a* forms narrower canals than *PIN1b* and is
271 likely involved in vein patterning [42]. A similar pattern was observed in wheat, where *PIN1b*
272 was expressed in a broader area than *PIN1a* (Fig. 3).

273 *PIN1a* and *PIN1b* were preferentially expressed in clusters c4 and c13 (Data S2, Fig. S13A),
274 which were annotated as vasculature clusters. *APL* was among the preferentially expressed genes
275 in both c4 and c13 (Data S2) but most of its expression was outside these clusters (Fig. S13B).
276 The expression of *CEN5* overlapped well with c13 (Fig. S13B) whereas the expression of *LAX4*
277 overlapped better with c4 (Fig. S13C). *PIN1a* showed extensive overlap with both clusters (Fig.
278 S13D). The early accumulation of *PIN1a* and *LAX4* (=*ZmLAX2*) in maize leaves marked the
279 initial procambium specification [22], suggesting that c13 and c4 represent early stages of
280 vascular development.

281 In maize, *LAX4* and *PIN1a* have antagonistic functions and are frequently co-expressed in pro-
282 vascular cells [22, 41], an overlap that was also apparent in wheat spikes at all three
283 developmental stages (Fig. 3). The phloem marker *APL* was detected at developing vascular
284 tissue of leaves at W1.5 and W2.5 (Fig. 3A-B), and in the central region of the spike at W3.5
285 (Fig. 3A).

286 In summary, an important function of the c4 and c13 vascular cells in the central region of the
287 spike is to provide an efficient vascular system that transports water and nutrients to the
288 developing spikelets. The central c3 cells are green from early developmental stages, suggesting
289 that they contribute photosynthetic products to the developing spikes (see section 2.5) and also
290 provide structural support to flowers and grains.

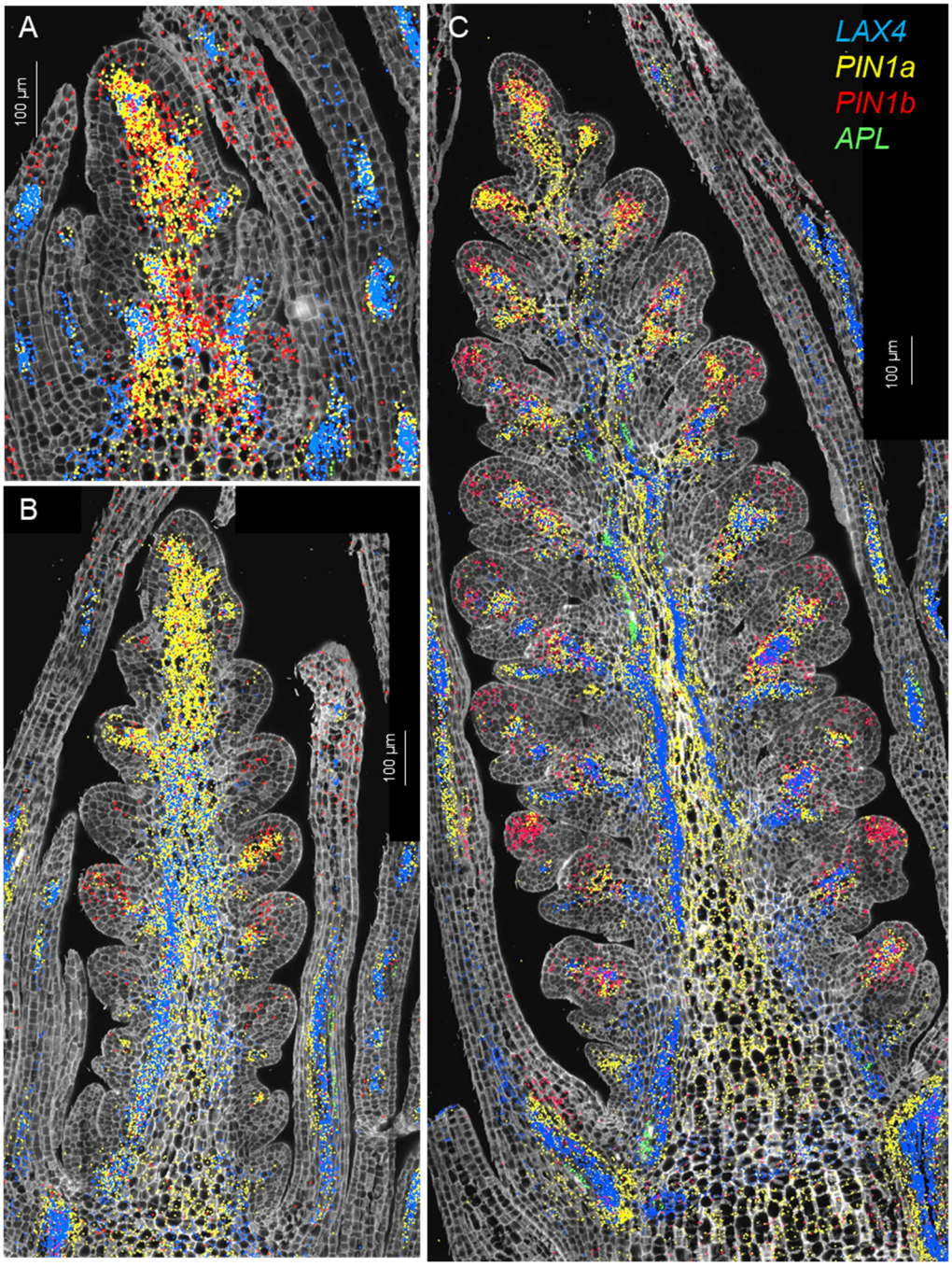

Fig. 3. Single-molecule fluorescence *in-situ* hybridization characterization of the spike central region (cluster c3). Spike central region (cluster c3). **A** W1.5, **B** W2.5, **C** W3.5. In all three panels, *LAX4* hybridization signal is shown in blue, *APL* in green, *PIN1a* in yellow, and *PIN1b*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

1.4. Region proximal to the spikelet meristem (SM)

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

301 ***1.4.1. Cluster c6 "suppressed bract".*** Cluster c6 includes the suppressed bract region (Fig. S14).
302 This cluster showed a relatively high correlation with cluster c19 in the transition zone (*R*= 0.71,
303 Fig. S14 and Data S3), which was also associated with leaf ridge repression. Both clusters
304 showed preferential expression of *SPL14* (Fig. S14A-B), *FUL2*, and *LFY* (Data S2), but cluster
305 c6 differed from cluster c19 by its preferential expression of *SPL13* (Fig. S14C-D) and lower
306 expression of *TB1* and *TB2*. In both c6 and c19 clusters, preferential expression of *SPL* genes
307 was associated with higher *VRN1* and *FUL2* expression (Fig. S14E-F). Combined *vrn1 ful2*
308 mutations resulted in bract outgrowth at the basal spikelets (Fig. S14G), suggesting a role of c6
309 and c19 cells in bract repression [3].

310 ***1.4.2. Cluster c9 "Spikelet base-center".*** Cluster c9 is located at the center of the region
311 proximal to the SM (Fig. 4) and showed high correlations (*R*= 0.75-0.80) with the surrounding
312 clusters c3, c7, c8, c14, and c16 (Data S3). This central cluster was enriched in *PIN1a* (Fig. 4A-
313 D), *PIN1b,* and *LAX4* (Data S2), suggesting a role in vascular development and the transport of
314 water and nutrients to developing spikelets. Additional genes preferentially expressed in c9
315 included *SOC1-1*, *SEP1-4, OSH1* and *TB1.*

316 ***1.4.3. Cluster c7 "Glume and lemma".*** The glume and lemma primordia, which became visible
317 at W3.5, were both included in cluster c7. Both organs showed preferential expression of *AP2L5*
318 (Fig. S6B), *FUL2* and *VRN1* (Fig. S14E-F) and reduced expression of *OSH1* (Fig. S12E) and
319 could be differentiated by the preferential expression of *SEP1-2* in lemmas and *SEP1-4* stronger
320 expression in glumes than lemmas (Figs. 4C-D, S15B). The latter result is consistent with
321 previous reports in rice [43]. Cluster c7 showed the highest correlation with the early meristem
322 cluster c8 (*R*= 0.81) from which glumes and lemmas originate, and also relatively high
323 correlations with the leaf cell clusters (*R*= 0.63-0.75, Data S3). This latter result supports the
324 hypothesis that glumes and lemmas are modified leaves [44]. Similar to leaves, glumes and
325 lemmas later turn green, contributing photosynthetic products to flowers and developing grains,
326 in addition to additional protection to the reproductive organs.

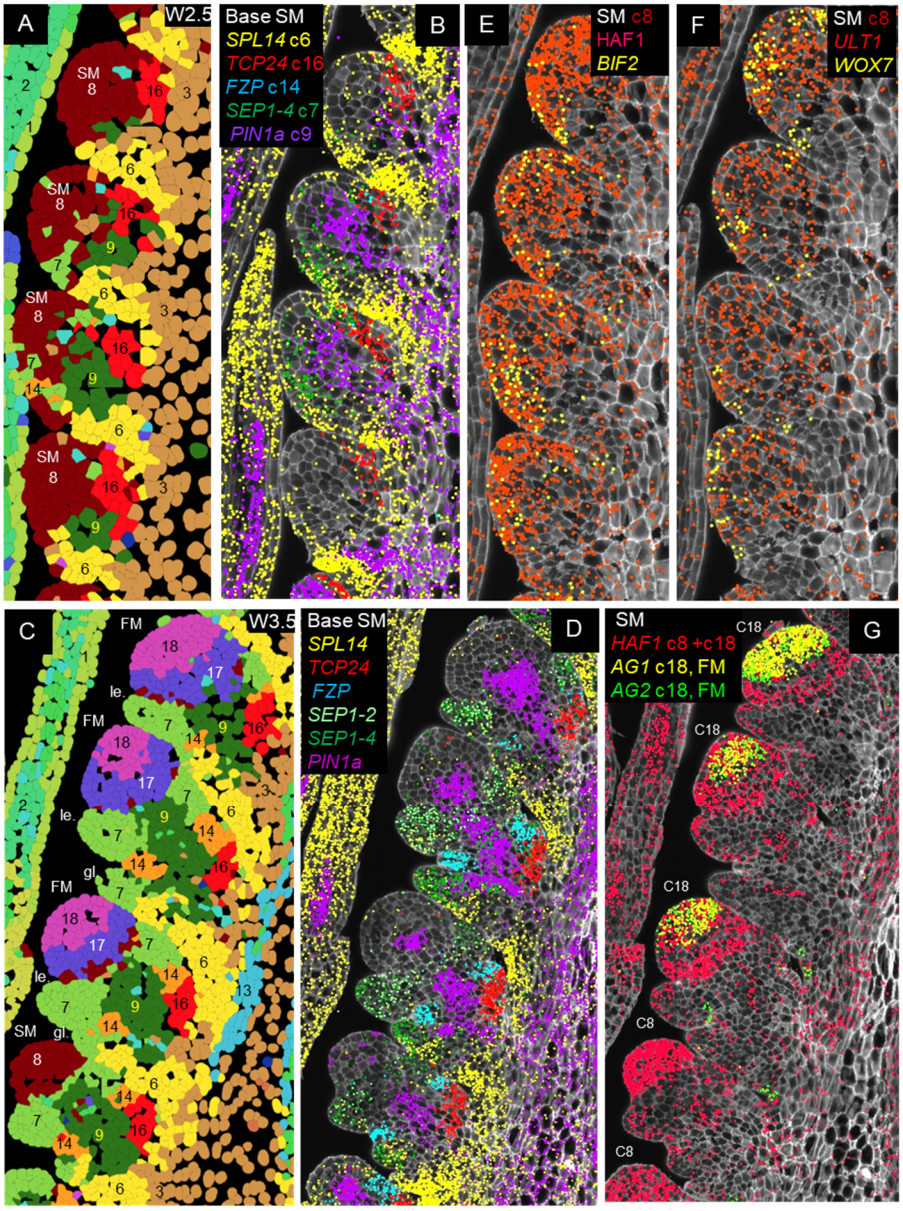

Fig. 4. Expression profiles of genes differentially expressed during spikelet development. **A-B, E-F** W2.5 (late double ridge stage). (**C-D, G**) W3.5 (floret primordia stage). **A, C** Cell segmentation and clustering. **B, D** Genes differentially expressed at the base of the SM. **E-G** Genes differentially 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 (stronger) and lemma, *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. S15-S18. SM= spikelet meristem, FM= floret meristem, le.= lemma, gl.= glume.

338 ***1.4.4. Cluster c16 "Adaxial boundary".*** This cluster was located adaxial to c9 (Figs. 4A-D, S15)
339 and was defined by the differential expression of the transcription factor *TCP24* (Fig. S15, Data
340 S2). *TCP24* and its closest paralogs *TCP22*, *TB1*, and *TB2* belong to the CYC/TB1 subclass of
341 TCP transcription factors (Fig. S11E-F) and can form homo and heterodimers [45]. In the maize
342 male inflorescence [46] and the *Brachypodium distachyon* spike [47], the orthologue of *TCP24,*
343 *BRANCH ANGLE DEFECTIVE 1* (*BAD1*), is preferentially expressed in the pulvinus and affects
344 branch or spikelet angle, respectively. The pulvinus is located at the axilla of the tassel branches
345 in maize and at the spikelet axilla in *Brachypodium* but is absent in the *Triticeae species*. Loss-
346 of-function mutations of *TCP24* in rice (=*OsTB2* = *RETARTDED PALEA 1*) are associated with
347 mild reductions in the size of the palea with no other obvious effects on inflorescence or flower
348 development [48]. By contrast, loss-of-function mutations in the barley ortholog *HvTB2* (also
349 known as *COMPOSITUM1*) result in spikes with branches in the basal region [45], suggesting
350 functional differentiation among grass species.

351 ***1.4.5. Cluster c14 "Glume axilla".*** Cluster c14 is defined by the expression of *FRIZZY*
352 *PANICLE* (*FZP,* Fig. S15B, Data S2), which is expressed at the axillae of the developing glumes
353 (Fig. 4C-D). A similar localization has been observed in previous *in situ* hybridization studies in
354 rice and maize [49, 50]. In spike longitudinal sections, c14 cells frequently appear as two small
355 cell clusters flanking c9, with the adaxial one adjacent to c16 (Figs. 4C-D, S15, S16). Cluster c14
356 also showed preferential expression of the homeobox domain gene *WOX7* (Fig. S16A) and the
357 PINOID-like serine/threonine kinase *BIF2* (Fig. S16B), although these genes are also expressed
358 outside of c14. Previous studies have shown that *FZP* plays an important role in the repression of
359 the glumes' axillary meristem. Natural or induced amino acid changes in FZP result in the
360 formation of branches at the base of the spikes in "miracle wheat" (*bh1* mutant) and barley
361 (*com2* mutant) [51, 52].

362 To characterize the function of *FZP* in tetraploid wheat, we generated loss-of-function mutations
363 in both *FZP* homeologs by CRISPR (Fig. S17). Eight independent T_0 lines with mutations in
364 both *FZP* homeologs showed spikelets replaced by dense arrays of nested glumes (Fig. S17B and
365 E). Dissection of these structures revealed new axes emerging from the axillae of the initial
366 glumes, each producing multiple secondary glumes. The axillary meristems of these secondary
367 glumes repeat the cycle but produce fewer and smaller tertiary glumes. The patterns repeat until
368 the meristems were exhausted (Fig. S17F). Poorly developed floral organs were occasionally

369 observed among these recurrent glumes (Fig. S17G). The rachilla continued its development but
370 only vestigial florets were observed in its nodes (Fig. S17H). A diagram describing the
371 development of this structure is presented in Fig S17I, and the dissection of a wildtype spikelet
372 and the corresponding diagram in Fig. S17J-K. In some distal nodes, we observed lemmas with
373 longer awns subtending more normal floral organs. These results indicate that *FZP* is required to
374 suppress the axillary meristems in the glumes to ensure normal floret development. Similar
375 phenotypes have been described for *fzp* loss-of-function mutants in rice and maize [49, 50] ,
376 suggesting a conserved function in grasses.

377 In summary, the different expression domains detected in the spikelet region proximal to the SM
378 is consistent with recent grass inflorescence models proposing that the basal region of the grass
379 spikelets as an important signaling center for spikelet development [53].

380 **1.5. Spikelet meristem (SM) and floret meristem (FM) regions**

381 At all three developmental stages, cells in the IM and early SMs were assigned to cluster c8. At
382 W2.5, this cluster exhibited preferential expression of *HAF1*, *BIF2*, (Figs. 4E, S18A), *ULT1* and
383 *WOX7* (Figs. 4F, S18C-D). *HAF1*, a C3HC4 RING domain-containing E3 ubiquitin ligase, has
384 no reported role in inflorescence development [54, 55], but the other three genes are well-
385 established regulators of meristem function. *BIF2* regulates auxin transport and is required for
386 axillary meristem initiation and lateral primordia development in maize and rice [56]. *ULT1*
387 encodes a SAND domain transcription factor that controls shoot and FM activity in Arabidopsis
388 [57, 58], whereas *WOX7* is a *WUSCHEL*-related homeobox gene that contributes to
389 inflorescence architecture in rice [59]. Based on their physical location and preferential gene
390 expression, we propose that c8 cells contribute to meristem function.

391 At W3.5, the more developed central spikelets showed the formation of floral meristems
392 (Fig.S1B) that were assigned to cluster c18 (FM, Figs. 1, S2). This cluster was defined by the
393 expression of floral homeotic genes *SEP3-1*, *SEP3-2* (Fig. S18E), *PI1*, *AP3* (Fig. S18F), *AG1*
394 and *AG2* (Figs. 4G, S18G, Data S2), strongly supporting a role in floral development. The SMs
395 are not visible in the sections of the central spikelets (Fig S1), either because the larger FMs
396 displaced the adjacent SMs from the section plane or because the section was not perfectly
397 centered (Fig. S1E).

398 Cluster c18 also showed preferential expression of *WAPO1*, a gene encoding a protein that
399 interacts with LFY (Fig. S18H). Loss-of-function mutations in wheat *WAPO1* or *LFY* result in
400 downregulation of floral homeotic genes and floral defects in wheat [13, 14] and other grasses
401 [60-62].

402 The floral meristems in c18 are separated from the lemma primordia (c7) by a group of cells
403 forming a distinct band assigned to cluster c17 (Figs. 1, 4C). This cluster was defined by the
404 preferential and high expression of *LFY* (Fig. S18H, Data S2). A previous study suggested that
405 the overlap between the *WAPO1* and *LFY* expression domains (Fig. S18H) provides critical
406 spatial information for correct flower development in wheat [13]. The dynamic changes in gene
407 expression during spike development are described later in section 1.7.

408 1.6. Changes within the inflorescence meristem (IM) during the transition to a terminal 409 spikelet (IM→TS) and *SPL14* functional validation

410 In wheat, the rate of lateral meristem formation and/or the timing of IM transition to a terminal
411 spikelet (IM→TS) determine SNS. Our previous studies showed that loss-of-function mutations
412 in *LFY* or *WAPO1* reduce the rate of SM formation and SNS [13]. The reductions in SNS were
413 comparable in the single and combined *lfy wapo1* mutants, suggesting that the LFY-WAPO1
414 protein complex is essential for establishing and/or maintaining a normal rate of SM generation.
415 However, *LFY* and *WAPO1* are co-expressed in the IM only during the earliest stages of spike
416 development (W1.5), suggesting that the early presence of the LFY-WAPO1 protein complex
417 may induce a persistent change in IM function [13].

418 Since changes in the timing of the IM→TS transition can also influence SNS, we focused on
419 genes involved in the early stages of this transition. We were unable to identify marker genes
420 that clearly distinguished IMs from early SMs (both grouped in cluster c8). Because spikelets are
421 "small spikes", SMs are also inflorescence meristem, making the identification of differential
422 markers between IMs and early SMs more challenging. However, as SMs begin to differentiate
423 into spikelets, an increasing number of genes become differentially expressed relative to the IM.

424 We manually delineated the apical meristem region and compared the hybridization density of
425 the 99 genes in this region before the IM→TS (W2.5) and at the early stages of the transition
426 (W3.5). These are within-gene comparisons and, therefore, were not influenced by probe

427 hybridization efficiency. For each stage, we analyzed four sections (four biological replications
428 for W2.5 and two for W3.5, each with two subsamples, see M&M). To verify that the selected
429 W3.5 sections had already initiated the IM→TS transition, we confirmed the expression of *FZP*
430 in the youngest lateral meristems (LM, Fig. S19). The selected region encompassed both the
431 terminal spike meristem (TM) and the two youngest LMs (henceforth, TM+2LM). At W2.5 the
432 TM corresponds to the IM and LMs to the SMs, whereas at W3.5 the TM corresponds to the SM
433 of the terminal spikelet and the LMs the future glumes. Using this approach, we previously
434 observed a significant decrease in *LFY* expression and a significant increase in *FUL2* expression
435 between W2.5 and W3.5 in the TM+2LM region [13].

436 In this study, we identified 11 additional genes with significantly altered hybridization densities
437 (spots/100 μm^2) in the selected region between W2.5 and W3.5 ($P < 0.05$, Fig. S20, Data S5).
438 Among them, *SPL14* exhibited a highly significant decrease in hybridization density during the
439 transition, (6.7-fold reduction, $P = 0.0031$, Figs. S19, S20, Data S5). We also identified two
440 additional MADS-box genes (*SOC1-2* and *SEP1-4*), four bZIP transcription factors (*bZIPC1*,
441 *bZIPC3*, *FDL2*, and *FDL6*), one bHLH transcription factor (*BC1*), one auxin efflux carrier
442 (*PIN1b*), and one receptor-like kinase (*CRN*) involved in the CLV-WUS negative feedback loop
443 (Fig. S20) that were significantly upregulated during the transition. Several of these genes have
444 been associated with inflorescence phenotypes. For example, mutations in *bZIPC1* in wheat
445 drastically reduce SNS [63], whereas mutations in the rice ortholog of *SEP1-4* (*OsMADS5*)
446 impact inflorescence branching [64].

447 To explore the phenotypic effect of reduced *SPL14* expression during the IM→TS transition, we
448 generated loss-of-function mutations in both *SPL-A14* and *SPL-B14* using CRISPR (Fig. S21A-
449 B). Plants with homozygous mutations for both homeologs showed a premature IM→TS
450 transition and a 42% reduction in SNS (Fig. S21C and E, Data S6). These results suggest that
451 *SPL14* plays an important role in delaying the timing of the IM→TS transition in durum wheat.
452 In addition, the *spl14* mutant plants were 12 cm taller (Fig. S21D and F), headed two days later
453 (Fig. S21G), and produced 2.5 more leaves than the wildtype (Fig. S21H, Data S6), indicating
454 that *SPL14* plays additional roles. Similar effects on SNS and heading time have been previously
455 reported in *spl14* CRISPR mutants in hexaploid wheat, and natural alleles have been identified
456 with favorable effects on wheat grain yield [4]

457 Since SPL proteins directly regulate the expression of *SQUAMOSA* MADS-box genes, we
458 hypothesize that the reduced *SPL14* expression may be linked to the correlated increase in *FUL2*
459 and *VRN1* expression in the IM observed during the IM→TS transition. Supporting this
460 hypothesis, we previously showed that a mutation in a predicted *SPL*-binding site in the *VRN1*
461 promoter was associated with a significant increase in *VRN1* expression during early spike
462 development [65]. In summary, we identified multiple wheat genes involved in the early
463 IM→TS transition and validated the association between *SPL14* and the final number of
464 spikelets per spike.

465 **1.7. Time and pseudotime changes in spike development**

466 Spike and spikelet developmental changes over time were explored across three developmental
467 stages (W1.5, W2.5 and W3.5), along the longitudinal axis of the spike, and through trajectory
468 analysis (Fig. S22). At the early double ridge stage (W1.5), the IM and the upper ridge cells were
469 included in cluster c8 (early meristem) and a few cells in cluster c6 (bract repression, Fig. 1A
470 and S5). These two cell clusters expanded at the late double ridge stage (W2.5, Fig. 1B), and a
471 gradient of spikelet differentiation started to form along the longitudinal axis, with the most
472 advanced spikelets located at the center of the spike. This developmental gradient became more
473 pronounced at W3.5, when well-developed floral meristems were observed at the central
474 spikelets (Fig. 1C, clusters c17 and c18).

475 These dynamic changes were captured in a trajectory analysis (Fig. S22), which used cluster c8
476 (IM and early SM meristems) as the root. From c8, two trajectories diverged in opposite
477 directions: one leading to the formation of floral meristems (c17 and c18) and the other
478 branching towards the formation of the transition zone (c19) , the repressed bract (c6) and the
479 central zone (c3). A third trajectory diverged into three branches: one leading to the central
480 spikelet (c9), another to the glume and lemma primordia (c7), and a third one to the boundary
481 clusters c14 and then c16 (Fig. S22). Cluster c9 was also connected to the suppressed bract
482 trajectory (Fig. S22).

483 These predicted trajectories reflected many of the developmental changes observed during spike
484 development but also showed a few discrepancies. Although the c14 cluster appeared earlier than
485 c16 in the predicted trajectory, c16 cells were observed at the base of the developing spikelets
486 earlier than those assigned to c14, both across developmental stages and along the longitudinal

487 gradient of the spike (Fig. 1 and S2). A closer examination of Fig. S22 showed that multiple c16
488 cells connect directly to the meristematic region (c8), suggesting an alternative trajectory from c8
489 to c16 independent of c14.

490 The comparison of the different spikelets within a spike provided a precise view of the dynamic
491 gene changes leading to the formation of the FM. The meristematic markers *HAF1* and *ULT1*
492 were expressed in both early and advanced meristems, whereas *BIF2* and *WOX7* expression in
493 the meristems was restricted to the youngest SMs (Fig. S18B and D). In more advanced SMs or
494 FMs, *BIF2* and *WOX7* were detected in c14 and other clusters (Fig. S16). In SMs slightly more
495 developed than those expressing *BIF2* and *WOX7*, we detected the MADS-box gene *AGL6*,
496 which is known to play a role in paleae and lodicule development [12]. Finally, in the more
497 developed central spikelets, *AGL6* expression was restricted to a region between the lemma
498 primordia and the FM that overlapped with c17, and the FM showed expression of floral
499 homeotic genes and *WAPO1* (Fig. S18E-H).

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

501 To identify additional genes regulating wheat spike development, we conducted single-cell
502 RNA-seq (scRNA-seq) using protoplasts isolated from *Kronos* spikes at the W2.5 and W3.5
503 stages (Fig. 5A). We optimized the cell wall digestion method to enhance meristem cell
504 recovery, following a previous single-cell study on maize inflorescence stem cell [66] (see
505 Methods). Raw scRNA-seq data are available in the NCBI Sequence Read Archive (SRA) under
506 BioProject accession number PRJNA1199502. Reads were mapped to the Kronos reference
507 genome v1.1 (annotation v1.0) [67].

508 In the initial analysis, we profiled 12,304 and 13,705 individual cells from three independent
509 replicates at W2.5 and W3.5 stages, respectively (Data S7). Good integration was observed
510 among both replicates and stages (Fig. S23A). We detected expression of 67,629 genes at W2.5
511 and 70,047 genes at W3.5 (Data S7), comparable to the number of genes detected in bulk RNA-
512 seq of Kronos spikes at the same developmental stages [68]. The mean and median unique
513 molecular indices (UMIs), the number of detected genes per cell, and other statistics are reported
514 in Fig. S23A and Data S7. Cell clustering of 26,009 cells from both stagers using Seurat v5 [69]
515 identified 25 distinct clusters in the initial analysis (Fig. S23A). Nine clusters were identified

with cells at either the G2/M or S phase, marked by *CYCB2* and *HIS2A1* expression, respectively (Fig. S23B-C).

To improve the resolution of biologically relevant clusters, cells from the nine cell-cycle clusters in the initial analyses were removed, and the remaining 17,283 cells were re-clustered into the final 23 clusters described in this study (sc0 to sc22, Fig. 5B). The final single-cell clusters were visualized using two-dimensional (2D) uniform manifold approximation and projection (UMAP) plots (Fig. 5B) and a 3D-UMAP (https://dubcovskylab.ucdavis.edu/content/kronos-scrna-3d-clusters). We also developed a Shiny App (https://dubcovskylab.ucdavis.edu/content/kronos-integrated-clusters) to visualize gene expression across clusters. To quantify the relationships among clusters, we calculated the average normalized gene expression across all cells for each cluster and used these expression profiles to calculate Pearson correlations for all cluster pairs (Data S8).

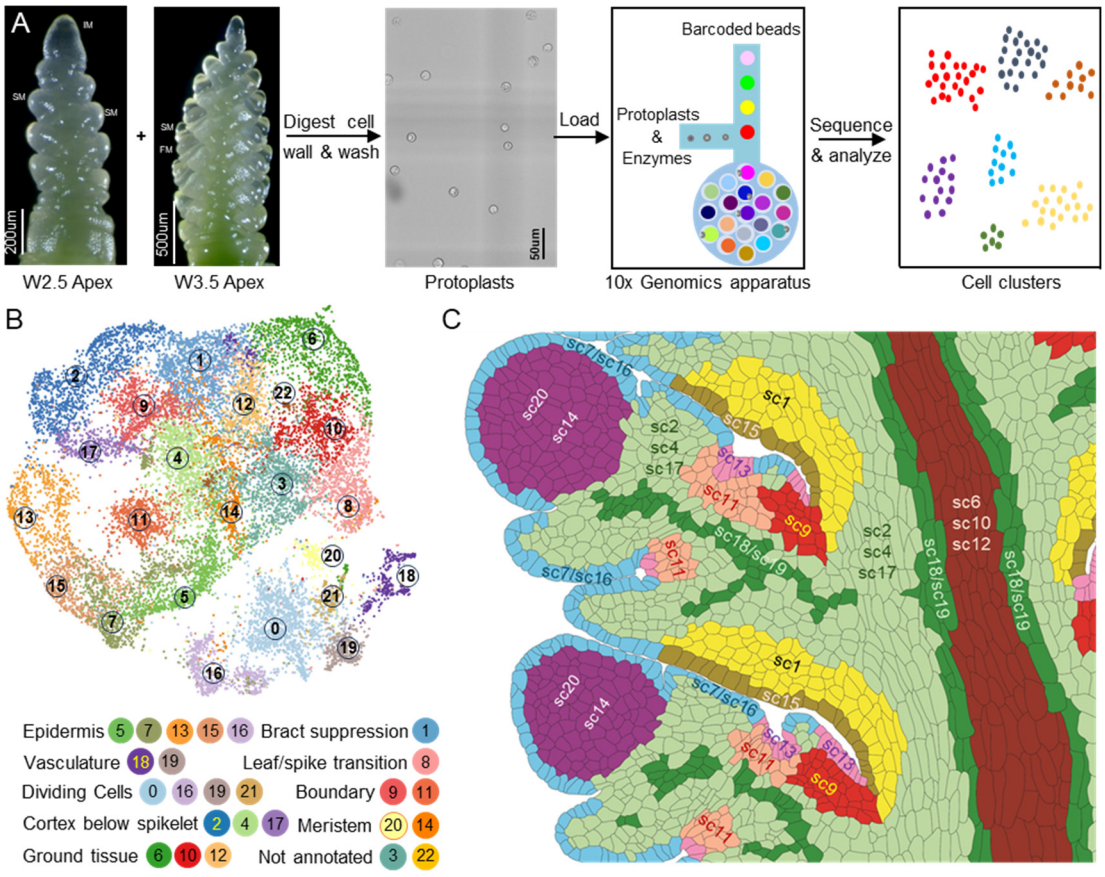

Fig. 5. Single-cell RNA-seq (scRNA-seq) analysis of wheat inflorescence spikes at the W2.5 and W3.5 stages. **A** Experimental workflow, including wheat protoplast isolation and subsequent scRNA-seq. **B** UMAP plots showing

531 cell clustering and their annotations, **C** Cartoon with inferred positions of the different scRNA-seq clusters based on
532 information discussed in the next sections. The cartoon is based on the same two central spikelets shown in Fig. 4C,
533 allowing for direct comparisons with the smFISH clusters.
534

535 **2.1. Characterization of scRNA-seq cell clusters and integration with spatial clusters**

536 To annotate the scRNA-seq clusters, we identified and annotated genes preferentially expressed
537 in each cluster (Data S9). We used this annotation to identify known marker genes (Data S1) and
538 to perform KEGG pathway analyses (Data S10). We also incorporated spatial information into
539 the scRNA-seq clusters using RNA-seq data from the basal, middle, and top sections of Kronos
540 developing spikes at W3.0 (Fig. S24 and Data S11), as well as from correlations with published
541 spatial transcriptomics clusters from hexaploid wheat (Data S13-15) [21].

542 To integrate the scRNA-seq and smFISH data, we used an imputation approach recently
543 published in barley that incorporates scRNA-seq information into smFISH-labelled cells while
544 retaining their spatial information [70]. Briefly, we analyzed the cells from both studies together
545 using the 91 common genes and for each smFISH-labelled cell we identified the five closest
546 scRNA-seq cells. We then used the average expression of these five cells (weighted by their
547 distances) to impute the scRNA-seq expression into the smFISH-labelled cell (see Materials and
548 Methods). We developed a web tool to visualize the spatial expression profiles of the 74,464
549 imputed genes in two spike sections at W2.5 and two at W3.5
550 (https://dubcovskylab.ucdavis.edu/imputed-genes). This tool greatly expanded the spatial
551 information available to annotate the scRNA-seq clusters.

552 We used a similar approach to project the smFISH cluster ID into the scRNA-seq cells (see
553 Materials and Methods). Since the number of genes in the smFISH clusters affected the number
554 of identified cells, we divided the number of projected cells by the total number of cells in the
555 smFISH clusters and generated a table including the "%-contribution" of each smFISH cluster to
556 each scRNA-seq cluster (Data S12).

557 The different sources of information used to annotate the scRNA-seq clusters are summarized in
558 Data S16 along with the final cluster annotations. The use in these annotations of similar names
559 as the smFISH clusters indicates a better overlap than with other clusters and not a perfect one-
560 to-one correspondence. The tentative physical locations of the scRNA-seq clusters are presented

561 schematically in Fig. 5C as a reference for the evidence discussed in the sections below, which
562 are grouped by similar functions.

563 **2.2. Cell cycle enriched clusters**

564 Although cells from the initial cell-cycle clusters were removed, four of the 23 final clusters
565 exhibited enrichment in cell-division marker genes combined with markers for other tissues. The
566 S-phase markers *HIS2A1-1* and *HIS2A1-2* were enriched in clusters sc0, sc16, sc19, and sc21
567 (Fig. S25). Cluster sc16 was also enriched in G2/M phase cyclin markers (*CYC7*), cyclin-
568 dependent kinases (*CDKB2*), and the *GROWTH REGULATING FACTORs GRF2*, *GRF4*, and
569 their cofactor *GIF1* (Fig. S25). Gene IDs and references for all marker genes are provided in
570 Data S1.

571 Among the genes preferentially expressed in cluster sc21 (Data S9) that exhibited significant
572 differences in the RNA-seq study, 65% were expressed at higher levels in the base of the spike
573 (Data S11-S16) suggesting a predominantly basal location for sc21. The smFISH cluster with the
574 highest contribution to sc21 was c3 (central spike) supporting the previous result (Data S12). The
575 basal and central region of the Kronos spike was characterized by a green coloration (Fig. 5A),
576 indicative of the presence of chlorophyll. This observation suggested that sc21 was enriched in
577 genes within the photosynthetic pathway, a hypothesis supported by the KEGG analysis (Data
578 S10). Based on these findings, cluster sc21 was annotated as "cell cycle / photosynthetic ground
579 tissue / basal region" (see section 2.5 for additional support).

580 Most of the genes preferentially expressed in sc0, sc16, and sc19 (Data S9) were expressed at
581 higher levels in the distal third of the developing spikes (Data S16). This enrichment was
582 strongest in sc0 (100% of the genes), intermediate in sc16 (78%), and weak but significant in
583 sc19 (47%). Based on its distal position and enrichment in meristematic markers (section 2.10),
584 sc0 was annotated as "cell cycle / meristematic / distal region". Among the smFISH clusters, the
585 one with the highest %-contribution to sc0 was c8 (early meristem) supporting the previous
586 result (Data S12). Cluster sc19 was also enriched in vascular markers (see section 2.4) , and
587 among the smFISH clusters projected into sc19 the highest value was for c13 (vasculature, Data
588 S12). Based on these results, sc19 was annotated as "cell cycle / vasculature / enriched in spike
589 tip". Finally, sc16 showed multiple epidermal markers (see section 2.3) and KEGG enrichment
590 in cutin, suberin, and wax biosynthesis (Data S10),which supported its annotation as "cell-cycle /

epidermis / enriched in spike tip". Additional information for sc16 is provided in the next section on epidermal markers.

2.3. Epidermal clusters

A similar enrichment in cutin, suberin, and wax biosynthesis was detected in the KEGG analysis for cluster sc15, supporting its annotation as an epidermal cluster. Clusters sc5, sc7, and sc13 were also annotated as epidermal clusters based on the preferential expression of known epidermal markers, including *ONI* (synonymous with *KCS10*), *PEL3*, *SCR*, *ROC2*, *ROC3*, *ROC7* (Data S1, Fig. S26). The Kronos imputed expression data confirmed the epidermal expression of these six genes (Fig. S26H). Since no epidermal markers were included in our smFISH experiment, these results exemplify the power of the imputation approach.

We compared the single-cell epidermal clusters to a recent hexaploid spatial transcriptomics study that identified two expression domains, ED1 and ED4, associated with epidermal cells [21]. These two epidermal expression domains had the highest correlations with sc7, sc13, sc15, and sc16 (Data S13-15), further supporting the epidermal annotation of these clusters. However, their correlations with sc5 were intermediate (Data S13-15). To identify markers that differentiate the five epidermal clusters, we performed a separate comparison including only these clusters and summarized the results in Data S17.

Epidermal clusters sc16 and sc7 were highly correlated (R= 0.79) and were both enriched in genes with higher expression levels at the spike tip (sc7= 50% and sc16= 69%). These clusters also showed high expression levels of ribosomal genes (Data S16), suggesting rapid cell growth and division. Cluster sc16 showed preferential expression of multiple cell-cycle genes (Fig. S25A-G, Data S17) and higher correlations with cell-cycle clusters sc0, sc19 and sc21 than all other epidermal clusters (Data S8). Cluster sc16 was also enriched in meristematic markers *WOX3a* and *WOX3b* (Data S17). The smFISH cluster c8 (early meristem) showed the highest %-contribution to sc16 relative to other smFISH clusters, supporting the previous result (Data S12). Taken together, these results justify the annotation of sc16 as "cell cycle / epidermis / enriched in spike tip".

Epidermal cluster sc7 was enriched in both meristematic and lateral organ markers. The enriched meristematic markers included *AGL6* (Fig. S18B and S27A), and the imputed *WOX3b* (Fig.

620 S27J) and *CLE33* (Fig. S27K). Cluster sc7 was also enriched in *YAB3* and *YAB4* genes (Fig.
621 S27C, Data S17), which were detected at the lower ridge and lateral organs in hexaploid wheat
622 [21]. This expression pattern was validated in our imputed expression data (Fig. 27L).
623 Correlations between epidermal clusters sc7 and sc16 and the hexaploid wheat expression
624 domains ED4 (meristem epidermis) and ED1 (glume and lemma epidermis) [21] were the
625 highest in both directions and in all three developmental stages (Data S13-S15). In addition, the
626 smFISH clusters with the highest %-contribution to sc7 were c7 (glume and lemma) and c8
627 (early meristem) relative to other smFISH clusters (Data S12). Taken together, these results
628 support the annotation of epidermal cluster sc7 as "epidermis / meristem and lateral organs"
629 (Figs. 5C).

630 Epidermal cluster sc15 showed the second highest correlations with ED1/ED4 after sc7/sc16 in
631 hexaploid wheat (Data S13-15). However, sc15 differed from sc7 by its lower expression of
632 genes encoding ribosomal proteins and in its more uniform distribution along the spike (Data
633 S16). Among the markers genes identified for sc15, *SPL13* and *SPL14* (Fig. S26G and S27D-E)
634 suggest that sc15 may correspond to the abaxial base of the spikelets including the suppressed
635 bract region (Fig. S14A-D). This hypothesis is also supported by the higher correlation observed
636 between sc15 and the suppressed bract cluster sc1 (R= 0.63) than with meristematic clusters sc14
637 and sc20 (Data S8). Consistent with the previous result, the smFISH cluster with the highest
638 relative %-contribution to sc15 was c6 (suppressed bract, Data S12). Based on these results, sc15
639 was annotated as "epidermis / suppressed bract" (Fig. 5C).

640 Epidermal cluster sc13 showed a relatively high correlation with sc15 (R= 0.77) but differed by
641 its preferential expression of *TCP24*, a marker gene for the adaxial boundary domain (Fig. 4A
642 and C). Cluster sc13 was enriched in additional boundary markers such as *TCP22, CUC3, LAX1*,
643 and *DP1*, which have been found to be preferentially expressed in the boundary domain in
644 hexaploid wheat [21] and in our imputed expression profiles (Figs. S27M-P, see also later
645 section 2.12). The sc13 cluster sowed higher correlations with the hexaploid wheat expression
646 domain ED12 (boundary domain) [21] than with epidermal clusters ED1/ED4 and all other
647 clusters (Data S13-15). Relative to other smFISH clusters, c16 (adaxial boundary) showed the
648 highest %-contribution to sc13 (Data S12). Based on these results, sc13 was annotated as
649 "Epidermis / adaxial boundary" Fig. 5C.

650 Finally, sc5 was the most divergent of the epidermal cluster, showing the lowest correlation with
651 other epidermal clusters, the lowest expression levels of epidermal markers *SCR*, *PEL* and *ONI*
652 (Fig. 26G), and the largest number of differentially expressed genes (Data S17) with two
653 examples in (Fig. S27H-I). The smFISH clusters with the highest contribution to sc5 was c19
654 (transition zone), suggesting an enrichment in epidermal cells from the transition zone. However,
655 based on sc5 intermediate/high levels of protoplasting induced genes (Data S16), high
656 correlations with unknown clusters sc3 and sc22 (Data S8), and uniformly low correlations with
657 spatial clusters in hexaploid wheat (Data S13-15), we designated sc5 as "Epidermis / unknown".

658 **2.4.Vascular clusters**

659 The vascular genes *APL*, *LAX4*, *CEN5*, *PIN1a*, and *PIN1b* characterized by smFISH (Fig. 3)
660 were all preferentially expressed in sc18 and sc19, along with additional known vascular marker
661 *SHR1*, phloem markers *JUL1* and *DOF19*, and xylem markers *TMAAT, TMO5L1, TMO5L3*,
662 *WAT1*, and *XCP1*, *LOG10b*, and *NAC52* (Fig. S28, Data S1). The vascular location of the last ten
663 markers was validated in the Kronos imputed expression profiles (Fig. S28H), together with *APL*
664 and *LAX4* controls characterized also by smFISH (Fig. 3). The smFISH cluster with the highest
665 contribution to sc18 and sc19 was c13 (vasculature, Data S12). These results support the
666 annotation of sc18 and sc19 as vascular clusters (Fig. 5C).

667 Both vascular clusters shared relatively high expression levels of *OSH1* and genes encoding
668 ribosomal proteins (Data S16). However, the genes preferentially expressed in sc19 (Data S9)
669 were expressed at higher levels at the tip of the spike (Data S16), whereas those preferentially
670 expressed in sc18 (Data S9) were expressed at higher levels at the base and middle of the
671 developing spikes (Data S16). Since sc19 was enriched in *HIS2A* expression and sc18 was not
672 (Fig. S25), sc19 was annotated as "cell cycle / vasculature / enriched in spike tip" and sc18 as
673 "vasculature / enriched in spike base and middle".

674 Cluster sc18 showed two distinct regions: one preferentially expressing phloem markers (Figs.
675 S28B and D) and the other expressing xylem markers (Figs. S28C, E-F). To compare these
676 regions, we further divided sc18 into three subclusters: sc18_0 with the highest correlation to the
677 ground tissue cluster c3 (R= 0.625), sc18_1 including phloem markers *JUL1* and *DOF19*, and
678 sc18_2 including xylem markers *WAT1*, *XCP1*, *TMO5L1*, *LOG10b* and *NAC52* (Fig. S28H).
679 Genes differentially expressed in sc18_1 and sc18_2 are available in Data S18.

680 Consistent with the key role of hormones in vascular development [71, 72], sc18 showed
681 preferential expression of both homeologs of the auxin signaling genes *PIN1*, and *ARF11*, and
682 the JA biosynthetic gene *OPR3*. Cluster sc18 also showed preferential expression of the *CPS1*
683 gene encoding a GA biosynthesis enzyme, and genes *CKX-B3*, *CKX-B5*, and *CKX-B11*, all
684 involved in cytokinin degradation (Fig. S28I).

685 Interestingly, the floral homeotic genes *AP3* and *PI1* (Fig. S18F) were also detected in cluster
686 sc18 (Figs. S29A-B). A closer examination of the smFISH profiles revealed that these two
687 MADS-box genes were co-expressed with the vascular marker *APL* at all three developmental
688 stages (Figs. S29C-E). Expression of class-B MADS-box genes in the developing vasculature
689 has been reported in a few eudicot plant species [73]. Our findings extend this conserved
690 function to monocots.

691 **2.5. Central spike clusters**

692 In addition to vascular and pro-vascular cells, the central region of the developing spike contains
693 ground tissue cells, which correspond to cluster c3 (Fig. 1 central spike) in our smFISH study,
694 and expression domains ED3 (developing rachis) and ED5 (rachilla) in the hexaploid wheat
695 study [21]. These ground tissue cells showed a green coloration indicative of chlorophyll
696 presence, which was stronger at the base of the spike (Fig. 5A). The KEGG analysis indicated
697 that genes preferentially expressed in single-cell clusters sc6 and sc10 were enriched in
698 photosynthesis pathways (Data S10). Eighty two percent of the top 50 genes preferentially
699 expressed in sc6 were annotated with photosynthesis-related functions (Data S9), including
700 *CAB2R* and *LHCB1.1*, which were detected in the central spike region in hexaploid wheat [21]
701 and were validated in the Kronos imputed expression data (Fig. S30A-H). The expression levels
702 of these photosynthetic genes across all clusters are presented in a heat-map in Fig. S31.

703 The highest average enrichment in photosynthetic genes was detected for sc6 followed by cell-
704 cycle cluster sc21 (Fig. S31, Data S16). This finding supported the previous annotation of sc21
705 in section 2.2 as "cell cycle / photosynthetic ground tissue / basal region". Enrichment in
706 photosynthetic genes was also observed in sc10 and sc12, but at decreasing levels relative to sc6
707 (Data S16 and Fig. S31). A large proportion of the genes preferentially expressed in sc10 (44%)
708 and sc12 (36%) were also preferentially expressed in sc6 (Fig. S32), supporting the connections
709 among these clusters.

710 The genes preferentially expressed in sc6, sc10, and sc12 showed higher levels of expression at
711 the basal region of the developing spikes based on the RNA-seq data (Data S11-S16). This
712 location was further supported by the enrichment of *FWL3* and *SVP1* in both sc6 and sc10,
713 which are expressed in the central and basal parts of the spike (Figs. S32A-C and H). In addition,
714 both sc10 and sc12 were enriched in *OSH1*, a gene highly expressed in the central spike region
715 (Fig. S12). Cluster sc6 was enriched in additional genes expressed in the central spike region,
716 including *AP3*, *bZIPC3*, *FDL15*, *FT2*, *OSH1*, *SPL3*, and *VRT2* (Figs. S12 and S30), as well as
717 *CAB2R* and *LHCB1.1* both in hexaploid wheat and in our imputed expression data (Fig. S30H).
718 The smFISH cluster with the highest relative %-contribution to sc6, sc10 and sc12 was c3 (spike
719 central, Data S12), supporting a central spike location for the cells of these three clusters. Based
720 on the above information, clusters sc6 and sc10 were annotated as "central spike / photosynthetic
721 ground tissue / spike base", and sc12 as "tentative ground tissue / enriched spike base" (Fig. 5C).

722 **2.6. Transition zone between leaves and spike**

723 The transition zone between the leaves and the spike is represented by the smFISH cluster c19
724 (Figs. 1, S5), which shared several preferentially expressed genes with scRNA-seq cluster sc8.
725 These genes included *TB1*, *TB2*, *SPL17*, *GNI2*, *IDD10*, and *GA20ox1* (Fig. S33A) from our
726 smFISH study. This region was also enriched in *LEC1* and *TRD1*, two genes identified in the
727 hexaploid wheat data [21] and validated in our imputed expression data (Fig. S33B-F). Cluster
728 sc8 also showed high enrichment in jasmonic acid signaling genes *MYC2* and *JAZ8*, although
729 both genes were also expressed in other clusters (Fig. S33E-F).

730 Among the genes preferentially expressed in sc8, *SPL17* and its downstream target *TRD1* [30]
731 (ortholog of *OsNL1* and *ZmTSH1*) were previously shown to play important roles in bract
732 suppression in the transition zone. Loss-of function mutations in *SPL17/SPL14* in hexaploid
733 wheat [4], and its orthologs in rice [30] and or maize [5] were associated with bract outgrowth at
734 the base of the inflorescences. A similar phenotype was observed in *TRD1* mutants in barley
735 [74], rice *OsNL1* [75], or maize *ZmTSH1* [76]. This phenotype was consistent with the
736 significantly higher expression at the base of the spike of *SPL17*, *TRD1* (Data S11). Among the
737 genes in the RNA-seq experiment with significant differences among the three sections of the
738 spike, 66% of the ones expressed in sc8 showed higher levels at the base of the spike (Data S16).
739 Among the smFISH clusters, the one with the highest contribution to sc8 was c19 (transition

740 leaf/spike) followed by c3 (central spike, Data S12). Based on these results, sc8 was annotated as
741 "transition leaf-spike / spike base".

742 **2.7. Suppressed bract cluster**

743 Cluster c6, which includes the suppressed bract cluster, was consistently detected in the smFISH
744 study at the base of each developing spikelet (Fig. 1). This cluster showed strong differential
745 expression of *SPL13* and *SPL14* (Data S2), two markers that were also enriched in the scRNA-
746 seq cluster sc1 (Fig. S34) and that supported its annotation as "suppressed bract region" (Fig.
747 5C). This annotation was also supported by smFISH cluster c6 (suppressed bract) showing the
748 highest contribution to sc1 among the different smFISH clusters (Data S12). The suppressed
749 bract clusters c6 and sc1 also showed high correlations with the transition zone clusters c19 and
750 sc8, respectively (Data S3 and S8), probably reflecting their shared function in repressing the
751 leaf ridge.

752 The overlap between c6 and sc1 is likely incomplete. Cluster sc1 also showed a relatively high
753 contribution of smFISH cluster c3 (Fig Data S12), suggesting that it may include cells from both
754 the suppressed bract and ground tissue. Conversely, the c6 cluster includes epidermal cells (Fig.
755 4C-D), which were included in the "epidermal / suppressed bract" cluster sc15 in the scRNA-seq
756 study (section 2.3).

757 **2.8. Boundary clusters**

758 Among the genes analyzed in the smFISH experiment, *FZP* and *TCP24* showed the most
759 specific expression domains (Figs. 4, S15, Data S2). *FZP* was expressed in the glume axilla
760 (c14), while *TCP24* was expressed adjacent to *FZP* at the adaxial boundary region below each
761 spikelet (c16, Fig. 4C-D and S15-16). In addition to *FZP,* the hexaploid wheat spatial expression
762 data showed enriched expression in the adaxial boundary for the *WUSCHEL-RELATED*
763 *HOMEOBOX9c* (*WOX9c*, Fig. S35A), *TCP22*, *LAX1*, and *DP1* (synonymous with *BAF1*) [21].
764 We validated this expression profiles in the Kronos imputed expression data (Fig. S27M, O-P).

765 Mutations in wheat *MULTI-FLORET SPIKELET 1* (*MFS1*, synonymous with *DUO* and
766 *ERF130*) wheat have been previously associated with mild supernumerary spikelets and
767 increased grain number per spike, highlighting its role in wheat spike development [77]. This
768 motivated us to include *MFS1* in the smFISH experiment, which showed high levels of

769 expression in clusters c14 and c16, and lower expression in other clusters (Fig. S35B-C). The
770 high expression in c14 and c16 correlated with high *MFS1* expression in the scRNA-seq clusters
771 sc9 and sc11 (Fig. S35D), prompting us to investigate a potential overlap among these clusters.

772 Three lines of evidence supported the annotation of single-cell cluster sc11 as "glume axilla"
773 (Fig. 5C). First, glume axillary markers *FZP* and *WOX9c* were both preferentially expressed in
774 sc11 (Fig. S36). Second, c14 (glume axilla) showed the highest %-contribution to sc11 among
775 the smFISH clusters, (Data S12). Finally, RNA-seq results showed that sc11 was significantly
776 enriched in genes with higher expression in the center of the spike (Data S16), mirroring the
777 expression patterns of *FZP* and *WOX9c*.

778 We annotated the scRNA-seq cluster sc9 as "adaxial boundary" based on multiple shared
779 markers with smFISH cluster c16. *TCP24* was preferentially expressed in sc9 and was also
780 abundant in sc13 (Fig. S37), which includes the epidermal cell layer of the adaxial boundary
781 domain (section 2.3). Other markers preferentially expressed in the adaxial domain, such as
782 *TCP22, LAX1,* and *DP1* were also preferentially expressed in sc9 and/or sc13 (Fig. S37). The
783 smFISH cluster c16 (adaxial boundary domain) %-contribution to sc9 was more than two-fold
784 higher than the second highest (c9 central-base spikelet), which was also more than two-fold
785 higher than the other smFISH cluster (Data S12).Taken together, these results support the
786 annotation of sc9 as adaxial boundary domain.

787 **2.9. Cortex cells below developing spikelets**

788 In the wheat spike, the cortex consists of a layer of cells situated between the vascular tissues and
789 developing spikelets (Fig. 1). The parenchymatic cells in this region are expected to resemble
790 those in the central ground tissue. As shown in Fig. 4C, the cell layer between the vasculature
791 and the SM base includes ground tissue cells from the central cluster c3 as well as the innermost
792 cells from clusters c6 (suppressed bract), c7 (glume – lemma), c9 (SM base-center), c16 (adaxial
793 boundary) and c14 (glume axilla). Since cortical cells show no greening and are not
794 meristematic, the scRNA-seq clusters including these cells are expected to show low enrichment
795 in photosynthetic genes and reduced expression of *OSH1* or ribosomal protein-encoding genes.

796 Clusters sc2, sc4, and sc17 exhibited high correlations with the spatial clusters located proximal
797 to the developing spikelets in hexaploid wheat (ED0, ED3, ED5, and ED11, Data S13-S15) [21].

798 In addition, we found relatively high %-contributions of the smFISH clusters c16 and c6 to sc2,
799 from clusters c9 and c7 to sc4 and from cluster c16 to sc17 (Data S12). They also exhibited low
800 expression levels of photosynthetic and meristematic genes (Data S16) and uniform expression
801 along the spike based on the RNA-seq data. Taken together, these results suggest that sc2, sc4,
802 and sc17 may correspond to cortical cells below the spikelets (Fig. 5C).

803 The annotation of sc2 as a cortex cluster was further supported by the preferential expression of
804 *LAX1* and *DEPRESSED PALEA1* (*DP1*) (Fig. S38), two genes previously shown to be
805 preferentially expressed in the adaxial boundary below the spikelets [78-81] and validated in the
806 Kronos imputed data (Fig. S27O-P). Two additional genes, *KAN1* and *RL9*, expressed in the
807 adaxial boundary region below the spikelet in hexaploid wheat [21], were also preferentially
808 expressed in sc2 (Fig. S38), and their imputed expression is presented in Fig. S38H. Additional
809 markers enriched in the correlated sc2 and sc13 clusters (Data S8) included *ABCG11* [82] and
810 *SOT17* [83] (Data S9) and were validated in Fig. S38H.

811 A more central location of sc4, relative to sc2 and sc17, was supported by the higher %-
812 contributions of the smFISH cluster c9 (central region below spikelet) to sc4 than to the other
813 two spatial clusters (Data S12). This hypothesis was also supported by the 2.4 to 3.4-fold higher
814 expression of *OSH1* in sc4 than in sc17 and sc2, respectively (Figs. S12, S39, Data S16). In
815 addition, the cortex marker gene *NPF3.1* (also known as *NPF6.4-LIKE2* in maize [84]) was
816 enriched in sc4 (Fig. S39F-G), suggesting that this cluster corresponds to cortical cells.

817 Overall, the scRNA-seq cluster sc17 showed lower correlations with the spatial clusters in
818 hexaploid wheat than sc2, but they were still higher for the contiguous cortex clusters than for
819 the rest of the spatial clusters (Data S13-15). The highest %-contribution from the smFISH
820 clusters to sc17 was from c16 (adaxial boundary, adjacent to cluster, Data S12). We annotated
821 sc17 as "tentative cortex region below spikelet lateral regions" due to the lack of validation with
822 marker genes of known spatial distribution. Additional genes preferentially expressed in sc4 and
823 sc17 are listed in Fig. S39 and Data S9.

824 **2.10. Meristem clusters**

825 The RNA-seq clusters sc14 and sc20 showed preferential expression of meristematic genes
826 *ULT1*, *AGL6* and *SEP3-1* (Fig. S40A-D) suggesting that they included meristematic cells. This

827 hypothesis received additional support from the differential expression of the two homeologs of
828 *ULT1.* The differences between *ULT-A1* and *ULT-B1* were evident in the bubble plot in Fig.
829 S40G, in the UMAPS for the separate homeologs (Fig. S40H-I), and in the imputed expression
830 profiles (Fig. S40J-K). In the bubble plots and UMAPS, *ULT-A1* showed high expression in
831 meristematic clusters sc14 and sc20 (Fig. S40H) whereas *ULT-B1* did not (*ULT-B1* showed
832 differential expression in the central clusters sc6 and sc10, Data S9). These differences were
833 reflected in the imputed expression profiles, with *ULT-A1* expression concentrated in the early
834 (c8) and floral meristems (c17 and c18) and *ULT-B1* expression concentrated in the central part
835 of the spike (Fig. S40J-K). The correspondence between the spatial location and *ULT1*
836 homeologs and their scRNA-seq cluster assignment in the UMAPS and bubble plots supported
837 the conclusion that sc20 and sc14 include meristematic cells. More generally, the *ULT1* results
838 suggest that the expression imputation method developed by Demesa-Arevalo et al. [70] in
839 diploid barley, can be also used differentiate expression profiles between wheat close homeologs
840 in polyploid species.

841 The presence of meristematic cells in sc20 and sc14 was also supported by high levels of
842 expression of ribosomal genes (Data S16), intermediate expression levels of *OSH1* (Fig. S40G)
843 and by the higher %-contributions of the smFISH cluster c8 to both sc20 (1.5% = 33 cells) and
844 sc14 (5.3% = 114 cells) relative to other smFISH clusters (Data S12). Taken together these
845 results support the annotation of sc14 and sc20 as meristematic clusters.

846 *AGL6* was detected in early meristems (c8) immediately before their transition to floral
847 meristems (Fig. S40L), after which it became more abundant at the base of the FM (c17, Fig.
848 S40L). This result is consistent with the role of *AGL6* role in paleae and lodicule development
849 [12]. *SEP3-1* showed a similar profile but its expression started slightly later than *AGL6* and was
850 detected in both the FM (c18) and its base (c17, Fig. S40M). Fig. S40N-P shows three additional
851 genes differentially expressed in sc14 (Data S9). *SEP1-2* is preferentially expressed in lemmas
852 and a few meristematic cells (*SEP1-2*, Fig. S40N and Fig. 4). The rice homolog of *SEP1-2*,
853 *LHS1/OsMADS1*, controls floret meristem specification [85]. The second gene, *JLO*, encodes a
854 LOB domain protein gene expressed in Arabidopsis at the base of the floral meristem (c17, Fig.
855 S40O) [86]. Finally, *FCP1* was expressed in the floral meristem (c17 and c18) and c14 at the
856 glume axilla (Fig. S40P). In rice, *FCP1* is a negative regulator of *WOX4* and of meristem
857 maintenance [87]. Based on these results sc14 was annotated as "meristem proximal region".

858 Cluster sc20 was enriched in genes differentially expressed in the distal one-third of the spike
859 (Data S16), so it was annotated as "meristem / enriched in spike tip".

860 Many of the genes enriched in sc14 and sc20, also show some enrichment in epidermal clusters
861 sc7 and sc16 (Fig. S40G, small arrows), supporting their previous annotation as "epidermis /
862 meristem and lateral organs". The imputed expression data showed enriched expression of
863 *WOX3b* and *CLE33* in c8 including the epidermal cells (Fig. S27J-K). Additional genes
864 preferentially expressed in sc14 and sc20 are listed in Data S9.

865 **2.11. Protoplasting affected clusters**

866 The last two scRNA-seq clusters, sc3 and sc22 (smallest cluster with only 163 cells), showed
867 evidence of being affected by the protoplasting. First, these clusters showed the highest
868 proportion of preferentially expressed genes that were not detected in RNA-seq studies of
869 Kronos developing spikes (this study and [68]). Approximately 19% of the preferentially
870 expressed genes in sc3 and 28% in sc22, were not detected at the same stages in Kronos spikes
871 under normal conditions (Data S16). The 235 genes detected only in the scRNA-seq were
872 enriched in stress-related GO terms (Data S19), further supporting their induction by
873 protoplasting. The preferentially expressed genes in sc3 and sc22 also showed a relatively high
874 overlap with a list of genes previously annotated as potentially affected by protoplasting [84]
875 (sc3= 11% and sc22= 13%, Data S16). The percentage of genes identified by these two criteria
876 across the 23 single-cell cluster showed a high correlation (R= 0.81), indicating a similar
877 underlying process. Cluster sc3 showed a high correlation with sc10 (R= 0.85, Data S8) and sc22
878 with cluster sc14 (R= 0.88, Data S8), suggesting that these clusters may contain cells related to
879 those clusters but affected by protoplasting. Based on these results, we annotated sc3 and sc22 as
880 "unknown".

881 **2.12. Co-expression analyses**

882 The generation of smFISH and scRNA-seq datasets from the same genotype, tissues, and
883 developmental stages provided the opportunity to identify additional genes co-expressed in
884 specific expression domains. We focused on four genes with specific expression domains —*FZP*
885 and *TCP24* in the boundary domains and *ULT1* and *AGL6* in the meristem— and calculated
886 Pearson correlations between each of them and all other genes across all cells in the scRNA-seq

887 dataset. The genes with *R* values exceeding four standard deviations above the mean (close to
888 zero as expected) were selected as co-expressed genes for further analyses (Data S20-S24).

889 ***2.12.1. Genes co-expressed with boundary genes FZP (c14)* and *TCP24 (c16)***

890 The genes with the highest co-expression values for *FZP* and *TCP24* were their respective
891 homeologs, serving as positive controls (Data S20-S21). We found 608 genes significantly co-
892 expressed with *FZP* and 606 with *TCP24*, with 226 shared (988 total).

893 Among the shared co-expressed genes, we identified the well-studied gene *CUP-SHAPED*
894 *COTYLEDON 3* (*CUC3*, Data S20-S21), a NAC-domain transcription factor involved in
895 meristem / organ boundary specification in rice [88] and Arabidopsis [89]. *In situ* and scRNA-
896 seq experiments in maize demonstrated that *CUC3* is expressed in both the *FZP* and *TCP24*
897 domains [81, 84, 90], a result validated by our imputed expression data (Fig. S27N). The *MFS1*
898 gene also showed a significant correlation with both *FZP* and *TCP24* (Data S20-S21) and was
899 co-expressed with both genes in the smFISH and imputed expression data (Fig. S35B-C).

900 Among the genes showing significant co-expression with *FZP* and not with *TCP24*, we
901 identified *WOX-9c* and *CLE33* (Data S20. Fig. S35). *WOX-9c* is a close paralog of *WOX7* [91,
902 92] with a similar expression profile as *FZP* (Fig. S35A). *CLE33* showed a similar expression
903 profile as *WOX7* (Figs S16 and S18C), changing from expression in the early meristems (Fig.
904 S27K) to an overlap with *FZP* in the more developed spikelets with FMs (Fig. S35E). In
905 Arabidopsis, RNA interference of *WOX9* in reproductive meristems resulted in the conversion of
906 flowers into inflorescence-like structures [93], resembling the conversion of spikelets into
907 branches observed in *fzp* natural mutants in wheat and other grass species [49-52]. The rice loss-
908 of function mutations in *WOX7* (*dwt1*) resulted in larger panicles, suggesting that these two
909 related *WOX* genes have functional roles in inflorescence development [59].

910 *TCP24* showed high co-expression with *TCP22* (the close paralog of *TCP24*), *DP1*, and *LAX1*
911 (Data S21), three genes previously confirmed to be expressed in the adaxial boundary domain in
912 hexaploid wheat [21] and validated in our imputed expression data (Fig. S27M-P). *DP1* is an
913 ortholog of the maize gene *BARREN STALK FASTIGIATE1* (*BAF1*, Data S21), which is
914 expressed in a narrow domain adaxial to developing axillary meristems in maize [90]. Mutations
915 in *BAF1* resulted in the absence of ears or their fusion to the main stalk, supporting its role in
916 demarcating the boundary of the developing axillary meristem. *LAX1* is required for the initiation

917 of aerial lateral meristems and is expressed in the adaxial boundary region of the spikelet
918 meristems in both rice and maize (*BARREN STALK1*, *ZmBA1*) [78, 79]. A strong correlation was
919 also observed between *TCP24* and *RAMOSA2* (*RA2*, Data S21), which is preferentially expressed
920 in cluster sc9 in the imputed expression data (Fig. S35F). In maize, *RA2* is expressed in a group
921 of cells that predict the position of the axillary meristem formation [94]. Additional genes co-
922 expressed with *TCP24* are listed in Data S21. Based on the physical location of *TCP24* and its
923 co-expressed wheat genes, together with the known functions of the maize homologs, we
924 hypothesize that c16 cells contribute to delimiting the wheat SM and act as an important
925 signaling center for spikelet development [95].

926 To investigate the interactions centered on *TCP24* and *FZP*, we generated a gene regulatory
927 network using GENIE3 [96], focusing on the 988 genes co-expressed with *FZP* (Data S20)
928 and/or *TCP24* (Data S21) in the scRNA-seq dataset. From the top-500 significant interactions,
929 we constructed a network visualized in Cytoscape 3.10.3 (Fig. S42). The 95 genes included in
930 this network were all transcription factors (Data S22) suggesting that *FZP* and *TCP24* are part of
931 a broad regulatory network including many other transcription factors. The analysis revealed
932 connections between the *TCP24* homeologs, which share multiple targets, including their closest
933 paralog *TCP22*. *TCP24* also showed regulatory links with *LAX1* homologs and *RA2*, supporting
934 their coordinated expression in the adaxial boundary (Fig. S42, Data S22). Similarly, the two
935 *FZP* homeologs were connected and shared common links including *WOX9c*. Additionally,
936 *TCP24* and *FZP* shared several downstream targets, such as the boundary regulator *CUC3*,
937 suggesting a potential functional convergence (Fig. S42, Data S22). This regulatory network
938 analysis provides a framework to investigate the genetic interactions that define the functions of
939 these boundary domains.

940 ***2.12.2. Genes co-expressed with meristem genes ULT1* and *AGL6***

941 The smFISH results for *ULT1* showed expression in the IM, SM and FM meristems as well as in
942 the central spike region (Fig. 4F and S18C-D). However, the co-expression analysis revealed
943 limited correlation between *ULT-A1* and *ULT-B1* homeologs (Data S23) supporting their
944 functional differentiation described in section 2.10 and Fig. S40H-K. Interestingly, among the
945 100 genes showing the highest correlations with *ULT-A1*, 82 encoded ribosomal proteins (Data
946 S23), suggesting a higher rate of protein production and rapid cell growth and division. This

947 enrichment was not observed for *ULT-B1*, supporting the conclusion that *ULT-A1* is the
948 homeolog predominantly expressed in meristematic cells. A heatmap based on the 45 ribosomal
949 genes with the highest correlation values with *ULT-A1* revealed that cluster sc20 exhibited the
950 highest expression level among the 23 single-cell clusters (Fig. S43). Cluster sc14 also showed
951 relatively high levels of ribosomal genes, but lower than those in sc20.

952 Co-expression analysis of *AGL6* revealed significant correlations with floral homeotic genes
953 *SEP3-1, SEP3-2*, *WAPO1*, *AG1*, and *PI1* (Data S24). Among the 100 genes showing the highest
954 co-expression values with *AGL6*, 43 encode ribosomal proteins (Data S24). A similar enrichment
955 of ribosomal protein encoding genes has been reported in a single-cell study of maize stem cells
956 [66]. This co-expression studies also revealed genes with previously uncharacterized roles in
957 inflorescence development (Data S23-S24). We explored the imputed expression of these genes
958 and prioritized three that were expressed in the spike meristematic regions for further functional
959 characterization: the remodeling genes *DDM1a* [97] (Fig. S41B) and *BTBN16* (Fig. S41C), and
960 the auxin response gene *SHI1* [98] (Fig. S41D).

961 **2.13. scRNA-seq trajectory analysis**

962 To summarize the relationships among the scRNA-seq clusters, we performed a trajectory
963 analysis (Fig. S44) starting from the early meristematic cluster sc14. In the lower region of the
964 graph, one branch of the trajectory connected sc14 with the vascular clusters sc18 and sc19 (and
965 the cell division clusters), and the other branch with the epidermal clusters (Fig. S44). The early
966 connection of sc14 with sc5 (epidermis unknown) and sc7 (epidermis meristem and lateral
967 organ) suggested that these clusters may correspond to early stages of differentiating epidermis
968 cells. The same trajectory later connected with sc15 (epidermis repressed bract) and sc13
969 (epidermis adaxial boundary).

970 In the opposite direction, sc14 was linked to two separate loops, each with multiple internal
971 connecting trajectories. The first loop included the ground tissue clusters sc6, sc10 and sc12 and
972 the transition zone sc8. The second interconnected loop included the cortex clusters sc2, sc4 and
973 sc17, and the related repressed bract cluster sc1 and adaxial boundary sc9. The glume axilla
974 cluster sc11 remained disconnected from the rest despite its proximity to cortex cluster sc4 and
975 epidermal cluster sc13. Overall, the predicted trajectories aligned relatively well with the

976 expected relationships among the different cell types derived from the meristematic cells in the
977 wheat spike.

978

979 **3. CONCLUSIONS**

980 The use of the same germplasm, tissues, and developmental stages in our smFISH and scRNA-
981 seq experiments allowed us to integrate the two datasets, which had positive synergistic effects.
982 The precise spatial information provided by the smFISH experiment was useful for improving
983 the annotation of the scRNA-seq clusters. However, the number of genes that can be included in
984 smFISH experiments is limited. We show here that this problem can be mitigated by using the
985 scRNA-seq data to impute the expression of a large number of genes into spatially anchored cells
986 from a spatial transcriptomics experiment. These imputed expression profiles can be used to
987 characterize the genes differentially expressed in each scRNA-seq cluster, thereby improving
988 their annotation. This imputation method can be also used to prioritize genes for functional
989 validation.

990 The method performed well both in diploid barley [70] and in tetraploid wheat, for which it
991 correctly predicted the locations of epidermal genes even when no epidermal markers were
992 included in the selected smFISH genes. The method also enabled us to distinguish expression
993 profiles between closely related wheat homeologs, confirming that this method can be useful for
994 studying expression in polyploid species with similar genomes.

995 The co-expression analyses, combined with smFISH and imputed expression profiles helped us
996 identify wheat genes with expression domains similar to those of their orthologs from rice and
997 maize, two crops with more extensive functional information than wheat. Comparative smFISH
998 studies of inflorescence development across different grass species using common sets of genes
999 can strengthen information bridges among crops. The analysis of co-expressed genes using gene
1000 networks provided another layer of information on gene relationships in a cell-type dependent
1001 manner. These localized networks can be useful for understanding the function of specific
1002 expression domains such as the glume axillae and the adaxial boundary domain. Finally, these
1003 tools can be used to identify genes with interesting expression profiles or interactions and
1004 prioritize them for functional characterization.

1005 In addition to providing useful datasets and tools, this study offered interesting biological
1006 insights into the function of *LFY*. For example, the discovery of *LFY* expression in the stem
1007 intercalary meristem motivated us to study its role in stem development and lead to the discovery
1008 of a previously unknown role of *LFY* in stem development. The smFISH experiment also
1009 revealed consistent and previously unknown expression domains in the region proximal to the
1010 developing wheat spikelets, which likely correspond to the "signaling centers" proposed in maize
1011 to regulate SM development and determinacy [95].

1012 In summary, the integrated smFISH and scRNA-seq dataset provides a dynamic view of gene
1013 expression during the early stages of wheat spike development with improved spatiotemporal
1014 resolution. All data and tools generated in this study are publicly available to empower future
1015 studies of wheat spike development and support comparative genomic studies across grass
1016 species.

1017

1018 **4. MATERIALS AND METHODS**

1019 **4.1. Spatial Expression**.

1020 ***4.1.1. Sample preparation***

1021 Developing wheat spikes from the tetraploid wheat Kronos were collected at three
1022 developmental stages: early transition from the vegetative to the reproductive stage (W1.5), late
1023 double ridge stage before the development of glume primordia (W2.5), and floret primordia
1024 stage, when the central spikelets initiated the transition to the floral meristem (W3.5). The
1025 different tissue sections used for W1.5 and W2.5 correspond to different plants, but the W3.5
1026 sections A1-1 and A2-1 correspond to the same spike and the sections A1-2 and A2-2 to a single
1027 spike from a different plant. Although the two sections from the same spike contain different
1028 cells, they were averaged and treated as technical replicates in the statistical analyses. After
1029 tissue collection, samples were fixed in 4% paraformaldehyde (PFA), dehydrated, and embedded
1030 in paraffin. Longitudinal sections (10 µm-thick) from the central plane of the developing spikes
1031 were positioned within the capture areas of the slides. The slides were dried overnight at 37°C,
1032 followed by a 10-minute bake at 50°C to enhance section adhesion. The sections were then
1033 deparaffinized, permeabilized, and refixed. After complete dehydration, they were mounted

1034 using SlowFade-Gold Antifade reagent, covered with a thin glass coverslip, and shipped on dry
1035 ice to Resolve BioSciences (Monheim am Rhein, Germany) for analysis.

1036 Spatial transcriptomics analyses were performed using Molecular Cartography™ and a 100-plex
1037 combinatorial single-molecule fluorescence in-situ hybridization. Upon arrival, tissue sections
1038 were washed twice in 1x PBS for two minutes, followed by one-minute washing in 50% and
1039 70% ethanol at room temperature. Ethanol was removed by aspiration, and DST1 buffer was
1040 added followed by tissue priming and hybridization. Tissues were primed for 30 minutes at
1041 37°C, followed by hybridization for 48 h using probes specific to the target genes. The following
1042 day, samples were washed to remove excess probes and fluorescently tagged in a two-step color
1043 development process. Regions of interest were imaged as described below, and fluorescent
1044 signals were removed during decolorization. Color development, imaging, and decolorization
1045 were repeated across multiple cycles to generate a unique combinatorial code for each target
1046 gene, derived from raw images as described below.

1047 ***4.1.2. Probe design***

1048 We used our published RNA-seq data [68] to select genes expressed at appropriate levels at the
1049 three selected spike developmental stages and prioritized those that showed dynamic changes
1050 across these stages. We also prioritized MADS-box genes and other transcription factors
1051 previously shown to be involved in inflorescence or floral development in other plant species.
1052 Finally, we also included marker genes to facilitate cluster annotations.

1053 For each of the selected tetraploid wheat genes, we selected the homeolog with higher expression
1054 based on published RNA-seq data for Kronos spike development [68]. We also provided Resolve
1055 Biosciences with the sequences of the lower-expressing homeolog to exclude them during probe
1056 specificity evaluations conducted against all the coding sequences of the wheat genome (Ref Seq
1057 v1.1, Data S1). Therefore, although probes were designed based on the sequence of the highest-
1058 expressing homeolog, they are not genome-specific and can detect both homeologs for each
1059 gene. Probes were compared with all the wheat expressed genes, and those that could amplify
1060 other genes were discarded. If no specific probes could be identified for a gene, it was excluded
1061 from the experiment.

1062 Probes for the 99 genes were designed using Resolve Biosciences' proprietary design algorithm
1063 and gene annotations from the Chinese Spring RefSeqv1.1 database. In brief, probe design was
1064 conducted at the gene level. For each targeted gene, all full-length protein-coding transcript
1065 sequences from the ENSEMBL database were utilized as design targets. To identify potential
1066 off-target sites, searches were confined to the coding regions. Each target sequence underwent a
1067 single scan for all k-mers, favoring regions with rare k-mers as seeds for full probe design. A
1068 probe candidate was generated by extending a seed sequence until reaching a certain target
1069 stability. Sequences known to cause problems experimentally were eliminated. Following these
1070 initial screens, each retained probe candidate was aligned with the background transcriptome,
1071 and probes with stable off-target hits were discarded. Specific probes were then scored based on
1072 the number of on-target matches (isoforms), giving preference to principal isoforms over others.
1073 A bonus was added if the binding site was within the protein-coding region. From the pool of
1074 accepted probes, the final set was selected from the highest-scoring probes.

1075 ***4.1.3. Imaging***

1076 Samples were imaged by Resolve Biosciences using a Zeiss Celldiscoverer 7 (CD7) with a 50×
1077 Plan Apochromat water immersion objective having an NA of 1.2 and a 0.5× magnification
1078 changer, resulting in a final magnification of 25×. Standard CD7 LED excitation light sources,
1079 filters, and dichroic mirrors were used in conjunction with customized emission filters optimized
1080 for detecting specific signals. The excitation time per image was fixed at 1,000 ms for each
1081 channel, 20 ms for DAPI, and 1 ms for Calcofluor White. A z-stack was captured at each region
1082 with a z-slice spacing according to the Nyquist-Shannon sampling theorem. A custom CD7
1083 CMOS camera (Zeiss Axiocam Mono 712, 3.45 µm pixel size) was employed. For each region, a
1084 z-stack was acquired for each of two fluorescent colors in each imaging round. A total of 8
1085 imaging rounds were performed for each position, resulting in 16 z-stacks per region. The
1086 entirely automated imaging process per round, which included water immersion generation and
1087 precise relocation of regions for imaging in all three dimensions, was implemented through a
1088 custom Python script utilizing the scripting API of the Zeiss ZEN software (Open application
1089 development). Imaging for the cell wall-specific stain, Calcofluor White, was conducted after the
1090 primary 8-round imaging.

1091 ***4.1.4. Spot segmentation and preprocessing***

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

In the initial image processing steps, all images underwent background fluorescence correction. The allowed number of maxima, representing the brightest points in the images, was determined based on the slice's area and an empirically optimized factor. Each image slice was processed independently. Maxima without neighboring counterparts in adjacent slices (referred to as z-groups) were discarded. An iterative filtering process adjusted thresholds for factors like absolute brightness, local background, and background of the periphery, all based on 3D image volume. Only maxima remaining in z-groups of at least 2 were retained. These maxima, with the highest absolute brightness within their z-groups, were documented as 3D-point cloud features.

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

Matched pixel profiles were assigned IDs, and neighboring pixels with the same ID formed groups. These pixel groups underwent further filtering, considering group size, adjacent pixel count, and dimensions with a size of two pixels. Local 3D-maxima in groups were identified as potential transcript locations. Maxima were filtered based on expectations from raw data images. The remaining maxima were evaluated by fitting to corresponding codes, and they were documented as transcripts. Specificity, assessed by the ratio of experiment-matching signals to non-experiment-matching signals, was used to estimate false positives. This comprehensive process enabled precise image analysis and transcript identification. Final image analysis was performed in ImageJ using the Polylux tool plugin from Resolve BioSciences to examine specific Molecular Cartography^{TM} signals. Images, programs, and gene expression results are deposited in Zenodo [99].

4.1.5. Cell segmentation, clustering, and cell trajectories

1122 We identified cell boundaries from the DAPI-stained images using the cell segmentation
1123 program QuPath [20]. To account for different cell sizes and shapes, we defined different regions
1124 in each image, applying region-specific parameter sets. Across all regions we used the
1125 parameters pixel size = 0, background radius = 8 µm, intensity threshold = 5, and opening by
1126 reconstruction with the "include cell nucleus" parameter disabled. For regular, well-defined cells,
1127 "split by shape" was activated, and the cell expansion was set to 7 µm, while for irregular,
1128 elongated cells, "split by shape" was turned off, and the cell expansion was set to 5 µm. The cell
1129 walls, visualized with Calcofluor White in the same sections, served as controls. The six images
1130 were merged into a single composite image and converted into a format suitable for clustering
1131 using ImageJ.

1132 Hybridization signals per cell and clustering were calculated using the Resolve Biosciences
1133 "Recognize" software with the following parameters: Min cell = 5, Min Features = 1, Min
1134 transcripts = 1, Dimensions = 20, Resolution = 0.75, and Selected genes = 99. "Recognize"
1135 utilizes Seurat v4 [100] as the backend, with the following methods: NormalizeData
1136 (normalization.method= "LogNormalize", scale.factor= 1), FindVariableFeatures, ScaleData,
1137 RunPCA, FindNeighbors (dims= 1:20), FindClusters (resolution= 0.75), RunUMAP (dims=
1138 1:20, n.components= 3, and FindAllMarkers. The R script to replicate the analysis using Seurat
1139 v5 can be found in Zenodo [99]. The average log2 fold changes of marker genes were generated
1140 locally using Seurat v5 [69], which implements a pseudocount at the group level instead of the
1141 cell level as in Seurat v4.

1142 Cells from the three developmental stages and two replications were clustered together. Clusters
1143 are represented with the same color across stages and replicates. To facilitate visualization of the
1144 individual cell clusters, we also present them separately in Fig. S5. The Seurat object was
1145 deposited in Zenodo [99].

1146 To find subclusters within scRNA-seq vasculature, we used the Seurat function
1147 "*FindSubCluster*" with parameter '*resolution* = 0.01, *graph.name* = "RNA_snn", *algorithm* = 2.
1148 We then used the Seurat function "*FindMarkers*" to find differentially expressed genes between
1149 the subclusters including xylem and phloem markers.

1150 We reconstructed cellular trajectories using the computer program Monocle 3 [101]. The 21 cell
1151 clusters from the spatial analysis and the 23 cell clusters from the scRNA-seq were leveraged to

constrain two separate trajectory inferences. The initial cluster object was transferred into a Monocle SingleCellExperiment object using the as.cell_data_set function. Dimensionality reduction was carried out using the UMAP embeddings previously computed with RunUMAP(dims=1:30) into the Monocle object. Then function "cluster_cells" was applied using the default random seed and UMAP as the reduction method to preserve cluster consistency, followed by "learn_graph" to construct the principal graph, and "order_cells" to infer pseudotime trajectories, specifying meristematic cluster c8 as the root cluster in the spatial trajectory and sc14 in the scRNA-seq trajectory.

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

Kronos spikes at the glume primordia stage (W3.0) were sectioned into three parts (Fig. S24) and placed into 500 µl of DNA/RNA Shield solution in 1.5-ml Eppendorf tubes. Approximately 20 to 25 apices were used per replicate, and three replicates were collected for each section. Samples in DNA/RNA Shield were stored at 4℃ prior to RNA extraction.

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

Data have been deposited in bioproject number PRJNA1217025. The three replicates from each region were clustered together in a principal component analysis, indicating high reproducibility among biological replicates (Fig. S24). Of the 49,786 detected genes, 13,024 exhibited significant differences among regions in Tukey tests (Data S11). For each region we calculated the average TMM per region and performed Tukey tests. The genes showing significant differences and the corresponding Tukey tests are listed in Data S11. Scripts for RNA-seq mapping, principal component plot, and Tukey tests were deposited in Zenodo [99].

4.3. Single-cell RNA sequencing

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

1181 We used a cell wall digestion method optimized to preserve meristem cells in maize
1182 inflorescence [66]. Briefly, the inflorescence spike was finely dissected (Fig. 1A) under a
1183 microscope and digested in a MOPS buffer-based enzyme mix (10 mM MOPS, pH 7.5; 10 mM
1184 L-Arginine HCl, pH 7.5; 0.65 M mannitol; 1.5% Cellulase RS; 1.5% Cellulase R10; 1.0%
1185 Macerozyme; 1.0% hemicellulose; 0.1% BSA; 5 mM β-mercaptoethanol; 1 mM $CaCl_2$) for 2
1186 hours. Before adding $CaCl_2$, β-mercaptoethanol, and BSA, the solution was heated to 55°C for
1187 10 minutes.

1188 The digested cells and enzyme mix were passed through a 30 µm pluriStrainer (pluriSelect,
1189 Catalog No. 43-50030-50) for filtration. The filtered protoplasts were pelleted by centrifugation
1190 at 350g for 2 minutes at 4°C and washed with wash buffer (10 mM MOPS, pH 7.5; 10 mM L-
1191 Arginine HCl, pH 7.5; 0.65 M mannitol). The protoplasts were pelleted again by centrifugation
1192 at 350g for 2 minutes at 4°C. The clean protoplast pellets were gently resuspended in wash
1193 buffer, and viability was checked using trypan blue. High-viability protoplasts (≥ 70%) were
1194 immediately loaded into a 10x Genomics Chromium System using Next GEM Single Cell 3' Kit
1195 v3.1 chemistry. The scRNA-seq libraries were sequenced on the Element Bio AVITI platform
1196 using the 10x Genomics recommended sequencing format, with 122 cycles used for Read 2
1197 (Data S7).

1198 ***4.3.2. Single-cell RNA-seq analysis and clustering***

1199 Single-cell RNA-seq reads were mapped to the Kronos genome assembly v1.1. (annotation v1.0)
1200 [67] (https://doi.org/10.5281/zenodo.11106422) using STARsolo [102, 103]. Since scRNA-seq
1201 reads were generated from the 3' end of the transcripts and not all Kronos genes have good
1202 annotations of 3' UTR regions, we extended the 3′ end of all gene models by 500 bp.

1203 We used the Seurat V5 pipeline [69] to cluster the cells. Briefly, genes expressed in at least three
1204 cells were retained. All cells from both stages were filtered to include those with 1,500 to 25,000
1205 expressed genes and 8,000 to 300,000 unique molecular identifiers (UMIs).

1206 Reads from individual samples were log-normalized, and 2,000 highly variable genes were
1207 selected for downstream analysis. The integrated matrix was scaled using the 'ScaleData'
1208 function. Principal component analysis (PCA) was performed with the 'RunPCA' function.
1209 Further integrative analysis of all W2.5 and W3.5 stage samples was conducted using the

1210 'IntegrateLayers' function with the anchor-based RPCA integration method. The top 35 principal
1211 components (PCs) were used. A resolution of 1 was applied with the 'FindClusters' function to
1212 identify clusters. Cell cycle clusters were identified (Fig. S23), and the remaining clusters were
1213 re-normalized, scaled, integrated using the RPCA method, and re-clustered. The top 35 PCs were
1214 used to generate the final clusters, at a resolution of 1 (Fig. 5B-D). Scripts used for mapping
1215 reads and clustering analysis and the final Seurat object files were deposited in Zenodo [99].

1216 ***4.3.3. Correlations among cell clusters***

1217 Normalized expression data for the 99 genes in the smFISH study were used to calculate Pearson
1218 correlation coefficients between pairs of spatial clusters (Data S3). Normalized expression data
1219 for the 200 genes used in the spatial transcriptomics study in hexaploid wheat were kindly
1220 provided by Dr. Cristobal Uauy (John Innes Centre, UK) [21]. The hexaploid wheat data was
1221 provided separately for the four different stages included in their study, but we only used the
1222 three early stages (late double ridge, lemma primordia and terminal spikelet) that overlapped
1223 with the stages of spike development used in our scRNAseq study.

1224 The average normalized gene expression across all cells for each single-cell cluster were
1225 obtained with the "AverageExpression" function in Seurat v5 in R after joining the six
1226 replications (three for each stage) of the scRNA-seq data. For this analysis, we used the final
1227 integrated data after removing the initial cell-cycle clusters and filtered the average normalized
1228 data per cluster using two criteria. First, we selected genes with coefficients of variation higher
1229 than 0.5 to eliminate genes with very similar expression levels across all clusters. We also
1230 eliminated genes with a total sum across clusters <0.1 or >20 to minimize the effect of genes
1231 with very high or very low expression levels (and multiple 0s). Using these criteria, we selected
1232 14,903 informative genes that were used to calculate Pearson correlation coefficients between all
1233 possible pairs of single-cell clusters using their average normalized expression across all
1234 common genes (Data S8).

1235 To calculate the correlations between the single-cell clusters and the spatial transcriptomics
1236 clusters from hexaploid wheat, we first identified the overlapping genes for each comparison.
1237 The coding regions of wheat homeologs are on average 97.5% identical, so most of the probes
1238 used in the wheat spatial transcriptomics experiment are expected to hybridize with the different
1239 homeologs. To make the single-cell data comparable, we summed the normalized expression of

1240 the three homeologs in hexaploid wheat, but separate correlation matrices were generated for the
1241 late double ridge (Data S13), lemma primordia (Data S14) and the terminal spikelet stage (Data
1242 S15).

1243 Clustering of the correlation matrices was done by the function "heatmap" from R package [104]
1244 with default parameters (clustering_distance_rows = "euclidean", clustering_distance_cols =
1245 "euclidean", clustering_method = "complete"). Scripts used to extract cluster averages and to
1246 perform the correlation analyses are deposited in Zenodo [99].

1247 ***4.3.4.* Imputation of scRNA-seq gene expression into spatial transcriptomics cells**

1248 To integrate the smFISH and scRNA-seq data without losing spatial information, we imputed the
1249 expression levels of all genes from the five nearest scRNA-seq cells into each cell from the
1250 spatial transcriptomic experiment using a recently published approach [70]. Briefly, we used
1251 SEURAT v5 [100] to calculate the cosine similarity (1 - cosine distance, varying from 0 for most
1252 dissimilar to 1 for the most similar) among the smFISH and scRNA-seq cells based on the
1253 expression of the 91 genes present in both studies (normalized counts per million). We identified
1254 the five nearest scRNA-seq cell neighbors to each smFISH-labelled cell and calculated a
1255 weighted average expression for each gene, using the cosine similarity as weight. The imputed
1256 gene expression was then incorporated into the smFISH-labelled cells.
1257 We integrated the smFISH and scRNA-seq for each stage (W2.5 and W3.5) separately and then
1258 combined the imputed data. We imputed 70,878 genes for W2.5 and 72,475 for W3.5, resulting
1259 in 74,464 for the combined analysis. We also developed a website for visualization of the
1260 combined data (https://dubcovskylab.ucdavis.edu/imputed-genes). The scripts for data
1261 imputation, the final imputed data, and the source data of the website are included in the Zenodo
1262 depository (folder imputed_data) [99].

1263 Using a similar procedure as in the imputation of scRNA-seq gene expression into spatial
1264 transcriptomics cells, after the integration, we found the five nearest smFISH cells for each
1265 scRNA-seq cell in W2.5 and W3.5 separately with one additional filtering: the nearest cell
1266 should have cosine similarity > 0.3 to exclude very different cells and at the same time keep
1267 enough number of cells for following analysis. We then assigned the scRNA-seq cell the most
1268 frequent smFISH cluster number within the 5 neighbors to the scRNA-seq cell. If there were no
1269 common smFISH cluster numbers (all 5 cluster numbers are unique), then this scRNA-seq cell

1270 had no smFISH cluster projection. After the projection, we calculated the frequency matrix
1271 between scRNA-seq original clusters and the projected smFISH clusters. The count matrix was
1272 normalized by the total number of cells used for each smFISH cluster (the frequency counts /
1273 total number of cells used in that smFISH cluster) to normalize for differences of cluster sizes.

1274 ***4.3.5. Annotation of scRNA-seq cell clusters***

1275 To annotate the scRNA-seq clusters, we used markers known to be expressed in specific cell
1276 types and/or previously associated with known functions in other species (Data S1) and validated
1277 their spatial expression in Kronos developing spikes using the imputation method described
1278 above. We supplemented this information with the correlations calculated between the single-cell
1279 and spatial expression clusters identified in the published spatial transcriptomics study in
1280 hexaploid wheat [21] (Data S13-15), and by the imputed expression of multiple differentially
1281 expressed genes in each scRNA-seq cluster using the methods described in the previous section.

1282 Additional information to annotate the scRNA-seq clusters was extracted from the genes
1283 preferentially expressed in each cluster (Data S9). These genes were selected based on their low
1284 abundance in other clusters (pct.2<0.10) relative to their abundance in the studied cluster
1285 (pct.1>0.10), adjusted *P* values <0.001, and a greater than 2-fold difference in expression levels
1286 (log2Fold>1, Data S9). Clusters sc1 and sc22 had few enriched genes, so we relaxed the filtering
1287 parameters (average log2FC>1, pct.1 no filtering, pct.2<0.2, adjusted *P* value <0.01) (Data S9).
1288 We subsequently determined the pathways enriched in these preferentially expressed genes using
1289 KEGG analyses (Kyoto Encyclopedia of Genes and Genomes, Data S10) through the DAVID
1290 webtool (https://davidbioinformatics.nih.gov/tools.jsp) [105]. We converted the Kronos v1.0
1291 gene IDs into Chinese Spring annotation v1.1 gene IDs, uploaded the CS gene IDs as
1292 "ENSEMBL_GENE_ID" (step 2: Select Identifier) and as "Gene List" (step 3: List Type), used
1293 the "*Triticum aestivum*" as background species, and started the analysis using "Functional
1294 Annotation Tool". In addition, we examined the average log2Fold expression across clusters for
1295 a subset of 41 photosynthetic genes to identify clusters associated with photosynthetic functions,
1296 and for 45 ribosomal protein-encoding genes to identify cells with high protein production,
1297 typically associated with elevated cell division rates. These results were visualized using heat
1298 maps.

1299 To incorporate additional spatial information into our cluster annotation, we used the RNA-seq
1300 data generated from the base, middle and top third of the spike described in section 4.2.
1301 (henceforth described as spike base, middle and tip). To determine if a cell cluster was enriched
1302 in any of these three spike regions, we identified genes with significant differences in Tukey tests
1303 among the preferentially expressed genes for each cluster and quantified the number of genes
1304 expressed at higher levels in each region. To determine if the distribution across the three regions
1305 was significantly different from a random uniform distribution, we performed chi-square tests for
1306 each cluster (Data S16).

1307 ***4.3.6.* Detection of clusters affected by the protoplasting process**

1308 The protoplasting process may trigger a stress response in some cells, potentially influencing the
1309 clustering process. We employed two independent approaches to assess the potential effects of
1310 protoplasting on gene expression in the different clusters. First, we used a list of 713 maize genes
1311 previously identified as potentially affected by the protoplasting process [84], identified the
1312 closest Kronos homologs, and determined the percentage of preferentially expressed genes in
1313 each cluster overlapping with this list (totaling 215 unique genes). The second approach, which
1314 identified 235 unique genes, involved the identification of preferentially expressed genes in each
1315 single-cell cluster that were not present in two RNA-seq data from Kronos developing spikes
1316 (this study and [68]). A GO-term analysis of the unique genes identified in both analyses
1317 revealed enrichment in stress-related annotations, particularly among the genes not detected in
1318 the two RNA-seq datasets (Data S19). These results suggest that many of these genes may have
1319 been induced by the protoplasting process and are typically not expressed under normal
1320 conditions in the developing wheat spike. The percentages of genes identified by these two
1321 independent methods across clusters were highly correlated (R= 0.813, P< 0.0001) suggesting a
1322 similar underlying phenomenon. Both methods identified clusters sc3 and sc22 as having the
1323 highest risk of being affected by the protoplasting process (Data S16).

1324 ***4.3.7.* scRNA-seq co-expression analysis**

1325 Using the cell x gene normalized count matrix, we calculated Pearson correlation coefficients (R)
1326 between genes of interest with known spatial distribution (*AGL6, FZP, TCP24*, *ULT-A1*) and all
1327 other detected genes in the scRNA-seq across all cells (26,009 cells. We focused on co-expressed

genes with correlation coefficients higher than the average plus four standard deviations. The script used for co-expression analysis was deposited in Zenodo [99].

4.3.8. Network analysis

To better understand the interactions between *TCP24* and *FZP*, we performed a gene regulatory network analysis using the program GENIE3 [96]. The analysis included the 988 genes co-expressed with *FZP* (Data S20) and/or *TCP24* (Data S21) in the scRNA-seq dataset. We extracted the top-ranked 500 links to generate and visualize the network using Cytoscape 3.10.3 (https://cytoscape.org/). In Data S22, we summarized the first neighbors of *TCP-A24*, *TCP-B24*, *FZP-A*, and *FZP-B* and identified shared and unique transcription factors among them.

4.3.9. Generation of loss-of-function mutants of FZP and SPL14 using CRISPR-Cas9

To validate the specific functions of *FZP* and *SPL14,* we used CRISPR-Cas9 to edit both homeologs of each target gene in the tetraploid wheat variety Kronos. We designed two guide RNAs for *FZP* (FZP-gRNA-1 and FZP-gRNA-2, Table S1) and one for *SPL14* (SPL14-gRNA, Table S1). Guide RNAs for each gene were cloned into a binary vector (Addgene Plasmid #160393) containing the *Cas9* gene and a *GRF4-GIF1* chimera, which enhances wheat regeneration efficiency during *Agrobacterium*-mediated transformation [106]. Transgenic plants were generated at the UC Davis Plant Transformation Facility using methods described previously [106]. Independent transgenic plants were screened for mutations using next-generation sequencing (NGS) as previously described [107] using primers described in Table S1. We identified eight independent T_0 lines with mutations in both *FZP* homeologs and further confirmed the edits by Sanger sequencing using primers listed in Table S1. For *SPL14*, we obtained 10 individual T_0 transgenic plants. This line was backcrossed with wildtype Kronos to segregate out the *Cas9* transgene, and we identified one line in the T_4 generation with successful edits in both the A and B genomes and without the transgene. We developed CAPS markers to track the *spl14* mutations across generations using primers described in Table S1.

ACKNOWLEDGEMENTS

The authors thank Chaozhong Zhang and Qiujie Liu (UC Davis) for their help with the selection of the genes included in the smFISH experiment, Samik Bhattacharya from Resolve Biosciences

for his continuous help with the spatial transcriptomics project, Cristobal Uauy from the John Innes Centre UK for his advice with the analysis of the spatial transcriptomics data and early access to his spatial transcriptomics data, Rüdiger Simon and Björn Usadel for their help with the expression imputation method, and Ksenia Krasileva from UC Berkeley for early access to the Kronos genome. JD acknowledges funding from the Howard Hughes Medical institute and the United States Department of Agriculture (USDA) – National Institute of Food and Agriculture (NIFA) grants 2022-68013-36439 (WheatCAP) and 2022-67013-36209. X.X acknowledges USDA-NIFA HATCH project (CA-D-PLB-2850-H) and UC Davis new faculty start-up funding.

AUTHORS CONTRIBUTIONS

XX, HL, and JD conceived the project. Experimental work was performed by HL, XX, FP, KL, 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 https://doi.org/10.5281/zenodo.14872904 [99]. The ImageJ version with the Polylux tool plugin, the high-resolution images for the six sections used in the paper (Calcofluor White Figs. S3, S4 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 [99]. We also developed a web tool to visualize individual cell clusters and heatmaps of the 99 genes in the six sections https://dubcovskylab.ucdavis.edu/JD99-wheat-spike-smFISH

and a similar tool to visualize the imputed genes https://dubcovskylab.ucdavis.edu/imputed-genes. These tools can be also downloaded from the same Zenodo link and run on individual computers. 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 [99].

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 https://doi.org/10.5281/zenodo.14872904.

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 [99].

REFERENCES

1. FAOSTAT: **Food and Agriculture Organization (FAO) of the United Nations. http://www.fao.org/faostat/en/#data.** 2023.
2. Kellogg EA: **Genetic control of branching patterns in grass inflorescences.** *Plant Cell* 2022, **34:**2518-2533.
3. Li C, Lin H, Chen A, Lau M, Jernstedt J, Dubcovsky J: **Wheat *VRN1*, *FUL2* and *FUL3* play critical and redundant roles in spikelet development and spike determinacy.** *Development* 2019, **146:**dev175398.
4. Chen H, Zhang X, Xu SH, Song CX, Mao HL: ***TaSPL17*s act redundantly with *TaSPL14*s to control spike development and their elite haplotypes may improve wheat grain yield.** *Frontiers in Plant Science* 2023, **14:**1229827.
5. Chuck G, Whipple C, Jackson D, Hake S: **The maize SBP-box transcription factor encoded by *tasselsheath4* regulates bract development and the establishment of meristem boundaries.** *Development* 2010, **137:**1243-1250.
6. Chuck GS, Brown PJ, Meeley R, Hake S: **Maize SBP-box transcription factors *unbranched2* and *unbranched3* affect yield traits by regulating the rate of lateral primordia initiation.** *Proc Natl Acad Sci U S A* 2014, **111:**18775-18780.

1417 7. Li K, Debernardi JM, Li C, Lin H, Zhang C, Jernstedt J, von Korff M, Zhong J,
1418 Dubcovsky J: **Interactions between SQUAMOSA and SVP MADS-box proteins**
1419 **regulate meristem transitions during wheat spike development.** *The Plant Cell* 2021,
1420 **33:**3621-3644.
1421 8. Adamski NM, Simmonds J, Brinton JF, Backhaus AE, Chen Y, Smedley M, Hayta S,
1422 Florio T, Crane P, Scott P, et al: **Ectopic expression of *Triticum polonicum VRT-A2***
1423 **underlies elongated glumes and grains in hexaploid wheat in a dosage-dependent**
1424 **manner.** *The Plant Cell* 2021, **33:**2296-2319.
1425 9. Debernardi JM, Lin H, Chuck G, Faris JD, Dubcovsky J: **microRNA172 plays a crucial**
1426 **role in wheat spike morphogenesis and grain threshability.** *Development* 2017,
1427 **144:**1966-1975.
1428 10. Greenwood JR, Finnegan EJ, Watanabe N, Trevaskis B, Swain SM: **New alleles of the**
1429 **wheat domestication gene *Q* reveal multiple roles in growth and reproductive**
1430 **development.** *Development* 2017, **144:**1959-1965.
1431 11. Debernardi JM, Greenwood JR, Jean Finnegan E, Jernstedt J, Dubcovsky J: ***APETALA***
1432 ***2*-like genes *AP2L2* and *Q* specify lemma identity and axillary floral meristem**
1433 **development in wheat.** *Plant Journal* 2020, **101:**171-187.
1434 12. Reinheimer R, Kellogg EA: **Evolution of *AGL6-like* MADS box genes in grasses**
1435 **(Poaceae): ovule expression is ancient and palea expression is new.** *Plant Cell* 2009,
1436 **21:**2591-2605.
1437 13. Paraiso F, Lin H, Li C, Woods DP, Lan T, Tumelty C, Debernardi JM, Joe A, Dubcovsky
1438 J: **LEAFY and WAPO1 jointly regulate spikelet number per spike and floret**
1439 **development in wheat.** *Development* 2024, **151:**dev202803.
1440 14. Kuzay S, Lin H, Li C, Chen S, Woods DP, Zhang J, Lan T, von Korff M, Dubcovsky J:
1441 ***WAPO-A1* is the causal gene of the 7AL QTL for spikelet number per spike in**
1442 **wheat.** *PLoS Genetics* 2022, **18:**e1009747.
1443 15. Kuzay S, Xu Y, Zhang J, Katz A, Pearce S, Su Z, Fraser M, Anderson JA, Brown-
1444 Guedira G, DeWitt N, et al: **Identification of a candidate gene for a QTL for spikelet**
1445 **number per spike on wheat chromosome arm 7AL by high-resolution genetic**
1446 **mapping.** *Theor Appl Genet* 2019, **132:**2689–2705.

1447 16. Muqaddasi QH, Brassac J, Koppolu R, Plieske J, Ganal MW, Roder MS: ***TaAPO-A1*, an**
1448 **ortholog of rice *ABERRANT PANICLE ORGANIZATION 1*, is associated with total**
1449 **spikelet number per spike in elite European hexaploid winter wheat (*Triticum***
1450 ***aestivum* L.) varieties.** *Scientific Reports* 2019, **9:**13853.
1451 17. Voss-Fels KP, Keeble-Gagnere G, Hickey LT, Tibbits J, Nagornyy S, Hayden MJ, Pasam
1452 RK, Kant S, Friedt W, Snowdon R, et al: **High-resolution mapping of rachis nodes per**
1453 **rachis, a critical determinant of grain yield components in wheat.** *Theoretical and*
1454 *Applied Genetics* 2019, **132:**2707-2719.
1455 18. Zhang J, Burguener GF, Paraiso F, Dubcovsky J: **Natural alleles of LEAFY and**
1456 **WAPO1 interact to regulate spikelet number per spike in wheat.** *Theor Appl Genet*
1457 2024, **137:**257.
1458 19. Waddington SR, Cartwright PM, Wall PC: **A quantitative scale of spike initial and**
1459 **pistil development in barley and wheat.** *Annals of Botany* 1983, **51:**119-130.
1460 20. Bankhead P, Loughrey MB, Fernandez JA, Dombrowski Y, McArt DG, Dunne PD,
1461 McQuaid S, Gray RT, Murray LJ, Coleman HG, et al: **QuPath: Open source software**
1462 **for digital pathology image analysis.** *Sci Rep* 2017, **7:**16878.
1463 21. Long KA, Lister A, Jones MRW, Adamski NM, Ellis RE, Chedid C, Carpenter SJ, Liu X,
1464 Backhaus AE, Goldson A, et al: **Spatial transcriptomics reveals expression gradients**
1465 **in developing wheat inflorescences at cellular resolution.** *bioRxiv* 2024,
1466 **https://doi.org/10.1101/2024.12.19.629411**.
1467 22. Perico C, Zaidem M, Sedelnikova O, Bhattacharya S, Korfhage C, Langdale JA:
1468 **Multiplexed in situ hybridization reveals distinct lineage identities for major and**
1469 **minor vein initiation during maize leaf development.** *Proceedings of the National*
1470 *Academy of Sciences of the United States of America* 2024, **121:**e2402514121.
1471 23. Bonke M, Thitamadee S, Mähönen AP, Hauser MT, Helariutta Y: **APL regulates**
1472 **vascular tissue identity in *Arabidopsis*.** *Nature* 2003, **426:**181-186.
1473 24. Zhang J, Li C, Zhang W, Zhang X, Mo Y, Tranquilli GE, Vanzetti LS, Dubcovsky J:
1474 **Wheat plant height locus *RHT25* encodes a PLATZ transcription factor that**
1475 **interacts with DELLA (RHT1).** *Proceedings of the National Academy of Sciences of*
1476 *the United States of America* 2023, **120:**e2300203120.
1477 25. McKim SM: **How plants grow up.** *J Integr Plant Biol* 2019, **61:**257-277.

1478 26. Wang J, Wang R, Wang Y, Zhang L, Zhang L, Xu Y, Yao S: ***Short and Solid***
1479 ***Culm/RFL/APO2* for culm development in rice.** *Plant J* 2017, **91:**85-96.
1480 27. Ndreca B, Huttly A, Bibi S, Bayon C, Lund G, Ham J, Alarcon-Reverte R, Addy J,
1481 Tarkowska D, Pearce S, et al: **Stacked mutations in wheat homologues of rice *SEMI-***
1482 ***DWARF1* confer a novel semi-dwarf phenotype.** *BMC Plant Biol* 2024, **24:**384.
1483 28. Pearce S, Vanzetti LS, Dubcovsky J: **Exogenous gibberellins induce wheat spike**
1484 **development under short days only in the presence of *VERNALIZATION1*.** *Plant*
1485 *Physiology* 2013, **163:**1433-1445.
1486 29. Sakamoto T, Morinaka Y, Ishiyama K, Kobayashi M, Itoh H, Kayano T, Iwahori S,
1487 Matsuoka M, Tanaka H: **Genetic manipulation of gibberellin metabolism in**
1488 **transgenic rice.** *Nat Biotechnol* 2003, **21:**909-913.
1489 30. Wang L, Ming L, Liao K, Xia C, Sun S, Chang Y, Wang H, Fu D, Xu C, Wang Z, et al:
1490 **Bract suppression regulated by the *miR156/529-SPLs-NL1-PLA1* module is required**
1491 **for the transition from vegetative to reproductive branching in rice.** *Mol Plant* 2021,
1492 **14:**1168-1184.
1493 31. Dong Z, Hu G, Chen Q, Shemyakina EA, Chau G, Whipple CJ, Fletcher JC, Chuck G: **A**
1494 **regulatory network controlling developmental boundaries and meristem fates**
1495 **contributed to maize domestication.** *Nat Genet* 2024, **56:**2528-2537.
1496 32. Lu Z, Yu H, Xiong G, Wang J, Jiao Y, Liu G, Jing Y, Meng X, Hu X, Qian Q, et al:
1497 **Genome-wide binding analysis of the transcription activator IDEAL PLANT**
1498 **ARCHITECTURE1 reveals a complex network regulating rice plant architecture.**
1499 *Plant Cell* 2013, **25:**3743-3759.
1500 33. Doebley J, Stec A, Hubbard L: **The evolution of apical dominance in maize.** *Nature*
1501 1997, **386:**485-488.
1502 34. Takeda T, Suwa Y, Suzuki M, Kitano H, Ueguchi-Tanaka M, Ashikari M, Matsuoka M,
1503 Ueguchi C: **The *OsTB1* gene negatively regulates lateral branching in rice.** *Plant J*
1504 2003, **33:**513-520.
1505 35. Ramsay L, Comadran J, Druka A, Marshall DF, Thomas WT, Macaulay M, MacKenzie
1506 K, Simpson C, Fuller J, Bonar N, et al: ***INTERMEDIUM-C*, a modifier of lateral**
1507 **spikelet fertility in barley, is an ortholog of the maize domestication gene**
1508 ***TEOSINTE BRANCHED 1*.** *Nat Genet* 2011, **43:**169-172.

1509 36. Dixon LE, Greenwood JR, Bencivenga S, Zhang P, Cockram J, Mellers G, Ramm K,
1510 Cavanagh C, Swain SM, Boden SA: ***TEOSINTE BRANCHED1* regulates inflorescence**
1511 **architecture and development in bread wheat (*Triticum aestivum*).** *Plant Cell* 2018,
1512 **30:**563-581.
1513 37. Volkman MM, Martin JM, Hogg AC, Wright L, Hale C, Carr PM, Giroux MJ: **Durum**
1514 **wheat *Teosinte Branched1* null mutations increase tillering.** *Crop Science* 2022,
1515 **62:**1522-1530.
1516 38. Huijser P, Klein J, Lonnig WE, Meijer H, Saedler H, Sommer H: **Bracteomania, an**
1517 **inflorescence anomaly, is caused by the loss of function of the MADS-box gene**
1518 ***squamosa* in *Antirrhinum majus*.** *Embo Journal* 1992, **11:**1239-1249.
1519 39. Klein J, Saedler H, Huijser P: **A new family of DNA binding proteins includes**
1520 **putative transcriptional regulators of the *Antirrhinum majus* floral meristem**
1521 **identity gene *SQUAMOSA*.** *Molecular and General Genetics* 1996, **250:**7-16.
1522 40. Reiser L, Sánchez-Baracaldo P, Hake S: **Knots in the family tree: evolutionary**
1523 **relationships and functions of homeobox genes.** *Plant Molecular Biology* 2000,
1524 **42:**151-166.
1525 41. Leiboff S, Li XR, Hu HC, Todt N, Yang JL, Li X, Yu XQ, Muehlbauer GJ, Timmermans
1526 MCP, Yu JM, et al: **Genetic control of morphometric diversity in the maize shoot**
1527 **apical meristem.** *Nature Communications* 2015, **6:**8974.
1528 42. O'Connor DL, Runions A, Sluis A, Bragg J, Vogel JP, Prusinkiewicz P, Hake S: **A**
1529 **division in PIN-mediated auxin patterning during organ initiation in grasses.** *Plos*
1530 *Computational Biology* 2014, **10:**e1003447.
1531 43. Zong J, Wang L, Zhu L, Bian L, Zhang B, Chen X, Huang G, Zhang X, Fan J, Cao L, et
1532 al: **A rice single cell transcriptomic atlas defines the developmental trajectories of**
1533 **rice floret and inflorescence meristems.** *New Phytol* 2022, **234:**494-512.
1534 44. Patterson EL, Richardson A, Bartlett M: **Pushing the boundaries of organ identity:**
1535 **Homology of the grass lemma.** *American Journal of Botany* 2023, **110:**e16161.
1536 45. Moraes TD, van Es SW, Hernández-Pinzón I, Kirschner GK, van der Wal F, da Silveira
1537 SR, Busscher-Lange J, Angenent GC, Moscou M, Immink RGH, van Esse GW: **The**
1538 **TCP transcription factor HvTB2 heterodimerizes with VRS5 and controls spike**
1539 **architecture in barley.** *Plant Reproduction* 2022, **35:**205-220.

1540 46. Bai F, Reinheimer R, Durantini D, Kellogg EA, Schmidt RJ: **TCP transcription factor,**
1541 **BRANCH ANGLE DEFECTIVE 1 (BAD1), is required for normal tassel branch**
1542 **angle formation in maize.** *Proceedings of the National Academy of Sciences of the*
1543 *United States of America* 2012, **109:**12225-12230.
1544 47. Poursarebani N, Trautewig C, Melzer M, Nussbaumer T, Lundqvist U, Rutten T,
1545 Schmutzer T, Brandt R, Himmelbach A, Altschmied L, et al: **COMPOSITUM 1**
1546 **contributes to the architectural simplification of barley inflorescence via meristem**
1547 **identity signals.** *Nat Commun* 2020, **11:**5138.
1548 48. Yuan Z, Gao S, Xue DW, Luo D, Li LT, Ding SY, Yao X, Wilson ZA, Qian Q, Zhang
1549 DB: ***RETARDED PALEA1*** **controls palea development and floral zygomorphy in**
1550 **rice.** *Plant Physiol* 2009, **149:**235-244.
1551 49. Komatsu M, Chujo A, Nagato Y, Shimamoto K, Kyozuka J: ***FRIZZY PANICLE*** **is**
1552 **required to prevent the formation of axillary meristems and to establish floral**
1553 **meristem identity in rice spikelets.** *Development* 2003, **130:**3841-3850.
1554 50. Chuck G, Muszynski M, Kellogg E, Hake S, Schmidt RJ: **The control of spikelet**
1555 **meristem identity by the *branched silkless1* gene in maize.** *Science* 2002, **298:**1238-
1556 1241.
1557 51. Dobrovolskaya O, Pont C, Sibout R, Martinek P, Badaeva E, Murat F, Chosson A,
1558 Watanabe N, Prat E, Gautier N, et al: ***FRIZZY PANICLE*** **drives supernumerary**
1559 **spikelets in bread wheat.** *Plant Physiology* 2015, **167:**189-199.
1560 52. Poursarebani N, Seidensticker T, Koppolu R, Trautewig C, Gawronski P, Bini F, Govind
1561 G, Rutten T, Sakuma S, Tagiri A, et al: **The genetic basis of composite spike form in**
1562 **barley and 'Miracle-Wheat'.** *Genetics* 2015, **201:**155-165.
1563 53. Bommert P, Whipple C: **Grass inflorescence architecture and meristem determinacy.**
1564 *Seminars in Cell & Developmental Biology* 2018, **79:**37-47.
1565 54. Yang Y, Fu D, Zhu C, He Y, Zhang H, Liu T, Li X, Wu C: **The RING-finger ubiquitin**
1566 **ligase HAF1 mediates Heading date 1 degradation during photoperiodic flowering**
1567 **in rice.** *Plant Cell* 2015, **27:**2455-2468.
1568 55. Zhu C, Peng Q, Fu D, Zhuang D, Yu Y, Duan M, Xie W, Cai Y, Ouyang Y, Lian X, Wu
1569 C: **The E3 ubiquitin ligase HAF1 modulates circadian accumulation of EARLY**

1570 **FLOWERING3 to control heading date in rice under long-day conditions.** *Plant Cell*
1571 2018, **30:**2352-2367.
1572 56. McSteen P, Malcomber S, Skirpan A, Lunde C, Wu XT, Kellogg E, Hake S: ***barren***
1573 ***inflorescence2*** **encodes a co-ortholog of the** ***PINOID*** **serine/threonine kinase and is**
1574 **required for organogenesis during inflorescence and vegetative development in**
1575 **maize.** *Plant Physiology* 2007, **144:**1000-1011.
1576 57. Carles CC, Choffnes-Inada D, Reville K, Lertpiriyapong K, Fletcher JC:
1577 ***ULTRAPETALA1*** **encodes a SAND domain putative transcriptional regulator that**
1578 **controls shoot and floral meristem activity in Arabidopsis.** *Development* 2005,
1579 **132:**897-911.
1580 58. Monfared MM, Carles CC, Rossignol P, Pires HR, Fletcher JC: **The** ***ULT1*** **and** ***ULT2***
1581 **trxG genes play overlapping roles in** ***Arabidopsis*** **development and gene regulation.**
1582 *Molecular Plant* 2013, **6:**1564-1579.
1583 59. Wang WF, Li G, Zhao J, Chu HW, Lin WH, Zhang DB, Wang ZY, Liang WQ: **DWARF**
1584 **TILLER1, a WUSCHEL-related homeobox transcription factor, is required for**
1585 **tiller growth in rice.** *Plos Genetics* 2014, **10:**e1004154.
1586 60. Bomblies K, Wang RL, Ambrose BA, Schmidt RJ, Meeley RB, Doebley J: **Duplicate**
1587 ***FLORICAULA/LEAFY*** **homologs** ***zfl1*** **and** ***zfl2*** **control inflorescence architecture and**
1588 **flower patterning in maize.** *Development* 2003, **130:**2385-2395.
1589 61. Ikeda-Kawakatsu K, Maekawa M, Izawa T, Itoh JI, Nagato Y: ***ABERRANT PANICLE***
1590 ***ORGANIZATION 2/RFL*****, the rice ortholog of Arabidopsis** ***LEAFY*****, suppresses the**
1591 **transition from inflorescence meristem to floral meristem through interaction with**
1592 ***APO1*****.** *Plant Journal* 2012, **69:**168-180.
1593 62. Selva C, Shirley NJ, Houston K, Whitford R, Baumann U, Li G, Tucker MR: ***HvLEAFY***
1594 **controls the early stages of floral organ specification and inhibits the formation of**
1595 **multiple ovaries in barley.** *Plant Journal* 2021, **108:**509-527.
1596 63. Glenn P, Woods DP, Zhang J, Gabay G, Odle N, Dubcovsky J: **Wheat bZIPC1**
1597 **interacts with FT2 and contributes to the regulation of spikelet number per spike.**
1598 *Theoretical and Applied Genetics* 2023, **136:**237.
1599 64. Zhu W, Yang L, Wu D, Meng Q, Deng X, Huang G, Zhang J, Chen X, Ferrandiz C,
1600 Liang W, et al: **Rice** ***SEPALLATA*** **genes** ***OsMADS5*** **and** ***OsMADS34*** **cooperate to**

1601 **limit inflorescence branching by repressing the *TERMINAL FLOWER1-like* gene**
1602 ***RCN4*.** *New Phytol* 2022, **233:**1682-1700.
1603 65. Zhang J, Xiong H, Burguener GF, Vasquez-Gross H, Liu Q, Debernardi JM, Akhunova
1604 A, Garland-Campbell K, Kianian SF, Brown-Guedira G, et al: **Sequencing 4.3 million**
1605 **mutations in wheat promoters to understand and modify gene expression.**
1606 *Proceedings of the National Academy of Sciences of the United States of America* 2023,
1607 **120:**e2306494120.
1608 66. Xu X, Passalacqua M, Rice B, Demesa-Arevalo E, Takebayashi MKY, Harris B,
1609 Sakakibara H, Gallavotti A, Gillis J, Jackson D: **Large-scale single-cell profiling of**
1610 **stem cells uncovers redundant regulators of shoot development and yield trait**
1611 **variation.** *bioRxiv* 2024, **doi.org/10.1101/2024.03.04.583414**.
1612 67. Seong K, Kumar R, Lunde C, Krasileva K: **Chromosome-level genome assembly of**
1613 ***Triticum turgidum* var 'Kronos' [Data set].** Zenodo.
1614 https://doi.org/10.5281/zenodo.11106422 2023.
1615 68. VanGessel C, Hamilton J, Tabbita F, Dubcovsky J, Pearce S: **Transcriptional**
1616 **signatures of wheat inflorescence development.** *Sci Rep* 2022, **12:**17224.
1617 69. Hao Y, Stuart T, Kowalski MH, Choudhary S, Hoffman P, Hartman A, Srivastava A,
1618 Molla G, Madad S, Fernandez-Granda C, Satija R: **Dictionary learning for integrative,**
1619 **multimodal and scalable single-cell analysis.** *Nat Biotechnol* 2024, **42:**293-304.
1620 70. Demesa-Arevalo E, Dörpholz H, Vardanega I, Maika JE, Pineda-Valentino I, Eggels S,
1621 Lautwein T, Köhrer K, Schnurbusch T, von Korff M, et al: **Imputation integrates**
1622 **single-cell and spatial gene expression data to resolve transcriptional networks in**
1623 **barley shoot meristem development.** *bioRxiv* 2025**:**2025.2005.2009.653223.
1624 71. Cho H, Dang TV, Hwang I: **Emergence of plant vascular system: roles of hormonal**
1625 **and non-hormonal regulatory networks.** *Curr Opin Plant Biol* 2017, **35:**91-97.
1626 72. Miyashima S, Sebastian J, Lee JY, Helariutta Y: **Stem cell function during plant**
1627 **vascular development.** *EMBO J* 2013, **32:**178-193.
1628 73. Skipper M: **Genes from the *APETALA3* and *PISTILLATA* lineages are expressed in**
1629 **developing vascular bundles of the tuberous rhizome, flowering stem and flower**
1630 **primordia of *Eranthis hyemalis*.** *Ann Bot* 2002, **89:**83-88.

1631 74. Houston K, Druka A, Bonar N, Macaulay M, Lundqvist U, Franckowiak J, Morgante M,
1632 Stein N, Waugh R: **Analysis of the barley bract suppression gene Trd1.** *Theor Appl*
1633 *Genet* 2012, **125:**33-45.
1634 75. Wang L, Yin H, Qian Q, Yang J, Huang C, Hu X, Luo D: ***NECK LEAF 1*, a GATA type**
1635 **transcription factor, modulates organogenesis by regulating the expression of**
1636 **multiple regulatory genes during reproductive development in rice.** *Cell Res* 2009,
1637 **19:**598-611.
1638 76. Whipple CJ, Hall DH, DeBlasio S, Taguchi-Shiobara F, Schmidt RJ, Jackson DP: **A**
1639 **conserved mechanism of bract suppression in the grass family.** *Plant Cell* 2010,
1640 **22:**565-578.
1641 77. Wang YG, Du F, Wang J, Wang K, Tian CH, Qi XQ, Lu F, Liu XG, Ye XG, Jiao YL:
1642 **Improving bread wheat yield through modulating an unselected AP2/ERF gene.**
1643 *Nature Plants* 2022, **8:**930-939.
1644 78. Oikawa T, Kyozuka J: **Two-step regulation of LAX PANICLE1 protein accumulation**
1645 **in axillary meristem formation in rice.** *Plant Cell* 2009, **21:**1095-1108.
1646 79. Galli M, Liu Q, Moss BL, Malcomber S, Li W, Gaines C, Federici S, Roshkovan J,
1647 Meeley R, Nemhauser JL, Gallavotti A: **Auxin signaling modules regulate maize**
1648 **inflorescence architecture.** *Proc Natl Acad Sci U S A* 2015, **112:**13372-13377.
1649 80. Jin Y, Luo Q, Tong H, Wang A, Cheng Z, Tang J, Li D, Zhao X, Li X, Wan J, et al: **An**
1650 **AT-hook gene is required for palea formation and floral organ number control in**
1651 **rice.** *Dev Biol* 2011, **359:**277-288.
1652 81. Gallavotti A, Zhao Q, Kyozuka J, Meeley RB, Ritter MK, Doebley JF, Pe ME, Schmidt
1653 RJ: **The role of *barren stalk1* in the architecture of maize.** *Nature* 2004, **432:**630-635.
1654 82. Panikashvili D, Shi JX, Bocobza S, Franke RB, Schreiber L, Aharoni A: **The**
1655 ***Arabidopsis* DSO/ABCG11 transporter affects cutin metabolism in reproductive**
1656 **organs and suberin in roots.** *Mol Plant* 2010, **3:**563-575.
1657 83. Douglas SJ, Li B, Kliebenstein DJ, Nambara E, Riggs CD: **A novel *Filamentous Flower***
1658 **mutant suppresses *brevipedicellus* developmental defects and modulates**
1659 **glucosinolate and auxin levels.** *PLoS One* 2017, **12:**e0177045.
1660 84. Xu XS, Crow M, Rice BR, Li F, Harris B, Liu L, Demesa-Arevalo E, Lu ZF, Wang LY,
1661 Fox N, et al: **Single-cell RNA sequencing of developing maize ears facilitates**

1662 **functional analysis and trait candidate gene discovery.** *Developmental Cell* 2021,
1663 **56:**557-568.
1664 85. Khanday I, Yadav SR, Vijayraghavan U: **Rice *LHS1/OsMADS1* controls floret**
1665 **meristem specification by coordinated regulation of transcription factors and**
1666 **hormone signaling pathways.** *Plant Physiol* 2013, **161:**1970-1983.
1667 86. Borghi L, Bureau M, Simon R: **Arabidopsis *JAGGED LATERAL ORGANS* is**
1668 **expressed in boundaries and coordinates *KNOX* and *PIN* activity.** *Plant Cell* 2007,
1669 **19:**1795-1808.
1670 87. Ohmori Y, Tanaka W, Kojima M, Sakakibara H, Hirano HY: **WUSCHEL-RELATED**
1671 **HOMEOBOX4 is involved in meristem maintenance and is negatively regulated by**
1672 **the CLE gene FCP1 in rice.** *Plant Cell* 2013, **25:**229-241.
1673 88. Wang J, Bao JL, Zhou BB, Li M, Li XZ, Jin J: **The *osa-miR164* target *OsCUC1***
1674 **functions redundantly with *OsCUC3* in controlling rice meristem/organ boundary**
1675 **specification.** *New Phytologist* 2021, **229:**1566-1581.
1676 89. Hibara K, Karim MR, Takada S, Taoka KI, Furutani M, Aida M, Tasaka M: **Arabidopsis**
1677 ***CUP-SHAPED COTYLEDON3* regulates postembryonic shoot meristem and organ**
1678 **boundary formation.** *Plant Cell* 2006, **18:**2946-2957.
1679 90. Gallavotti A, Malcomber S, Gaines C, Stanfield S, Whipple C, Kellogg E, Schmidt RJ:
1680 **BARREN STALK FASTIGIATE1 is an AT-hook protein required for the**
1681 **formation of maize ears.** *Plant Cell* 2011, **23:**1756-1771.
1682 91. Li Z, Liu D, Xia Y, Li Z, Jing D, Du J, Niu N, Ma S, Wang J, Song Y, et al:
1683 **Identification of the WUSCHEL-related homeobox (*WOX*) gene family, and**
1684 **interaction and functional analysis of TaWOX9 and TaWUS in wheat.** *Int J Mol Sci*
1685 2020, **21:**1581.
1686 92. Fang F, Ye S, Tang J, Bennett MJ, Liang W: **DWT1/DWL2 act together with**
1687 **OsPIP5K1 to regulate plant uniform growth in rice.** *New Phytol* 2020, **225:**1234-
1688 1246.
1689 93. Hendelman A, Zebell S, Rodriguez-Leal D, Dukler N, Robitaille G, Wu XL, Kostyun J,
1690 Tal L, Wang PP, Bartlett ME, et al: **Conserved pleiotropy of an ancient plant**
1691 **homeobox gene uncovered by *cis*-regulatory dissection.** *Cell* 2021, **184:**1724-1739.

1692 94. Bortiri E, Chuck G, Vollbrecht E, Rocheford T, Martienssen R, Hake S: ***ramosa2***
1693 **encodes a LATERAL ORGAN BOUNDARY domain protein that determines the**
1694 **fate of stem cells in branch meristems of maize.** *Plant Cell* 2006, **18:**574-585.
1695 95. Whipple CJ: **Grass inflorescence architecture and evolution: the origin of novel**
1696 **signaling centers.** *New Phytologist* 2017, **216:**367-372.
1697 96. Huynh-Thu VA, Irrthum A, Wehenkel L, Geurts P: **Inferring regulatory networks**
1698 **from expression data using tree-based methods.** *PLoS One* 2010, **5:**e12776.
1699 97. Tan F, Zhou C, Zhou Q, Zhou S, Yang W, Zhao Y, Li G, Zhou DX: **Analysis of**
1700 **chromatin regulators reveals specific features of rice DNA methylation pathways.**
1701 *Plant Physiol* 2016, **171:**2041-2054.
1702 98. Baylis T, Cierlik I, Sundberg E, Mattsson J: ***SHORT INTERNODES/STYLISH* genes,**
1703 **regulators of auxin biosynthesis, are involved in leaf vein development in**
1704 ***Arabidopsis thaliana*.** *New Phytol* 2013, **197:**737-750.
1705 99. Xu X, Lin H, Zhang J, Burguener G, Paraiso F, Li K, Tumelty C, Li C, Liu Y,
1706 Dubcovsky J: **Spatial and single-cell expression analyses reveal complex expression**
1707 **domains in early wheat spike development [Data set]** In *Zenodo*
1708 https://doi.org/10.5281/zenodo.14872904; 2025.
1709 100. Hao Y, Hao S, Andersen-Nissen E, Mauck WM, 3rd, Zheng S, Butler A, Lee MJ, Wilk
1710 AJ, Darby C, Zager M, et al: **Integrated analysis of multimodal single-cell data.** *Cell*
1711 2021, **184:**3573-3587 e3529.
1712 101. Cao J, Spielmann M, Qiu X, Huang X, Ibrahim DM, Hill AJ, Zhang F, Mundlos S,
1713 Christiansen L, Steemers FJ, et al: **The single-cell transcriptional landscape of**
1714 **mammalian organogenesis.** *Nature* 2019, **566:**496-502.
1715 102. Kaminow B, Yunusov D, Dobin A: **STARsolo: accurate, fast and versatile**
1716 **mapping/quantification of single-cell and single-nucleus RNA-seq data.** *bioRxiv*
1717 2021**:**2021.2005.2005.442755.
1718 103. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M,
1719 Gingeras TR: **STAR: ultrafast universal RNA-seq aligner.** *Bioinformatics* 2013,
1720 **29:**15-21.
1721 104. Kolde R: **pheatmap: Pretty Heatmaps. https://CRAN.R-**
1722 **project.org/package=pheatmap.** 2019.

1723 105. Huang DW, Sherman BT, Lempicki RA: **Systematic and integrative analysis of large**
1724 **gene lists using DAVID bioinformatics resources.** *Nature Protocols* 2009, **4:**44-57.
1725 106. Debernardi JM, Tricoli DM, Ercoli MF, Hayta S, Ronald P, Palatnik JF, Dubcovsky J: **A**
1726 **GRF-GIF chimeric protein improves the regeneration efficiency of transgenic**
1727 **plants.** *Nature Biotechnology* 2020, **38:**1274-1279.
1728 107. Zhang J: *Check CRISPR editing events in transgenic wheat with next-generation*
1729 *sequencing.* New York, NY: Springer US; 2022.
1730