

# Plant stem cell organization and differentiation at single-cell resolution

James W. Satterlee<sup>1</sup>, Josh Strable<sup>1</sup>, Michael J. Scanlon<sup>1\*</sup>

<sup>1</sup>School of Integrative Plant Science, Cornell University, Ithaca, NY, USA

\*Correspondence to: [mjs298@cornell.edu](mailto:mjs298@cornell.edu)

**Plants maintain populations of pluripotent stem cells in shoot apical meristems (SAMs), which continuously produce new aboveground organs. We used single-cell RNA sequencing to achieve an unbiased characterization of the transcriptional landscape of the maize shoot stem-cell niche and its differentiating cellular descendants. Stem cells housed in the SAM tip are engaged in genome integrity maintenance and exhibit a low rate of cell division, consistent with their contributions to germline and somatic cell fates. Surprisingly, we find no evidence for a canonical stem cell organizing center subtending these cells. In addition, we use trajectory inference to trace the gene expression changes that accompany cell differentiation. These data provide a valuable scaffold on which to better dissect the genetic control of plant shoot morphogenesis.**

Unlike animals where organogenesis is typically completed in juvenile stages, plants initiate new organs throughout the lifespan via the persistence of pluripotent stem-cell populations long after embryogenesis. These stem cells are housed within the shoot apical meristem (SAM), which gives rise to all of the above ground organs of the plant (*I*). Canonical descriptions of SAM organization in flowering plants include a stem-cell niche within the central zone at the SAM tip, subtended by the stem-cell organizing center, and a peripheral zone surrounding the SAM flank

24 that provides initial cells for organogenesis. The mechanisms that underlie stem-cell niche  
25 organization and maintenance, and how cells attain differentiated fates remain fundamental  
26 questions in plant development.

27 Class I *KNOTTED-LIKE HOMEODOMAIN* (*KNOX*) genes broadly promote indeterminate  
28 cell identity in vascular plants (2); *KNOX* downregulation is a marker of cell differentiation and  
29 comprises an initial step in lateral organ identity acquisition at the SAM periphery. In  
30 *Arabidopsis*, stem-cell homeostasis is achieved via a negative feedback loop wherein the activity  
31 of the stem-cell-organizing transcription factor WUSCHEL (*WUS*) is repressed by binding of the  
32 small secreted peptide CLAVATA3 (*CLV3*) to the CLAVATA1 (*CLV1*) receptor (3). The  
33 canonical *CLV-WUS* signaling pathway and other receptor-ligand signaling complexes  
34 regulating *WUS*-mediated control of shoot meristem size are identified across the flowering  
35 plants (3).

36 In order to better understand the spatial organization of the SAM and the process of cell  
37 differentiation during plant development, we took a single-cell transcriptomic approach to  
38 achieve an unbiased sampling of cell types from the maize (*Zea mays* spp. *mays*) SAM and  
39 seedling shoot apex unimpeded by prior histological assumptions. Improved protocols for the  
40 isolation of living plant stem cells enabled this first, to our knowledge, single-cell transcriptomic  
41 analysis of a vegetative shoot meristem in any plant. Two zones of cell identity are identified  
42 within the maize SAM: (1) a slowly-dividing stem-cell domain at the SAM tip expressing genes  
43 with functions in genome integrity, and (2) a subtending population of cells undergoing transit-  
44 amplifying divisions. Surprisingly, although the *CLV-WUS* stem-cell homeostatic pathway is  
45 well described in a diverse array of angiosperm SAMs and in the inflorescence and floral  
46 meristems of maize (3), we do not find evidence for a stem-cell organizing center expressing

47 *WUS* in the maize SAM (4). In addition, we use trajectory inference to identify dynamic gene  
48 expression patterns correlated with cell differentiation and ultimate cell fate in the seedling  
49 shoot. We find that during both early and later stages of seedling shoot ontogeny, a similar set of  
50 genes is deployed to mediate the transition from indeterminate to determinate cell fate during  
51 plant development.

52

## 53 **Results**

### 54 **Single-cell transcriptomic approach for the analysis of maize vegetative SAM cells**

55 Single-cell transcriptomic analyses of plant cells require the preparation of protoplasts,  
56 viable cells whose rigid, cellulosic cell walls are enzymatically removed. Previously, failures in  
57 recovery of viable protoplasts from SAM-enriched plant tissues has presented an obstacle to  
58 scRNAseq analyses of shoot meristems (5). To achieve a higher rate of viable cell recovery, we  
59 supplemented the protoplasting solutions with L-Arginine (L-Arg), which modestly enhanced  
60 cell viability (Fig. S1A). This finding was consistent with previous reports of enhanced cell  
61 viability of oat (*Avena sativa*) coleoptile protoplasts cultured in media supplemented with L-Arg  
62 (6). Increasing the pH of the media further enhanced protoplast viability (Fig. S1B), in keeping  
63 with prior studies showing that the *in vivo* pH of SAM tissue in the herbaceous plant  
64 *Chenopodium rubrum* is two orders of magnitude more alkaline than typical plant protoplasting  
65 solutions (7). Together, these modifications to our protoplasting protocol improved cell viability  
66 between 10-30 fold, depending on the tissue.

67 To capture cells from the microscopic seedling SAM, we manually harvested protoplasts  
68 from dissected apices comprising the SAM plus the two most recently-initiated leaf primordia  
69 (SAM + P2). After filtering, six biological replicates captured a total of 327 cells for single-cell

70 RNAseq (scRNA-seq) analyses (Fig. 1A, S2). We first used *k*-means clustering to classify  
71 transcriptionally similar cells, adding a number where appropriate to reflect subpopulations  
72 within a given cell type. Next, we performed dimensionality reduction using Uniform Manifold  
73 Approximation and Projection (UMAP), which plotted the seven resulting clusters corresponding  
74 to major cell types derived from the epidermis, primordia, and vasculature, along with  
75 indeterminate cell types from the SAM (Fig. 1B, S3). Owing to the abundance of cycling cells in  
76 the SAM + P2 tissues (49% in S/G2/M phase), we regressed out variation contributed by the cell  
77 cycle on cell clustering (Table S1). Intriguingly, instead of forming discrete, isolated cell  
78 clusters, the majority of SAM-enriched cells exhibited a continuum of intermediate identity  
79 states (Fig. 1B), suggesting that differentiation is highly dynamic and continuous throughout the  
80 maize shoot apex.

### 81 **Characterization of the maize stem-cell niche**

82 We first attempted to identify signatures of the stem cell population of the maize SAM.  
83 The tip of the maize SAM is thought to house the stem-cells (4, 8), which are essential for the  
84 generating the above-ground somatic tissue of the maize plant as well as cells that give rise to the  
85 germline. Our cell clustering analysis independently identified a transcriptionally distinct cell  
86 population in which *DYNAMIN* (*DYN*), a previously-identified marker of tip cells in the maize  
87 SAM, was upregulated (4). We therefore designated cells belonging to this cluster as the putative  
88 stem cells residing within the SAM tip and used differential expression (DE) analysis to identify  
89 89 genes preferentially expressed within this population (Fig. 1C-D, Table S2) (4). Among these  
90 were genes with confirmed or predicted roles in intercellular signaling, small RNA biogenesis,  
91 DNA maintenance, response to the plant hormones auxin and cytokinin, and transcriptional  
92 regulation (Fig. 1E). Closer inspection of the numerous genes (FDR for GO term enrichment =

93 0.05; Table S3) involved in RNA biogenesis suggested that these cells are engaged in RNA-  
94 dependent gene silencing activities. For example, the stem cell-enriched *SUPPRESSOR OF GENE*  
95 *SILENCING3-LIKE (SGS3-LIKE)*, *RNA POLYMERASE IV/V SUBUNIT2*, and *ARGONAUTE4a*  
96 (*AGO4a*) all encode members of the RNA-directed DNA methylation pathway that maintains  
97 heterochromatin at repetitive, retrotransposon-enriched, maize genomic regions (9). Indeed,  
98 maintenance of heterochromatin is likewise essential for the genomic stability and homeostasis  
99 of stem-cell populations in animals (10). However, unlike animals where germline cells are  
100 specified and sequestered during early embryonic stages, plants lack a segregated germ cell  
101 lineage during vegetative stages of development (11). Upregulation of genes involved in DNA  
102 repair related processes, such as a *PROTECTION OF TELOMERES1-LIKE (POT1-LIKE)* and a  
103 *DNA-DAMAGE BINDING2-LIKE* gene may reflect the advantage of maintaining high genomic  
104 fidelity among cells that have the potential for both somatic and germ cell fate (12). Collectively,  
105 these data suggest that cells in the maize SAM tip are engaged in genome protective functions  
106 consistent with their plant stem-cell identity.

### 107 **Divergence in SAM stem-cell regulation**

108 We next sought to analyze the cell-specific expression patterns of regulators of stem-cell  
109 maintenance within the SAM + P2 tissues. *FON2-LIKE CLE PROTEIN1 (FCP1)* was the only  
110 *CLV3*-like ligand-encoding transcript detected in meristem tip cells (Fig. 1F). RNA *in situ*  
111 hybridization identified weak expression of *FCP1* just below the SAM tip, as well as the  
112 originally reported expression in the SAM periphery and leaf primordia (13). The FCP1 peptide-  
113 ligand is perceived by the FEA3 receptor to repress stem-cell identity (13). *FEA3* transcripts  
114 show low and diffuse accumulation in the SAM periphery, and heightened expression in leaf  
115 primordia (Fig. 1G). Other maize transcripts encoding predicted leucine-rich repeat (LRR)

116 receptors exhibited similar accumulation patterns (Table S4), with higher expression in the SAM  
117 periphery and primordia but a lack of a strong SAM-specific expression profile as is seen for  
118 *CLV1* in Arabidopsis (14). This may reflect a role for LRR receptors in inhibiting stem-cell  
119 identity outside of the SAM tip domain, as described previously for the FCP1-FEA3 ligand-  
120 receptor system (13). Notably, mutations in the *CLV1* and *CLV2* homologs of maize cause  
121 enlarged inflorescence meristems, but do not affect vegetative SAM size (3).

122 In Arabidopsis, the stem-cell promoting transcription factor *WUS* is negatively regulated  
123 by *CLV1-CLV3* function, to control the size of the meristematic stem-cell pool (3). *WUS* is  
124 mobile and is expressed in the organizing center below the stem-cell domain from where it is  
125 trafficked to promote stem-cell fate in the SAM tip. A similar, stem-cell organizing *ZmWUS*  
126 function is described in the maize inflorescence meristem (3). However, no *ZmWUS*-expressing  
127 cells were identified in the SAM, although transcripts of several maize *WUS* homologs including  
128 *ZmWOX3a*, *ZmWOX9b* and *ZmWOX9c* were detected in the SAM (Fig. 1H). The single  
129 Arabidopsis *WOX9* homolog promotes cell proliferation in meristematic tissues upstream of  
130 *WUS* function, but belongs to a more ancient, functionally-divergent *WOX* clade that lacks the  
131 repressive *WUS*-box (15, 16). Overall, this suggests that the maize co-orthologs *WOX9b* and  
132 *WOX9c* are unlikely to be functionally homologous to *WUS*. Moreover, although *WOX3A* does  
133 encode a *WUS*-box and is detected in the SAM (Fig. 1H), its expression pattern is not cell-type  
134 specific. Thus, our data identify no candidate *WOX* gene expressed in the maize SAM that is  
135 likely to function as a stem-cell organizing center, homologous to *WUS* in Arabidopsis and  
136 *ZmWUS* in the maize inflorescence meristem. Together, these results suggest that the canonical  
137 *CLV1-CLV3-WUS* pathway has been bypassed in the maize SAM, with changes in the spatial  
138 expression patterns of corresponding maize paralogs as a defining feature.

## 139 **Characterization of a SAM core region**

140 To further investigate the organization of the maize SAM we examined gene expression  
141 patterns among cells derived from a recently-reported domain situated in the center, or ‘core’  
142 region, of the SAM, which is marked by the expression of *GRMZM2G049151*, a gene of  
143 unknown function (4). We identified cells within the core region by transcript accumulation of  
144 *GRMZM2G049151*, and used differential expression (DE) analysis to characterize their  
145 expression profiles (Fig. 2A-B, Table S5) (4). To reiterate our results above, no maize *WOX*  
146 genes are DE in the SAM core. In addition, SAM core cells show upregulated expression of the  
147 *PLASTOCHRON1 (PLA1)* gene. *PLA1* promotes cell division and growth in an auxin-dependent  
148 manner, and is also expressed within multiple maize organs and tissues (17). The auxin-  
149 promoted dormancy-associated gene *DRM1* (18) and the maize *HAIRY MERISTEM3 (HAM3)*  
150 homolog *GRAS33* were also upregulated in the SAM core, as confirmed by RNA *in situ*  
151 hybridizations (Fig. 2C-D, S4B-C). Arabidopsis *HAM* genes are expressed in the organizing  
152 center, SAM periphery, and in leaf primordia. *AtHAM* genes promote SAM maintenance through  
153 their physical interaction with *WUS*, and also activate the formation and maintenance of axillary  
154 meristems (AMs) that give rise to lateral branches (19, 20). Indeed, both *DRM1* and *GRAS33*  
155 show expanded expression in maize AMs (Fig. 2E-F). To determine if *GRAS33* activity in the  
156 core domain has a conserved role in maintenance of the maize SAM, we generated *gras32*  
157 *gras33* double mutant seedlings and analyzed SAM size (Fig. S4A). Compared to the wild-type  
158 siblings, *gras32/+ gras33* seedlings and *gras32 gras33* double mutants displayed shorter SAMs  
159 (Fig. 2G-I). Notably, only *gras32 gras33* SAMs possessed shorter AMs, likely owing to genetic  
160 redundancy between these factors in these branch meristems (Fig. 2J). These data suggest that  
161 maize *GRAS32* and *GRAS33* may, like their Arabidopsis homologs, have roles in regulating

162 SAM homeostasis from a stem cell-surrounding region that overlaps the maize SAM core  
163 domain. These results further suggest that unlike in Arabidopsis, the maize SAM tip is not  
164 subtended by a *WOX*-expressing, stem-cell organizing center. Rather, this SAM core region may  
165 be engaged in auxin response and functionally akin to the tip-subtending SAM regions  
166 expressing *HAM* genes in Arabidopsis, with maize *HAM* genes having a WUS-independent SAM  
167 regulatory function.

### 168 **Rates of cell division are kept low in the SAM stem-cell niche**

169 We next asked whether different maize SAM domains exhibit equivalent cell division  
170 rates. Estimations of cell-cycle stage generated during cell-cycle regression (see Methods)  
171 indicate that the fraction of SAM + P2 cells in G1 phase decreases as differentiation progresses,  
172 indicative of higher rates of cell division (Fig. 3A). For example, the SAM tip population  
173 contains a larger fraction of cells in G1-phase than the remainder of the meristem, leaf primordia,  
174 or vasculature cell populations, suggesting a lower cell division rate among the stem cells. In  
175 order to test this hypothesis, we performed RNA *in situ* hybridization on *HISTONE H3* and  
176 *CYCLIN1* transcripts that accumulate in cells at S-phase and G2/M-phase, respectively. We next  
177 subdivided medial SAM sections into five proximodistal bins of equal height along the  
178 proximodistal axis, and inferred cell division levels by image thresholding on *HISTONE H3* and  
179 *CYCLIN1* staining (Fig. 3B-E). Cells in bin 1 that comprises the tip-most region of the SAM  
180 consistently had the lowest number of dividing cells. When considered together with the  
181 previously described accumulation of transcripts promoting genomic and epigenomic stability in  
182 the SAM tip (Fig. 1E), a low cell division rate among the stem-cell population may explain how  
183 plants avoid unfavorable increases in genetic load over successive generations in the absence of a  
184 segregated germline early in ontogeny. A reduced rate of cell division at the SAM tip is likewise

185 consistent with the low number of cell divisions that are predicted to occur between formation of  
186 the maize zygote and the gametophytes contributing to the next generation (21). Cells in the  
187 remainder of the SAM, beyond the tip, showed higher rates of cell division, similar to transit  
188 amplifying cell divisions found in animal stem cell niches (22). These proliferative cell divisions  
189 in transit-amplifying cells generate the anlagen for determinate lateral organs, obviating the  
190 requirement for high levels of stem-cell divisions. Finally, we observed that in AMs, the highest  
191 concentration of dividing cells typically occurs closer to the AM tip as compared to the SAM  
192 (Fig. 3D). This may reflect changes in the size of the meristem core domain in the AM, in line  
193 with our observations of expanded core marker gene expression in AMs relative to the SAM.

#### 194 **Cell differentiation follows a continuum of transcriptional states**

195 Given that the individual transcriptomes of cells within the maize shoot apex fall  
196 overwhelmingly along a continuum of cell differentiation states (Fig. 1B), we aimed to  
197 determine the dynamic changes in gene expression that accompany this developmental  
198 progression. We applied a principal graph algorithm to identify a branching path among the  
199 embedded cellular coordinates in the UMAP projection, which we used to infer the  
200 differentiation trajectory of cells in the SAM + P2 dataset. Each cell was then assigned a  
201 pseudotime value based on its distance along the resulting path, relative to a specified  
202 pseudotime start position within the SAM tip cell population (Fig. 4A). As expected, we found  
203 that pseudotime progression is associated with the transition of cells from indeterminate to  
204 determinate cell fates; *KN1* and other markers of indeterminate meristematic identity are highly  
205 expressed early in pseudotime, whereas genes correlated with lateral organ development such as  
206 *GA2OX1* and the YABBY-family transcription factor-encoding gene *DROOPING LEAF1*  
207 (*DRL1*) (23) show increased expression levels as pseudotime progresses (Fig. 4B). To survey the

208 transcriptional changes associated with cell differentiation over time, we performed differential  
209 gene expression analysis to identify transcript accumulation patterns that significantly correlate  
210 with pseudotime progression. In total, over 2000 genes exhibited significant changes in  
211 expression over pseudotime (Table S6). Hierarchical clustering grouped each transcript  
212 according to expression pattern and identified several patterns of transcript accumulation that  
213 correspond to particular stages of cell differentiation (Fig. 4C).

214 Early in pseudotime, genes enriched for stem-cell functions in RNA metabolism and  
215 chromatin organization give way to genes enriched for glutathione transferase and cation-binding  
216 activities as well as expression of *DRM1* and *GRAS33*, which are expressed among transit-  
217 amplifying cells. Next, cells progress through a putative boundary domain identity, marked by  
218 upregulation of an *ARABIDOPSIS THALIANA HOMEBOX GENE1-LIKE (ATH1-LIKE)* gene,  
219 *LIGULELESS3 (LG3)*, and *GA2OX1*. *ATH1* promotes organ boundary formation in *Arabidopsis*  
220 and antagonizes activity of the growth phytohormone gibberellic acid, while *GA2OX1*  
221 catabolizes gibberellic acid (24, 25). In addition, *LG3* is expressed in specific boundaries during  
222 leaf development (26). After progressing to this putative shoot boundary domain, the cellular  
223 transcriptomes of SAM + P2 cells resolve into either epidermal, or ground and vascular cell  
224 identities. The lack of a transcriptionally distinct lineage of undifferentiated epidermal cells  
225 (protoderm) early in the trajectory is notable, given that the cell lineage of the outer protodermal  
226 layer is separate from that of underlying cells, even within the SAM tip (1). Together, this  
227 suggests that despite their cell lineage differences, the distinctive transcriptional profiles among  
228 cell types in the SAM tip become detectable only after exiting the stem cell niche.

229 As expected, we found that epidermal cell differentiation correlates with upregulation of  
230 the OUTER CELL LAYER homeodomain leucine zipper IV transcription factor-encoding genes

231 that promote epidermal cell identity (27). On the other hand, cells fated to become leaf primordia  
232 and/or vasculature exhibit upregulation of auxin response genes and transcripts associated with  
233 cell wall, chloroplast, and organ development. In addition, primordia and vascular cells are  
234 significantly enriched for transcripts related to translation, suggesting a large burst in protein  
235 synthesis accompanies leaf initiation and expansion. Selected genes significantly associated with  
236 pseudotime progression were examined by RNA *in situ* hybridization to validate their expression  
237 patterns along the developmental trajectory (Fig. 4D-H). Two genes expressed at the start of the  
238 pseudotime trajectory, *ZINC FINGER DOMAIN-LIKE (ZNF-LIKE)* and *RS21-C6*, showed  
239 expression in the SAM tip (Fig. 4D,E). Meanwhile, *FLOWERING PROMOTING FACTOR1-*  
240 *LIKE (FPF1-LIKE)* and *WALLS ARE THINI-LIKE (WAT1-LIKE)* transcripts are upregulated  
241 later in pseudotime and accumulate in the differentiating cells of leaf primordia (Fig. 4F,G).  
242 Lastly, *HYBRID PROLINE-RICH PROTEIN1 (HYPI)* transcripts are upregulated in the  
243 pseudotime boundary cluster, albeit detected at later stages of ontogeny (Fig. 4H). Together,  
244 these results support the notion that the transcriptomes of differentiating cells are highly dynamic  
245 across a continuum defined by pseudotime progression.

#### 246 **Redeployment of genes patterning determinate and indeterminate cells across ontogeny**

247 After leaf initiation at the SAM periphery, cells continue to proliferate in the leaf  
248 proximal region until beyond the P6 stage (i.e. the sixth leaf from the SAM), necessitating  
249 continued patterning of indeterminate and determinate zones at the junction of the leaf and stem  
250 across ontogeny (28). We hypothesized that the transcriptomic signatures of this developmental  
251 process would be similar at both early and late stages of leaf development, reflecting the iterative  
252 and modular patterning of the plant shoot system. To test this, we isolated cells derived from  
253 tissues comprising approximately 3 mm of the maize shoot apex, dissected to include the 6 most

254 recently-initiated leaf primordia plus the SAM (SAM + P6; Fig. S5A). In total, we captured the  
255 transcriptomes of over 10,000 protoplasts (in two biological replicates) using microfluidic  
256 droplet capture. To identify broad classes of cells, we performed dimensionality reduction and  
257 found clusters corresponding to epidermal, vascular, leaf primordial, indeterminate, and cell  
258 cycle states (Fig. S5B, S6, S7, Tables S7-9). Cells within our SAM + P6 dataset are  
259 overwhelmingly derived from later stages of shoot ontogeny (i.e. P4 to P6), owing to the  
260 markedly increased size of these older leaf primordia and associated stems. For example,  
261 although we identified indeterminate cells based on their expression of the transcription factor  
262 gene *KN1*, its duplicate paralogous genes *ROUGH SHEATH1 (RS1)* and *GNARLEY1 (GNI)*, as  
263 well as the KN1-direct gene target (*GIBERELLIC ACID 2-OXIDASE1*) *GA2OX1* (Fig. S7C)  
264 (25, 29) we did not identify cells with transcriptomic signatures of the seedling SAM stem cell  
265 niche in our SAM + P6 dataset (Fig. 1E).

266 We therefore selected a subset of cells spanning this indeterminate-determinate junction  
267 in the SAM + P6 dataset to represent cell differentiation at later stages of seedling leaf ontogeny  
268 (Fig. S9A). Among these, indeterminate cells expressing *KN1* and a *BLH14 (BELLI-LIKE*  
269 *HOMEODOMAIN14)* gene give way to cells occupying a boundary domain, characterized by  
270 expression of the maize homolog of *CUP-SHAPED COTYLEDONS2 (ZmCUC2)* (26, 30) (Fig.  
271 S9B). Intriguingly, boundary-region cells were also positive for expression of *WOX9c*, an  
272 observation supported by RNA *in situ* hybridization, which suggests a potential non-meristem  
273 related function for *WOX9c* (Fig. S9C). Determinate cells were characterized by upregulated  
274 expression of the *YABBY*-family transcription factor genes *YAB15* and *DRL1* (Fig. S9C). We  
275 again used trajectory inference to assign cells a pseudotime value reflective of their position in  
276 the indeterminate-determinate transition and utilized DE analysis to identify genes with

277 pseudotime-correlated expression patterns (Fig. 5A). Overall, we identified approximately 3000  
278 genes that met a stringent significance cutoff (adj.  $p < 1E-100$ ) and compared them to  
279 pseudotime-correlated genes in the SAM + P2 dataset. We found that approximately one third of  
280 the transcripts showed a significant correlation with pseudotime in both datasets  
281 (Hypergeometric test  $p$ -value  $< 1E-100$ ), suggesting a core module of genes controls the  
282 indeterminate-determinate cell transition across ontogeny (Fig. 5B). In addition, pseudotime  
283 expression curves for identical genes in both datasets have a lower mean Fréchet distance than  
284 non-identical genes, indicating similar expression behavior across ontogeny (Fig. 5C). Among  
285 the 1003 shared genes, GO functions related to translation, cell wall, organ polarity, auxin, and  
286 gibberellin-related processes were enriched, likely reflecting the roles of auxin and gibberellic  
287 acid hormones in promoting differentiation in opposition to *KN1* that imposes indeterminacy (2,  
288 25). Furthermore, *BARELY ANY MERISTEM2-LIKE (BAM2-LIKE)*, a *CLV1* paralog that  
289 regulates meristem and leaf development in Arabidopsis, shows upregulated expression among a  
290 subset of indeterminate cells as well as specific determinate cell populations, suggesting context-  
291 dependent signaling functions during maize shoot ontogeny (Fig. S9D-F) (31).

292 Strikingly, the pseudotime trajectory analysis revealed a cell subpopulation at the  
293 transition between indeterminate and determinate cell fates, consistent with the persistence of a  
294 transcriptionally distinct developmental boundary (Fig S9A). Strong signatures of this boundary  
295 are evident in a subcluster marked by upregulated expression of several *CUC2*-like genes, which  
296 are known to promote organ separation in plants (Figure 5C) (32). Other genes specifically  
297 expressed in the subcluster include a putative *SHORT HYPOCOTYL2-LIKE (SHY2-LIKE)* gene  
298 that is boundary specific, which may balance auxin and cytokinin signaling to maintain  
299 undifferentiated fates (33). Likewise, *WOX9c* and a *LATE MERISTEM IDENTITY1-LIKE*

300 (*LMII-LIKE*) gene were also highly specific for boundary cells, along with a *LATERAL ORGAN*  
301 *BOUNDARIES DOMAIN-LIKE (LBD-LIKE)* and *LIGHT SENSITIVE HYPOCOTYLS-LIKE*  
302 (*LSH-LIKE*) genes. Taken together, the patterning of indeterminate to determinate cell identities  
303 in the shoot employs a shared set of gene functions during early and late phases of leaf ontogeny.  
304 Furthermore, the persistent boundary regions identified here transcriptionally partition  
305 indeterminate from determinate cell identities across ontogeny.

## 306 **Discussion**

307 Here we present the first, to our knowledge, single-cell transcriptomic survey of a plant  
308 vegetative shoot apex including the stem cells housed within the SAM. This technique provides  
309 the key advantage of uncovering cell differentiation states and transitions in an unbiased fashion,  
310 without strict reliance on histological or genetic markers. Critically, we identify known and  
311 novel markers of the putative SAM stem-cell niche and show that it is characterized by increased  
312 expression of DNA methylation and DNA repair-related genes, as well as a low cell division  
313 rate. This observation supports the notion that only a subset of cells at the SAM tip have the  
314 specialized properties of stem cells, and may underlie the ability of plants to maintain high  
315 intergenerational genetic fidelity despite the absence of embryonic segregation of the germline as  
316 is found in animals (8). In addition, a low cell division rate and the upregulated expression of  
317 genes with genome protective functions highlights a shared convergently evolved solution to  
318 maintaining stem cells in plant and animals (22, 34). Questions are raised regarding the lineage  
319 contributions of these stem cells to the developing plant. For example, do cells comprising the  
320 SAM tip make significant contributions to lateral organs, or are they preferentially destined to  
321 contribute to germ fates as described in the “*meristem d’attente*” (“*meristem in waiting*”) model  
322 proposed by early histological studies (1)? In the latter view, plant lateral organs are mostly

323 derived from the transit-amplifying cell population that subtends the stem-cell domain, and the  
324 SAM tip makes minimal contributions to organogenesis. Emerging cell lineage tracing  
325 methodologies may shed light on these questions (35). Meanwhile, we find a lack of evidence for  
326 a WUS-expressing organizing center in the maize vegetative SAM, unlike what is modeled in  
327 other angiosperms, suggesting that genes with other, non-canonical, stem-cell organizing  
328 functions may be at play. Finally, by leveraging the continuum of cell states ranging from  
329 indeterminate to determinate cell identities, we reconstruct differentiation trajectories for cells of  
330 the seedling vegetative shoot apex. Many of the genes that are dynamically expressed along  
331 these pseudotime trajectories show similar accumulation patterns at early and later stages of leaf  
332 ontogeny, indicating the iterative redeployment of developmental programs.

### 333 **Methods**

334 *Plant materials and growth conditions.* Plants for single-cell RNA-Seq (scRNA-Seq), *in situ*  
335 hybridization, and phenotyping were grown in 72-well trays in a Percival A100 growth chamber  
336 under 16 hr days, a day temperature of 29.4 C, a night temperature of 23.9 C, and a relative  
337 humidity of 50%. Soil consisted of a 1:1 mixture of Turface MVP and LM111. The maize inbred  
338 B73 was used for scRNA-Seq and *in situ* hybridization analyses. The *gras32* and *gras33* alleles  
339 were obtained from the Maize Genetics Co-op Center (Urbana, IL, USA) in the W22 inbred  
340 background. Crosses were performed at a field site in Aurora, NY.

341  
342 *In situ hybridization.* RNA was isolated by Trizol extraction from liquid nitrogen-ground 2-  
343 week-old maize seedling shoot apices. Total RNA was DNase I treated and cDNA was prepared  
344 using polyT-primed SuperScript III reverse transcriptase (Thermo Fisher Scientific, MA, USA)  
345 according to the manufacturer's instructions. cDNA was then used as a template to amplify

346 probe sequences, which were TA-cloned into the pCR4-TOPO vector backbone encoding  
347 flanking T3 and T7 polymerase promoters (Thermo Fisher Scientific) (See Table S10 for PCR  
348 primers). Following the verification of probe sequence and orientation by Sanger sequencing,  
349 antisense RNA probes were generated using a DIG-labeling kit and the resulting probes  
350 hydrolyzed as previously described (Roche Diagnostics, IN, USA) (36). An LNA probe was used  
351 for the *WOX9c in situ* presented in Figure 1H and was ordered directly from Qiagen (Hilden,  
352 Germany). LNA probe hybridization was carried out using 10  $\mu$ M probe concentration and a 55  
353  $^{\circ}$ C hybridization temperature according to published methods (37).

354 Tissues for *in situ* hybridization were prepared and processed as previously described  
355 (36). Briefly, 2-week-old maize shoot apices were fixed overnight at 4 $^{\circ}$ C in FAA solution (3.7%  
356 formalin, 5% acetic acid and 50% ethanol in water) and dehydrated through an ethanol series,  
357 cleared through increasing concentrations of Histo-Clear II (National Diagnostics, GA, USA),  
358 and then embedded in paraplast. 10  $\mu$ m Sections were prepared using a Leica RM2235  
359 microtome (Leica Biosystems, Wetzlar, Germany) and adhered to Probe-on-Plus microscope  
360 slides (Thermo Fisher Scientific). Sectioned tissues were then deparaffinized and rehydrated  
361 through a reverse ethanol series prior to Proteinase K treatment, refixation, and acetic anhydride  
362 treatment. Dehydrated tissues were then hybridized with probe overnight and washed, RNase A  
363 treated, and incubated with AP-conjugated anti-DIG antibody (Roche Diagnostics). A  
364 colorimetric NBT/BCIP (Roche Diagnostics) reaction was then allowed to proceed until  
365 sufficient signal developed at which point the reaction was stopped in TE, slides were  
366 dehydrated, washed in Histo-Clear II, and mounted with Permount (Thermo Fisher Scientific).  
367 Images were obtained using an Axio Imager.Z10 (Carl Zeiss Microscopy, LLC, Thornwood,  
368 NY) microscope equipped with an AxioCam MRc5 camera.

369

370 *gras32 gras33 mutant analysis.* Genomic DNA was extracted from the leaf tissue of F2 seedlings  
371 segregating for exonic *gras32* and *gras33* Mu insertion alleles. DNA extraction was performed as  
372 previously described (38). Plants were genotyped by PCR (see Table S10 for primers) and  
373 Paraffin-embedded FAA-fixed shoot apex tissues (see *in situ* hybridization) were longitudinally  
374 sectioned to 10  $\mu\text{m}$  and adhered to Probe-on-Plus slides. Tissues were deparaffinized and  
375 rehydrated through an ethanol series and then equilibrated in 1% sodium borate (w/v). Tissues  
376 were then stained in a 0.5% solution of *o*-Toluidine (TBO) in 1% sodium borate for 5 min  
377 followed by an ethanol dehydration series, washing in Histo-Clear II, and mounting in Permount.  
378 Samples were then imaged (see *in situ* hybridization). SAM width and height were determined in  
379 medial sections using ImageJ.

380

381 *Cell cycle quantification.* Medial SAM sections probed by *in situ* hybridization for expression of  
382 the S-phase and G2/M-phase upregulated H3 and CYC1 genes, respectively, were imaged and  
383 imported into ImageJ. Images were converted to 16-bit format and processed using the  
384 thresholding tool such that stained areas could be differentiated from non-stained areas. For each  
385 SAM, five sections of equal height were measured and the percent ratio of above threshold  
386 (stained) area relative to the area of each bin was calculated.

387

388 *Generation and collection of protoplasts.* Two-week-old maize seedlings were hand dissected to  
389 either a  $\sim 3$  mm portion of the stem and plastochron 6 primordia (SAM + P6) or the SAM and  
390 plastochron 2 (SAM + P2). Dissected tissue was briefly macerated and placed immediately in  
391 protoplasting solution, which consisted of 0.65 M mannitol, 1.5% Cellulase R10, 1.5% Cellulase

392 RS, 1.0% Macerozyme, 1.0% hemicellulase (Sigma-Aldrich), 10 mM MOPS pH 7.5, 10 mM L-  
393 Arginine HCl pH 7.5, 1 mM CaCl<sub>2</sub>, 5 mM β-mercaptoethanol, and 0.1% BSA. Prior to the  
394 addition of CaCl<sub>2</sub>, β-mercaptoethanol, and BSA, the solution was heated to 55°C for 10 min to  
395 facilitate enzyme solubilization. All tissue collection was completed within 30-45 min. Tissue  
396 digestion was carried out with gentle shaking at 29°C for 2 hrs. After digestion, fluorescein  
397 diacetate (FDA) was added to the cell suspension at a concentration of 5 μg/mL and cells were  
398 allowed to incubate in darkness for 5 min. The cell suspension was then filtered using a 40 μm  
399 nylon filter and the cells were centrifuged at 250 x g for 3 min at 4°C. Cells were resuspended in  
400 washing buffer consisting of 0.65 M mannitol, 10 mM MOPS pH 7.5, and 10 mM L-Arginine  
401 pH 7.5 and washed 3 times using the same centrifugation conditions. Cell viability and  
402 concentration were assessed using a hemocytometer and a fluorescent Axio Imager.Z10  
403 microscope equipped with a 488 nm filter for FDA staining detection.

404 Cells isolated from SAM + P6 tissue were suspended at a concentration of ~10,000  
405 cells/mL and loaded onto the 10X Genomics Chromium Controller using v3 reagents following  
406 the manufacturer's instructions to target approximately 10,000 cells. Cells from the SAM + P2  
407 tissue were resuspended in 1 mL of wash buffer and a 100-200 uL aliquot was transferred to the  
408 well of a clear-bottom CoStar plate kept over ice. 200 uL of wash buffer was distributed to other  
409 wells on the plate. A Leica M205 FCA microscope equipped with a 488 nm fluorescent filter  
410 was then used to transfer individual viable (FDA+) cells. Each cell was carried in 0.1 μL  
411 volumes through three wash buffer wells. Following washing, cells were transferred to a 96-well  
412 LoBind (Eppendorf) plate containing reagents for reverse transcription (see library construction  
413 and sequencing), kept on dry ice. Cell collection was completed in <1 hr. Plates were sealed with  
414 adhesive film and transferred to a -80°C freezer prior to further processing.

415

416 *Cell viability assays.* Protoplasts were obtained as described (see generation and collection of  
417 protoplasts) with the exception of the pH and concentration of buffer components used. For the  
418 pH 5.7 condition, 10 mM MES buffer was used whereas for the pH 6.5, 7.0, and 7.5. conditions  
419 10 mM MOPS buffer was used. To quantify cell viability, the ratio of fluorescing (living)  
420 protoplasts to dead (non-fluorescing) and large debris (> 5  $\mu\text{m}$ ) were quantified in each of the  
421 four larger corners of a hemocytometer using a fluorescent Axio Imager.Z10 microscope  
422 equipped with a 488 nm filter.

423

424 *SAM + P2 single-cell RNA isolation and amplification.* Single-cell RNA-Seq library  
425 construction was performed using an adaptation of the Cel-Seq2 protocol (39). 96-well LoBind  
426 plates were prepared with each well containing 0.22  $\mu\text{L}$  25 ng/ $\mu\text{L}$  Cel-Seq2 RT primer (1s – 96s,  
427 Supplementary Table 10), 0.11  $\mu\text{L}$  10 mM dNTPs, and 0.77  $\mu\text{L}$  nuclease-free  $\text{H}_2\text{O}$  such that each  
428 well of the plate contained a unique cell barcoded RT primer. After cell collection and storage at  
429  $-80^\circ\text{C}$ , plates were thawed briefly on ice and centrifuged at 2000 x g for 2 min at  $4^\circ\text{C}$ . Next,  
430 plates were incubated at  $65^\circ\text{C}$  for 5 min and again centrifuged using the same settings. For  
431 reverse transcription, 0.54  $\mu\text{L}$  First Strand Buffer, 0.27  $\mu\text{L}$  0.1 M DTT, 0.135  $\mu\text{L}$  RNaseOUT  
432 (Thermo Fisher Scientific), 0.034  $\mu\text{L}$  SuperScript III, and 0.52  $\mu\text{L}$  nuclease-free  $\text{H}_2\text{O}$  were added  
433 to each well and plates were incubated for 1 hr at  $42^\circ\text{C}$  followed by RT inactivation for 10 min at  
434  $70^\circ\text{C}$ . cDNA was pooled horizontally into the eight wells at the end of each plate. Then, 2.5  $\mu\text{L}$   
435 10X Exonuclease I buffer and 2.1  $\mu\text{L}$  Exonuclease I were added to each well and incubated at  
436  $37^\circ\text{C}$  for 15 min followed by heat inactivation at  $80^\circ\text{C}$  for 15 min. cDNA was purified using  
437 Ampure Clean XP beads according to the manufacturer's instructions and resuspended in 7  $\mu\text{L}$

438 nuclease-free H<sub>2</sub>O. Second strand synthesis was performed by adding 2.31 μL Second Strand  
439 Buffer, 0.23 μL dNTPs, 0.08 μL *E. coli* DNA Ligase, 0.3 μL *E. coli* DNA Polymerase, and 0.08  
440 μL RNase H and incubated at 16°C for 2 hrs. Pooled dsDNA from each of the eight reactions  
441 was pooled and then purified using Ampure XP beads according to the manufacturer's  
442 instructions followed by resuspension in 6.4 μL nuclease-free H<sub>2</sub>O. *In vitro* transcription was  
443 performed by the addition of 1.6 μL each of the A, G, C, U dNTPs, 10X T7 Polymerase Buffer,  
444 and T7 Polymerase (Ambion, TX, USA) followed by incubation at 37°C for 13 hrs. Amplified  
445 RNA was then treated with ExoSAP-IT (Thermo Fisher Scientific) for 15 min at 37°C followed  
446 by RNA fragmentation by addition of 5.5 μL fragmentation buffer (200 mM Tris pH 8.0, 150  
447 mM MgCl<sub>2</sub>) and incubation for 3 min at 94°C. Fragmentation was stopped by transfer to ice and  
448 the immediate addition of 2.75 μL 0.5 M EDTA pH 8.0.

449  
450 *Library construction and sequencing.* For reverse transcription, 5 μL of the amplified RNA was  
451 added to 1 μL randomhexRT primer and 0.5 μL dNTPs, incubated at 65°C for 5 min, and chilled  
452 on ice. Next, 2 μL First Strand Buffer, 1 μL 0.1 M DTT, 0.5 μL RNaseOUT, and 0.5 μL  
453 SuperScript III were added and the samples incubated for 10 min at 25°C followed by a 1 hr  
454 incubation at 42 °C. Half of the completed RT reaction was subjected to PCR by addition of 5.5  
455 μL nuclease-free H<sub>2</sub>O, 12.5 μL Phusion High-Fidelity PCR Master Mix with HF Buffer (New  
456 England Biolabs, MA, USA), RNA PCR Primer1 (RP1), and 1 μL of RNA PCR Primer X  
457 (RPIX). PCR conditions were as follows: 30 s at 98°C, 11 cycles of: [10 s at 98°C, 30 s at 60°C,  
458 30 s at 72°C], and 10 min at 72°C. The pooled cDNA from each plate received a unique RPIX.  
459 Samples were then purified and size-selected via two rounds of AMPure XP bead treatment  
460 using a 1:1 ratio of beads-to-sample and the final library resuspended in 10 μL nuclease-free

461 H<sub>2</sub>O. 1 µL of the library was submitted for fragment analysis using a Bioanalyzer to confirm a  
462 target library size between 200-400 bp. An additional 1 uL was used for concentration  
463 measurement using a Qubit. If the library concentration was suboptimal, the second unused half  
464 of the RT reaction was amplified using up to 15 PCR cycles. Libraries were then sequenced  
465 using a single lane on an Illumina NextSeq 500 instrument using the small RNA chemistry.  
466 Paired-end sequencing was performed with 15 and 77 bp obtained for read 1 and read 2,  
467 respectively. The libraries generated from the biological replicates of SAM + P6 cells were also  
468 sequenced using a NextSeq 500 instrument, with each replicate allocated a single lane of  
469 sequencing.

470

471 *Single-cell RNA-Seq read processing and cell filtering.* SAM + P2 FASTQ files were processed  
472 using the default settings in the celseq2 pipeline (<https://github.com/yanailab/celseq2>), which  
473 includes read trimming, alignment, and UMI counting steps to generate a UMI count matrix (39).  
474 Reads were aligned to version 3 of the B73 reference genome. SAM + P6 reads were trimmed,  
475 aligned, and UMI count matrices generated using the CellRanger version 3.1.0 pipeline under the  
476 default settings. Reads were aligned to version 3 of the B73 reference genome. The UMI count  
477 matrices for individual biological replicates were merged prior to further analysis. For the SAM  
478 + P2 dataset, cells with fewer than 500 genes detected were removed while in the SAM + P6  
479 dataset, cells with fewer than 2500 genes detected were removed. In both datasets, cells with  
480 over 1% of transcripts encoded by the mitochondrial genome were removed.

481

482 *Dimensionality reduction, cell type classification, and differential expression analysis.* Cell type  
483 analysis and clustering were performed using Seurat v3.0 (40). The merged UMI count matrices

484 were converted to Seurat objects. Normalization and variance stabilization were performed using  
485 SCTransform and the 3000 genes with the highest expression variability were used for the  
486 calculation of principal components. Uniform Manifold Approximation and Projection (UMAP)  
487 was then used to embed cells in lower dimensional space for data visualization. For projection of  
488 SAM + P2 cells, UMAP was run using  $\text{dim} = 1:5$ ,  $\text{n.neighbors} = 15$ ,  $\text{min.dist} = 0.1$ , and  $\text{spread} =$   
489  $5$ . For the projection of SAM + P6 cells, UMAP was run using  $\text{dim} = 1:25$ ,  $\text{n.neighbors} = 25$ ,  
490  $\text{min.dist} = 0.01$ , and  $\text{spread} = 1$ . For the SAM+P6 subset cells, cells belonging to clusters 5 and 0  
491 were re-clustered in isolation using the same parameters as for the full dataset. Cells were  
492 assigned to clusters using k-means hierarchical clustering. All differential expression analyses  
493 used to compare gene expression on a per-cluster basis were performed using Wilcoxon ranked-  
494 sum tests. Gene ontology (GO) enrichment analysis was done using a Fisher's exact test  
495 implemented in AgriGO v2 (41). Cell cycle regression was used to reduce the effects of the cell  
496 cycle on cell clustering in the SAM + P2 dataset. Differentially expressed genes among cells  
497 belonging to cell clusters with S-phase and G2/M-phase marker gene expression were first  
498 identified (adjusted p-value < 0.05). Genes that were highly specific for these clusters were  
499 identified using a ratio of the number of cells expressing a given differentially expressed gene  
500 within the cluster to those in all clusters. Those with a ratio greater than 2 were deemed phase  
501 specific and their expression was used to calculate a numeric cell cycle score and a cell cycle  
502 stage (G1, S, G2/M). SCTransform was run again on the raw UMI count matrix with cell cycle  
503 score as a variable to regress, followed by PCA and UMAP dimensionality reduction.

504

505 *Trajectory inference and pseudotime analysis.* Trajectory inference and pseudotime analysis was  
506 performed using Monocle3 (<https://cole-trapnell-lab.github.io/monocle3>) (42). A principal graph

507 was generated using the `learn_graph` function and cells were assigned a pseudotime value using  
508 `order_cells` with a pseudotime start or “root” position manually selected. Genes that were  
509 differentially expressed along the inferred trajectories were identified using the `graph_test`  
510 function, which applies a Moran’s  $I$  test to detect spatial autocorrelation. For the SAM+P2  
511 dataset, individual Moran’s  $I$  tests were performed on cell subsets to better identify genes with  
512 branch-specific expression patterns. This involved two tests on a common population of cells  
513 derived from the Tip, Meristem 1, and Meristem 2 clusters merged with cells from the Primordia  
514 and Vasculature clusters or the Epidermis 1 and 2 clusters. The significantly differentially  
515 expressed genes along both trajectories were then further analyzed. For the visualization and  
516 analysis of pseudotime-dependent gene expression patterns, cubic smoothing splines were fit to  
517 each gene using the R `smooth.spline` function with a `spar` parameter of 1.1. To compare the  
518 expression behavior of genes in the SAM+P2 and SAM+P6 datasets, smoothed expression  
519 profiles for each gene were averaged in 10 pseudotime bins and z-z-scaled values for each bin  
520 were calculated. The Fréchet distances between the curves of identical genes and all non-  
521 identical genes were calculated using the `SimilarityMeasures` R package (43).

522

523 *Phylogenetic analysis.* For the phylogenetic analysis of GRAS proteins, amino acid sequences of  
524 all Arabidopsis GRAS proteins and HAM-LIKE homologs in rice (*Oryza sativa*) and maize were  
525 downloaded from Phytozome. Amino acid sequences were then aligned using Clustal Omega.  
526 Maximum likelihood phylogenetic tree construction was performed using PhyML. The Jones-  
527 Taylor-Thornton (JTT) amino acid substitution model was selected based on its Akaike  
528 Information Criterion (AIC) calculated using Smart Model Selection (SMS) implemented in

529 PhyML (44). Branch support values were calculated using the aLRT SH-LIKE fast likelihood-  
530 based method (45).

531 **References:**

- 532 1. T. A. Steeves, I. M. Sussex, *Patterns in Plant Development* (Prentice-Hall, Englewood  
533 Cliffs, New Jersey, 1972).
- 534 2. A. Hay, M. Tsiantis, *KNOX* genes: versatile regulators of plant development and  
535 diversity. *Development*. 137, 3153–3165 (2010).
- 536 3. M. Somssich, B. I. Je, R. Simon, D. Jackson, CLAVATA-WUSCHEL signaling in the  
537 shoot meristem. *Development*. 143, 3238–3248 (2016).
- 538 4. S. Knauer, M. Javelle, L. Li, X. Li, X. Ma, K. Wimalanathan, S. Kumari, R. Johnston, S.  
539 Leiboff, R. Meeley, P. S. Schnable, D. Ware, C. Lawrence-Dill, J. Yu, G. J. Muehlbauer,  
540 M. J. Scanlon, M. C. P. Timmermans, A high-resolution gene expression atlas links  
541 dedicated meristem genes to key architectural traits. *Genome Res*. 29, 1962–1973 (2019).
- 542 5. C. Luo, A. R. Fernie, J. Yan, Single-cell genomics and epigenomics; technologies and  
543 applications in plants. *Trend in Plant Science*. doi.org/10.1016/j.tplants.2020.04.016
- 544 6. A. Altman, R. Kaur-Sawhney, A. W. Galston, Stabilization of Oat Leaf Protoplasts  
545 through Polyamine-mediated Inhibition of Senescence. *Plant Physiol*. 60, 570–574  
546 (1977).
- 547 7. J. T. P. Albrechtova, S. Heilscher, L. Leske, P. Walczysko, E. Wagner, Calcium and pH  
548 patterning at the apical meristem are specifically altered by photoperiodic flower  
549 induction in *Chenopodium* spp. *Plant Cell Environ*. 26, 1985–1994 (2003).
- 550 8. T. Laux, The Stem Cell Concept in Plants: A Matter of Debate. *Cell*, 281–283 (2003).
- 551 9. M. A. Matzke, R. A. Mosher, RNA-directed DNA methylation: an epigenetic pathway of  
552 increasing complexity. *Nat. Rev. Genet*. 15, 394–408 (2014).
- 553 10. R. Ren, A. Ocampo, G.-H. Liu, J. C. Izpisua Belmonte, Regulation of Stem Cell Aging  
554 by Metabolism and Epigenetics. *Cell Metab*. 26, 460–474 (2017).
- 555 11. R. Lanfear, Do plants have a segregated germline? *PLoS Biol*. 16, e2005439 (2018).
- 556 12. C. N. McDaniel, R. S. Poethig, Cell-lineage patterns in the shoot apical meristem of the  
557 germinating maize embryo. *Planta*. 175, 13–22 (1988).
- 558 13. B. I. Je, J. Gruel, Y. K. Lee, P. Bommert, E. D. Arevalo, A. L. Eveland, Q. Wu, A.  
559 Goldshmidt, R. Meeley, M. Bartlett, M. Komatsu, H. Sakai, H. Jönsson, D. Jackson,  
560 Signaling from maize organ primordia via FASCIATED EAR3 regulates stem cell  
561 proliferation and yield traits. *Nat. Genet*. 48, 785–791 (2016).
- 562 14. S. E. Clark, R. W. Williams, E. M. Meyerowitz, The *CLAVATA1* Gene Encodes a  
563 Putative Receptor Kinase That Controls Shoot and Floral Meristem Size in *Arabidopsis*.  
564 *Cell*. 89, 575–585 (1997).

- 565 15. X. Wu, T. Dabi, D. Weigel, Requirement of Homeobox Gene *STIMPY/WOX9* for  
566 Arabidopsis Meristem Growth and Maintenance. *Curr. Biol.* 15, 436–440 (2005).
- 567 16. E. van der Graaff, T. Laux, S. A. Rensing, The WUS homeobox-containing (WOX)  
568 protein family. *Genome Biol.* 10, 248 (2009).
- 569 17. X. Sun, J. Cahill, T. Van Hautegeem, K. Feys, C. Whipple, O. Novák, S. Delbare, C.  
570 Versteede, K. Demuynck, J. De Block, V. Storme, H. Claeys, M. Van Lijsebettens, G.  
571 Coussens, K. Ljung, A. De Vlieghe, M. Muszynski, D. Inzé, H. Nelissen, Altered  
572 expression of maize *PLASTOCHRON1* enhances biomass and seed yield by extending  
573 cell division duration. *Nat. Commun.* 8, 14752 (2017).
- 574 18. G. M. Rae, K. David, M. Wood, The Dormancy Marker *DRM1/ARP* Associated with  
575 Dormancy but a Broader Role *In Planta*. *Dev. Biol.* 2013, 1–12 (2013).
- 576 19. E. M. Engstrom, C. M. Andersen, J. Gumulak-Smith, J. Hu, E. Orlova, R. Sozzani, J. L.  
577 Bowman, Arabidopsis Homologs of the *Petunia HAIRY MERISTEM* Gene Are Required  
578 for Maintenance of Shoot and Root Indeterminacy. *Plant Physiol.* 155, 735–750 (2011).
- 579 20. Y. Zhou, A. Yan, H. Han, T. Li, Y. Geng, X. Liu, E.M. Meyerowitz, HAIRY  
580 MERISTEM with WUSCHEL confines CLAVATA3 expression to the outer apical  
581 meristem layers. *Science* 361, 502–506 (2018).
- 582 21. S. P. Otto, V. Walbot, DNA methylation in eukaryotes: kinetics of demethylation and de  
583 novo methylation during the life cycle. *Genetics.* 124, 429–437 (1990).
- 584 22. R. Heidstra, S. Sabatini, Plant and animal stem cells: similar yet different. *Nat. Rev. Mol.*  
585 *Cell. Biol.* 15, 301–312 (2014).
- 586 23. J. Strable, J. G. Wallace, E. Unger-Wallace, S. Briggs, P. J. Bradbury, E. S. Buckler, E.  
587 Vollbrecht, Maize YABBY Genes drooping leaf1 and drooping leaf2 Regulate Plant  
588 Architecture. *Plant Cell.* 29, 1622–1641 (2017).
- 589 24. C. Gómez-Mena, R. Sablowski, *ARABIDOPSIS THALIANA HOMEBOX GENE1*  
590 Establishes the Basal Boundaries of Shoot Organs and Controls Stem Growth. *Plant Cell.*  
591 20, 2059–2072 (2008).
- 592 25. N. Bolduc, S. Hake, The Maize Transcription Factor KNOTTED1 Directly Regulates the  
593 Gibberellin Catabolism Gene *ga2ox1*. *Plant Cell.* 21, 1647–1658 (2009).
- 594 26. R. Johnston, M. Wang, Q. Sun, A. W. Sylvester, S. Hake, M. J. Scanlon, Transcriptomic  
595 Analyses Indicate That Maize Ligule Development Recapitulates Gene Expression  
596 Patterns That Occur during Lateral Organ Initiation. *Plant Cell.* 26, 4718–4732 (2014).
- 597 27. M. Javelle, C. Klein-Cosson, V. Vernoud, V. Boltz, C. Maher, M. Timmermans, N.  
598 Depège-Fargeix, P. M. Rogowsky, Genome-wide characterization of the HD-ZIP IV  
599 transcription factor family in maize: preferential expression in the epidermis. *Plant*  
600 *Physiol.* 157, 790–803 (2011).
- 601 28. Poethig, R. S. (1984). Cellular parameters of leaf morphogenesis in maize and tobacco. In  
602 *Contemporary Problems of Plant Anatomy* (eds. R. A. White and W. C. Dickinson), pp.  
603 235–238. New York: Academic Press.

- 604 29. D. Jackson, B. Veit, S. Hake, Expression of maize *KNOTTED1* related homeobox genes  
605 in the shoot apical meristem predicts patterns of morphogenesis in the vegetative shoot.  
606 *Development*. 120, 405–413 (1994).
- 607 30. K. Tsuda, M. J. Abraham-Juarez, A. Maeno, Z. Dong, D. Aromdee, R. Meeley, T.  
608 Shiroishi, K. Nonomura, S. Hake, KNOTTED1 Cofactors, BLH12 and BLH14, Regulate  
609 Internode Patterning and Vein Anastomosis in Maize. *Plant Cell*. 29, 1105–1118 (2017).
- 610 31. B. J. DeYoung, K. L. Bickle, K. J. Schrage, P. Muskett, K. Patel, S. E. Clark, The  
611 CLAVATA1-related BAM1, BAM2 and BAM3 receptor kinase-like proteins are  
612 required for meristem function in Arabidopsis. *Plant J*. 45, 1–16 (2006).
- 613 32. A. Richardson, S. Hake, Drawing a Line: Grasses and Boundaries. *Plants*. 8, 4 (2018).
- 614 33. R. D. Ioio, K. Nakamura, L. Moubayidin, S. Perilli, M. Taniguchi, M. T. Morita, T.  
615 Aoyama, P. Costantino, S. Sabatini, A Genetic Framework for the Control of Cell  
616 Division and Differentiation in the Root Meristem. *Science*. 322, 1380–1384 (2008).
- 617 34. U. Weissbein, N. Benvenisty, U. Ben-David, Genome maintenance in pluripotent stem  
618 cells. *J. Cell Biol*. 204, 153–163 (2014).
- 619 35. A. McKenna, G. M. Findlay, J. A. Gagnon, M. S. Horwitz, A. F. Schier, J. Shendure,  
620 Whole-organism lineage tracing by combinatorial and cumulative genome editing.  
621 *Science*. 353, aaf7907 (2016).
- 622 36. M. Javelle, C. F. Marco, M. Timmermans, In situ hybridization for the precise  
623 localization of transcripts in plants. *J. Vis. Exp.*, e3328 (2011).
- 624 37. M. Javelle, M. C. P. Timmermans, In situ localization of small RNAs in plants by using  
625 LNA probes. *Nat Protoc*. 7, 533–541 (2012).
- 626 38. K. Edwards, C. Johnstone, C. Thompson, A simple and rapid method for the preparation  
627 of plant genomic DNA for PCR analysis. *Nucl. Acids Res*. 19, 1349–1349 (1991).
- 628 39. T. Hashimshony, N. Senderovich, G. Avital, A. Klochendler, Y. de Leeuw, L. Anavy, D.  
629 Gennert, S. Li, K. J. Livak, O. Rozenblatt-Rosen, Y. Dor, A. Regev, I. Yanai, CEL-Seq2:  
630 sensitive highly-multiplexed single-cell RNA-Seq. *Genome Biol*. 17, 77 (2016).
- 631 40. A. Butler, P. Hoffman, P. Smibert, E. Papalexi, R. Satija, Integrating single-cell  
632 transcriptomic data across different conditions, technologies, and species. *Nat*.  
633 *Biotechnol*. 36, 411–420 (2018).
- 634 41. T. Tian, Y. Liu, H. Yan, Q. You, X. Yi, Z. Du, W. Xu, Z. Su, agriGO v2.0: a GO analysis  
635 toolkit for the agricultural community, 2017 update. *Nucl. Acids Res*. 45, W122–W129  
636 (2017).
- 637 42. J. Cao, M. Spielmann, X. Qiu, X. Huang, D. M. Ibrahim, A. J. Hill, F. Zhang, S.  
638 Mundlos, L. Christiansen, F. J. Steemers, C. Trapnell, J. Shendure, The single-cell  
639 transcriptional landscape of mammalian organogenesis. *Nature*. 566, 496–502 (2019).
- 640 43. H. Alt, M. Godau, Computing the Frechet Distance Between Two Polygonal Curves. *Int*.  
641 *J. Comput. Geom. Appl*. 05, 75–91 (1995).
- 642 44. V. Lefort, J.-E. Longueville, O. Gascuel, SMS: Smart Model Selection in PhyML. *Mol*.  
643 *Biol. Evol*. 34, 2422–2424 (2017).

644 45. S. Guindon, J.-F. Dufayard, V. Lefort, M. Anisimova, W. Hordijk, O. Gascuel, New  
645 algorithms and methods to estimate maximum-likelihood phylogenies: assessing the  
646 performance of PhyML 3.0. *Syst. Biol.* 59, 307–321 (2010).

647

648 **Acknowledgments:** We thank the Cornell University Biological Resource Center for assistance  
649 in single-cell genomics and sequencing techniques as well as J. Cammarata and A. Roeder for  
650 critical feedback on the manuscript. This work was funded by National Science Foundation  
651 grants IOS-2016021 and IOS-1238142 (to M.J.S) and a Schmittau-Novak Small Grant (to  
652 J.W.S.).

653 **Author contributions:** J.W.S. and M.J.S. conceptualized the study, J.W.S analyzed resulting  
654 data, J.W.S. and J.S. performed experiments, J.W.S. and M.J.S wrote the manuscript with J.S.  
655 contributing to review and editing.

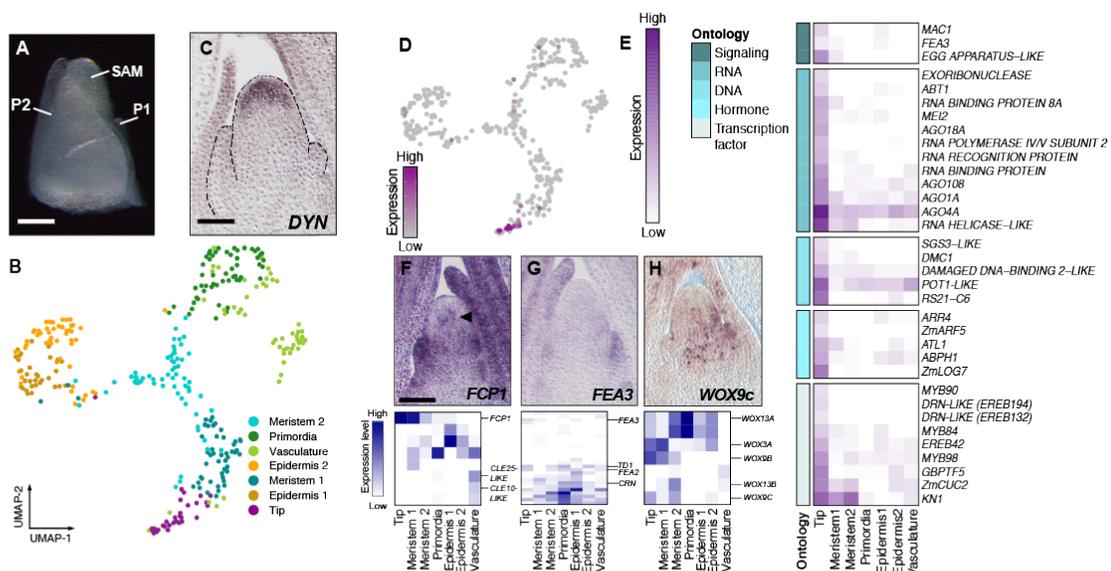
656 **Data Availability:** Sequencing data generated in this work are available at the NCBI Short  
657 Reads Archive (accession no. PRJNA637882).

658 **Additional Information / Supplementary Materials:**

659 Supplementary Figures S1-S9

660 Supplementary Datasets S1-S10

661



662

663 **Fig. 1. Transcriptomic signatures of stem-cell identity and maintenance in the maize SAM.**

664 (A) Cells were isolated from the SAM plus two most recently initiated leaf primordia (SAM +

665 P2). (B) Dimensionality reduction and cell classification for cells in the SAM + P2 dataset. (C)

666 RNA *in situ* hybridization with antisense probe to *DYNAMIN* in medial longitudinal section of

667 the SAM. (D) *DYN*, a marker for the SAM tip and putative stem cell domain, in the SAM + P2

668 dataset correlates with the putative tip cell population. (E) Heatmap of select differentially

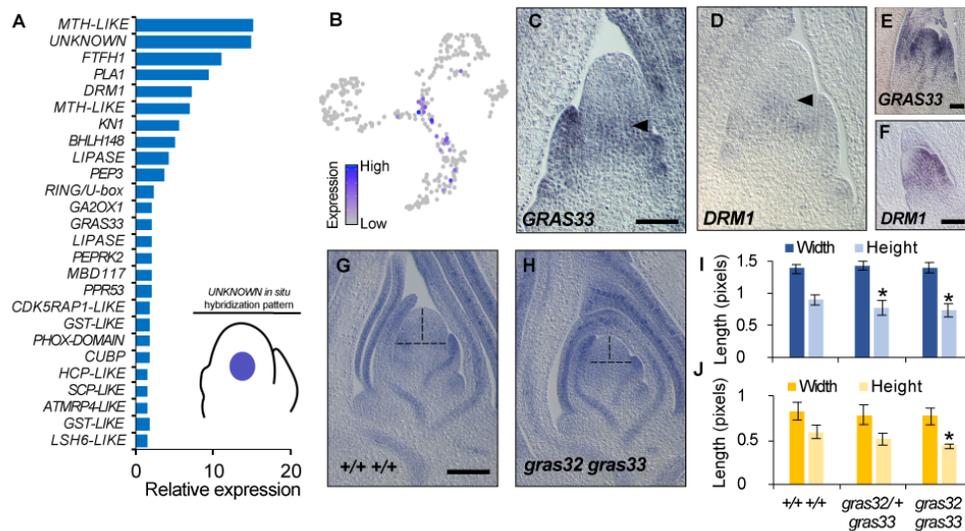
669 expressed genes in the tip domain grouped based on functional ontology. (F-H) RNA *in situ*

670 hybridization with antisense probes to *FCP1* (F), *FEA3* (G) and *WOXC9c* (H) (top) and DE

671 analysis along with paralog gene expression in the SAM + P2 dataset (bottom). Scale bars = 100

672  $\mu\text{m}$ .

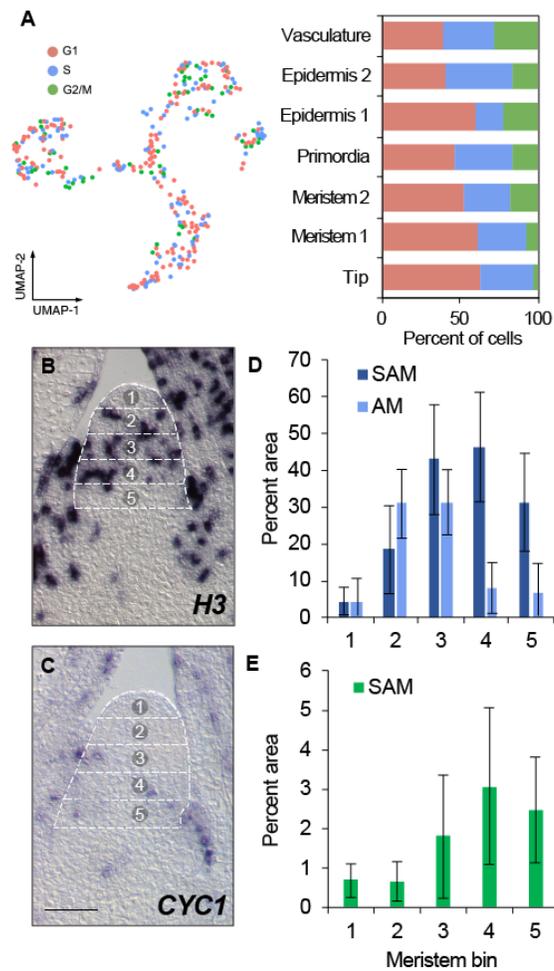
673



674

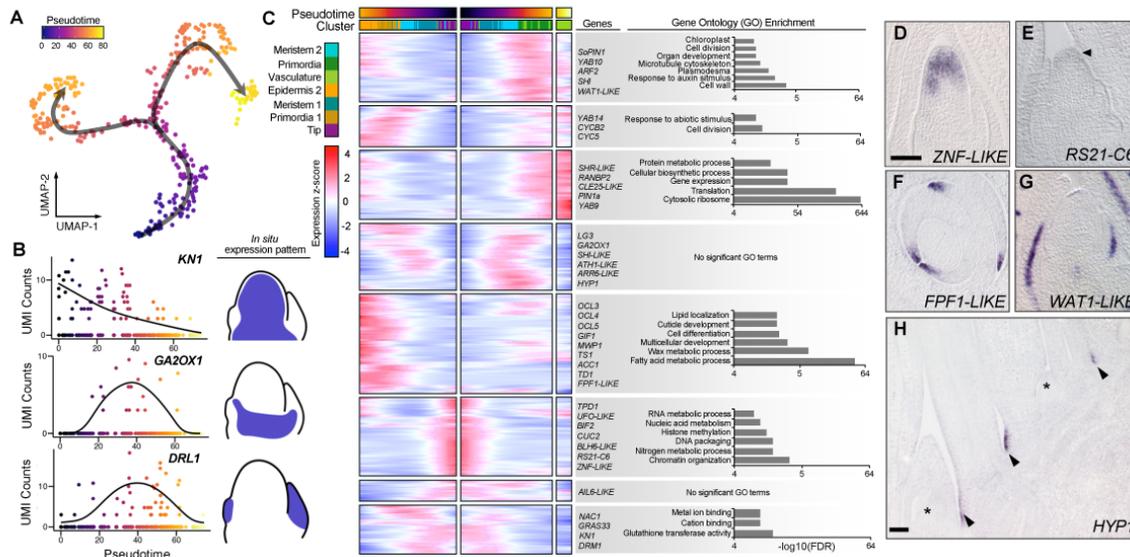
675 **Fig. 2. Characterization of the SAM core.** (A) Expression of select marker genes in cells positive for  
676 expression of the *UNKNOWN* core marker gene. The illustration reflects the published  
677 *UNKNOWN* gene expression pattern (4). (B) Cells expressing the *UNKNOWN* marker gene in the  
678 SAM + P2 dataset. (C-F) RNA *in situ* hybridization of antisense probes to *GRAS32* (C,E) and  
679 *DRM1* (D,F) shows transcript accumulation patterns in SAM (C,D) and AM (E,F) medial  
680 longitudinal sections. (G,H) Toluidine Blue-O stained medial longitudinal sections of the SAM  
681 from normal siblings and *gras32 gras33* double mutants. Vertical and horizontal dashed lines  
682 indicate SAM height and width, respectively. (I, J) Quantification of SAM ( $n = 5-8$ ) (I) and AM  
683 ( $n = 3$ ) (J) height and width in normal siblings and *gras32 gras33* double mutants (two-tailed  
684 Student's t-test, \*  $p < 0.05$ ). Scale bars = 100  $\mu\text{m}$ .

685



686 **Fig. 3. Cell division dynamics throughout the maize SAM.** (A) Estimated cell division stage  
 687 of cells in the SAM + P2 datasets. Bar charts show fraction of cells in each stage among cell  
 688 clusters. (B,C) RNA *in situ* hybridization with antisense probe to *H3* (B; SAM,  $n = 9$ ; AM,  $n = 4$ )  
 689 and *CYC1* (C; SAM,  $n = 7$ ) in medial longitudinal sections of the SAM showing bins (outlined in  
 690 a dotted line grid and numbered) used for quantification of the proportion of cells in (D) S-phase  
 691 and (E) G2/M-phase. Scale bars = 100  $\mu\text{m}$ .

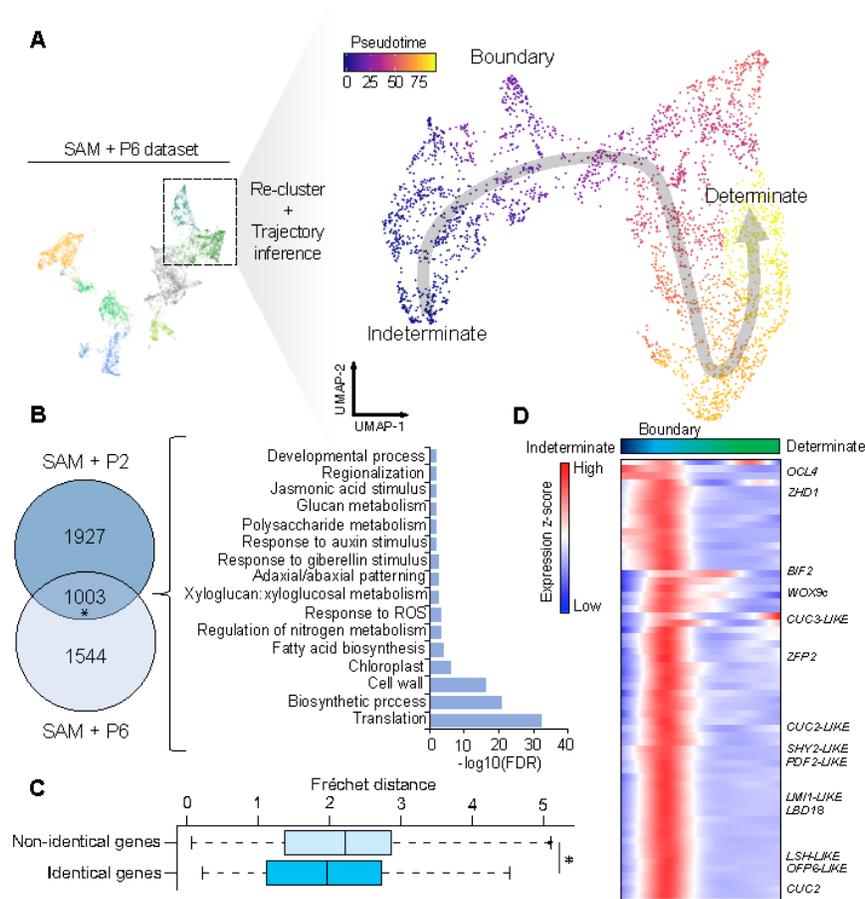
692



693 **Fig. 4. Tracing the gene expression patterns associated with cell differentiation.** (A)  
 694 Pseudotime values and trajectories for cells in the SAM + P2 dataset. (B) Gene expression  
 695 patterns of published marker genes of cell differentiation over pseudotime and illustrations of  
 696 their associated transcript accumulation patterns from RNA *in situ* hybridization studies. (C)  
 697 Heatmap of approximately 2000 genes that show correlated changes in gene expression along the  
 698 inferred trajectory clustered based on their expression patterns. Cells are mirrored along the  
 699 central axis prior to the trajectory branch point. Representative genes and significant GO term  
 700 enrichments for each cluster are shown. (D-H) Transcript accumulation patterns for trajectory-  
 701 correlated genes showing high expression levels at early, intermediate and late points in the  
 702 trajectory. RNA *in situ* hybridization of antisense probes to *ZNF-LIKE* and *RS21-C6* in medial  
 703 longitudinal sections of the SAM show early trajectory expression (D,E), *FPF1-LIKE* in  
 704 transverse section above the SAM and *WAT1-LIKE* in medial longitudinal section of the SAM  
 705 show late trajectory expression (F,G), and *HYP1* in longitudinal section below the SAM shows

706 intermediate trajectory expression (H). Arrowheads in (E,H) and asterisks in (H) represent  
707 transcript accumulation and AMs, respectively.

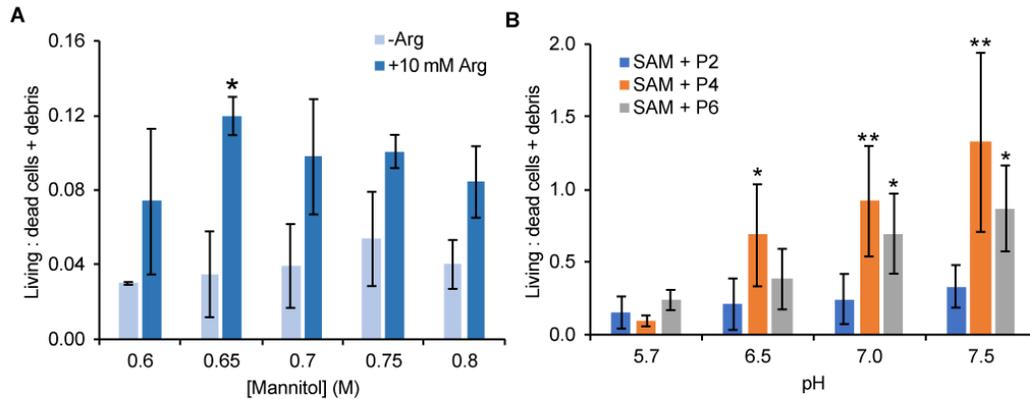
708



709 **Fig. 5. Identifying general features of the indeterminate-determinate cell fate transition**  
 710 **across ontogeny.** (A) A subset of cells from the SAM + P6 dataset (see Fig. S3B for inferred  
 711 cluster identities) were re-clustered and trajectory inference was used to assign cell pseudotime  
 712 scores along a transition from indeterminate to determinate cell fates. (B) Overlap of trajectory-  
 713 correlated genes from the SAM + P2 and SAM + P6 datasets and their GO term enrichment (\*,  
 714 Hypergeometric test,  $p$ -value =  $1.2E-320$ ). (C) Average Fréchet distance for identical and non-  
 715 identical genes in the SAM + P2 and SAM + P6 datasets (\*, Wilcoxon rank-sum test,  $p = 5.2E-$   
 716  $9$ ). (D) Heatmap of genes with high specificity for a boundary region that delimits more  
 717 indeterminate from more determinate cell fates with cells ordered by their pseudotime values.  
 718

719

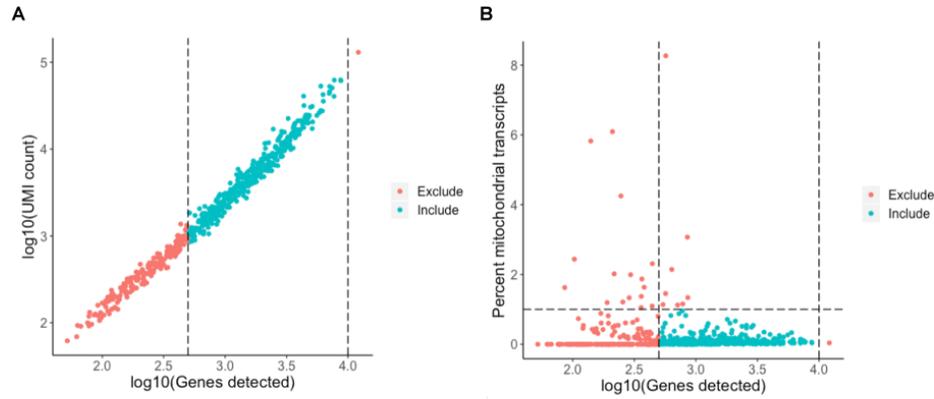
## Supplementary Figures



720

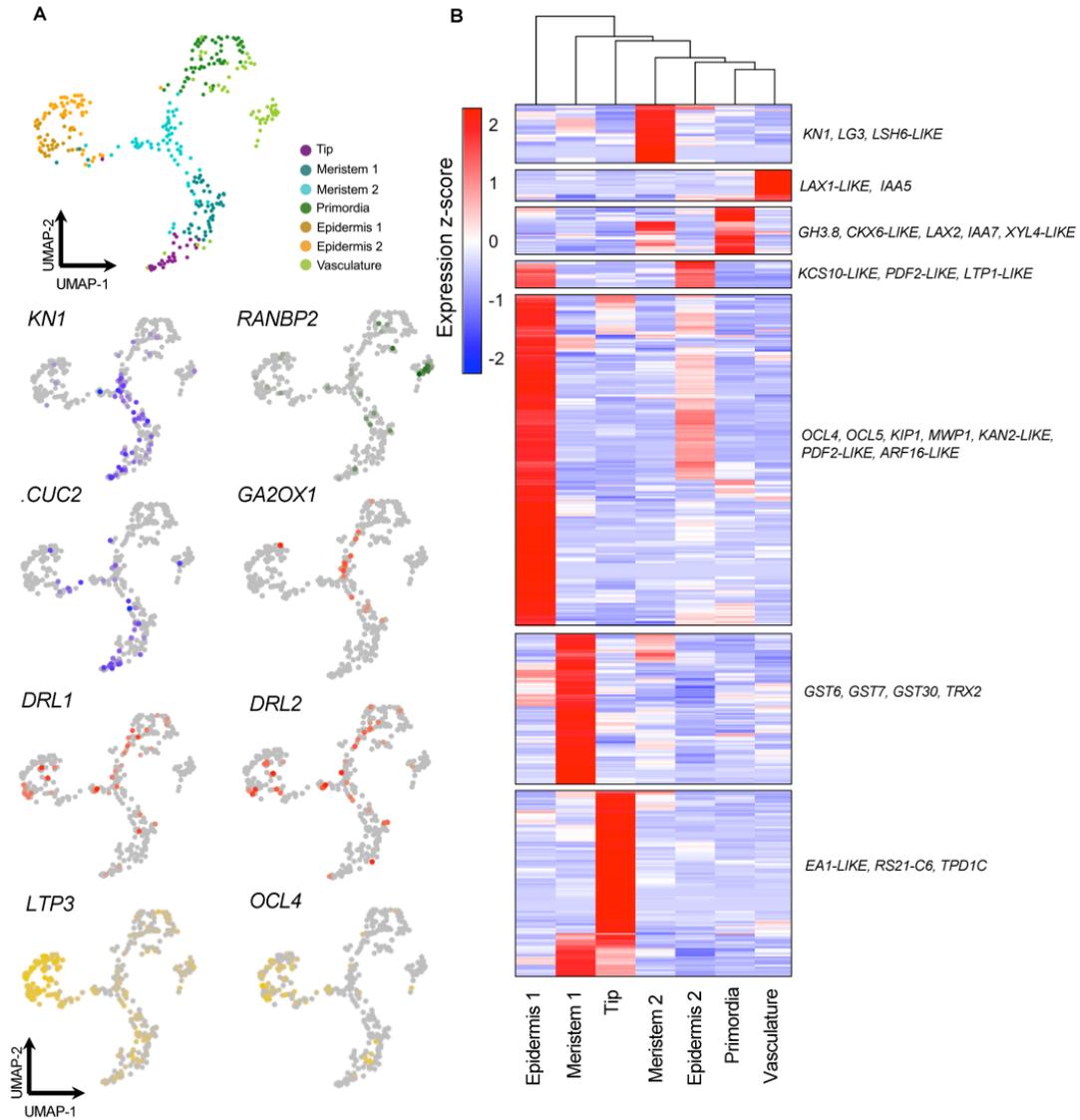
721 **Fig. S1. L-Arginine supplementation and increased pH improve maize protoplast viability.**  
722 (A) Ratio of FDA+ (viable) protoplasts relative to dead cells and large (> 5  $\mu\text{m}$ ) cellular debris in  
723 protoplasts isolated from SAM+P4 tissue carried out in solutions with varying mannitol and 10  
724 mM L-arginine supplementation ( $n = 2$ ) Asterisks signify the result of a Student's t-test  
725 comparing the mean of -Arg and +10 mM Arg samples at each pH value (\*,  $p < 0.05$ ). (B)  
726 Protoplast viability assayed in solutions buffered at varying pH levels and supplemented with 10  
727 mM L-arginine from three different tissue types ( $n = 2-3$ ). Asterisks signify the result of a  
728 Student's t-test comparing the mean of each tissue type at a given pH compared to its mean at pH  
729 5.7 (\*,  $p < 0.05$ ; \*\*,  $p < 0.01$ ).

730

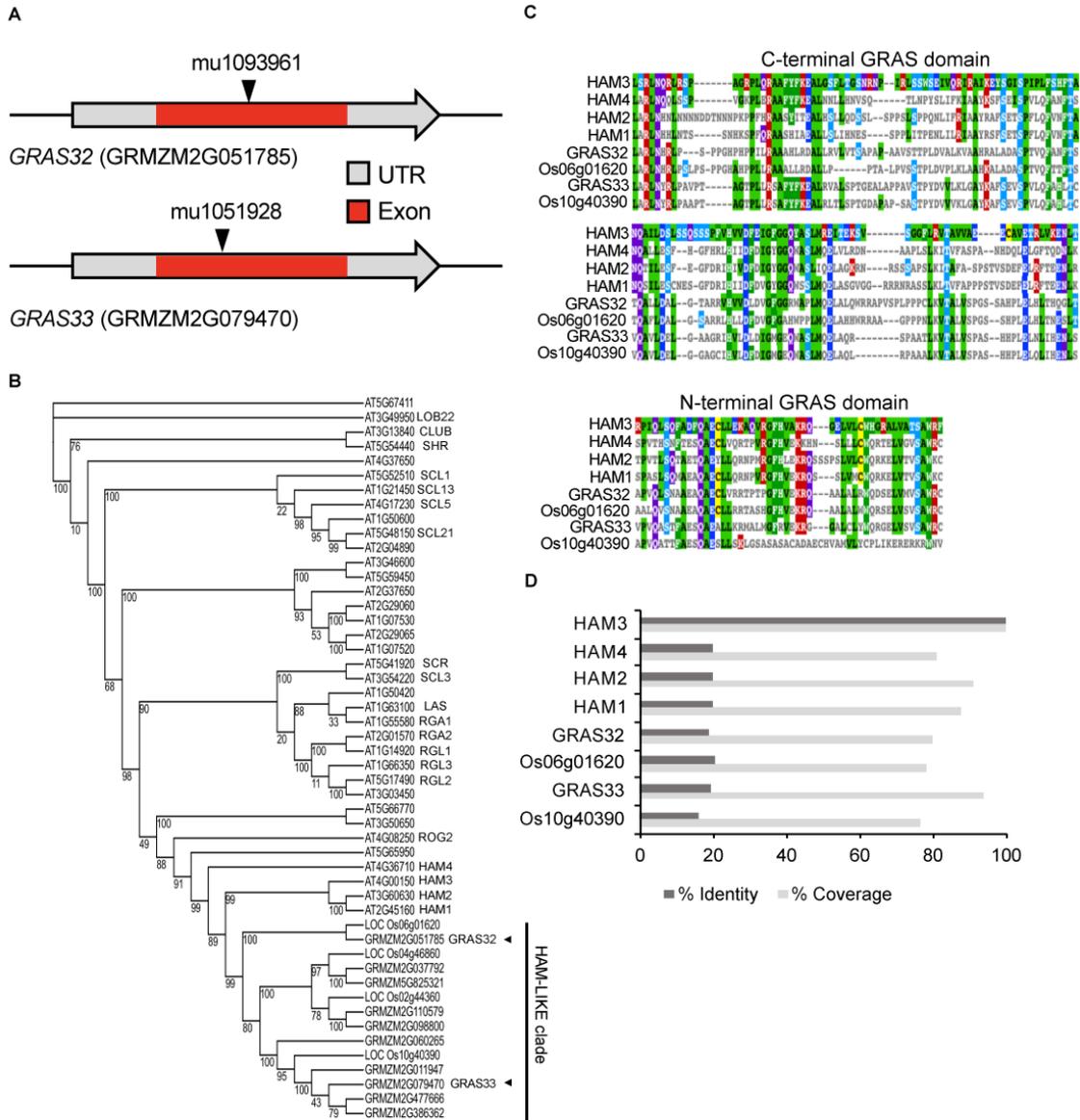


731  
732  
733  
734  
735  
736  
737

**Fig. S2. Cell filtering for the SAM+P2 dataset.** (A) Relationship between the number of genes and transcripts detected per cell in the SAM+P2 dataset. The dashed line indicates the genes per cell cutoff used to filter low quality cells. (B) Relationship between the number of genes detected and the percentage of mitochondrial transcripts per cell in the SAM+P2 dataset. Dashed lines indicate the genes per cell and percent mitochondrial transcripts filtering cutoffs.

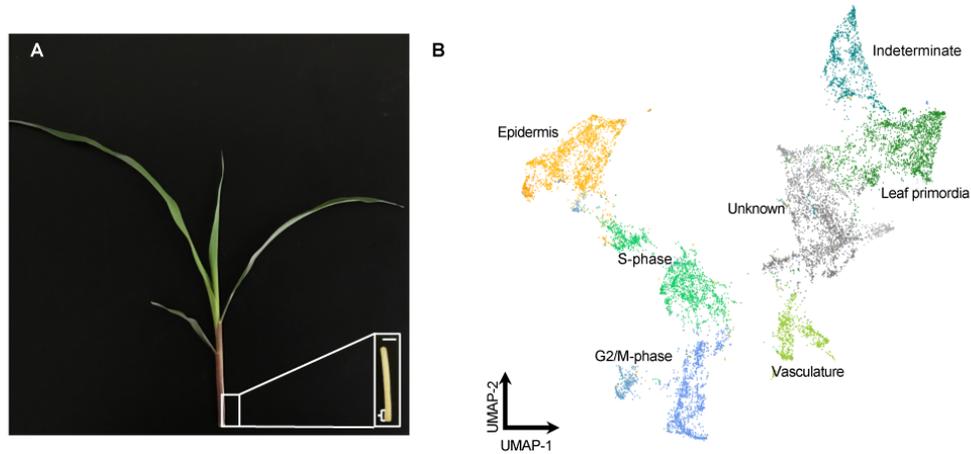


738 **Fig. S3. Expression patterns of markers used for cell type inference.** (A) Dimensionality  
 739 reduction of the SAM + P2 dataset and associated marker genes used to infer cell identity. (B)  
 740 Heatmap of all identified marker genes for each cell type cluster with representative markers of  
 741 each group indicated to the right.  
 742

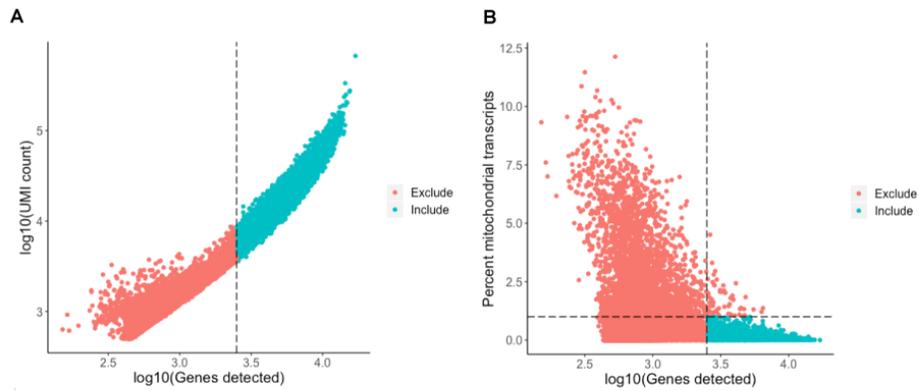


743 **Fig. S4. Maize *GRAS32* and *GRAS33* are *HAM-LIKE* genes.** (A) Gene models for maize  
744 *GRAS32* and *GRAS33* showing the *Mutator* transposon insertion sites for the mutant alleles used  
745 in this study. (B) Maximum likelihood phylogenetic tree of Arabidopsis GRAS family  
746 transcription factor amino acid sequences with maize and rice *HAM-LIKE* homologs. Branch  
747 support values are the result of the Approximate Likelihood-Ratio test. (C) Amino acid  
748 alignment showing regions of high similarity in the N- and C-terminal regions of the GRAS  
749 domain between select maize, rice, and Arabidopsis *HAM* proteins. (D) Amino acid percent  
750 identity and percent coverage of *HAM* homologs relative to Arabidopsis *HAM3*.

751



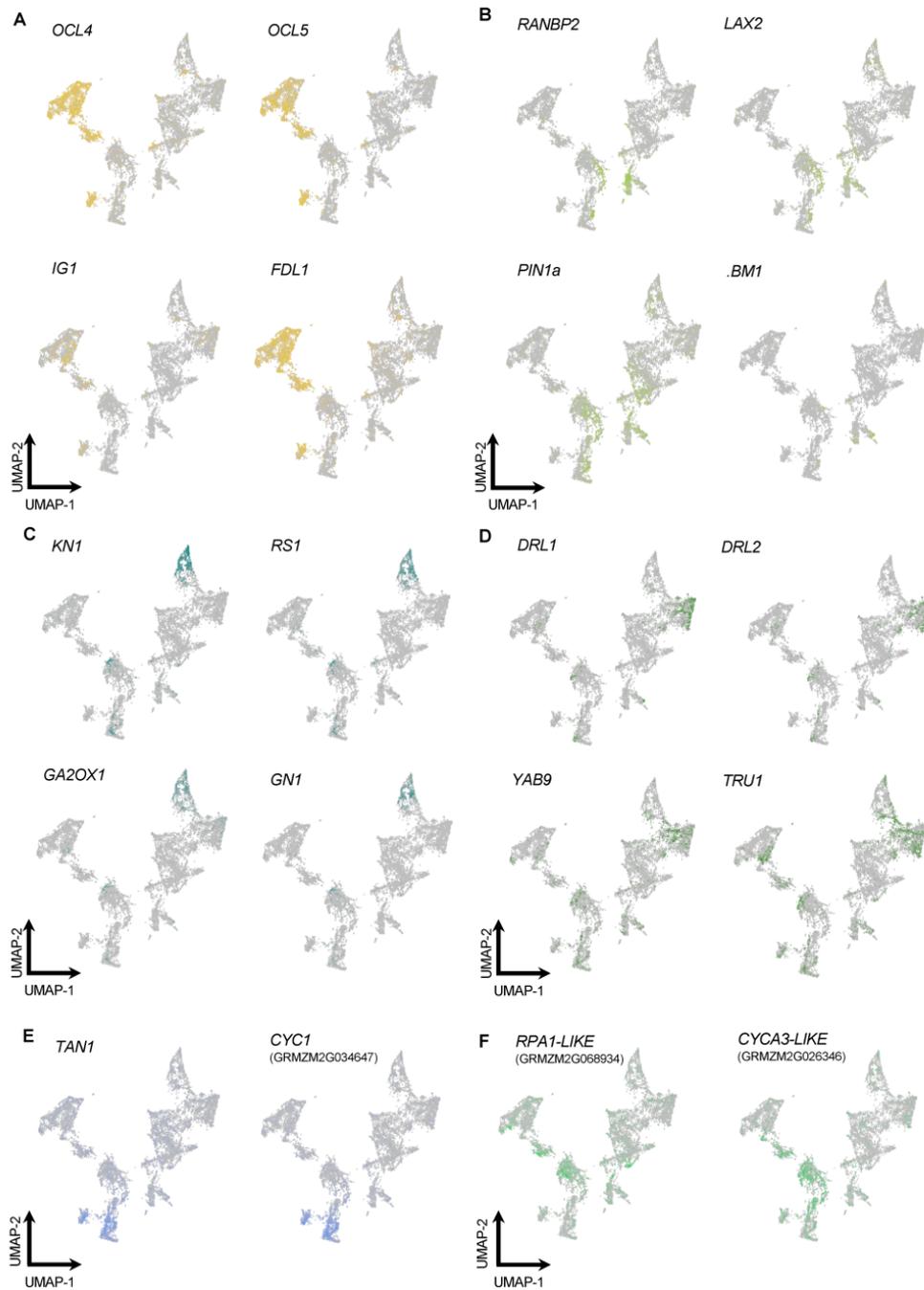
752 **Fig. S5. Tissue collection and cell type representation in the SAM+P6 dataset.** (A) Two-  
753 week-old inbred B73 maize seedling. (Inset) Dissection to plastochron 6 (P6) showing the tissue  
754 targeted for analysis consisting of approximately 3 mm of stem and P6 tissue (scale bar = 5 mm).  
755 (B) Dimensionality reduction of the filtered cellular transcriptomes represented in the SAM + P6  
756 dataset along with their inferred cluster identities.  
757



758

759 **Fig. S6. Cell filtering for the SAM+P6 dataset.** (A) Relationship between the number of genes  
760 and transcripts detected per cell in the SAM+P6 dataset. The dashed line indicates the genes per  
761 cell cutoff used to filter low quality cells. (B) Relationship between the number of genes detected  
762 and the percentage of mitochondrial transcripts per cell in the SAM+P6 dataset. Dashed lines  
763 indicate the genes per cell and percent mitochondrial transcripts filtering cutoffs.

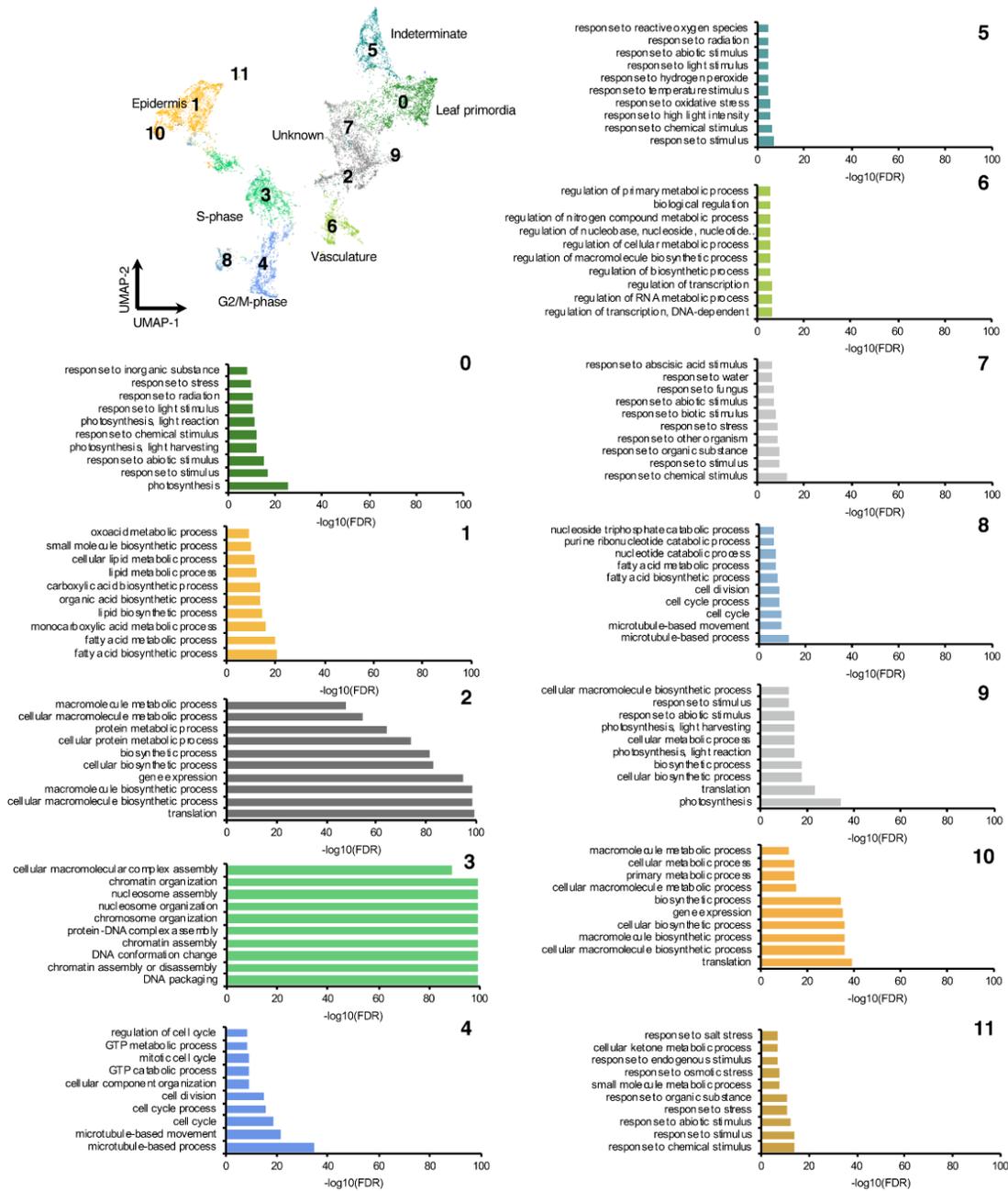
764



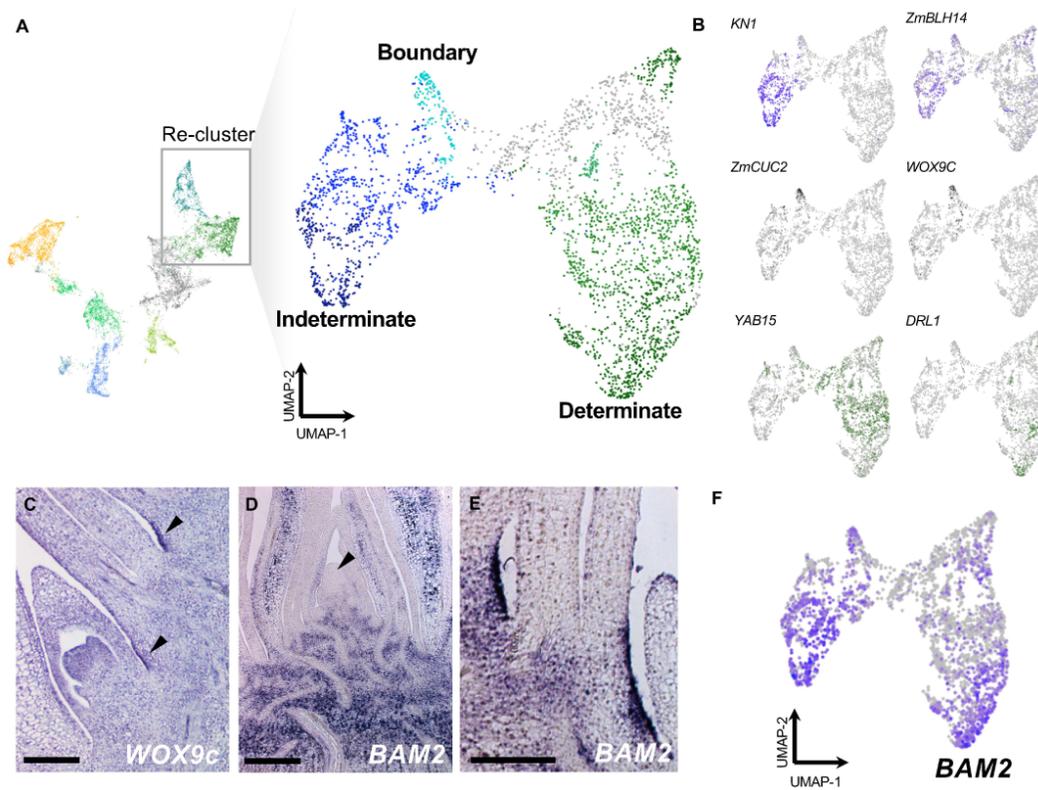
765

766 **Fig. S7. Expression patterns of markers used for cell type inference.** Selected marker genes  
767 for: (A) epidermis; (B) vasculature; (C) indeterminate (stem tissue); (D) leaf primordia; (E)  
768 G2/M-phase; (F) S-phase.

769



770 **Fig. S8. Gene ontology (GO) enrichment for each of the identified clusters in the SAM + P6**  
 771 **dataset.** Clusters identified by hierarchical clustering were numbered. Bar plots display the top  
 772 10 most significant GO enrichments for genes that mark each cluster.  
 773



774 **Fig. S9. Marker gene expression among cells with determinate and indeterminate identity.**  
775 (A) Re-clustering and cluster assignment of cells identified in the SAM + P6 dataset. (B)  
776 Expression patterns of marker genes among re-clustered cells. (C) *WOX9c* transcript  
777 accumulation patterns in boundary regions of the seedling shoot system (arrowheads). (D,E)  
778 *BAM2* expression patterns in the seedling shoot showing expression in indeterminate and  
779 determinate cell populations (with the SAM indicated by an arrowhead) (D) as well as in  
780 boundary regions (E). (F) *BAM2* expression in the SAM + P6 dataset. Scale bars = 250  $\mu$ m.  
781

782

**Table S1. Genes used for cell cycle regression.**

783

**Table S2. Differentially expressed genes for all clusters in the SAM + P2 dataset.** Statistical results ( $p$ -value) reflect the output of a Wilcoxon ranked sum test followed by a Bonferonni correction (Adjusted  $p$ -value) performed in Seurat v3.

784

785

786

**Table S3. Gene Ontology (GO) term enrichment for differentially expressed genes in each SAM + P2 cluster.** Differentially expressed genes (see Table S1,  $p$ . adj < 0.05) were analyzed for GO Term enrichment using AgriGo v2.

787

788

789

**Table S4. Average cluster-based expression of potential maize CLV1-CLV3-WUS pathway genes.** Values reflect average gene expression in each cluster (UMI counts).

790

791

**Table S5. Differentially expressed genes in cells positive for expression of the core marker gene (GRMZM2G049151) in the SAM + P2 dataset.** Statistical results ( $p$ -value) reflect the output of a Wilcoxon ranked sum test followed by a Bonferonni correction (Adjusted  $p$ -value) performed in Seurat v3.

792

793

794

795

**Table S6. Genes with dynamic expression over pseudotime in the SAM + P2 dataset.**

Statistical results reflect the output of a Moran's  $I$  test implemented in Monocle v3 for epidermal and primordia/vasculature trajectories.

796

797

798

**Table S7. Differentially expressed genes for all clusters in the SAM + P6 dataset.** Statistical results ( $p$ -value) reflect the output of a Wilcoxon ranked sum test followed by a Bonferonni correction (Adjusted  $p$ -value) performed in Seurat v3 (.).

799

800

801

**Table S8. GO Term enrichment for heatmap clusters of pseudotime-correlated genes.**

Differentially expressed genes (see Table S7,  $p$ . adj < 0.05) were analyzed for GO Term enrichment using AgriGo v2.

802

803

804

**Table S9. Gene Ontology (GO) term enrichment for differentially expressed genes in each SAM + P6 cluster.** Differentially expressed genes (see Supplementary Table 1,  $p$ . adj < 0.05) were analyzed for GO Term enrichment using AgriGo v2.

805

806

807

**Table S10. Oligonucleotide and primer sequences used in this study.**

808