Single-cell transcriptome reveals the redifferentiation trajectories of the early stage of *de novo* shoot regeneration in *Arabidopsis thaliana*

Authors:

Guangyu Liu^{1,#}, Jie Li^{1,2,#}, Ji-Ming Li^{1,#}, Zhiyong Chen¹, Peisi Yuan¹, Ruiying Chen^{1,2}, Ruilian Yin^{1,2}, Zhiting Liao¹, Xinyue Li^{1,2}, Ying Gu^{1,3,*}, Hai-Xi Sun^{1,2,*}, Keke Xia^{1,*}

¹ BGI-Shenzhen, Shenzhen 518083, Guangdong, China

² College of Life Sciences, University of Chinese Academy of Sciences, Beijing 100049 , China

³ Guangdong Provincial Key Laboratory of Genome Read and Write, BGI-Shenzhen, Shenzhen 518120, Guangdong, China

[#] These authors contributed equally to this work.

*Corresponding authors:

Keke Xia, (xiaoke@genomics.cn)

Hai-Xi Sun, (sunhaixi@genomics.cn)

Ying Gu, (guying@genomics.cn)

Author list:

Guangyu Liu, (liuguangyu@genomics.cn)

Jie Li, (lijie8@genomics.cn)

Ji-Ming Li, (lijiming@genomics.cn)

Zhiyong Chen, (chenzhiyong@genomics.cn)

Peisi Yuan, (yuanpeisi@genomics.cn)

Ruiying Chen, (chenruiying@genomics.cn)

Ruilian Yin, (yinruilian@genomics.cn)

Zhiting Liao, (liaozhiting@genomics.cn)

Xinyue Li, (lixinyue1@genomics.cn)

Ying Gu, (guying@genomics.cn)

30 Hai-Xi Sun, (sunhaixi@genomics.cn)

31 Keke Xia, (xiakeke@genomics.cn)

32

33

34 Abstract

35 *De novo* shoot regeneration from a callus plays a crucial role in both plant biotechnology and
 36 the fundamental research of plant cell totipotency. Recent studies have revealed many
 37 regulatory factors involved in this developmental process. However, our knowledge of the
 38 cell heterogeneity and cell fate transition during *de novo* shoot regeneration is still limited.
 39 Here, we performed time-series single-cell transcriptome experiments to reveal the cell
 40 heterogeneity and redifferentiation trajectories during the early stage of *de novo* shoot
 41 regeneration. Based on the single-cell transcriptome data of 35,669 cells at five-time points,
 42 we successfully determined seven major cell populations in this developmental process and
 43 reconstructed the redifferentiation trajectories. We found that all cell populations resembled
 44 root identities and undergone gradual cell-fate transitions. In detail, the totipotent callus cells
 45 differentiated into pluripotent QC-like cells and then gradually developed into less
 46 differentiated cells that have multiple root-like cell identities, such as pericycle-like cells.
 47 According to the reconstructed redifferentiation trajectories, we discovered that the canonical
 48 regeneration-related genes were dynamically expressed at certain stages of the
 49 redifferentiation process. Moreover, we also explored potential transcription factors and
 50 regulatory networks that might be involved in this process. The transcription factors detected
 51 at the initial stage, QC-like cells, and the end stage provided a valuable resource for future
 52 functional verifications. Overall, this dataset offers a unique glimpse into the early stages of
 53 *de novo* shoot regeneration, providing a foundation for a comprehensive analysis of the
 54 mechanism of *de novo* shoot regeneration.

55

56 Keywords:

57 Single-cell transcriptome, redifferentiation trajectories, transcription factors, *de novo* shoot
 58 regeneration, *Arabidopsis*

59

60

61

62 Introduction

63 In contrast to animals, plants have a remarkable regeneration capability, ensuring
64 developmental plasticity in various environmental conditions. Until now, two modes of
65 regeneration have been described, including indirect *de novo* shoot regeneration and direct
66 shoot organogenesis¹. Indirect shoot regeneration includes two steps: pluripotent callus
67 induction on an auxin-rich callus inducing medium (CIM) and shoot regeneration on a
68 cytokinin-rich shoot inducing medium (SIM)¹. Based on plant regeneration ability, *in vitro*
69 tissue culture techniques have been widely used in agriculture². Although many studies have
70 been carried out on plant regeneration, the mechanistic details of shoot regeneration are
71 largely unknown. Shoot regeneration is still a major hurdle in tissue culture^{3,4}. Therefore, a
72 comprehensive and deeper understanding of the plant regeneration process is urgently needed.

73
74 During *de novo* shoot regeneration, cells undergo complex cell fate transition and
75 reprogramming, and four phases have been defined recently, including pluripotency
76 acquisition (phase I), shoot pro-meristem formation (phase II), the establishment of the
77 confined shoot progenitor (phase III), and shoot outgrowth (phase IV)⁵. These phases involve
78 highly dynamic changes in the transcriptome and reprogramming of cell fate. Recently,
79 several genes and regulatory networks involved in regulating shoot regeneration have been
80 identified. For example, *WUSCHEL* (*WUS*), a core transcriptional factor required for shoot
81 regeneration, is regulated by the dynamic expression of Type-B *ARABIDOPSIS RESPONSE*
82 *REGULATORs* (*ARRs*), including *ARR1*, *ARR2*, *ARR10*, and *ARR12*^{6,7}. The *WOX11-LBD16*
83 pathway, *PLT3, 5, 7-PLT1, 2* pathway, and *WOX5-PLT1, 2-TAA1* pathway promote
84 pluripotency acquisition in callus cells⁸⁻¹⁰, and *PLT3, 5, 7-CUC2* pathway promotes *de novo*
85 shoot formation⁹. What's more, *CUP-SHAPED COTYLEDON 1* (*CUC1*) and *CUC2* are
86 controlled by *ENHANCER OF SHOOT REGENERATION 1* (*ESR1*) for the establishment of
87 shoot pro-meristem¹¹⁻¹³. However, the knowledge on the regulatory network of *de novo* shoot
88 regeneration is still limited.

89

90 In addition, the interaction of multiple cell types and the spatiotemporal dynamic expression

of genes also play important roles in controlling *de novo* shoot regeneration^{7,14,15}, such as the mutually exclusive and spatiotemporal specific distribution pattern of auxin and cytokinin response cells¹⁴. Auxin and cytokinin are uniformly distributed at the edge region of the callus cultured on CIM. After transferring callus to SIM, auxin response cells progressively translocate and form the ‘auxin ring’, whereas the cytokinin response cells are restricted to the center of the auxin ring, overlapping with the region of *WUS* expression. These findings suggest that a variety of cell types with unknown characteristics appear progressively during *de novo* shoot regeneration.

Despite extensive studies, many questions related to *de novo* shoot regeneration remain to be answered. For example, to what extent does heterogeneity present in a callus? Which cell type in a callus can regenerate new shoots? How do cell populations change during *de novo* shoot regeneration? What are the expression patterns of genes involved in shoot regeneration? These unsolved questions are largely due to technical limitations. For instance, bulk RNA-seq has been widely used in characterizing molecular changes during *de novo* shoot regeneration, but information related to callus heterogeneity is lost^{15,16}. Therefore, to answer these questions mentioned above, new techniques with higher resolution are needed.

Recent advances in single-cell RNA sequencing (scRNA-seq) technology provide unprecedented opportunities for systematically identifying cellular and molecular differentiation trajectories of plant regeneration at the single-cell level^{17,18}. The scRNA-seq technique has been widely used in plant development studies, facilitating reconstructing the development trajectories and establishing gene expression regulatory networks of root epidermis^{19,20}, lateral root development²¹⁻²⁸, shoot apex^{29,30}, leaf phloem cells^{31,32}, guard cells^{33,34}, seedlings³⁵ and callus¹⁰, demonstrating its advantages in discovering regulatory elements³⁶. Up to now, scRNA-seq has not been used to analyze the complex shoot regeneration process in *Arabidopsis*.

Here, by performing time-series single-cell transcriptome experiments, we revealed the early

redifferentiation process of *de novo* shoot regeneration at the single-cell level. First, we successfully reconstructed the early redifferentiation trajectories of *de novo* shoot regeneration. Next, we carried out the cell identities analysis and found seven major cell populations that resembled QC cells and other root cell types. Moreover, we discovered the dynamic expression patterns of canonical regeneration-related genes and also explored potential transcription factor regulatory networks involved in *de novo* shoot regeneration, which provided valuable resources for a deeper understanding of the mechanism of shoot organogenesis.

Results

Constructing the early redifferentiation trajectories during *de novo* shoot regeneration from callus cells

During the early period of *de novo* shoot regeneration, we could observe distinct morphological changes. For example, the callus gradually turned compact and green after the callus was transferred to SIM for around 10 days (Supplementary Fig. 1a), indicating that the cell identities had already been reprogrammed. To comprehensively understand cell fate transition during the early stage of *de novo* shoot regeneration, we performed scRNA-seq using root-derived callus cultured on SIM for 0, 2, 4, 6, 10 day(s), respectively (Fig. 1a). For each time point, three replicates were included. After data filtering, we obtained a total of 35,669 cells from the five time points. The cell population from each time point was designated as SIM0, SIM2, SIM4, SIM6, and SIM10, which had 7,571, 9,755, 4,893, 7,795, and 5,655 cells, respectively. In total, 21,129 genes were detected with an average of 1,652 genes per cell.

To check the reproducibility of the data, we performed correlation analysis among all samples. We found that samples from the same time point were clustered together and showed high similarity (Supplementary Fig. 1b), indicating reproducibility of the data. In addition, to assess data consistency at the gene level among cell populations from different time points, we checked the gene number, the unique molecular identifier (UMI) number, and the

149 proportion of mitochondrial and chloroplast genes. We observed similar numbers for different
150 time points for each of the tested parameters (Supplementary Fig. 1c). Lastly, a comparison of
151 gene expression between scRNA-seq and bulk RNA-seq at day 0, day 2, day 4, and day 6,
152 respectively, also showed the reliability of our dataset (Supplementary Fig. 1d). Taken
153 together, we obtained reliable and reproducible data which were used for further analysis.

154

155 As depicted in Supplementary Fig. 1b, the correlation of cell populations changed gradually
156 from SIM0 to SIM10, demonstrating that the gene expression patterns of these cells had
157 changed during *de novo* shoot regeneration. To reconstruct the redifferentiation trajectories
158 during *de novo* shoot regeneration, we first used well-known tools, including monocle 2 and
159 monocle 3³⁷. However, we found that the resulting pseudo-time trajectories contradicted with
160 timeline (Supplementary Fig. 2a-e). Thus, monocle 2 and monocle 3 might be not suitable for
161 this dataset. In this study, the root-derived pluripotent callus would eventually develop into
162 the more differentiated shoot apex and shoot tissues. Therefore, we decided to manually select
163 pluripotency- and differentiation-related genes for re-constructing the developmental
164 trajectories.

165

166 First, we compared SIM0 scRNA-seq data with the previously reported scRNA-seq data
167 obtained from root³⁸ and shoot apex²⁹, respectively. We found that there were 2,623 and 2,632
168 genes up-regulated in callus compared with root and shoot apex, respectively. Then the 1,888
169 overlapped genes of these two up-regulated genes lists were obtained and designated as pre-
170 pluripotency-related genes (Fig. 1b,c and Supplementary Table 1). Then, we further selected
171 genes that might be involved in pluripotency or differentiation processes by assessing the
172 expression patterns of these 1,888 genes in both bulk RNA-seq datasets and the pseudo-bulk
173 RNA-seq dataset. Among them, 721 genes that had decreasing expression patterns in both the
174 bulk RNA-seq dataset and the pseudo-bulk RNA-seq dataset were found and defined as
175 pluripotency-related genes (Fig. 1d and Supplementary Fig. 3c). Similarly, we obtained 287
176 genes that had increased expression patterns in both the bulk RNA-seq dataset and the
177 pseudo-bulk RNA-seq dataset and designated these genes as differentiation-related genes (Fig.
178 1e and Supplementary Fig. 3c). In addition, we also included 61 (6 genes were repetitive with

selected pluripotency- and differentiation-related genes) previously reported genes that are involved in shoot regeneration as regeneration-related genes (Supplementary Fig. 3d,e). Finally, a total of 1,063 genes, including 721 pluripotency-related genes, 287 differentiation-related genes, and 55 regeneration-related genes, were selected as input for cell clustering using the Seurat package (Supplementary Table 1). As shown in Fig. 1f, the 35,669 cells were clustered into 7 clusters. Excitingly, expression of pluripotency-related genes in different clusters showed a gradual decrease in the order of cell population 5, 0, 4, 6, 3, 2, and 1 (Fig. 1g), and the order of cell populations largely corresponded to the main populations in each time point (see below). Thus, we reconstructed the redifferentiation trajectories according to cell pluripotency and differentiation, and we defined the redifferentiation trajectories as the following: the cell population 5 were highly pluripotent, which gradually differentiated into cell population 0, 4, 6, 3, 2 and finally developed into more differentiated cell population 1 during the early stage of *de novo* shoot regeneration (Fig. 1h). So far, the scRNA-seq technology facilitated the construction of the landscape of redifferentiation trajectories, which could offer insights into the characteristics of the cell populations during *de novo* shoot regeneration.

195

196 **Cell populations resemble QC cells and other known root cell types during the early** 197 **stage of *de novo* shoot regeneration**

198 Recently, Shin et al. has reconstituted the shoot regeneration process with four phases,
199 including pluripotency acquisition (phase I), shoot pro-meristem formation (phase II),
200 establishment of the confined shoot progenitor (phase III), and shoot outgrowth (phase IV)⁵.
201 Here, to determine the developmental stage of shoot regeneration in our dataset, we assessed
202 the expression pattern of phase-specific genes at the five time points. As expected, phase I-
203 specific genes were expressed, including *WOX5*, *SCR*, and *PLT1/2* (Fig. 2a and
204 Supplementary Fig. 4a). In addition, we detected the expression of phase II-specific genes,
205 such as *PINI*, *CUC2*, and *ESR1* (Fig. 2a and Supplementary Fig. 4a). However, we did not
206 detect the expression of phase III-specific genes, such as *WUS* and *CLV3*, and these genes are
207 key shoot meristem marker genes. Moreover, the expression of *WUS* was also not detected in
208 our bulk RNA-seq data (Data not shown). These results demonstrated that the shoot

organogenesis process in this study preceded phase III.

Next, we explored whether the identified populations showed any quantitative affinities to the known cell types. It has been shown that callus has root identity³⁹, and root identity marker genes, such as *WOX5*, *SCR*, and *PLT1/2*, were also expressed in our dataset. As *de novo* shoot organogenesis eventually leads to the formation of the shoot, we systematically explored known cell identities, including both root and shoot cell types (Supplementary Table 2). It turned out that cell populations in our dataset were more similar to root cell types rather than leaf cell types (Fig. 2b). For example, cell populations 0, 2, and 6 had the highest endodermis scores, while cell populations 4 and 5 resembled root epidermis and cortex, respectively. Surprisingly, cell population 3 presented the highest score for quiescent center (QC) cells. Consistently, QC-like cells have also been found in hypocotyl-derived callus in a recent study¹⁰. Lastly, cell populations changed dramatically at day 10, with the majority of cells belonging to cell population 1 (Fig. 2c,d). We found that cell population 1 had an affinity to pericycle identity (Fig. 2b). It has been shown that the pericycle cells have the organogenetic capacity and are the source of lateral roots⁴⁰, suggesting that the pericycle identity of population 1 might play an important role in the shoot regeneration process. In addition to pericycle cells, cell population 1 also showed high scores for other root cell types, such as the cortex, columella, and endodermis cells, suggesting that the composition of cell population 1 might be complicated.

To further validate cell population identities, we determined the overlap between cluster-specific markers in our dataset and the reference cell type-specific marker genes (Supplemental Table 2). And then we generated the heatmap using the top 5 overlapped marker genes. As shown in Fig. 2e, the top 5 overlapped genes for cell clusters 5, 3, and 1 had relatively high specific expression patterns. We also observed that the marker genes of cell cluster 0 were also highly expressed in cell population 5, suggesting that these two populations were similar. In addition, we found *AT4G14620* and *AT4G14622* were specifically and highly expressed in cluster 6. In conclusion, these specifically expressed marker genes not only confirmed the reliability of the cell type identification but also

provided potential marker genes and candidate genes for future functional studies, such as *AT4G14620* and *AT4G14622*.

Then we investigated gene ontology (GO) enrichment in the top 100 gene markers (Supplemental Table 2) for each population and found specific features (Supplementary Fig. 4b and Supplemental Table 2). Interestingly, multiple plant hormone pathways were specifically enriched in certain cell populations (Supplementary Fig. 4b). It is well-known that phytohormones, such as auxin, JA, SA, and ethylene, play important roles during plant regeneration^{5,41}, when and where these phytohormones play their roles remain unclear. We found that the GO terms associated with populations 5 and 0 were highly similar, and the auxin biosynthesis process was specifically enriched in these cell populations. For cell population 3 (QC-like cells), biological processes related to jasmonic acid (JA) and salicylic acid (SA) were specifically enriched (Fig. 2f). In addition, the GO terms related to the ethylene metabolic process were enriched in cell population 1 (Fig. 2f). Here, based on the re-constructed trajectories using scRNA-seq, we were able to demonstrate the spatiotemporal roles of these phytohormones during shoot regeneration.

In addition to the enrichment of phytohormone-related biological processes, other GO terms were also enriched in specific cell populations. For example, the amine metabolism process in cell populations 5 and 0 (Supplementary Fig. 4c), suggesting that the amine metabolism process might be involved in regulating the formation of callus. Interestingly, we also observed stress-related process was specifically enriched in cell population 1 (Supplementary Fig. 4b). Recently, Zeng et. al found that the stress-related signal was overrepresented in the shoot stem cell niche, and it was necessary to maintain *Arabidopsis* shoot stem cells⁴². Here, the enriched stress response processes in cell population 1 appeared before the formation of the shoot apex, indicating the stress signaling might play a role in regulating the formation of shoot stem cells.

Dynamic expression patterns of regeneration-related genes along the redifferentiation trajectory

269 So far, many key regulatory genes have been discovered during the shoot regeneration
270 process^{5,43}, but few studies have focused on the dynamic expression patterns of these genes.
271 In this study, we detected the expression of 75 previously reported regeneration-related genes
272 (Supplementary Table 3), and then we investigated the expression patterns of these genes
273 along the reconstructed redifferentiation trajectory. As shown in Fig. 3a,b, we observed 3
274 distinct dynamic expression patterns. Genes in patterns 1 and 2 had gradually increased and
275 decreased expression trend, respectively, while genes in pattern 3 had specifically high
276 expression in cell populations 3 (Fig. 3a,b).

277
278 Thirty-nine genes were present in pattern 1, and these genes were highly expressed in cell
279 population 1. GO analysis showed that the genes related to the primary shoot apical meristem
280 specification, meristem initiation, and shoot system morphogenesis were highly enriched in
281 pattern 1 (Fig. 3c), and the finding was consistent with the fact that the regenerating callus
282 would eventually form the shoot. Among the 39 genes in pattern 1, some genes have been
283 confirmed as key factors for shoot pro-meristem formation, such as cytokinin receptor genes
284 *AHK4*, *WOX4*, and *PIN1*^{44,45} (Fig. 3d).

285
286 For pattern 2, 22 genes were present, and they were highly expressed on day 0 (Fig. 3b). We
287 found that the genes related to biosynthesis, polar transportation, receptor, and response factor
288 of auxin were included in pattern 2, such as *ARF5*, *TIR1*, *LBD16*, *YUC4*, and *PIN2* (Fig. 3f),
289 which was consistent with the enrichment of auxin biosynthetic process in cell populations 5
290 and 0 in Fig. 2e,f and Supplementary Fig. 5a. In addition, *ARF5*, *LBD16*, and *TIR1* are known
291 as key factors in regulating callus formation^{46,47}. We also detected the expression of *WOX13*
292 in Pattern 2 (Supplementary Fig. 5a), and this gene has been verified to modify cell walls for
293 efficient callus formation⁴⁸. Considering many genes in pattern 2 have been reported to be
294 involved in regulating callus formation, we speculated that these genes in pattern 2 could
295 mainly play roles in regulating the formation of callus, but not in the shoot regeneration
296 process^{46,47}.

297
298 As for gene expression pattern 3, 17 genes were highly expressed in the identified QC-like

cells, and the stem cell population maintenance GO term was especially enriched in this expression pattern (Fig. 3g). We found that genes involved in maintaining the quiescent central niche were included in this pattern, such as *PLT1/5*, and *WOX5* (Fig. 3h), indicating that these 17 genes might promote the acquisition of cell pluripotency during the redifferentiation process^{49,50}. Taken together, here we demonstrate that the 75 canonical regeneration-related genes were specifically expressed at certain stages during shoot regeneration, and these findings shed light on the shoot regeneration mechanisms.

Inference of cell type-specific transcription factor regulatory networks

To understand the roles of transcription factors (TFs) during the early stage of *de novo* shoot regeneration, we analyzed the expression of TFs in our dataset. Among the detected 1254 TFs we screened 101 TFs for further analysis (Supplementary Table 4), which were highly expressed in more than 33% cells of each cell population.

First, we investigated the expression patterns of the 101 transcription factors along the re-constructed redifferentiation trajectories. As shown in Fig. 4a, we observed three main expression patterns. For example, some TFs were highly expressed in cell populations 5 and 0, while other TFs were highly expressed in cell populations 3 and 1, respectively (Fig. 4a), indicating that different developmental stages require different sets of TFs during *de novo* shoot regeneration. Consistent with the above-mentioned results in this study (Fig. 3a and Supplementary Fig. 4b), we observed the high similarity between cell populations 5 and 0. Then we checked which TF was present in these patterns. We found that TFs such as *RAP2.6*, *NAP*, *ATMYBR1*, and *HY5* were highly expressed in cell populations 5 and 0, and *WRKY70*, *WRKY18*, and *WRKY75* were predominantly expressed in cell population 3, and *WRKY18* and *RRTF1* were highly expressed in the cell population 1. As far as we know, these TFs have not been proved to be involved in the regeneration process. Based on the dynamic and specific expression patterns, these TFs could play important roles in regulating the cell fate transition during *de novo* shoot regeneration, and further mutant analysis is required in future studies

To further explore the TFs and their target genes, we performed TF regulatory network

(TFRN) analysis. By combining the regulatory relationships analysis results and the co-expression analysis with the verified *Arabidopsis* cistrome⁵¹, we were able to show the TFRNs (Fig. 4b). 31 core TFs and 642 target genes were revealed (Supplementary Table 4). Five and 17 TFRNs were identified in cell populations 5 and 0, respectively, and 3 TFRNs were shared between these two cell populations, such as *WRKY45*, *WRKY65*, and *HY5*. For the QC-like cell population 3, 8 TFRNs were identified, and *ATAF1*, *ERF13*, *bZIP53*, *WRKY33*, *WRKY70*, *WRKY18*, *WRKY75*, and *AT1G74840* were included. Five TFRNs were identified in cell population 1, such as *ATAF1*, *ZAT10*, *WRKY18*, *RRTF1*, and *AT1G19210*. Among these TFRNs, *RAP2.6L* is required for *de novo* shoot regeneration^{52,53}, and we discovered that there were 3 candidate TFs that directly regulate *RAP2.6L* expression, including *RAP2.6*, *MYB44*, and *VNI2*. In addition, *ARR5*, a transcription repressor that mediates the negative feedback loop in cytokinin signaling, showed a significantly increased expression pattern when callus was transferred to SIM (Supplementary Fig. 5b), which was consistent with the role of *ARR5* in pluripotency acquisition¹⁰. We found that the *WRKY18* might directly promote the expression of *ARR5* during *de novo* shoot regeneration. The predicted TFRNs provided valuable resources for future studies on the mechanism of *de novo* shoot regeneration.

Discussion

Here using scRNA-seq, we characterized the early stage of *de novo* shoot regeneration from a root-derived callus and reconstructed the redifferentiation trajectories. The characteristics of the seven major cell populations along the redifferentiation trajectories would provide valuable resources for a comprehensive and deeper understanding of the mechanism of shoot organogenesis.

To construct the developmental trajectory of callus cells during *de novo* shoot regeneration, we innovatively selected pluripotency-related genes as input for cell clustering⁵⁴. We successfully constructed the redifferentiation trajectory which corresponded to the biological developmental process. Although we tried to apply tools, such as monocle 2 and 3, for

358 constructing the developmental trajectory, we did not obtain reasonable results
359 (Supplementary Fig. 2a-e), indicating that suitable tools or methods need to be selected for
360 different datasets.

361

362 Along the developmental trajectory, we found that cells had relatively higher pluripotency at
363 the starting points compared with cells at a later developmental stage. Interestingly, we
364 identified the QC-like cells at the intermediate stage of redifferentiation, and the considerably
365 increased expression of well-known QC regulator, *WOX5*, in this cell cluster further
366 supported the QC-like cell identity. Consistently, QC-like cells, which have been suggested to
367 be primarily responsible for organ regeneration¹⁰, have also been identified during lateral root
368 formation²¹. QC cells are the core cells of Root apical meristem (RAM) and could generate
369 diverse root cell types⁵⁵. Correspondingly, we detected the pericycle-like cells appeared at the
370 later stage. In addition, cell cluster 1 also showed similarity to several other root cell types,
371 indicating that the redifferentiation directions might diverge at this point and further studies
372 are needed to uncover the complex developmental processes. Moreover, we found that the JA
373 and SA-related biological processes were enriched and several TFs were highly expressed in
374 the QC-like cells, which provided valuable insights for further research to reveal the roles of
375 QC-like cells during *de novo* shoot regeneration.

376

377 In addition to phytohormones, light is another important factor for regulating *de novo* shoot
378 regeneration. In this study, the adding of light treatment was an essential factor for promoting
379 shoot regeneration after callus was transferred to SIM. Here we found that *HY5*, a basic
380 leucine zipper (bZIP) transcription factor and a positive regulator of photomorphogenesis,
381 was highly expressed at the beginning of redifferentiation trajectories, which was consistent
382 with the finding that the loss of function in *HY5* leads to the defect of callus formation⁵⁶.
383 However, the expression of *HY5* was gradually decreased in the following stages, indicating
384 that *HY5* might play its role during the formation of callus. In addition to *HY5*, we also found
385 other light-related genes, including *miR163* and its target *PXMT1*, which might also be
386 involved in the *de novo* shoot regeneration process. Previously, it has been reported that light-
387 inducible *miR163* targets *PXMT1* transcripts to regulate *Arabidopsis* seed germination and

primary root elongation⁵⁷. In this study, when the callus was transferred to SIM and cultured on the light condition, *miR163* was expressed immediately, while its target *PXMT1* that highly expressed at the previous stage and could not be detected (Supplementary Fig. 5c). The inversely correlated expression pattern of these two genes indicated that the expression of *miR163* is induced by light and it further targets *PXMT1* transcripts to regulate the process of *de novo* shoot regeneration. Collectively, we speculate that light-inducible *miR163* and its targets *PXMT1* are involved in regulating the *de novo* shoot regeneration process, and further verifications are needed.

The reconstructed redifferentiation trajectories, the dynamic gene expression, and the complex regulatory networks at single-cell resolution provided great insights and valuable resources for future studies on *de novo* shoot regeneration. In this study, we only revealed the early stage of *de novo* shoot regeneration. To have an overall view of *de novo* shoot regeneration, single-cell transcriptome profiling of cells after day 10 is still needed. Considering the friable callus gradually changes into compact tissues, the collection of single cells by protoplasting is no longer infeasible. Therefore, using single-nucleus RNA-seq should be a better solution. Moreover, epigenetic regulations play important roles in regulating the regeneration process, and the single-cell sequencing assay for transposase-accessible chromatin (scATAC-seq) is the state-of-the-art technology for analyzing genome-wide chromatin accessibility at the single-cell level during *de novo* shoot regeneration.

Methods

Plant growth conditions and *de novo* shoot regeneration culturing

A. thaliana Col-0 was used as the wild-type. Seeds were sterilized with (0.05% (v/v) Tween-20 solution for 10 minutes) and subsequently sown on 1/2 Murashige and Skoog (MS) medium (Sigma M5519) at 22°C in long-day conditions (16 h light/8 h dark). Harvested root explants from 5-days (d)-old seedlings were cultured on callus-induction medium (CIM) with MS medium, 3% w/v sucrose, 0.3% w/v Phytigel, 0.22 µM N-(Phenylmethyl)-9H-purin-6-amine and 4.52 µM 2,4-Dichlorophenoxyacetic acid, buffered to PH 5.7, at 25°C in dark

417 conditions for callus induction, and sub-culture was performed every two weeks. Two months
418 later, calli were collected and transferred to shoot induction medium (SIM) with MS medium,
419 1% w/v sucrose, 0.05% w/v MES (Sigma, M3671), 0.3% w/v Phytigel, 0.1% v/v Gambor's
420 B5 vitamin, 9.12 μ M zeatin, 4.1 μ M biotin and 1.97 μ M 3-indolebutyric acid, buffered to PH
421 5.7, at 22°C in 16 h light/8 h dark conditions for shoot regeneration.

422

423 **Protoplasting**

424 Protoplast isolation was performed with a modified version of the polyethylene glycol (PEG)-
425 mediated method as described previously⁵⁸. Samples were collected from two-month-old
426 CIM-induced calli in 0, 2, 4, 6, and 10 days after being transferred into SIM medium. All the
427 calli were digested in the protoplast isolation solution containing 1% w/v Cellulase 'Onozuka'
428 R-10 (Yakult Pharmaceutical), 0.2% w/v macerozyme R-10 (Yakult Pharmaceutical), 0.1%
429 w/v Pectinase (Sigma, 17389), 20 mM MES (pH 5.7, Sigma, M3671), 10 mM CaCl₂ (Sigma,
430 C5670), 10 mM KCl (Sigma, P5405), 0.4 M D-mannitol (Sigma, M1902), and 0.1% BSA
431 (Sigma, V900933) for approximately 2 h on a shaker at 60 rpm in the dark respectively. To
432 remove enzyme solution and tissue debris, the digested tissue suspensions were filtered
433 through 40 μ m nylon mesh and washed by W5 buffer with the components of 154 mM NaCl
434 (Sigma, S5886), 5 mM KCl (Sigma, P5405), 125 mM CaCl₂, and 2 mM MES (pH 5.7). The
435 mixed solutions were gently centrifuged at 100 g for 5 min to collect Protoplasts, which
436 were then washed and resuspended with W5 buffer at a concentration of 2×10^6 cells/mL,
437 counted using a hemocytometer.

438

439 **scRNA-seq**

440 Single-cell RNA-seq libraries were prepared with a DNB C4 scRNA-seq kit following the
441 published protocol^{59,60}. In brief, 100 μ l W5 buffer containing protoplasts was supplemented
442 with 6% ficoll 400, 0.04% BSA, 2U/ μ l RNase inhibitor as single-cell suspension of every
443 sample, which was loaded into the chip with Barcoded mRNA capture beads and oil for
444 droplet generation. After incubating at room temperature for 20 minutes, collected droplets
445 were broken by the bead filter and the bead pellets were resuspended with 100 μ l RT mix in
446 the thermal cycle as follows: 42 °C for 90 minutes, 10 cycles of 50 °C for 2 minutes, 42 °C

for 2 minutes. The bead pellets were then resuspended in 200 µl of exonuclease mix for digestion and amplified by PCR as follows: 95 °C for 30 s, 15 cycles of 98 °C for 20 s, 58 °C for 20 s, 72 °C for 30 s, and finally 72 °C for 5 minutes. Amplified cDNA was purified and subsequently fragmented for indexed sequencing libraries according to the manufacturer's protocol of the DNB C4 scRNA-seq kit. All libraries were further prepared based on the BGISEQ-T1 sequencing platform. The DNA nanoballs (DNBs) were loaded into the patterned nanoarrays and sequenced on the BGISEQ-T1 sequencer using the following read length: 41 bp for read 1, 100 bp for read 2, and 10 bp for sample index.

Bulk RNA-seq data processing

Total RNA was extracted from two-month-old CIM-induced calli in 0, 2, 4, and 6 days after being transferred into SIM medium using miRNeasy Mini Kit (QIAGEN, 217004). High-quality RNA was used to prepare sequencing libraries with the MGIEasy RNA library preparation kit. All libraries were sequenced on the BGISEQ-T1 sequencer using the following read length: 100 bp for read 1, 100 bp for read 2, and 10 bp for sample index. Then, analysis of the bulk RNA-seq data was performed. First, the bulk RNA-seq data were filtered using SOAPnuke (v1.5.2) with parameters `-l 5 -q 0.2 -n 0.0561`. After removing low-quality bases RNAs, the clean data were mapped to the *Arabidopsis thaliana* genome (TAIR10) by using HISAT2⁶² with parameters `"-k 1 -p 4 -x --min-intronlen 40 --max-intronlen 5000 --no-unal --dta --un-gz --rna-strandness RF"`. Then the expression levels of each gene were calculated by the transcripts per kilobase of exon model per million mapped reads (TPM) by using StringTie⁶³ with parameters `"-t -C -e -B -A --rf"`. The final TPM matrix of all samples was used for subsequent analysis.

Single-cell RNA-seq data processing

First, single-cell data were mapped to *Arabidopsis thaliana* genome (TAIR10) by using STAR⁶⁴ with parameters `"--outStd SAM --outSAMunmapped Within --runThreadN 4 --alignIntronMin 40 --alignIntronMax 5000"`. Then, the gene expression level in each cell is calculated by the unique molecular identifiers (UMI) using PISA (<https://github.com/pisa-engine/pisa>).

477 In addition, we also downloaded the 10x single-cell RNA-seq data of *Arabidopsis thaliana*
478 root (GSM4466788 and GSM4466787 of GSE122687)²⁷ and shoot (CRS098237)²⁹ from GEO
479 and GSA, respectively. First, fastq-dump was used to convert SRA data of root into fastq
480 format. Then these 10X sequencing data were mapped to the *Arabidopsis thaliana* genome
481 (TAIR10) and generated the UMI matrix of each cell by using Cell Ranger (version 5.0.1)
482 count pipeline.

483

484 **Identification of genes related to pluripotency and differentiation**

485 Seurat's "findMarkers" function was used to identify the differentially expressed genes (DEGs)
486 between root and shoot and D0, respectively. These differential genes were defined as root-up,
487 D0d-up1, shoot-up, and D0d-up2, respectively. The overlapped genes of D0d-up1 and D0d-
488 up2 were defined as pre-pluripotency-related genes. And the shoot-up genes were defined as
489 pre-differentiation-related genes. At the same time, some key genes in the regeneration
490 process were collected, and these genes were defined as key genes. The Mfuzz⁶⁵ function was
491 also used to find gene sets that showed the same trend from Day 0 to Day 6/10 in bulk and
492 single-cell data from the three gene sets. On Day 0 to Day 6/10, the gene set with a decreasing
493 trend was defined as pluripotency-related genes, while the gene set with an increasing trend
494 was defined as differentiation-related genes. Finally, the four genes were combined to obtain
495 the genes related to callus development.

496

497 **Clustering**

498 The cells with less than 500 genes, more than 5,000 genes, and more than 10 mitochondria
499 were filtered out. Through cell quality control, 35,669 high-quality cells were finally obtained
500 for subsequent analysis. The genes related to callus development were selected for linear
501 dimensionality reduction using the function "RunPCA". The "FindNeighbors" and
502 "FindClusters" functions were used for cell clustering respectively, with parameters of 30 PC
503 and 0.2 resolution. Finally, the "RunUMAP" function was used for visualization.

504

505 **Gene signature scores**

506 First, each Gene in a given gene set was normalized by the "scale" function in all cells. Then,

the normalized scores of gene set in each cell were averaged, which was defined as Gene signature scores.

Differential expression analysis

The cell identity scores were calculated using Hypergeometric Distribution Enrichment Analysis (HDEA). Specifically, the root cell type-specific markers from the online database³⁸ and the leaf cell type-specific markers from Arabidopsis shoot apex²⁹ were used as reference (Supplementary Table 2). The cluster-specific markers were calculated using Seurat “FindAllMarkers” function log-fold change ≥ 0.25 , present in $\geq 25\%$ cells. The z-score of genes from specific GO terms was calculated.

GO enrichment analysis

GO enrichment analysis was performed using R package clusterProfiler⁶⁶ with TAIR10 annotation as the background. The smaller the P-value is, the more the GO term is significantly enriched. The GO terms enriched in the top 100 markers (using all markers if markers were less than 100) for each cluster were performed.

Correlation analysis

The “cor” function was used to perform correlation analysis on the whole transcriptome of each single cell sample. The distance used was Euclidian, and the cluster method was “kendall” from the normalized data for each sample calculated by Seurat.

Transcription factor (TF) regulatory network analysis

Firstly, genes with more than 33% gene expression in the marker list of each cluster were defined as candidate genes, and then the transcription factor list of each cluster was screened from these genes. The GENIE3⁶⁷ was used to predict the regulatory relationships of these transcription factors and candidate genes from the normalized data for each cell calculated by Seurat with parameters “nTrees = 1000, nCores = 8” and the random seed was 123. Then the top 10,000 regulatory relationships were screened for subsequent analysis. And the known regulatory relationships between transcription factors (TFs) and target genes of *Arabidopsis*

537 were downloaded from the Plant Cistrome Database⁵¹. Finally, the regulatory relationships
538 present in both the database and the top 10,000 regulatory relationships predicted by
539 GENIE3⁶⁷ were obtained.

540

541 **Trajectory analysis**

542 Trajectory analysis was performed using R packages Monocle2 and Monocle3³⁷, respectively.
543 Among them, the key genes used in Monocle2 were the top 770 highly variable genes. The
544 default parameters were used for analysis in Monocle2 and Monocle3.

545

546 **Statistical analysis**

547 Statistical analysis was performed using R (version 4.0.4). Graphs were generated using R
548 package (ggplot2, mfuzz, ComplexHeatmap, and cytoscape). The Wilcoxon rank-sum test
549 was used to calculate differential genes between different cell types or samples, and $p < 0.05$
550 was considered significant. Fisher's Exact Test was used for gene function annotation.

551

552 **Acknowledgments**

553 This research was supported by Guangdong Provincial Key Laboratory of Genome Read and
554 Write (No. 2017B030301011). We also sincerely thank the support provided by China
555 National Gene Bank (CNGB).

556

557 **Author contributions**

558 K.X. and H.-X.S. designed and supervised the study. K.X. and G.L., designed the experiment,
559 K.X., G.L., Z.C., P.Y., R.Y., and X.L. performed the protoplast isolation, library preparation,
560 and sequencing. H.-X.S., J.L., J.-M.L., and R.C. performed bioinformatics analysis. Z.C., P.Y.,
561 and Z.L. performed the tissue culture experiment. G.L., J.-M.L., and J.L. wrote the
562 manuscript. K.X., H.-X.S., and Y.G. participated in the manuscript editing and discussion. All
563 authors edited and approved the manuscript.

564

565 Competing interests

566 The authors declare that there is no conflict of interests.

567

568 Data availability

569 The sequencing data of *Arabidopsis thaliana* that support the findings of this study have been
570 deposited into CNGB Sequence Archive (CNSA)⁶⁸ of China National GeneBank DataBase
571 (CNGBdb)⁶⁹ with accession number CNP0002343.

572

573 Figure legends

574 **Fig. 1 | The developmental trajectory during the early stage of *de novo* shoot**
575 **regeneration. a**, Schematic representation of the scRNA-seq procedure. **b**, Volcano plots
576 showing the differentially expressed genes (DEGs) in the root (left) and shoot (right)
577 compared with SIM0. The red dots: the up-regulated genes in Day 0, the yellow dots: the up-
578 regulated genes in the shoot, and the blue dots: the up-regulated genes in the root. The gray
579 dots: genes with no significant difference. **c**, The overlap of the up-regulated DEGs in SIM0
580 compared with the root and shoot. **d**, The pre-pluripotency-related genes with decreased
581 expression patterns in our bulk RNA-seq and scRNA-seq data of samples in Day 0, 2, 4, 6,
582 and 10. **e**, The pre-differentiation-related genes with increased expression patterns in our bulk
583 RNA-seq and scRNA-seq data of samples in Day 0, 2, 4, 6, and 10. **f**, Clustering results of
584 35,669 cells. **g**, Signature scores of pluripotency-related genes at the single-cell level. **h**, The
585 development trajectory of cell clusters. SIM0 represents the cell population on Day 0. Root-
586 up, shoot-up, and SIM0-up represent up-regulated genes in the root, shoot, and SIM0,
587 respectively. B-p and Sc-p represent patterns in bulk RNA-seq and scRNA-seq, respectively.
588

589 **Fig. 2 | Cell type identification and characterization. a**, Expression of phase-specific genes
590 at single-cell level during *de novo* shoot regeneration, including *WOX5*, *SCR*, *PINI*, and
591 *CUC2*. **b**, Hypergeometric Distribution Enrichment Analysis (HDEA) of cell populations. **c**,
592 Profiles of cell populations in Day 0, 2, 4, 6, and 10. **d**, The proportion of cluster-based cells
593 in Day 0, 2, 4, 6, and 10. **e**, The expression patterns of the top 5 overlapped genes for each

594 cluster between cluster-specific markers and the reference cell type-specific marker genes.
 595 Percent Expressed means percent of gene-expressed cell number in all cells of specific
 596 clusters. The average expression means the average expression of genes at the single-cell
 597 level in specific clusters. One overlapped marker gene was detected. **f**, Hormone-related GO
 598 terms in specific cell clusters, including auxin biosynthetic process in clusters 0 and 5,
 599 response to jasmonic acid, and salicylic acid metabolic process in cluster 3 and ethylene
 600 metabolic process in cluster 1.

601

602 **Fig. 3 | Dynamic expression patterns of canonical genes during *de novo* shoot**
 603 **regeneration. a**, The heatmap of 75 canonical genes in *de novo* shoot organogenesis. Blue
 604 and red represent z score < 0 and > 0, respectively. **b**, Dynamic expression patterns of 75
 605 canonical genes in *de novo* shoot organogenesis along the development trajectory. **c**, Top 15
 606 GO terms in pattern 1. **d**, Dynamic expression of represented *AHK4*, *WOX4*, and *PIN1* genes
 607 in pattern 1. **e**, Top 15 GO terms in pattern 2. **f**, Dynamic expression of represented *ARF5*,
 608 *TIR1*, and *LBD16* genes in pattern 2. **g**, Top 15 GO terms in pattern 3. **h**, Dynamic expression
 609 of represented *PLT1/5* and *WOX5* genes in pattern 3. Relative Percent Expressed means the
 610 percentage of the number of gene expression cells in all cells of a specific cluster relative to
 611 the percentage of expression in cluster 5. The average expression means the average
 612 expression of genes at the single-cell level in specific clusters.

613

614 **Fig. 4 | Transcription factor (TF) regulatory network analysis at the early stage of *de***
 615 ***nov* shoot organogenesis. a**, Heatmap of upregulated marker TFs for each cluster. **b**, The
 616 regulatory networks of 31 core TFs. Marker TFs in clusters 5, 0, 4, 6, 3, 2, and 1 were
 617 represented by red, orange, yellow, canary yellow, grass green, green, and blue circles,
 618 respectively.

619

620 **Supplementary Information**

621 **Supplementary Fig. 1 | The data quality analysis of callus at the early stage of *de novo***
 622 **shoot regeneration. a**, The 2-month-old root-derived calli grew on SIM for 26 days. **b**,
 623 Correlation analysis of 15 scRNA-seq samples. **c**, Gene number, UMI, percentage of

mitochondrial and chloroplast genes at the single-cell level at five time points. **d**, Correlation analysis of gene expression between bulk RNA-seq and scRNA-seq data in Day 0, 2, 4, and 6.

Supplementary Fig. 2 | Developmental trajectory analysis using Monocle2 and Monocle3. **a-b**, Development trajectory using Monocle2. **c**, Developmental trajectory split by time points using Monocle2. **d-e**, Development trajectory using Monocle3.

Supplementary Fig. 3 | Selecting genes for reconstructing the developmental trajectory. **a**, Relatively increased expression pattern (p) of the pre-pluripotency-related genes in the bulk (b) RNA-seq and scRNA-seq (sc) data in Day 0, 2, 4, 6 and 10. **b**, Relatively decreased expression patterns (p) of the pre-pluripotency-related genes in the bulk (b) RNA-seq and scRNA-seq (sc) data in Day 0, 2, 4, 6 and 10. **c**, Genes with both decreased expression patterns in both bulk RNA-seq and scRNA-seq were designated as pluripotency-related genes, while genes with both increased expression pattern in both bulk RNA-seq and scRNA-seq were designated as differentiation-related genes. **d**, Regeneration-related genes with decreased expression patterns in both bulk RNA-seq and scRNA-seq were selected. **e**, Regeneration-related genes with relatively increased expression patterns in both bulk RNA-seq and scRNA-seq were chosen.

Supplementary Fig. 4 | Callus cell type characterization. **a**, Expression of phase-specific genes at single-cell level during *de novo* shoot regeneration, including *PLT1*, *PLT2*, and *ESR1* marker genes. **b**, Heatmap of top 100 GO terms in 7 cell clusters. **c**, The specific hormone-related GO term in cluster 0 and 5.

Supplementary Fig. 5 | Dynamic expression patterns of specific genes. **a**, Dynamic expression of *WOX13*, *YUC4*, and *Pin2* in pattern 2 of Fig. 3b. **b**, Dynamic expression pattern of *ARR5*. **c**, Dynamic expression pattern of *PXMT1* and *ath-MIR163*. Relative Percent Expressed means the percentage of the number of gene expression cells in all cells of a specific cluster relative to the percentage of expression in cluster 5. The average expression means the average expression of genes at the single-cell level in specific clusters.

654 **References**

655

- 656 1 Radhakrishnan, D. *et al.* Shoot regeneration: a journey from acquisition of
657 competence to completion. *Curr Opin Plant Biol* **41**, 23-31,
658 doi:10.1016/j.pbi.2017.08.001 (2018).
- 659 2 Efferth, T. Biotechnology applications of plant callus cultures. *Engineering* **5**, 50-59,
660 doi:10.1016/j.eng.2018.11.006 (2019).
- 661 3 Maher, M. F. *et al.* Plant gene editing through de novo induction of meristems. *Nat*
662 *Biotechnol* **38**, 84-89, doi:10.1038/s41587-019-0337-2 (2020).
- 663 4 Debernardi, J. M. *et al.* A GRF–GIF chimeric protein improves the regeneration
664 efficiency of transgenic plants. *Nature Biotechnology*, doi:10.1038/s41587-020-0703-
665 0 (2020).
- 666 5 Shin, J., Bae, S. & Seo, P. J. De novo shoot organogenesis during plant regeneration.
667 *J Exp Bot* **71**, 63-72, doi:10.1093/jxb/erz395 (2020).
- 668 6 Zhang, T.-Q. *et al.* A two-step model for de novo activation of WUSCHEL during plant
669 shoot regeneration. *The Plant Cell* **29**, 1073-1087, doi:10.1105/tpc.16.00863 (2017).
- 670 7 Meng, W. J. *et al.* Type-B ARABIDOPSIS RESPONSE REGULATORS specify the
671 shoot stem cell niche by dual regulation of WUSCHEL. *The Plant Cell* **29**, 1357-1372,
672 doi:10.1105/tpc.16.00640 (2017).
- 673 8 Liu, J. *et al.* The WOX11–LBD16 pathway promotes pluripotency acquisition in callus
674 cells during de novo shoot regeneration in tissue culture. *Plant and Cell Physiology*
675 **59**, 739-748, doi:10.1093/pcp/pcy010 (2018).

- 676 9 Kareem, A. *et al.* PLETHORA genes control regeneration by a two-step mechanism.
677 *Current Biology* **25**, 1017-1030, doi:10.1016/j.cub.2015.02.022 (2015).
- 678 10 Zhai, N. & Xu, L. Pluripotency acquisition in the middle cell layer of callus is required
679 for organ regeneration. *Nature Plants* **7**, 1453-1460, doi:10.1038/s41477-021-01015-
680 8 (2021).
- 681 11 Aida, M., Ishida, T. & Tasaka, M. Shoot apical meristem and cotyledon formation
682 during Arabidopsis embryogenesis: interaction among the CUP-SHAPED
683 COTYLEDON and SHOOT MERISTEMLESS genes. *Development* **126**, 1563-1570,
684 doi:10.1242/dev.126.8.1563 (1999).
- 685 12 Matsuo, N., Makino, M. & Banno, H. Arabidopsis ENHANCER OF SHOOT
686 REGENERATION (ESR) 1 and ESR2 regulate in vitro shoot regeneration and their
687 expressions are differentially regulated. *Plant science* **181**, 39-46,
688 doi:10.1016/j.plantsci.2011.03.007 (2011).
- 689 13 Banno, H., Ikeda, Y., Niu, Q.-W. & Chua, N.-H. Overexpression of Arabidopsis ESR1
690 induces initiation of shoot regeneration. *The Plant Cell* **13**, 2609-2618,
691 doi:10.1105/tpc.010234 (2001).
- 692 14 Cheng, Z. J. *et al.* Pattern of Auxin and Cytokinin Responses for Shoot Meristem
693 Induction Results from the Regulation of Cytokinin Biosynthesis by AUXIN
694 RESPONSE FACTOR3 *Plant Physiology* **161**, 240-251, doi:10.1104/pp.112.203166
695 (2013).
- 696 15 Ikeuchi, M., Ogawa, Y., Iwase, A. & Sugimoto, K. Plant regeneration: cellular origins
697 and molecular mechanisms. *Development* **143**, 1442, doi:10.1242/dev.134668 (2016).

698 16 Chupeau, M. C. *et al.* Characterization of the early events leading to totipotency in an
699 Arabidopsis protoplast liquid culture by temporal transcript profiling. *Plant Cell* **25**,
700 2444-2463, doi:10.1105/tpc.113.109538 (2013).

701 17 Mironova, V. & Xu, J. A single-cell view of tissue regeneration in plants. *Current*
702 *Opinion in Plant Biology* **52**, 149-154, doi:https://doi.org/10.1016/j.pbi.2019.09.003
703 (2019).

704 18 Zheng, H.-x., Wu, F.-h., Li, S.-m., Zhang, X. S. & Sui, N. Single-cell profiling lights
705 different cell trajectories in plants. *aBIOTECH* **2**, 64-78, doi:10.1007/s42994-021-
706 00040-7 (2021).

707 19 Efroni, I. *et al.* Root Regeneration Triggers an Embryo-like Sequence Guided by
708 Hormonal Interactions. *Cell* **165**, 1721-1733,
709 doi:https://doi.org/10.1016/j.cell.2016.04.046 (2016).

710 20 Zhang, T.-Q., Chen, Y., Liu, Y., Lin, W.-H. & Wang, J.-W. Single-cell transcriptome
711 atlas and chromatin accessibility landscape reveal differentiation trajectories in the
712 rice root. *Nature Communications* **12**, 2053, doi:10.1038/s41467-021-22352-4 (2021).

713 21 Serrano-Ron, L. *et al.* RECONSTRUCTION OF LATERAL ROOT FORMATION
714 THROUGH SINGLE-CELL RNA-SEQ REVEALS ORDER OF TISSUE INITIATION.
715 *Molecular Plant*, doi:10.1016/j.molp.2021.05.028 (2021).

716 22 Gala, H. P. *et al.* A single cell view of the transcriptome during lateral root initiation in
717 Arabidopsis thaliana. *The Plant Cell*, doi:10.1093/plcell/koab101 (2021).

718 23 Denyer, T. *et al.* Spatiotemporal developmental trajectories in the Arabidopsis root
719 revealed using high-throughput single-cell RNA sequencing. *Developmental cell* **48**,

720 840-852. e845 (2019).

721 24 Jean-Baptiste, K. *et al.* Dynamics of gene expression in single root cells of
722 *Arabidopsis thaliana*. *The Plant Cell* **31**, 993-1011 (2019).

723 25 Liu, Q. *et al.* Transcriptional landscape of rice roots at the single-cell resolution.
724 *Molecular Plant* **14**, 384-394 (2021).

725 26 Ryu, K. H., Huang, L., Kang, H. M. & Schiefelbein, J. Single-cell RNA sequencing
726 resolves molecular relationships among individual plant cells. *Plant physiology* **179**,
727 1444-1456 (2019).

728 27 Shulse, C. N. *et al.* High-throughput single-cell transcriptome profiling of plant cell
729 types. *Cell reports* **27**, 2241-2247. e2244 (2019).

730 28 Zhang, T.-Q., Xu, Z.-G., Shang, G.-D. & Wang, J.-W. A single-cell RNA sequencing
731 profiles the developmental landscape of *Arabidopsis* root. *Molecular plant* **12**, 648-
732 660 (2019).

733 29 Zhang, T.-Q., Chen, Y. & Wang, J.-W. A single-cell analysis of the *Arabidopsis*
734 vegetative shoot apex. *Developmental Cell* **56**, 1056-1074.e1058,
735 doi:10.1016/j.devcel.2021.02.021 (2021).

736 30 Satterlee, J. W., Strable, J. & Scanlon, M. J. Plant stem-cell organization and
737 differentiation at single-cell resolution. *Proceedings of the National Academy of*
738 *Sciences* **117**, 33689-33699, doi:10.1073/pnas.2018788117 (2020).

739 31 Bezruczyk, M. *et al.* Evidence for phloem loading via the abaxial bundle sheath cells
740 in maize leaves. *The Plant Cell* **33**, 531-547 (2021).

741 32 Kim, J.-Y. *et al.* Distinct identities of leaf phloem cells revealed by single cell

742 transcriptomics. *The Plant Cell* **33**, 511-530 (2021).

743 33 Lopez-Anido, C. B. *et al.* Single-cell resolution of lineage trajectories in the
744 Arabidopsis stomatal lineage and developing leaf. *Developmental Cell* **56**, 1043-
745 1055.e1044, doi:10.1016/j.devcel.2021.03.014 (2021).

746 34 Liu, Z. *et al.* Global Dynamic Molecular Profiles of Stomatal Lineage Cell
747 Development by Single-Cell RNA Sequencing. *Molecular Plant* **13**, 1178-1193,
748 doi:10.1016/j.molp.2020.06.010 (2020).

749 35 Yu, W., Qh, A., Ke, L. A. & Wqa, B. Single-cell transcriptome atlas of the leaf and root
750 of rice seedlings. *Journal of Genetics and Genomics*, doi:10.1016/j.jgg.2021.06.001
751 (2021).

752 36 Wendrich, J. R. *et al.* Vascular transcription factors guide plant epidermal responses
753 to limiting phosphate conditions. *Science* **370**, eaay4970,
754 doi:10.1126/science.aay4970 (2020).

755 37 Qiu, X. *et al.* Single-cell mRNA quantification and differential analysis with Census.
756 *Nature methods* **14**, 309-315, doi:10.1038/nmeth.4150 (2017).

757 38 Chen, H. *et al.* PlantscRNAdb: A database for plant single-cell RNA analysis.
758 *Molecular Plant*, doi:10.1016/j.molp.2021.05.002 (2021).

759 39 Sugimoto, K., Jiao, Y. & Meyerowitz, E. M. Arabidopsis regeneration from multiple
760 tissues occurs via a root development pathway. *Developmental cell* **18**, 463-471,
761 doi:10.1016/j.devcel.2010.02.004 (2010).

762 40 Lavenus, J. *et al.* Lateral root development in Arabidopsis: fifty shades of auxin.
763 *Trends Plant Sci* **18**, 450-458, doi:10.1016/j.tplants.2013.04.006 (2013).

- 764 41 Skoog, F. & Miller, C. O. Chemical regulation of growth and organ formation in plant
765 tissues cultured in vitro. *Cheminform* **27**, 118, doi:DOI: 10.1146/annurev-arplant-
766 050718-100434 (1996).
- 767 42 Achard, P. *et al.* The plant stress hormone ethylene controls floral transition via
768 DELLA-dependent regulation of floral meristem-identity genes. *Proceedings of the*
769 *National Academy of Sciences* **104**, 6484-6489, doi:10.1073/pnas.0610717104
770 (2007).
- 771 43 Su, Y. H., Tang, L. P., Zhao, X. Y. & Zhang, X. S. Plant cell totipotency: Insights into
772 cellular reprogramming. *Journal of Integrative Plant Biology* **63**, 228-243,
773 doi:<https://doi.org/10.1111/jipb.12972> (2021).
- 774 44 Kubiasová, K. *et al.* Cytokinin fluoroprobe reveals multiple sites of cytokinin
775 perception at plasma membrane and endoplasmic reticulum. *Nature communications*
776 **11**, 1-11, doi:10.1038/s41467-020-17949-0 (2020).
- 777 45 Hao, Z. *et al.* Conserved, divergent and heterochronic gene expression during
778 Brachypodium and Arabidopsis embryo development. *Plant Reproduction*, 1-18,
779 doi:10.1007/s00497-021-00413-4 (2021).
- 780 46 Smit, M. E. *et al.* Specification and regulation of vascular tissue identity in the
781 Arabidopsis embryo. *Development* **147**, dev186130, doi:10.1242/dev.186130 (2020).
- 782 47 Goh, T. *et al.* Lateral root initiation requires the sequential induction of transcription
783 factors LBD16 and PUCHI in Arabidopsis thaliana. *New Phytologist* **224**, 749-760,
784 doi:10.1242/dev.186130 (2019).
- 785 48 Momoko, I. *et al.* Wound-inducible WUSCHEL-RELATED HOMEBOX 13 is required

786 for callus growth and organ reconnection. *Plant Physiology*,
787 doi:10.1093/plphys/kiab510 (2021).

788 49 Wang, Z. *et al.* Salicylic acid promotes quiescent center cell division through ROS
789 accumulation and down-regulation of PLT1, PLT2, and WOX5. *J Integr Plant Biol* **63**,
790 583-596, doi:10.1111/jipb.13020 (2021).

791 50 Wein, A., Le Gac, A.-L. & Laux, T. Stem cell ageing of the root apical meristem of
792 *Arabidopsis thaliana*. *Mechanisms of ageing and development* **190**, 111313,
793 doi:10.1016/j.mad.2020.111313 (2020).

794 51 Niwa, H. The principles that govern transcription factor network functions in stem cells.
795 *Development* **145**, dev157420, doi:10.1242/dev.157420 (2018).

796 52 Yang, S., Poretska, O. & Sieberer, T. ALTERED MERISTEM PROGRAM1 restricts
797 shoot meristem proliferation and regeneration by limiting HD-ZIP III-mediated
798 expression of RAP2. 6L. *Plant physiology* **177**, 1580-1594, doi:10.1104/pp.18.00252
799 (2018).

800 53 Che, P., Lall, S., Nettleton, D. & Howell, S. H. Gene expression programs during
801 shoot, root, and callus development in *Arabidopsis* tissue culture. *Plant physiology*
802 **141**, 620-637, doi:10.1104/pp.106.081240 (2006).

803 54 Schiebinger, G. *et al.* Optimal-transport analysis of single-cell gene expression
804 identifies developmental trajectories in reprogramming. *Cell* **176**, 928-943. e922,
805 doi:10.1016/j.cell.2019.01.006 (2019).

806 55 Drisch, R. C. & Stahl, Y. Function and regulation of transcription factors involved in
807 root apical meristem and stem cell maintenance. *Frontiers in Plant Science* **6**, 505,

808 doi:10.3389/fpls.2015.00505 (2015).

809 56 Cluis, C. P., Mouchel, C. F. & Hardtke, C. S. The Arabidopsis transcription factor HY5
810 integrates light and hormone signaling pathways. *The Plant Journal* **38**, 332-347,
811 doi:10.1111/j.1365-313X.2004.02052.x (2004).

812 57 Chung, P. J. *et al.* Light-inducible miR163 targets PXMT1 transcripts to promote seed
813 germination and primary root elongation in Arabidopsis. *Plant physiology* **170**, 1772-
814 1782, doi:10.1104/pp.15.01188 (2016).

815 58 Yoo, S. D., Cho, Y. H. & Sheen, J. Arabidopsis mesophyll protoplasts: a versatile cell
816 system for transient gene expression analysis. *Nature protocols* **2**, 1565-1572,
817 doi:10.1038/nprot.2007.199 (2007).

818 59 Zhu, L. *et al.* Single-cell sequencing of peripheral mononuclear cells reveals distinct
819 immune response landscapes of COVID-19 and influenza patients. *Immunity* **53**, 685-
820 696. e683, doi:10.1016/j.immuni.2020.07.009 (2020).

821 60 Ding, R. *et al.* Single-cell transcriptome analysis of the heterogeneous effects of
822 differential expression of tumor PD-L1 on responding TCR-T cells. *Theranostics* **11**,
823 4957, doi:10.7150/thno.55075 (2021).

824 61 Wang, Z., Gerstein, M. & Snyder, M. RNA-Seq: a revolutionary tool for
825 transcriptomics. *Nature reviews genetics* **10**, 57-63, doi:10.1038/nrg2484 (2009).

826 62 Kim, D., Langmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner with low
827 memory requirements. *Nature methods* **12**, 357-360, doi:10.1038/nmeth.3317 (2015).

828 63 Pertea, M. *et al.* StringTie enables improved reconstruction of a transcriptome from
829 RNA-seq reads. *Nature Biotechnology* **33**, 290-295, doi:10.1038/nbt.3122 (2015).

830 64 Magoč, T. & Salzberg, S. L. FLASH: fast length adjustment of short reads to improve
831 genome assemblies. *Bioinformatics* **27**, 2957-2963, doi:10.1093/bioinformatics/btr507
832 (2011).

833 65 Futschik, M. E. & Carlisle, B. Noise-robust soft clustering of gene expression time-
834 course data. *Journal of bioinformatics and computational biology* **3**, 965-988,
835 doi:0.1142/S0219720005001375 (2005).

836 66 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and
837 dispersion for RNA-seq data with DESeq2. *Genome biology* **15**, 1-21,
838 doi:0.1186/s13059-014-0550-8 (2014).

839 67 Huynh-Thu, V. A., Irrthum, A., Wehenkel, L. & Geurts, P. Inferring regulatory
840 networks from expression data using tree-based methods. *PloS one* **5**, e12776,
841 doi:10.1371/journal.pone.0012776 (2010).

842 68 Guo, X. *et al.* CNSA: a data repository for archiving omics data. *Database* **2020**,
843 doi:10.1093/database/baaa055 (2020).

844 69 Chen, F. Z. *et al.* CNGBdb: China National GeneBank DataBase. *Yi chuan =*
845 *Hereditas* **42**, 799-809, doi:10.16288/j.ycz.20-080 (2020).

846

Figure 1

bioRxiv preprint doi: <https://doi.org/10.1101/2022.01.01.474510>; this version posted January 2, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

a

b

c

d

e

f

g

h

Figure 2

bioRxiv preprint doi: <https://doi.org/10.1101/2022.01.01.474510>; this version posted January 2, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 3

a

b

c

d

e

f

g

h

Figure 4

a

b