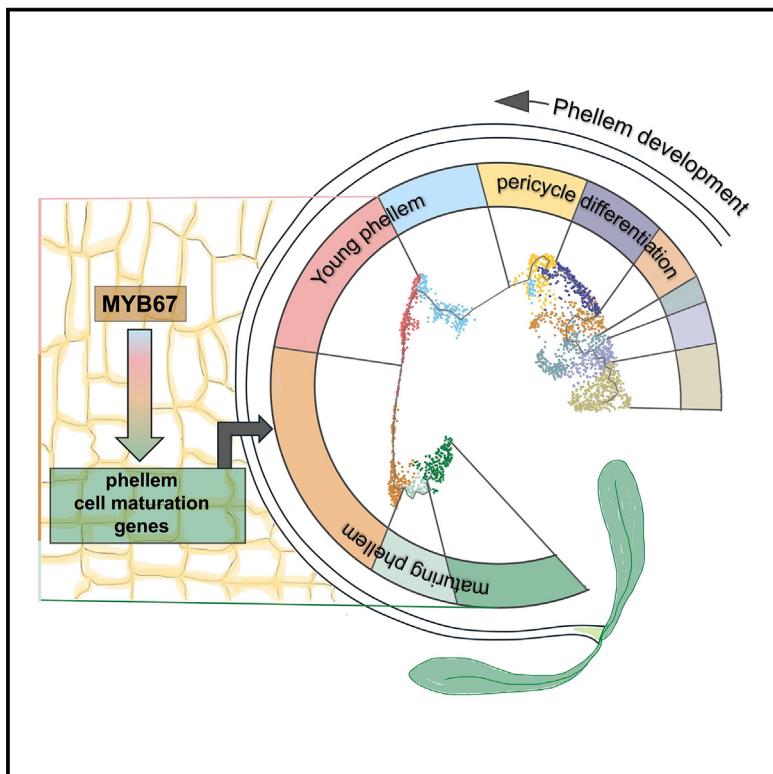


Developmental Cell

A single-nuclei transcriptome census of the *Arabidopsis* maturing root identifies that *MYB67* controls phellem cell maturation

Graphical abstract



Authors

Charlotte N. Miller,
Sean Jarrell-Hurtado,
Manisha V. Haag, ..., Todd P. Michael,
Julie A. Law, Wolfgang Busch

Correspondence

wbusch@salk.edu

In brief

Using small nuclear RNA sequencing (snRNA-seq), Miller et al. produced a gene expression census of the maturing *Arabidopsis* root and, by exploring the development of the suberizing phellem, identify that *MYB67* controls phellem cell maturation. This census may serve as resource for engineering future plants with improved stress resilience and carbon capture capacity.

Highlights

- scRNA-seq produced a comprehensive expression census of the maturing *Arabidopsis* root
- Distinct phellem developmental stages are captured across multiple nuclei populations
- Dynamic shifts in gene expression and biological processes drive phellem development
- *MYB67*, identified in the maturing root census, controls phellem cell maturation



Resource

A single-nuclei transcriptome census of the *Arabidopsis* maturing root identifies that *MYB67* controls phellem cell maturation

Charlotte N. Miller,¹ Sean Jarrell-Hurtado,¹ Manisha V. Haag,¹ Y. Sara Ye,¹ Mathew Simenc,¹ Paloma Alvarez-Maldonado,¹ Sara Behnami,¹ Ling Zhang,¹ Joseph Swift,¹ Ashot Papikian,¹ Jingting Yu,² Kelly Colt,¹ Joseph R. Ecker,^{1,3,4} Todd P. Michael,¹ Julie A. Law,^{1,5} and Wolfgang Busch^{1,6,*}

¹Plant Molecular and Cellular Biology Laboratory, The Salk Institute for Biological Studies, La Jolla, CA 92037, USA

²Integrative Genomics and Bioinformatics Core, The Salk Institute for Biological Studies, La Jolla, CA 92037, USA

³Genomic Analysis Laboratory, The Salk Institute for Biological Studies, La Jolla, CA 92037, USA

⁴Howard Hughes Medical Institute, The Salk Institute for Biological Studies, La Jolla, CA 92037, USA

⁵Division of Biological Sciences, University of California, San Diego, La Jolla, CA 92093, USA

⁶Lead contact

*Correspondence: wbusch@salk.edu

<https://doi.org/10.1016/j.devcel.2024.12.025>

SUMMARY

The periderm provides a protective barrier in many seed plant species. The development of the suberized phellem, which forms the outermost layer of this important tissue, has become a trait of interest for enhancing both plant resilience to stresses and plant-mediated CO₂ sequestration in soils. Despite its importance, very few genes driving phellem development are known. Employing single-nuclei sequencing, we have generated an expression census capturing the complete developmental progression of *Arabidopsis* root phellem cells, from their progenitor cell type, the pericycle, through to their maturation. With this, we identify a whole suite of genes underlying this process, including *MYB67*, which we show has a role in phellem cell maturation. Our expression census and functional discoveries represent a resource, expanding our comprehension of secondary growth in plants. These data can be used to fuel discoveries and engineering efforts relevant to plant resilience and climate change.

INTRODUCTION

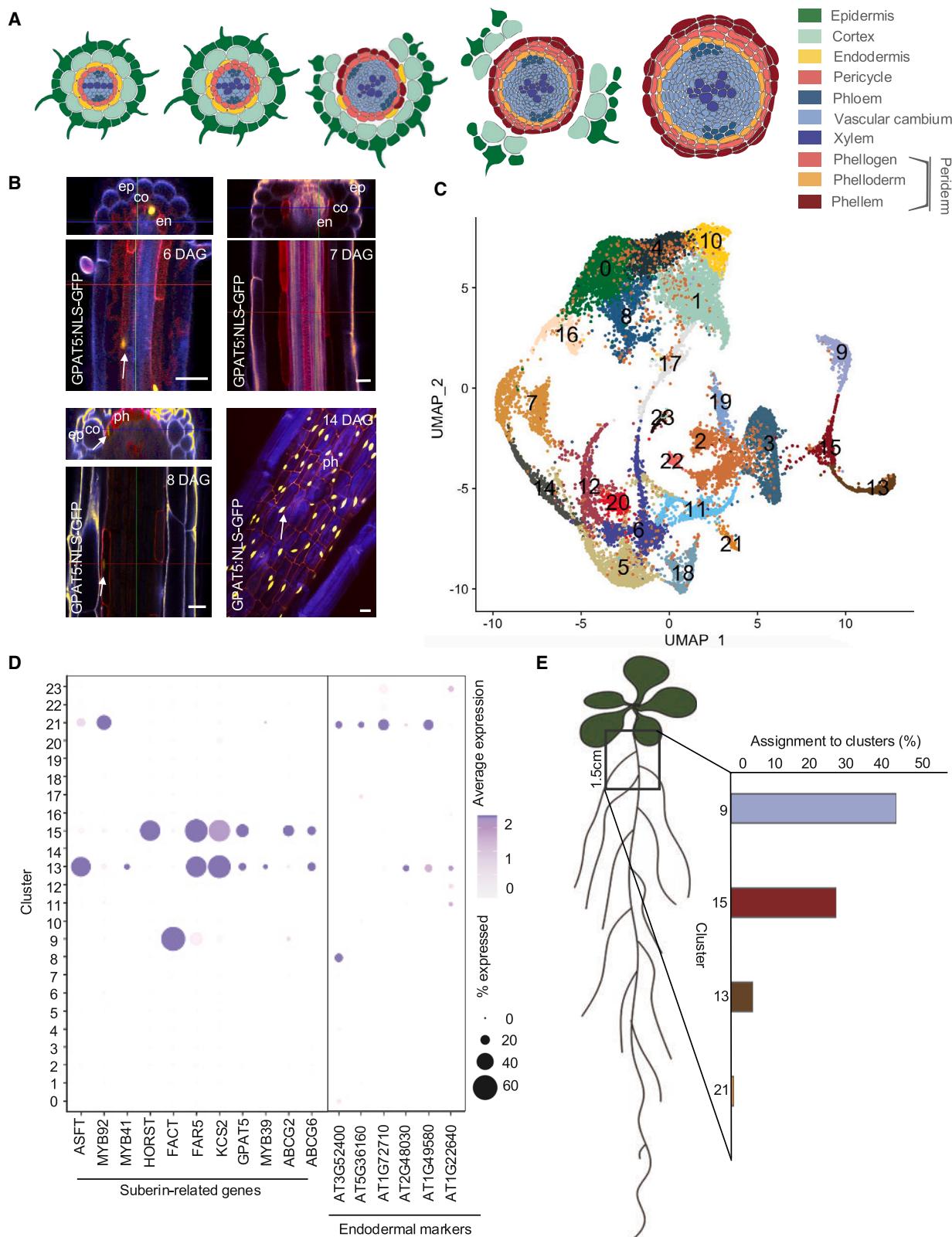
Plants live in a constantly changing and often unpredictable environment but lack motility. To survive, plants have evolved ways to dynamically adjust their physiology in response to their environment. Key for this is the establishment and maintenance of meristematic cell populations that can develop into differentiated cells when needed. In the young root, the primary root meristem actively divides to increase root length, giving rise to several cell types. Plant roots also have secondary meristems that allow for radial thickening. An example is the dividing phellogen, which differentiates to form two cell types, the phelloidem and the phellem. Together, the phellogen, phellem, and phelloidem make up the periderm. The periderm provides a kind of armor to the plant body, which is, at least in part, due to the polymers deposited in the walls of phellem cells that are made up of a highly recalcitrant, carbon-rich suberin complex.¹ In addition to regulating water movement,^{2,3} suberin deposition has been associated with tolerance to salt stress,³ cadmium toxicity,^{4,5} and ammonium stress.⁵ Suberin also plays an important role in defense against pathogens. For example, in soybean, suberin deposition increases tolerance to *Phytop-*

thora sojae, which causes root rot.⁶ Furthermore, in recent years, we have seen interest in making use of suberin as an approach for increasing carbon sequestration capacity in plant roots to help mitigate climate change.^{7,8} Understanding the process of suberization, and the development of the suberizing phellem, could thus have important implications in crop improvement and provide tools that could help us mitigate the ongoing climate crisis.

Although the periderm has been most widely studied in Cork oak (*Quercus suber*), recent research has explored this tissue in the model plant *Arabidopsis thaliana* (*Arabidopsis*) where Wunderling et al. observed periderm formation in the hypocotyl and root system. In the *Arabidopsis* root, this process begins when pericycle cells^{9,10} in the mature root divide to form the secondary meristem, the phellogen. As these cells proliferate, nearby endodermal cells undergo programmed cell death (PCD), and the phellogen forms two cell layers: the phelloidem and suberized phellem. With maturation, the outer cortex and epidermis abscise, making the periderm the outermost root layer.¹⁰ This process is illustrated in Figure 1A.

Alongside histological insights, progress has been made in understanding the genes driving periderm development. In





(legend on next page)

Arabidopsis, for example, *WUSCHEL-RELATED HOMEOBOX 4* (*WOX4*) and *BREVIPEDICELLUS* (*BP*) control phellogen activity in an auxin-dependent manner.¹¹ Numerous genes essential for suberin synthesis in phellem cells have also been identified. These include genes for monomer precursor synthesis (such as *LONG-CHAIN ACYL-COA SYNTHETASE 2* (*LACS2*)¹²), glycerol and ferric acid esterification (including various members of the *GLYCEROL-3-PHOSPHATE* (*GPAT*)¹³ family and *ALIPHATIC SUBERIN FERULOYL-TRANSFERASE* (*ASFT*)¹⁴, respectively), monomer transport (including *ABCG2*, *ABCG6*, and *ABCG20*¹⁵), and the building of the suberin polymer at the wall (requiring the polymerization action of *GDSL LIPASE-LIKE PROTEIN 22* (*GELP22*), *GDSL LIPASE-LIKE PROTEIN 38* (*GELP38*), and their homologs¹⁶).

Insights into *Arabidopsis* root phellem development were gained through a translating ribosome affinity purification (TRAP)-seq study that isolated and sequenced *GLYCEROL-3-PHOSPHATE sn-2-ACYLTRANSFERASE 5* (*GPAT5*) -expressing cells in the upper root, identifying genes involved in lipid storage, suberin biosynthesis, transcriptional regulation, and cell wall biogenesis.¹⁷ However, key phellem-related transcriptional events might occur in cells without *GPAT5* expression, and bulk analysis could obscure specific developmental events in distinct *GPAT5*-expressing cell populations. Further expression studies are needed for a comprehensive understanding of phellem development.

In recent years, we have seen the first use of cutting-edge, single-cell sequencing technologies in plants, offering unprecedented transcriptional resolution to explore individual cell types throughout development. This method identifies unique cell populations and subpopulations that bulk analyses may miss. We have seen several studies successfully employing this technology to further our understanding of root development, including a comprehensive census of the *Arabidopsis* root tip by Shahan et al.,¹⁸ as well as investigations into lateral root development¹⁹ and root response to phosphate deficiency.²⁰ However, these previous studies were restricted to time points that precede periderm development, which occurs only in more mature roots. Applying this technology to older roots presents an exciting opportunity to investigate the periderm at unparalleled resolution.

Here, we employ single-nuclei sequencing to identify the transcriptional landscape of developing phellem cells in the maturing *Arabidopsis* root. With this, we describe multiple, distinct developmental stages of phellem cells as they emerge from the pericycle, undergo suberization, and enter a phase of maturation. With these data, we identify key genes driving these processes, including *MYB DOMAIN PROTEIN 67* (*MYB67*), which controls

phellem cell maturation. These findings provide a valuable resource for understanding periderm development and offer insights that could support engineering plants with enhanced resilience to stress and improved soil carbon sequestration.

RESULTS

Single-nuclei sequencing of the maturing root identifies a putative phellem developmental trajectory

Previous studies indicate that periderm development begins around 6 days after germination (DAG), with phellem cells visible by 8 DAG in maturing *Arabidopsis* roots.¹⁷ To verify this developmental window under our growth conditions, we conducted a confocal time course using a *GPAT5* reporter^{10,21} and Nile Red staining to track suberization (Figure 1B). At 6 DAG, *GPAT5* and Nile red signals were detectable within endodermal cells only. By 7 DAG, these signals were seen to dissipate in the oldest endodermal cells found just below the hypocotyl. Shortly after this, at 8 DAG, we could see both Nile red and *GPAT5* signal in the first phellem cells forming in this upper root region. These signals were then seen to spread rapidly along the root thereafter and, by 14 DAG, phellem cells covered roughly one-third of the primary root. This confirms 6–14 DAG as a key period for studying the initiation and progression of periderm development in the *Arabidopsis* root.

Using this developmental window, we employed the 10× Genomics platform for single-nuclei sequencing, yielding 19,755 high-quality nuclei across the time course and identifying 24 clusters representing the maturing *Arabidopsis* root (Figure 1C; see Data S1A). With cell-type-specific markers from Shahan et al.,¹⁸ we identified clusters for 8 major cell types (trichoblast, atrichoblast, cortex, endodermis, pericycle, xylem, phloem, and quiescent center (QC)) across various developmental stages, accounting for 21 of the 24 clusters (Data S1B). Clusters 2, 9, and 15 lacked Shahan marker gene enrichment. To locate nuclei related to phellem development, we next analyzed the expression of suberization-related genes (Figure 1D), identifying clusters 9, 13, 15, and 21. Clusters 13 and 21 were also enriched in endodermal marker genes, indicating endodermal-related suberin signals. We hypothesized that the remaining suberin-gene-enriched clusters, 9 and 15, may relate to the developing phellem. The lack of Shahan et al.¹⁸ marker gene enrichment in these two clusters may in itself indicate that these clusters relate to phellem, a cell type not expected to be captured in a root tip census.

Unlike the endodermis, which develops and undergoes suberization relatively early in the *Arabidopsis* root, the phellem

Figure 1. Single-nuclei sequencing of mature *Arabidopsis* root identifies potential phellem nuclei populations

- (A) Periderm development begins in the pericycle, forming phellogen cells, which differentiate into phelloidem and phellem.
- (B) Using a *GPAT5* reporter, we tracked periderm development across roots at 6–14 DAG. *GPAT5* signal is indicated with white arrows. The images in the horizontal rectangle boxes show the cross-section, X section. The red lines in the vertical rectangle image boxes indicates the position of these digital cross-sections. Endodermal and phellem cells are stained with Nile red; Calcofluor white is used as a counterstain. Phellem (ph); cortex (co); epidermis (ep); endodermis (en). Scale bar, 20 μm.
- (C and D) Single-nuclei sequencing over this period yielded a UMAP with 24 clusters (C), with suberin-related gene expression in four clusters. Clusters 21 and 13 were enriched in endodermal markers, whereas clusters 9 and 15 represented putative phellem nuclei. The average gene expression level for a given gene is indicated in the dot plot by the dot color intensity. The size of each dot provides information regarding the proportion (%) of nuclei in a given cluster that showed expression of that gene (D).
- (E) An additional single-nuclei experiment on the upper 1.5 cm of the root at 14 DAG further supported clusters 9 and 15 as phellem-related, with most suberin-enriched nuclei from this region clustering here (E). The schematic in (A) shows periderm development as illustrated by Wunderling et al.¹⁰
- See also Data S1.

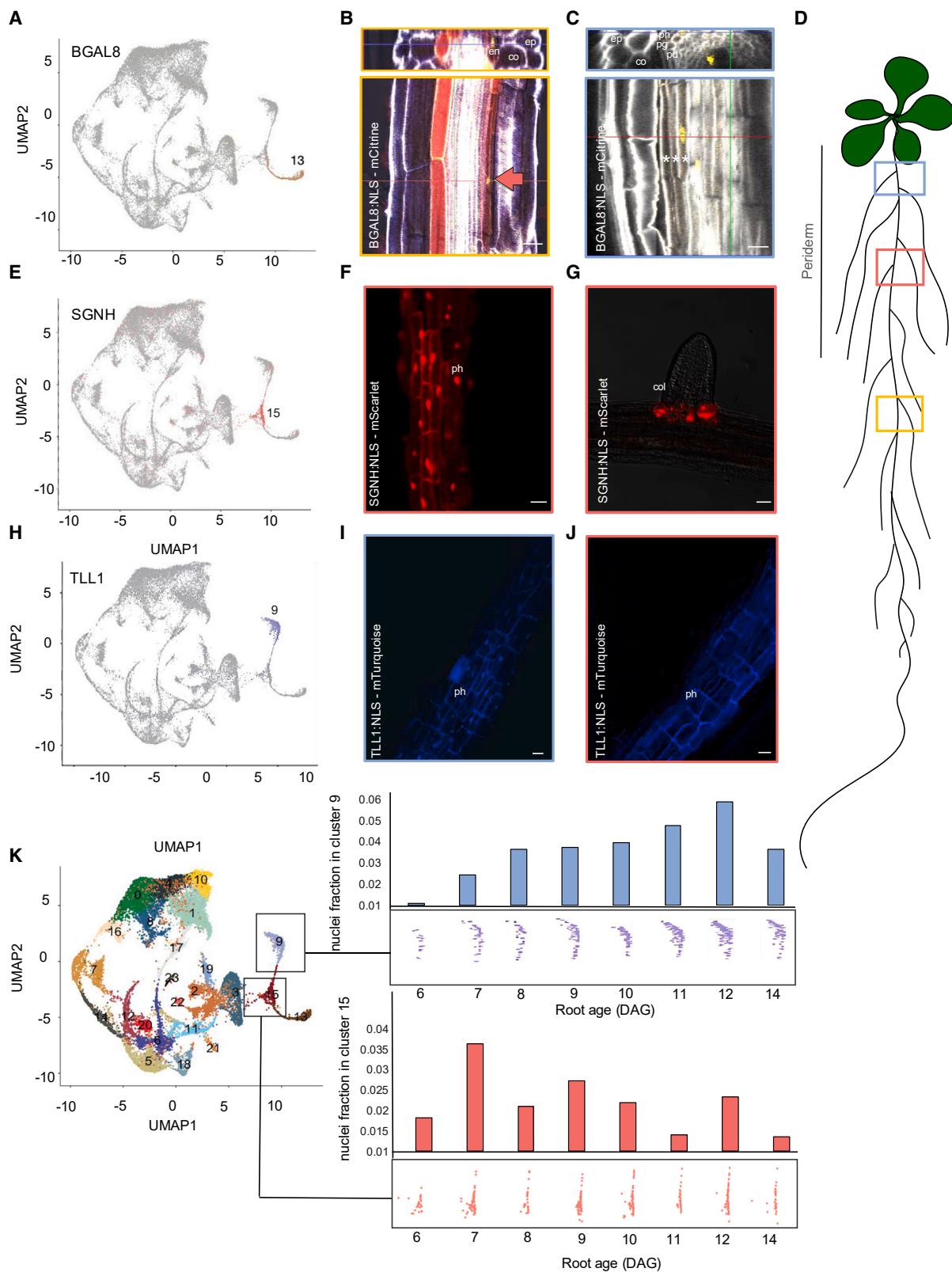


Figure 2. Validation of suberizing cell-type clusters

(A and B) BGAL8, a marker for cluster 13 (A), was expressed in the suberizing endodermis at 14 DAG, colocalizing with Nile red (red arrow) (B). (C) BGAL8 was also observed in the third periderm layer, the phellogen (asterisks indicate the periderm layers).

(legend continued on next page)

develops in the maturing, upper root just below the hypocotyl. Making use of this difference in developmental time, we dissected the top 1.5 cm of the root for additional single-nuclei sequencing and integrated the data with our time course uniform manifold approximation and projection (UMAP) clustering to add spatial context to our suberin-related clusters. In doing this, we could see that, of the upper root nuclei contributing to the suberin-gene-enriched clusters, 97% were assigned to clusters 9 and 15 (Figure 1E). This provides strong evidence that these clusters carry important information regarding phellem cell development.

Phellem cells form when pericycle cells divide to form a secondary meristem, the phellogen. Identifying cluster 3 as pericycle-related (Data S1B) and noting its connection to clusters 9 and 15, we propose that clusters 3, 15, and 9 may represent a developmental lineage of phellem cells as they emerge from the pericycle.

Reporter gene expression validates the proposed phellem developmental trajectory

Using single-nuclei sequencing, we have identified clusters that we propose may capture the development of phellem cells in the *Arabidopsis* root. To explore this further, we next took an unbiased approach to identify cluster-specific marker genes that could be used to generate transcriptional reporter lines. To do this, we selected differentially expressed genes (DEGs) that were both highly and specifically expressed in just one cluster of interest. Data S1C summarizes all DEGs detected across the 24 clusters.

Based on marker gene expression, we hypothesize that cluster 13 contains information describing the suberizing endodermis. Because these markers come from younger roots where suberization has not yet been initiated, we identified *BETA-GALACTOSIDASE 8* (*BGAL8*; AT2G28470) as a highly enriched gene to confirm this identity (Figure 2A). The promoter of this gene was used to drive the expression of mCitrine. A nuclear localisation signalling (NLS) peptide was added to aid in visualisation. In the younger root tissue of transgenic plants, *BGAL8:NLS-mCitrine* expression was seen in the suberizing endodermis, co-localizing with Nile red staining (Figures 2B red arrow and 2D for root locations). In the upper root, mCitrine signal was also seen in the third periderm layer, expected to be the phelloidem (Figure 2C). This suggests that cluster 13 may represent a mixture of nuclei taken from both suberizing endodermis and the very understudied peridermal cell layer, the phelloidem.

We hypothesize that cluster 15 corresponds to the developing phellem. DEG analysis identified AT5G37690, encoding a *SGNH* hydrolase (from here on called “*SGNH*”), as highly expressed in this cluster (Figure 2E). In line with our hypothesis, *SGNH*-promoter-driven mScarlet showed clear expression in phellem cells (Figure 2F). We were also able to see mScarlet signal in the so-called “collar cells” that form at the base of lateral roots (Figure 2G).

For cluster 9, we identified *TRIACYLGLYCEROL LIPASE-LIKE 1* (*TLL1*; AT1G45201) as a marker (Figure 2H). *TLL1:NLS-mTurquoise* was also observed in phellem cells (Figure 2I). However, unlike *SGNH*, *TLL1* signal was only observed in the phellem cells found directly below the hypocotyl and was absent in younger phellem along the root (Figure 2J). This may suggest that cluster 9 represents a more mature phellem nuclei population (additional confocal images of the *TLL1* signal in the mature root are given in Figure S1). In support of this, as would be expected for cells developing only in the mature root and that are reaching a final state of differentiation, we could see that very few nuclei were assigned to cluster 9 at 6 DAG and that the proportion of nuclei making up this cluster showed a general increase across our time course. The proportion of nuclei contributing to cluster 15, however, was seen to peak earlier along our time course at 7 DAG and to then show a slight flux over time. This, again, would be expected and is in line with our hypothesis that nuclei assigned to this population are taken from phellem cells earlier in development and at a stage where cells are transitioning from one developmental stage and into another (Figure 2K).

We next wanted to explore whether our UMAP clustering could guide us through the earlier stages of development, capturing cells making their transition away from pericycle cell identity and just beginning their developmental journey toward the young phellem cell identity that we proposed to be described by cluster 15. We identified cluster 3 as the pericycle-related nuclei population, directly connected to cluster 15. We hypothesized that genes bridging these clusters could be useful in studying the transition from pericycle to early phellem development. One such gene is *DEFECTIVELY ORGANIZED TRIBUTARIES 1* (*DOT1*; AT2G36120), known to control vascular patterning.²² *DOT1* is expressed in both the pericycle cluster and in a subset of nuclei making up cluster 15 (this region of interest is circled with a dotted line in Figure 3A). Visualization of *DOT1:NLS-mCitrine* in the young root confirmed expression in the pericycle cells (Figure 3B) and in mature roots showed expression in what are expected to be phellogen cells, found adjacent to the phellem (Figure 3C). In line with this, we could also see *DOT1* signal in cells close to the root perimeter that had just undergone cell division (shown with an orange arrow) with pre dividing cells seen just upstream (shown with a white arrow) (Figure 3D). Root locations of these images are shown in Figure 3E. Given that *DOT1* expression was seen to bridge the pericycle and the young phellem cluster, we next wanted to explore possible coexpression of *DOT1* and *SGNH*. In a small subset of nuclei in our UMAP this coexpression was evident (highest coexpression is indicated with a gray arrow in Figure 3F). Based on the expression of *DOT1* observed in the phellogen and that of *SGNH* observed in the young suberizing phellem, we hypothesized that we could make use of this coexpression to identify pericycle cells transitioning toward phellem cell identity. A transgenic line

(D) The schematic indicates root positions for each image.

(E–G) *SGNH* (AT5G37690), a marker for cluster 15 (E), showed expression in phellem (F) and suberized lateral root bases (G).

(H–K) Cluster 9 marker *TLL1* (AT1G45201) (H) appeared only in mature phellem closer to the hypocotyl (I and J). The proportion of nuclei assigned to cluster 15 and 9 were plotted across the time course. This showed an increase over time for cluster 9, in line with it relating to a maturing cell type, while the young phellem cells fractions (cluster 15) were seen peak at 7 DAG and then to flux overtime (K). Phellem (ph); phellogen (pg); phelloidem (pd); cortex (co); epidermis (ep); endodermis (en); pericycle (pe). Scale bar, 20 μ m.

See also Data S1 and Figure S1.

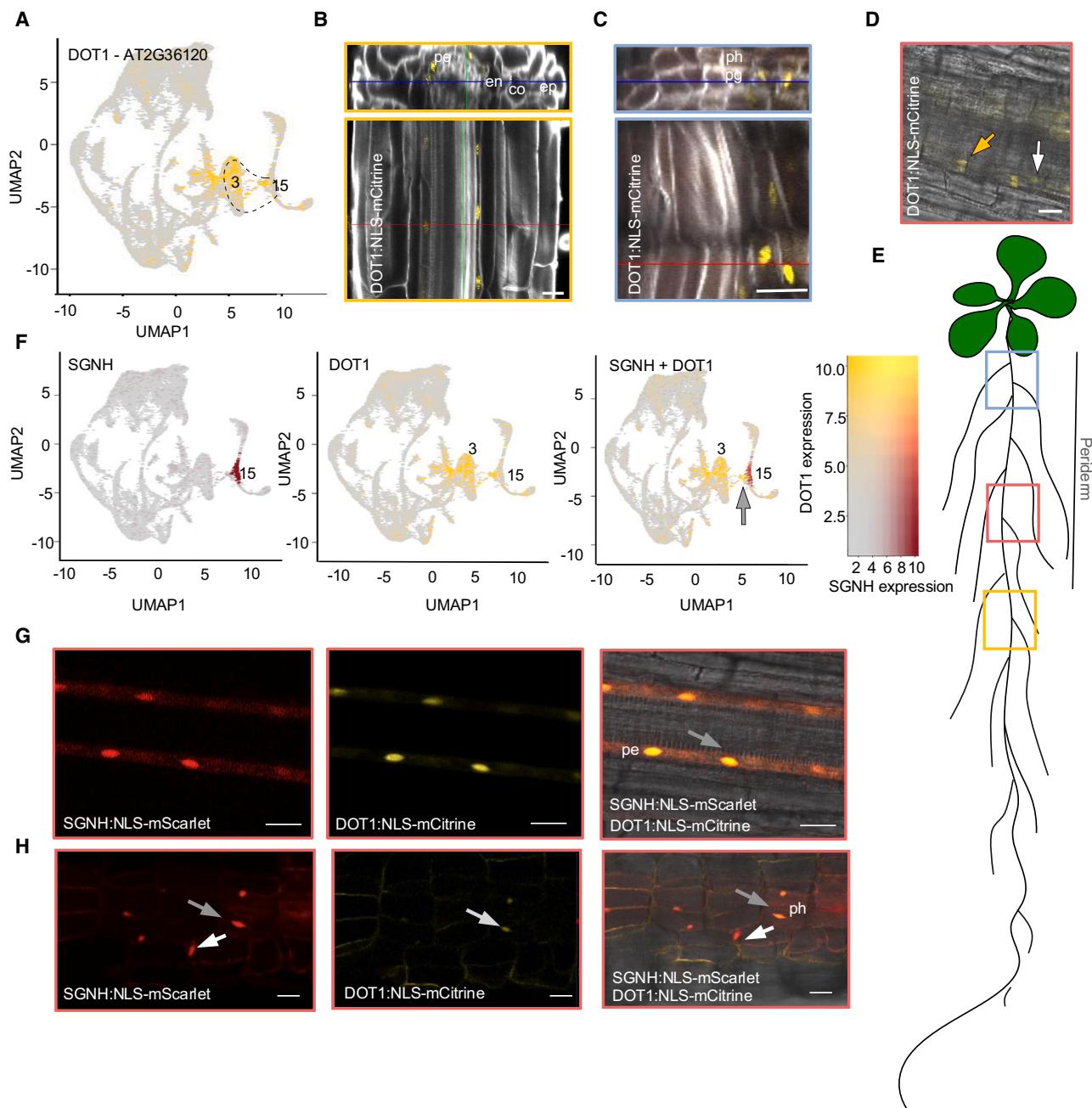


Figure 3. Colocalization of *DOT1* and *SGNH* as markers for early phellem development

(A) In the UMAP, *DOT1* expression is detected in the pericycle (cluster 3) and in the young phellem-related cluster (cluster 15). This region is highlighted with a dotted line (A). These expression patterns were validated *in planta* using a transgenic line expressing *DOT1:NLS-mCitrine*.

(B and C) In the young root, mCitrine signal was detected in the pericycle (B) and in the mature root, signal was observed in the phellogen (C).

(D) *DOT1* signal was also observed in dividing cells, which is in line with the phellogen expression pattern observed.

(E) The schematic shows root positions for each image.

(F) UMAP analysis indicated colocalization of *DOT1* and *SGNH* in a subset of nuclei (F), confirmed using transgenic lines expressing both *DOT1:NLS-mCitrine* and *SGNH:NLS-mScarlet*.

(G and H) Colocalization was seen in pericycle cells just preceding developing phellem (G) and in phellem cells (H). Gray arrows indicate colocalization, white arrows indicate *SGNH*-only nuclei. Phellem (ph); phellogen (pg); cortex (co); epidermis (ep); endodermis (en); pericycle (pe). Scale bar, 20 μ m.

See also Data S1.

expressing both *SGNH:NLS-mScarlet* and *DOT1:NLS-mCitrine* reporters indicates that this is the case: coexpression was seen in pericycle cells in the region of the root just preceding the development of phellem (Figure 3G) and could be seen to persist in cells showing clear phellem cell identity (Figure 3H, gray arrows). This is consistent with our UMAP interpretation.

In a completely independent dataset collected in day-14 roots, we were able to observe very similar clustering relating to the phellem and the endodermis. This is shown in Figure S2A, where the marker genes used in our validation work for the endodermis (Figure S2B) and the phellem (Figure S2C) provide context. In this biological replicate, we could see a seamless trajectory for the non-suberized and suberized endodermis (Figure S2D). As seen for the results of our time-course analysis, for the phellem, we could observe nuclei populations connecting *DOT1*-expressing nuclei with *SGNH*-expressing nuclei and with a population showing high differential expression of *TLL1* in close proximity to nuclei expressing the latter (Figure S2E). This shows that the methods used are robust and that the results that we obtained are reproducible.

Dynamic shifts in gene expression guide the multiple developmental processes that underlie phellem cell identity

Single-nuclei sequencing identified key nuclei populations relating to phellem development. We are confident that phellem development begins in cluster 3, associated with the pericycle. In support of this, we have shown that pericycle cells expressing *DOT1*, which is expressed in cluster 3, go on to form phellem cells. However, cluster 3 is relatively large and *DOT1* is widely expressed across this nuclei population. Furthermore, given the known pluripotency of the pericycle, we expect that cluster 3 may describe multiple developmental events, some of which are likely unrelated to the phellem. Given this, we wanted to identify a more specific route of pericycle differentiation within this cluster that might describe the very earliest stages of phellem development. To do this, we performed pseudotime analysis, a computational method that allows the ordering of nuclei along a developmental gradient according to their expression profiles. To avoid making assumptions about how this path might travel through the pericycle population, we performed pseudotime analysis in reverse, making use of our proposed mature phellem cluster as pseudotime zero. This allowed us to then work backward to identify a likely route of differentiation through the pericycle nuclei population. Re-clustering pericycle, young, and mature phellem nuclei produced a phellem-focused UMAP (Figure 4A), with cell-type markers *DOT1*, *SGNH*, and *TLL1* providing context (Figure 4B). The UMAP was divided into pseudobins to allow for the downstream DEG analysis. The pericycle cluster-related pseudobins are labeled P1–P6. The young phellem-related pseudobins are labeled YPH1 and YPH2 and the mature phellem pseudobins are labeled MPH1–3. Pseudotime analysis showed a single differentiation path from mature phellem, progressing through *SGNH*-expressing clusters, then *SGNH* and *DOT1* coexpressing cells, and finally connecting with pericycle nuclei, where a distinct route of differentiation was observed (this route is shown with a black line in Figure 4A).

With a possible phellem developmental trajectory defined, we next wanted to cross-check our pseudotime analysis with a

marker gene that was discovered independently of our work. Previous studies have shown that *WOX4* plays an important role in the maintenance of the vascular meristem.²³ More recently, this transcription factor has been identified as a regulator of phellogen activity in the *Arabidopsis* root.¹¹ *WOX4* is also known to be highly expressed in the phellogen cells of Cork oak.²⁴ Given that this gene represents one of the only known marker genes for phellogen cells in *Arabidopsis*, we queried where along our pseudotime trajectory *WOX4* expression could be identified. By plotting *WOX4* across pseudotime, we were able to identify a clear and very distinct subset of nuclei showing high *WOX4* expression at a specific point along the pseudotimeline (Figure 4C), with peak *WOX4* expression observed in pseudobin “P4” in our phellem-focused UMAP. This region likely marks the onset of phellem differentiation from the pericycle. In line with this being a pseudobin relating to the onset of pericycle differentiation, we could see that the pseudobins preceding P4 along our pseudotimeline (P1–P3), showed enrichment of genes more in line with the known transport and xylem loading functions of pericycle cells,^{25–27} including “ion homeostasis” and “transport” (Data S2B).

To identify genes and processes driving the pericycle-to-phellem development, we conducted DEG analysis on each pseudobin (Data S2A) and plotted the top 200 DEGs as a heatmap over pseudotime (Figure 4D), focusing on the pericycle-related pseudobins P4–P6 that are most relevant to phellem development. With pseudotime in reverse, later pseudobins represent younger cells. Using DEGs, we performed Gene Ontology (GO) analysis^{28,29} to identify underlying processes (key GO terms are shown in Figure 4E and a full list in Data S2B). Consistent with *WOX4*’s role in regulating phellogen, the *WOX4*-enriched P4 pseudobin displayed GO terms linked to cell division and morphogenesis, such as “microtubule movement” involving kinesin-like proteins (*KINESIN-LIKE 12C* (*KIN12C*), *KINESIN-LIKE 7I* (*KIN7I*) and *KINESIN-LIKE 7L* (*KIN7L*)). Kinesin-related proteins are known to be involved in cell expansion and cell division.³⁰ “Response to hormone” and “auxin transport” also shows high enrichment in this nuclei population, with genes including *PIN-FORMED 2* (*PIN2*), *PIN-FORMED 4* (*PIN4*), and *LIKE AUX1 3* (*LAX3*) as well as *INDOLE-3-ACETIC ACID INDUCIBLE 28* (*IAA28*), the latter of which is known to be expressed in phellogen cells in the *Arabidopsis* root.¹¹ “Anatomical structure morphogenesis”-related genes also show high enrichment here. This includes the known phellogen marker gene *WOX4* and several genes related to expansin proteins, including *EXPANSIN 18* (*EXPA18*), *EXPANSIN 14* (*EXPA14*), and *EXPANSIN B1* (*EXPB1*). The transcriptional regulator, *AGAMOUS-LIKE 14* (*AGL14*), also shows differential expression in this population of nuclei. *AGAMOUS*-like transcription factors have been found previously to be highly expressed in phellogen cells in Cork oak.²⁴

Pseudobin “P5” shows enriched GO terms, such as “pectin catabolic process,” involving pectinesterase (*PECTIN METHYL-ESTERASE 2* (*PME2*) and *PECTIN METHYLESTERASE 41* (*PME41*)) and pectin lyase (*PECTIN LYASE-LIKE 21* (*PLL21*)) genes, key to cell wall remodeling.³¹ P5 is also enriched in genes for camalexin and glucosinolate biosynthesis, both of which are secondary metabolites known to have roles in the plant defense response.³² This might provide important clues as to how the

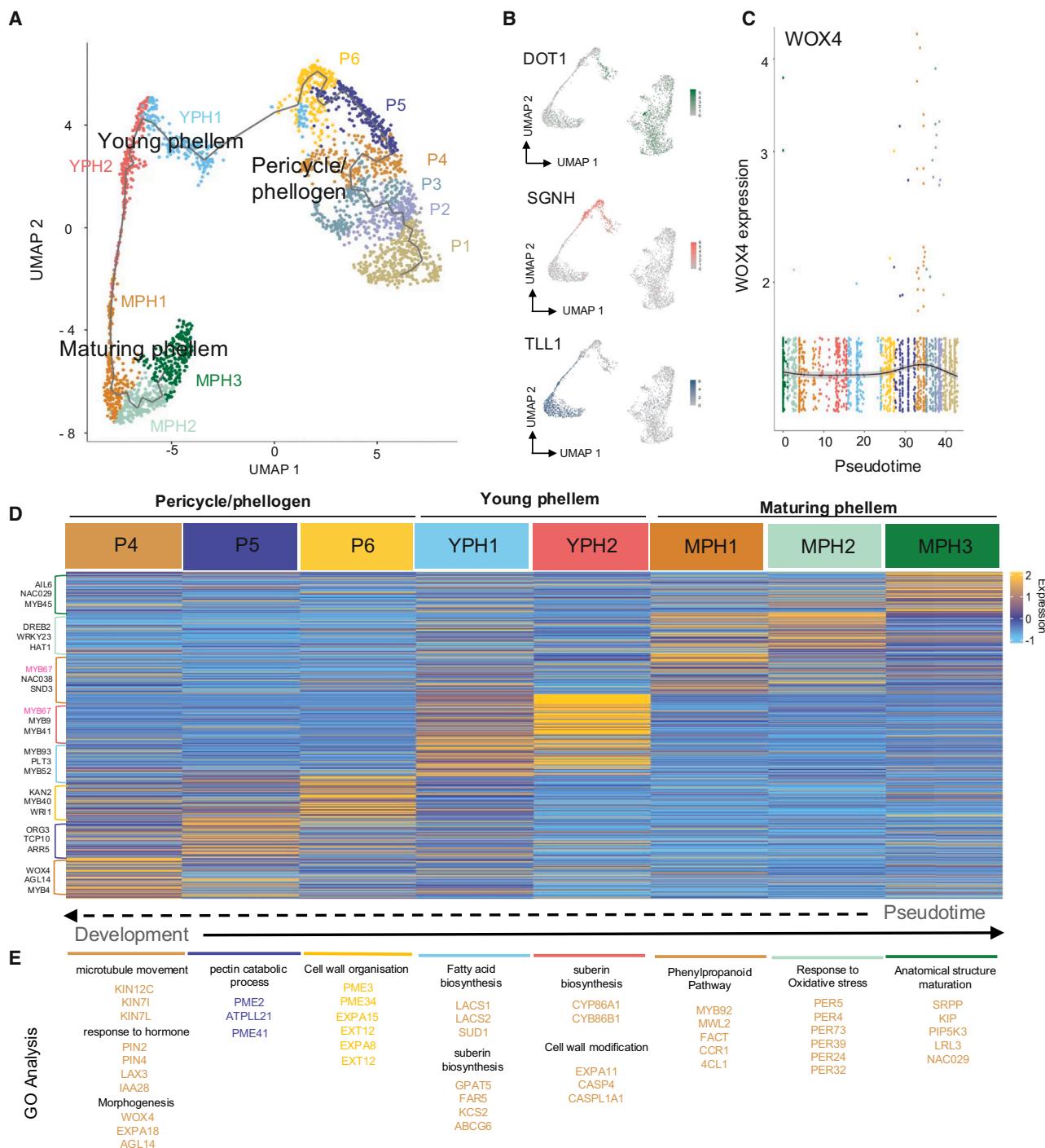


Figure 4. Key biological processes in phellem development identified through differentially expressed genes across pseudotime

(A) A phellem-focused UMAP was generated by subclustering pericycle, young phellem, and mature phellem nuclei. Using mature phellem as pseudotime zero, pseudotime analysis was performed in reverse with the pseudotimeline divided into equal-sized pseudobins (A).

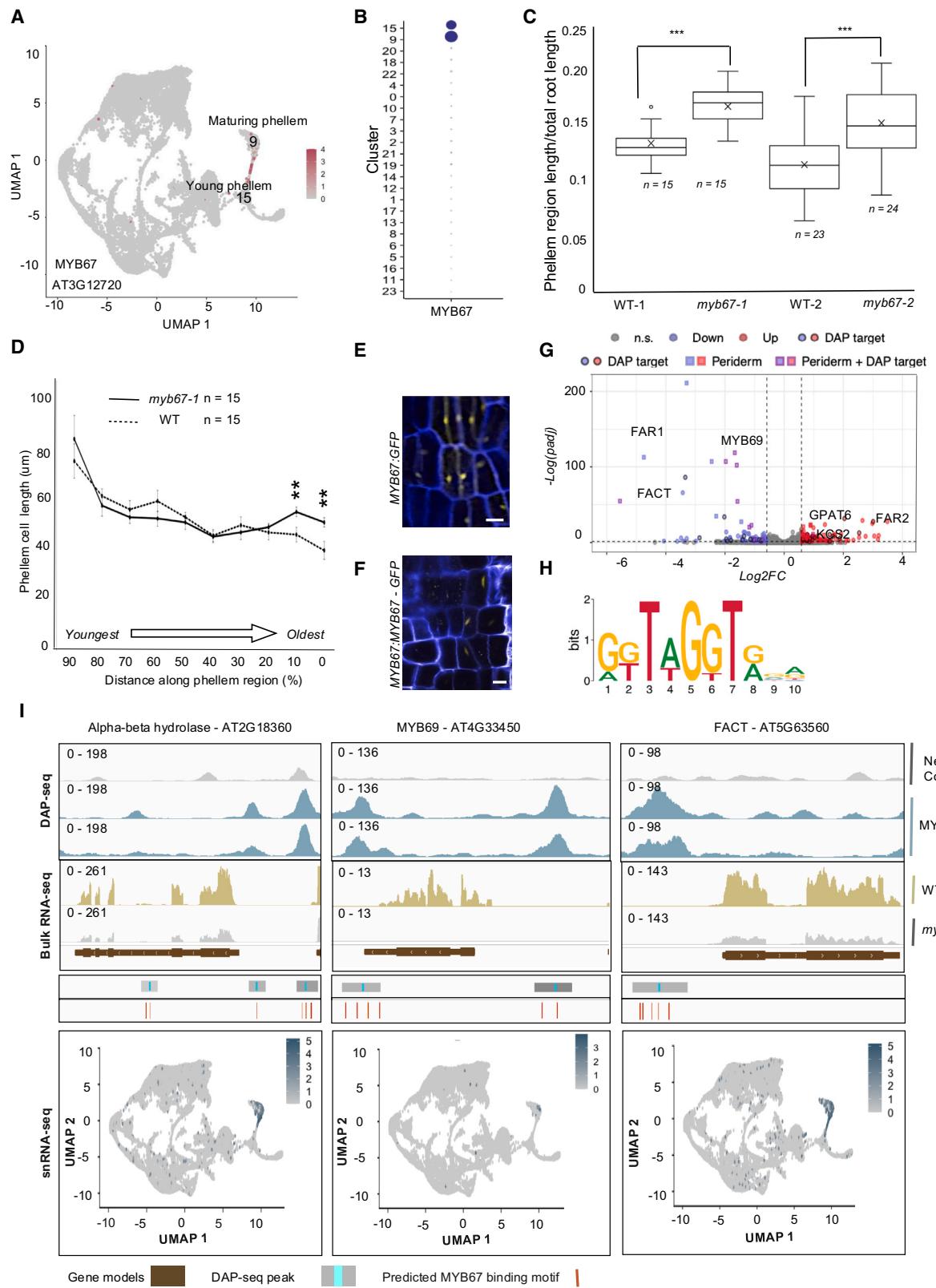
(B) Feature plots for validated genes are shown.

(C) WOX4 expression identified pseudobin P4 as the likely start of phellem differentiation from the pericycle.

(D) DEG analysis from P4 onward highlighted the top 200 genes per bin, displayed in a heatmap (D), with bins P4–P6 representing pericycle, YPH1–YPH2 young phellem, and MPH1–MPH3 mature phellem.

(E) GO analysis of DEGs was conducted for each pseudobin.

See also Data S2.



(legend on next page)

developing phellem provides protection to the maturing plant body. In P6, cell wall organization genes include additional pectinesterases (*PECTIN METHYLESTERASE 34 (PME34)* and *PECTIN METHYLESTERASE 3 (PME3)*), expansins (*EXPANIN 15 (EXPA15)* and *EXPANSIN 8 (EXPA8)*), and extensins (*EXTENSIN 12 (EXT12)*), all associated with cellular growth.

The early phellem cluster, YPH1, identified by *SGNH* and *DOT1* coexpression, is the first along our pseudotimeline to show fatty acid and suberin biosynthesis gene enrichment. This includes *LONG-CHAIN ACYL-COA 1 (LACS1)* and *LONG-CHAIN ACYL-COA 2 (LACS2)*, essential for long-chain fatty acid synthesis, and *SUPPRESSOR OF DRY2 DEFECTS 1 (SUD1)*, an E3 ligase linked to cuticular wax biosynthesis. Identification of this gene here suggests that this might also represent a post-translational regulator involved in the suberization process. Suberin-related genes in YPH1 include *GPAT5*, *FATTY ACID REDUCTASE 5 (FAR5)*, *3-KETOACYL-COA SYNTHASE 2 (KCS2)*, and *ABCG6*, all key to suberin monomer synthesis and transport.

The expression of these suberin-related genes continues in cluster YPH2 with the addition of *CYTOCHROME P450 86A1 (CYP86A1)* and *CYTOCHROME P450 86B1 (CYP86B1)*. This nuclei population also shows the expression of CASP-related genes, *CASPARIAN STRIP MEMBRANE DOMAIN PROTEIN 4 (CASP4)* and *CASPARIAN STRIP MEMBRANE DOMAIN PROTEIN-LIKE 1A1 (CASP1A1)*, indicating that CASP-related genes have roles to play in the development of both the suberizing endodermis³³ and the suberizing phellem.

In the mature phellem cluster MPH1, we observe activation of the phenylpropanoid pathway, including *MYB DOMAIN PROTEIN 92 (MYB92)*, *CINNAMOYL COA-REDUCTASE 1 (CCR1)*,³⁴ *MODIFYING WALL LIGNIN 2 (MWL2)*,³⁵ and *4-COUMARATE: COA LIGASE 1 (4CL1)*,³⁶ key for lignin synthesis, a component of the suberin structure. The following cluster, MPH2, shows genes related to “response to oxidative stress,” such as peroxidase genes *PEROXIDASE 4 (PER4)* and *PEROXIDASE 5 (PER5)*.³⁷

The final cluster along our pseudotimeline shows highest enrichment for genes involved in “anatomical structure maturation.” This GO term is in line with our marker validation work, which indicates that this cluster represents maturing phellem cells. The associated gene list here includes genes known to be involved in seed and root hair development and includes *SEED AND ROOT HAIR PROTEC-*

TIVE PROTEIN (SRPP),³⁸ *1-PHOSPHATIDYLINOSITOL-4-PHOSPHATE-5-KINASE 3 (PIP5K3)*,³⁹ and *ROOTHAIRLESS 1 (RjRHL1)*.⁴⁰ In addition, the transcriptional regulator *NAC DOMAIN-CONTAINING PROTEIN 29 (NAC29)* is included in this GO term. This gene is known to induce senescence in leaves through the activation of abscisic acid (ABA) signaling.⁴¹ This may suggest that a similar process is involved in the regulation of phellem maturation.

These analyses identified dynamic transcriptional shifts and biological processes driving phellem development. Notably, numerous transcription factors are differentially expressed along the proposed developmental trajectory (a subset is listed on the y axis in Figure 4D and full list in Data S2C).⁴² This highlights the potential complexity of the transcriptional control underlying phellem cell development and offers a toolset that could be used to effectively and precisely engineer phellem development at specific stages.

MYB67 controls phellem cell maturation

We have identified nuclei populations that describe the transcriptional landscape of phellem cells. Our results indicate that these nuclei populations describe phellem cells captured at different stages of development. We next wanted to see whether we could make use of these data to identify genes that control the progression of phellem cell development from one stage to another. Related to this, one gene that shows an interesting expression pattern in our phellem developmental trajectory, is the transcription factor, *MYB67* (AT3G12720). *MYB67* was identified as a DEG for pseudobins YPH2 and MPH1 (highlighted in pink in Figure 4D). Closer inspection identified *MYB67* expression in a small nuclei population forming a “bridge” between younger (cluster 15) and older (cluster 9) phellem-cell-related nuclei (Figure 5A). Expression was highly specific to these clusters (Figure 5B). This pattern suggests that *MYB67* may control the transition of phellem cells from younger to older stages.

To test this hypothesis, we assessed phellem development in two independent *myb67* knockout lines using Fluorol yellow staining and microscopy. The *myb67-1* line (GK-765A08)⁴³ has a T-DNA insertion in the first exon (Figure S3A), whereas *myb67-2* was generated by deleting the *MYB67* coding sequence with CRISPR-Cas9 (Figures S3B–S3D). Both knockout lines showed an increase in the proportion of the root composed of phellem cells compared with wild type (WT)

Figure 5. MYB67 as a regulator of phellem development

(A and B) *MYB67* is expressed within a small population of nuclei spanning two clusters relating to young phellem and maturing phellem. (C) Phellem region length was significantly greater in two *myb67* mutant alleles compared with WT roots (C) where “n” relates to the number of individual roots measured. Cell length was quantified along the length of the phellem region at 10% intervals (with 0% being at the hypocotyl junction and 90% being in some of the youngest phellem cells). For each sample, at least 100 phellem cells were measured. A significant increase in phellem cell length was observed in the oldest phellem in *myb67-1* (*n* = 15) relative to WT (*n* = 15) where “n” relates to number of roots phenotyped. (D) No differences between genotypes were observed in the younger phellem cells found closer toward the root tip. All root phenotype data are presented as the mean ± SEM. The assigned *p* values are indicated by asterisks, where **p* < 0.05, ***p* < 0.001, and ****p* < 0.0001. These statistical significance values were determined by two-tailed t test.

See also Data S3, S4, and S5 and Figures S3–S5.

(E and F) Transcriptional (E) and translational (F) reporter lines showed the expression of *MYB67* in the phellem.

(G and H) (G) Bulk RNA analysis of WT and *myb67-1* roots identified 184 DEGs (fold change ≥ 1.5, FDR < 0.05 calculated using the Benjamini-Hochberg method), visualized in a volcano plot where some key genes of interest are annotated (G). To identify which of these genes might represent direct targets of *MYB67*, DAP-seq was performed. This identified a likely *MYB67* binding motif (H). A significant number of downregulated DEGs were identified as direct DAP-seq targets. These are highlighted in (G) and a subset of these are shown in greater detail in (I), which shows the DAP-seq and mRNA-seq tracks as well as the expression pattern of these genes in the context of the single-nuclei UMAP. For each track the data ranges are shown in the upper left corner.

(Figure 5C; raw data in Data S3A), suggesting mis-regulation of phellem development in the absence of a functional MYB67. Because *MYB67* is expressed in nuclei bridging populations expected to represent cells transitioning from young to mature phellem cells, we hypothesized that its loss of function might disrupt the phellem cell maturation process. Phellem cell length, known to decrease with maturation,¹⁰ was measured at intervals along the phellem region. In *myb67-1*, mature phellem cells near the hypocotyl were significantly longer than those in WT (Figure 5D), indicating delayed maturation. No size difference was observed in younger phellem cells. These phenotypes were confirmed in *myb67-2* (Data S3B).

To further explore *MYB67* function, we next generated transcriptional and translational reporter lines. For both lines, as expected based on the expression pattern observed in our single-nuclei census, *MYB67* signal was observed in the phellem (Figures 5E and 5F).

To further understand *MYB67*'s role in phellem development, we performed RNA sequencing (RNA-seq) on the upper phellem-containing region of *myb67-1* and WT roots (read summary and read counts are provided in Data S4A and S4B). DEG analysis identified 184 genes with significant differential expression (fold change [FC] = |1.5|, false discovery rate [FDR] < 0.05; Figure 5G; Data S4C–S4F), with 93 upregulated and 91 downregulated genes in *myb67-1*. Notably, several suberin-related genes, including *FATTY ALCOHOL:CAFFEOYL COA TRANSFERASE* (*FACT*)¹⁴ and *FATTY ACID REDUCTASE 1* (*FAR1*)⁴⁴ (downregulated) and *GLYCEROL-3-PHOSPHATE sn-2-ACYL-TRANSFERASE 6* (*GPAT6*)⁴⁵, *KCS2*,¹⁵ and *FATTY ACID REDUCTASE 2* (*FAR2*)³ (upregulated), showed altered expression, as did *MYB DOMAIN PROTEIN 69* (*MYB69*), a regulator of lignin biosynthesis¹⁸ (these genes are annotated in Figure 5G). GO analysis of downregulated genes highlighted “lipid transport” enrichment, aligning with *MYB67*'s role in phellem development, as suberin deposition is key to this process.

To identify potential direct targets of *MYB67*, we conducted DNA affinity purification sequencing (DAP-seq),⁴⁶ identifying 5,390 enriched peaks across two *MYB67* replicates compared with controls (FDR < 10⁻⁵; Data S5A–S5C). *De novo* motif discovery showed a motif in over 96% of peaks that showed high similarity to other *MYB*-related motifs and that was previously identified for *MYB67* in an independent study.⁴⁷ This motif is shown in Figure 5H.

To identify which of the genomic regions bound by *MYB67* *in vitro* represent the most promising and potentially biologically relevant direct targets for *MYB67* *in vivo*, we compared the DAP-seq peaks to the 184 DEGs identified in the *myb67-1* mutant roots. In agreement with our finding that downregulated genes in *myb67-1* roots were significantly associated with a role in lipid transport—a GO term consistent with genes involved in the regulation of a suberizing cell type—a statistically significant number of downregulated DEGs ($n = 22/91$), but not upregulated DEGs ($n = 17/93$), were identified as *MYB67* DAP-seq targets (i.e., *myb67-1* DEGs associated with DAP-seq peaks) (Data S5D; hypergeometric test, $p = 0.01$). These 22 downregulated DEGs with *MYB67* DAP-seq peaks are marked in Figure 5G as blue squares/circles with purple/black outlines. Taken together, these genetic and genomic findings indicate that *MYB67* works

as a transcriptional activator in the mature root to control phellem development.

Given that the data presented thus far are indicative of a role for *MYB67* in regulating phellem cell maturation, we next queried the expression of the *myb67-1* DEGs within the context of our single-nuclei RNA-seq clustering to assess enrichment within the confirmed phellem-related nuclei populations. This analysis showed that, although many of the DEGs showed very low and/or non-cell-type-specific expression in our single-nuclei data, a subset of 18 downregulated genes were specifically enriched in nuclei populations related to phellem development (clusters 9 and 15), with particularly high expression observed in the maturing phellem cluster (cluster 9) (Figure S4A). Furthermore, of these 18 phellem-expressed DEGs, 9 were identified as direct *MYB67* DAP-seq targets (bold with asterisk in Figure S4A and blue squares with purple outlines in Figure 5G), 8 of which were also found to be *MYB67* targets by a previous study⁴³ (Data S5E). This represents a more than 3-fold enrichment for *MYB67* DAP-seq peaks across this phellem-expressed DEG set (fold enrichment 3.4) as compared with all *myb67-1* downregulated DEGs (fold enrichment 1.6) (Data S5D). As an example, Figure 5I shows the DAP-seq and mRNA-seq tracks together with the single-nuclei RNA-seq expression patterns for 3 of these genes, with the remaining 6 shown in Figure S5. All 9 of these genes are likely direct targets of *MYB67* and are expressed in nuclei populations that are relevant to the phenotypes observed in the *myb67* mutant, we hypothesize that they could help mediate *MYB67*'s role in regulating phellem development. This gene list includes genes known to play roles in suberin and lignin biosynthesis (i.e., *FAR1*, *FACT*, and *MYB69*), representing two key processes that underlie phellem development. For the other target genes identified, their roles in phellem development are less clear. Examples include *ATNPF3* (*AT1G68570*),⁴⁸ which is involved in gibberellin acid transport, and *PRECURSOR OF PEPTIDE 1* (*PROPEP1*) (*AT5G64900*),⁴⁹ which is known to have a role in regulating plant immunity. Exploring the role of these genes in phellem development represents an exciting opportunity for future research. Several *myb67* DEGs not identified as direct *MYB67* targets suggest indirect regulation, possibly through other transcription factors like *MYB69*, identified as a direct *MYB67* target.

With phellem cell maturation impaired in the *myb67-1* mutant, we hypothesized that phellem cells in this mutant might be suspended at a younger phellem cell developmental stage. To explore this hypothesis, we assessed how genes expressed in the young phellem cluster were mis-regulated in the *myb67-1* mutant. Interestingly, unlike what was observed for the mature phellem cluster, the vast majority of DEGs expressed in the young phellem cluster were upregulated in the mutant (Figure S4B). This is in line with our hypothesis that phellem cells in *myb67* roots are maintained in a younger phellem cell state. Taken together, we propose a model where *MYB67* controls phellem cell maturation through the activation of a suite of phellem cell maturation genes, including *FACT*, *STRUBBELIG RECEPTOR FAMILY 9* (*SUB*), and *MYB69*. In the absence of a functional *MYB67*, transcription of these gene is not correctly induced and thus phellem cell maturation is impaired.

DISCUSSION

We have seen several studies in recent years making use of single-cell sequencing to study the *Arabidopsis* root. None, however, have yet made use of such technologies to study the maturing root. Yet, many of the processes that contribute to plant fitness, such as efficient soil exploration, can more accurately be studied in maturing root systems. In dicots, another example of such processes is secondary growth and can include the development of the periderm, a complex tissue composed of the phellogen, phelloiderm, and phellem.

Here, we report the construction of a UMAP that intuitively guides us through the transcriptional shifts underlying phellem development in a way that reflects what happens in the *Arabidopsis* root. With this, we propose that phellem development can be dissected into multiple, transcriptionally distinct but interconnected phases. We show that, making use of these data, we are not only able to explore these distinct phases but also have the power to study the transcriptional signals that coordinate transitions between them. The assessment of known marker genes for phellogen activity and for suberization, together with observations made in our transcriptional reporter and mutant lines, validate this and provide confidence that these data can be used to explore fundamental developmental questions relating to this understudied cell type.

Our dataset's resolution enables targeted inquiries into phellem cell development. For instance, based on *MYB67*'s expression, which was shown to span nuclei populations relating to young and maturing phellem, we proposed and confirmed its role in regulating phellem cell maturation. These findings underscore the dataset's value for advancing phellem development research.

Phellem has recently drawn interest for its potential to enhance carbon capture and storage via increased suberin production. Engineering plants to boost suberin levels within their phellem cells, and/or to increase the number of phellem cells produced, could serve this purpose. Our study identifies nuclei populations that highlight gene sets involved in both phellem suberization and the preceding developmental processes, including pericycle cell division and early phellem growth, providing multiple pathways through which engineering goals can be achieved.

Although centered on phellem development, our single-nuclei census also offers insights into other periderm processes. For instance, we identified *BGAL8* as a phelloiderm marker, providing a potential tool for studying this cell type. In addition, capturing endodermal nuclei across multiple developmental stages, our data may provide insights into endodermal PCD, an event that seems to be intimately related to periderm development. Similarly, our time-course data could also help clarify whether abscission of outer cortex and epidermal cells in later stages is mechanically driven by maturing phellem cells or involves specific transcriptional regulation.

Our single-nuclei time-course dataset also offers an opportunity to study developmental processes beyond the periderm. Using cell type markers from Shahan et al.,¹⁸ we identified nuclei populations linked to developmental trajectories of multiple major cell types (in addition to the phellem). Although the younger parts of these developmental trajectories were captured in the elegant work described by Shahan et al.,¹⁸ our data, which

also included the upper, more maturing root, offers an opportunity to explore these cell types as they move through later developmental stages. In addition to periderm development, there are many other processes that take place as roots mature, including lateral root development (which is initiated at around 7 DAG) as well as vascular cambium development, the onset of which is tightly linked with periderm development.¹¹ Our nuclei census may also be of value in studying such developmental processes.

Our dataset offers a valuable resource for studying key developmental processes in maturing roots. Focusing on the phellem, a protective barrier crucial for stress resilience and proposed as a target for soil carbon sequestration, our single-nuclei census and findings could aid in developing climate-resilient crops with suberin-rich roots to enhance carbon capture, leveraging plants to combat climate change.

Limitations of the study

Although we have made numerous efforts to cross-validate the clusters marking unique cell states during periderm development, further independent validation is necessary to fully realize the potential of this census. Additional validation will help confirm the transitions between different clusters and highlight biologically meaningful subpopulations of cells or distinct cell states within clusters. For example, cluster 13 likely comprises a mixture of endodermal and phelloiderm cells, and we cannot exclude the possibility that other clusters may also contain mixed cell types. Moreover, although two independently generated *myb67* mutants clearly show a role for *MYB67* in phellem differentiation, a complementation line would provide a better sense of the impact of *MYB67*. Additionally, the lack of *MYB67* does not fully block phellem differentiation, suggesting that other genes likely play important roles and may partially compensate for *MYB67*'s absence. Finally, it remains unclear what the functional consequences of *MYB67*-mediated phellem maturation are in terms of plant fitness. Further studies will be needed to elucidate these aspects.

RESOURCE AVAILABILITY

Lead contact

Requests for further information and resources and reagents should be directed to and will be fulfilled by the lead contact, Wolfgang Busch (wbusch@salk.edu).

Materials availability

Seeds and plasmid relating to the *DOT1:NLS-mCitrine*, *SGNH:NLS-mScarlet*, *TL1:NLS-mTurquoise*, and *BGAL8:NLS-mCitrine*, as well as *myb67* CRISPR alleles described in his paper, are available from Wolfgang Busch upon request.

Data and code availability

- Single-nucleus RNA-seq data have been deposited at NCBI Gene Expression Omnibus under accession number NCBI: GSE261441 and are publicly available as of the date of publication. Bulk RNA-seq data have been deposited at NCBI Gene Expression Omnibus under accession number NCBI: GSE261437 and are publicly available as of the date of publication. DAP-seq data have been deposited at NCBI Gene Expression Omnibus under accession number NCBI: GSE261431 and are publicly available as of the date of publication.
- All original code is available in this paper's **STAR Methods** section.
- Any additional information required to reanalyze the data reported in this paper is available from the **lead contact** upon request.

ACKNOWLEDGMENTS

This work was supported by gifts to the Salk Institute's Harnessing Plants Initiative (HPI) from the Bezos Earth Fund, the Hess Corporation, and through the TED Audacious Project. This work was supported by the NGS Core Facility of the Salk Institute with funding from NIH-NCI CCSG: P30 CA01495, NIH-NIA San Diego Nathan Shock Center P30 AG068635, the Chapman Foundation, and the Helmsley Charitable Trust. This work was supported by the Waitt Advanced Biophotonics Core Facility of the Salk Institute (RRID:SCR_014838) with funding from NIH-NCI CCSG P30 CA014195, NIH-NIA San Diego Nathan Shock Center P30 AG068635, The Henry L. Guenther Foundation, and the Waitt Foundation. Core personnel include Daniela Boassa, Sammy Weiser-Novak, and Elsie Quansah. This work was supported by the Flow Cytometry Core Facility of the Salk Institute (RRID:SCR_014839) with funding from NIH-NCI CCSG P30 CA014195, shared instrumentation grants S10-OD023689 (Aria Fusion cell sorter), and S10 OD034268 (Thermo Fisher Bigfoot). Core personnel include Carolyn (Caz) O'Connor, Michelle Liem, Mikayla Marrin, and Cynthia Rose. This work was supported by the Razavi Newman Integrative Genomics and Bioinformatics Core Facility of the Salk Institute (RRID:SCR_014842 and SCR_014846) with funding from NIH-NCI CCSG P30 CA014195, NIH-NIA San Diego Nathan Shock Center P30 AG068635, the NIH-NIA Liver Cancer P01 AG073084-04, the Howard and Maryam Newman Family Foundation, and the Helmsley Trust. Core personnel include April Williams, Jingting Yu, Kat Lande, Garrett Evensen, Ling Ouyang, and Tzuvwen Wang.

AUTHOR CONTRIBUTIONS

C.N.M., W.B., and J.A.L. conceived and designed experiments. C.N.M. conducted experiments and wrote the paper. S.J.-H., M.V.H., Y.S.Y., P.A.-M., S.B., K.C., A.P., and J.S. conducted experiments. M.S., J.Y., and L.Z. performed bioinformatic analyses. J.A.L. and W.B. edited the manuscript. W.B., J.A.L., T.P.M., and J.R.E. supervised work and contributed resources and funding.

DECLARATION OF INTERESTS

W.B., J.A.L., T.P.M., and J.R.E. are co-founders of Cuesta, a company that works on crop root growth and carbon sequestration.

DECLARATION OF GENERATIVE AI AND AI-ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

During the preparation of this work, the authors used ChatGPT to aid in the reduction of the manuscript word count. After using this tool/service, content was reviewed and edited as needed by the authors, who take full responsibility for the content of the published article.

STAR METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
- METHOD DETAILS
 - Nuclei extraction and FACS
 - Generation and selection of transgenic lines
 - Generation of CRISPR/Cas9 k.o. lines for MYB67
 - Nile Red and Calcofluor White staining and confocal imaging
 - Fluorol Yellow Staining, Imaging, and Analysis
 - RNA-seq
 - DAP-seq, cloning, immunoprecipitation, library construction and sequencing
- QUANTIFICATION AND STATISTICAL ANALYSIS
 - Single-nuclei data bioinformatic work-flow
 - DEG analysis
 - Mutant phenotype quantification

- Bulk RNA-seq data processing workflow
- DAP-seq data processing workflow
- Predicting MYB67 binding motifs

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.devcel.2024.12.025>.

Received: March 13, 2024

Revised: September 10, 2024

Accepted: December 11, 2024

Published: January 9, 2025

REFERENCES

1. Serra, O., Mähönen, A.P., Hetherington, A.J., and Ragni, L. (2022). The Making of Plant Armor: The Periderm. *Annu. Rev. Plant Biol.* 73, 405–432. <https://doi.org/10.1146/annurev-aplant-102720-031405>.
2. Ranathunge, K., and Schreiber, L. (2011). Water and solute permeabilities of *Arabidopsis* roots in relation to the amount and composition of aliphatic suberin. *J. Exp. Bot.* 62, 1961–1974. <https://doi.org/10.1093/jxb/erq389>.
3. de Silva, N.D.G., Murmu, J., Chabot, D., Hubbard, K., Ryser, P., Molina, I., and Rowland, O. (2021). Root suberin plays important roles in reducing water loss and sodium uptake in *Arabidopsis thaliana*. *Metabolites* 11, 735. <https://doi.org/10.3390/metabo11110735>.
4. Lux, A., Vaculík, M., Martinka, M., Lišková, D., Kulkarni, M.G., Stirk, W.A., and Van Staden, J. (2011). Cadmium induces hypodermal periderm formation in the roots of the monocotyledonous medicinal plant *Merwilla plumbea*. *Ann. Bot.* 107, 285–292. <https://doi.org/10.1093/aob/mcq240>.
5. Ranathunge, K., Schreiber, L., Bi, Y.-M., and Rothstein, S.J. (2016). Ammonium-induced architectural and anatomical changes with altered suberin and lignin levels significantly change water and solute permeabilities of rice (*Oryza sativa* L.) roots. *Planta* 243, 231–249. <https://doi.org/10.1007/s00425-015-2406-1>.
6. Ranathunge, K., Thomas, R.H., Fang, X., Peterson, C.A., Gijzen, M., and Bernards, M.A. (2008). Soybean root suberin and partial resistance to root rot caused by *Phytophthora sojae*. *Phytopathology* 98, 1179–1189. <https://doi.org/10.1094/PHYTO-98-11-1179>.
7. Eckardt, N.A., Ainsworth, E.A., Bahuguna, R.N., Broadley, M.R., Busch, W., Carptia, N.C., Castrillo, G., Chory, J., DeHaan, L.R., Duarte, C.M., et al. (2023). Climate change challenges, plant science solutions. *Plant Cell* 35, 24–66. <https://doi.org/10.1093/plcell/koac303>.
8. Busch, W., and Miller, C. (2022). The greenest revolution – harnessing the power of plants to help combat climate change. *Biochemist (Lond.)* 44, 13–18. https://doi.org/10.1042/bio_2022_113.
9. Smetana, O., Mäkiilä, R., Lyu, M., Amiryousefi, A., Sánchez Rodríguez, F., Wu, M.-F., Solé-Gil, A., Leal Gavarrón, M., Siligato, R., Miyashima, S., et al. (2019). High levels of auxin signalling define the stem-cell organizer of the vascular cambium. *Nature* 565, 485–489. <https://doi.org/10.1038/s41586-018-0837-0>.
10. Wunderling, A., Ripper, D., Barra-Jimenez, A., Mahn, S., Sajak, K., Targem, M.B., and Ragni, L. (2018). A molecular framework to study periderm formation in *Arabidopsis*. *New Phytol.* 219, 216–229. <https://doi.org/10.1111/nph.15128>.
11. Xiao, W., Molina, D., Wunderling, A., Ripper, D., Vermeer, J.E.M., and Ragni, L. (2020). Pluripotent pericycle cells trigger different growth outputs by integrating developmental cues into distinct regulatory modules. *Curr. Biol.* 30, 4384–4398.e5. <https://doi.org/10.1016/j.cub.2020.08.053>.
12. Ayaz, A., Saqib, S., Huang, H., Zaman, W., Lü, S., and Zhao, H. (2021). Genome-wide comparative analysis of long-chain acyl-CoA synthetases (LACSSs) gene family: A focus on identification, evolution and expression profiling related to lipid synthesis. *Plant Physiol. Biochem.* 161, 1–11. <https://doi.org/10.1016/j.plaphy.2021.01.042>.

13. Beisson, F., Li, Y., Bonaventure, G., Pollard, M., and Ohlrogge, J.B. (2007). The acyltransferase GPAT5 is required for the synthesis of suberin in seed coat and root of *Arabidopsis*. *Plant Cell* 19, 351–368. <https://doi.org/10.1105/tpc.106.048033>.
14. Molina, I., Li-Beisson, Y., Beisson, F., Ohlrogge, J.B., and Pollard, M. (2009). Identification of an *Arabidopsis* feruloyl-coenzyme A transferase required for suberin synthesis. *Plant Physiol.* 151, 1317–1328. <https://doi.org/10.1104/pp.109.144907>.
15. Yadav, V., Molina, I., Ranathunge, K., Castillo, I.Q., Rothstein, S.J., and Reed, J.W. (2014). ABCG transporters are required for suberin and pollen wall extracellular barriers in *Arabidopsis*. *Plant Cell* 26, 3569–3588. <https://doi.org/10.1105/tpc.114.129049>.
16. Ursache, R., De Jesus Vieira Teixeira, C., Déneraud Tendon, V., Gully, K., De Bellis, D., Schmid-Siegert, E., Grube Andersen, T., Shekhar, V., Calderon, S., Pradervand, S., et al. (2021). GDSL-domain proteins have key roles in suberin polymerization and degradation. *Nat. Plants* 7, 353–364. <https://doi.org/10.1038/s41477-021-00862-9>.
17. Leal, A.R., Barros, P.M., Parizot, B., Sapeta, H., Vangheluwe, N., Andersen, T.G., Beeckman, T., and Oliveira, M.M. (2022). Translational profile of developing phellem cells in *Arabidopsis thaliana* roots. *Plant J.* 110, 899–915. <https://doi.org/10.1111/tpj.15691>.
18. Shahan, R., Hsu, C.-W., Nolan, T.M., Cole, B.J., Taylor, I.W., Greenstreet, L., Zhang, S., Afanassiev, A., Vlot, A.H.C., Schieberer, G., et al. (2022). A single-cell *Arabidopsis* root atlas reveals developmental trajectories in wild-type and cell identity mutants. *Dev. Cell* 57, 543–560.e9. <https://doi.org/10.1016/j.devcel.2022.01.008>.
19. Gala, H.P., Lanctot, A., Jean-Baptiste, K., Guiou, S., Chu, J.C., Zemke, J.E., George, W., Queitsch, C., Cuperus, J.T., and Nemhauser, J.L. (2021). A single-cell view of the transcriptome during lateral root initiation in *Arabidopsis thaliana*. *Plant Cell* 33, 2197–2220. <https://doi.org/10.1093/plcell/koab101>.
20. Wendrich, J.R., Yang, B., Vandamme, N., Verstaen, K., Smet, W., Van de Velde, C., Minne, M., Wybauw, B., Mor, E., Arents, H.E., et al. (2020). Vascular transcription factors guide plant epidermal responses to limiting phosphate conditions. *Science* 370, eaay4970. <https://doi.org/10.1126/science.aay4970>.
21. Naseer, S., Lee, Y., Lapierre, C., Franke, R., Nawrath, C., and Geldner, N. (2012). Casparyan strip diffusion barrier in *Arabidopsis* is made of a lignin polymer without suberin. *Proc. Natl. Acad. Sci. USA* 109, 10101–10106. <https://doi.org/10.1073/pnas.1205726109>.
22. Petricka, J.J., Clay, N.K., and Nelson, T.M. (2008). Vein patterning screens and the defectively organized tributaries mutants in *Arabidopsis thaliana*. *Plant J.* 56, 251–263. <https://doi.org/10.1111/j.1365-313X.2008.03595.x>.
23. Ji, J., Strable, J., Shimizu, R., Koenig, D., Sinha, N., and Scanlon, M.J. (2010). WOX4 promotes procambial development. *Plant Physiol.* 152, 1346–1356. <https://doi.org/10.1104/pp.109.149641>.
24. Fernández-Piñán, S., Boher, P., Soler, M., Figueras, M., and Serra, O. (2021). Transcriptomic analysis of cork during seasonal growth highlights regulatory and developmental processes from phellogen to phellem formation. *Sci. Rep.* 11, 12053. <https://doi.org/10.1038/s41598-021-90938-5>.
25. Huang, S., Yamaji, N., Sakurai, G., Mitani-Ueno, N., Konishi, N., and Ma, J.F. (2022). A pericycle-localized silicon transporter for efficient xylem loading in rice. *New Phytol.* 234, 197–208. <https://doi.org/10.1111/nph.17959>.
26. Takano, J., Noguchi, K., Yasumori, M., Kobayashi, M., Gajdos, Z., Miwa, K., Hayashi, H., Yoneyama, T., and Fujiwara, T. (2002). *Arabidopsis* boron transporter for xylem loading. *Nature* 420, 337–340. <https://doi.org/10.1038/nature01139>.
27. Vakhmistov, D.B. (1981). Specialization of root tissues in ion transport. *Plant Soil* 63, 33–38. <https://doi.org/10.1007/BF02374254>.
28. Thomas, P.D., Ebert, D., Muruganujan, A., Mushayahama, T., Albou, L.P., and Mi, H. (2022). PANTHER: making genome-scale phylogenetics accessible to all. *Protein Sci.* 31, 8–22. <https://doi.org/10.1002/pro.4218>.
29. Tian, T., Liu, Y., Yan, H., You, Q., Yi, X., Du, Z., Xu, W., and Su, Z. (2017). agriGO v2.0: a GO analysis toolkit for the agricultural community, 2017 update. *Nucleic Acids Res.* 45, W122–W129. <https://doi.org/10.1093/nar/gkx382>.
30. Zhu, C., and Dixit, R. (2012). Functions of the *Arabidopsis* kinesin superfamily of microtubule-based motor proteins. *Protoplasma* 249, 887–899. <https://doi.org/10.1007/s00709-011-0343-9>.
31. Kumar, R., Meghwanshi, G.K., Marcianò, D., Ullah, S.F., Bulone, V., Toffolatti, S.L., and Srivastava, V. (2023). Sequence, structure and functionality of pectin methylesterases and their use in sustainable carbohydrate bioproducts: a review. *Int. J. Biol. Macromol.* 244, 125385. <https://doi.org/10.1016/j.ijbiomac.2023.125385>.
32. Stotz, H.U., Sawada, Y., Shimada, Y., Hirai, M.Y., Sasaki, E., Krischke, M., Brown, P.D., Saito, K., and Kamiya, Y. (2011). Role of camalexin, indole glucosinolates, and side chain modification of glucosinolate-derived isothiocyanates in defense of *Arabidopsis* against *Sclerotinia sclerotiorum*. *Plant J.* 67, 81–93. <https://doi.org/10.1111/j.1365-313X.2011.04578.x>.
33. Roppolo, D., De Rybel, B., Déneraud Tendon, V., Pfister, A., Alassimone, J., Vermeer, J.E.M., Yamazaki, M., Stierhof, Y.-D., Beeckman, T., and Geldner, N. (2011). A novel protein family mediates Casparyan strip formation in the endodermis. *Nature* 473, 380–383. <https://doi.org/10.1038/nature10070>.
34. Xue, J., Luo, D., Xu, D., Zeng, M., Cui, X., Li, L., and Huang, H. (2015). CCR1, an enzyme required for lignin biosynthesis in *Arabidopsis*, mediates cell proliferation exit for leaf development. *Plant J.* 83, 375–387. <https://doi.org/10.1111/tpj.12902>.
35. Mewalal, R., Mizrahi, E., Coetze, B., Mansfield, S.D., and Myburg, A.A. (2016). The *Arabidopsis* Domain of unknown function 1218 (DUF1218) containing proteins, MODIFYING WALL LIGNIN-1 and 2 (AT1G31720/MWL-1 and AT4G19370/MWL-2) function redundantly to alter secondary cell wall lignin content. *PLoS One* 11, e0150254. <https://doi.org/10.1371/journal.pone.0150254>.
36. Li, Y., Kim, J.I., Pysh, L., and Chapple, C. (2015). Four isoforms of *Arabidopsis thaliana* 4-coumarate: CoA ligase (4CL) have overlapping yet distinct roles in phenylpropanoid metabolism. *Plant Physiol.* 169, 2409–2421. <https://doi.org/10.1104/pp.15.00838>.
37. Li, Z., Zhang, D., Shi, P., Htwe, Y.M., Yu, Q., Huang, L., Zhou, H., Liu, L., and Wang, Y. (2023). Cell wall lignification may be necessary for somatic embryogenesis of areca palm (*Areca catechu*). *Sci. Hortic.* 307, 111538. <https://doi.org/10.1016/j.scienta.2022.111538>.
38. Tanaka, N., Uno, H., Okuda, S., Gunji, S., Ferjani, A., Aoyama, T., and Maeshima, M. (2017). SRPP, a cell-wall protein is involved in development and protection of seeds and root hairs in *Arabidopsis thaliana*. *Plant Cell Physiol.* pcx008. *Plant Cell Physiol.* 58, 760–769. <https://doi.org/10.1093/pcp/pcx008>.
39. Kato, M., Tsuge, T., Maeshima, M., and Aoyama, T. (2019). *Arabidopsis* PC aP2 modulates the phosphatidylinositol 4,5-bisphosphate signal on the plasma membrane and attenuates root hair elongation. *Plant J.* 99, 610–625. <https://doi.org/10.1111/tpj.14226>.
40. Bruex, A., Kainkaryam, R.M., Wieckowski, Y., Kang, Y.H., Bernhardt, C., Xia, Y., Zheng, X., Wang, J.Y., Lee, M.M., Benfey, P., et al. (2012). A gene regulatory network for root epidermis cell differentiation in *Arabidopsis*. *PLoS Genet.* 8, e1002446. <https://doi.org/10.1371/journal.pgen.1002446>.
41. Guo, Y., and Gan, S. (2006). AtNAP, a NAC family transcription factor, has an important role in leaf senescence. *Plant J.* 46, 601–612. <https://doi.org/10.1111/j.1365-313X.2006.02723.x>.
42. Jin, J., Tian, F., Yang, D.-C., Meng, Y.-Q., Kong, L., Luo, J., and Gao, G. (2017). PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic Acids Res.* 45, D1040–D1045. <https://doi.org/10.1093/nar/gkw982>.
43. Kleinboelt, N., Huep, G., Kloetgen, A., Viehoever, P., and Weisshaar, B. (2012). GABI-Kat SimpleSearch: new features of the *Arabidopsis thaliana* T-DNA mutant database. *Nucleic Acids Res.* 40, D1211–D1215. <https://doi.org/10.1093/nar/gkr1047>.

44. Domergue, F., Vishwanath, S.J., Joubès, J., Ono, J., Lee, J.A., Bourdon, M., Alhattab, R., Lowe, C., Pascal, S., Lessire, R., et al. (2010). Three *Arabidopsis* fatty acyl-coenzyme A reductases, FAR1, FAR4, and FAR5, generate primary fatty alcohols associated with suberin deposition. *Plant Physiol.* 153, 1539–1554. <https://doi.org/10.1104/pp.110.158238>.
45. Feng, T., Wu, P., Gao, H., Kosma, D.K., Jenks, M.A., and Lü, S. (2022). Natural variation in root suberization is associated with local environment in *Arabidopsis thaliana*. *New Phytol.* 236, 385–398. <https://doi.org/10.1111/nph.18341>.
46. Bartlett, A., O’Malley, R.C., Huang, S.C., Galli, M., Nery, J.R., Gallavotti, A., and Ecker, J.R. (2017). Mapping genome-wide transcription-factor binding sites using DAP-seq. *Nat. Protoc.* 12, 1659–1672. <https://doi.org/10.1038/nprot.2017.055>.
47. O’Malley, R.C., Huang, S.C., Song, L., Lewsey, M.G., Bartlett, A., Nery, J.R., Galli, M., Gallavotti, A., and Ecker, J.R. (2016). Cistrome and Epicistrome features shape the regulatory DNA landscape. *Cell* 165, 1280–1292. <https://doi.org/10.1016/j.cell.2016.04.038>.
48. David, L.C., Berquin, P., Kanno, Y., Seo, M., Daniel-Vedele, F., and Ferrario-Méry, S. (2016). N availability modulates the role of NPF3.1, a gibberellin transporter, in GA-mediated phenotypes in *Arabidopsis*. *Planta* 244, 1315–1328. <https://doi.org/10.1007/s00425-016-2588-1>.
49. Huang, Y., Cui, J., Li, M., Yang, R., Hu, Y., Yu, X., Chen, Y., Wu, Q., Yao, H., Yu, G., et al. (2023). Conservation and divergence of flg22, pep1 and nlp20 in activation of immune response and inhibition of root development. *Plant Sci.* 331, 111686. <https://doi.org/10.1016/j.plantsci.2023.111686>.
50. Stuttmann, J., Barthel, K., Martin, P., Ordon, J., Erickson, J.L., Herr, R., Ferik, F., Kretschmer, C., Berner, T., Keilwagen, J., et al. (2021). Highly efficient multiplex editing: one-shot generation of 8× *Nicotiana benthamiana* and 12× *Arabidopsis* mutants. *Plant J.* 106, 8–22. <https://doi.org/10.1111/tpj.15197>.
51. Engler, C., Youles, M., Gruetzner, R., Ehnert, T.-M., Werner, S., Jones, J.D.G., Patron, N.J., and Marillonnet, S. (2014). A Golden Gate modular cloning toolbox for plants. *ACS Synth. Biol.* 3, 839–843. <https://doi.org/10.1021/sb4001504>.
52. Andrews, S. (2010). FastQC: A quality control tool for high throughput sequence data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
53. Wingett, S.W., and Andrews, S. (2018). FastQ Screen: A tool for multi-genome mapping and quality control. *F1000Res* 7, 1338. <https://doi.org/10.12688/f1000research.15931.2>.
54. Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
55. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21. <https://doi.org/10.1093/bioinformatics/bts635>.
56. Thorvaldsdóttir, H., Robinson, J.T., and Mesirov, J.P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinform.* 14, 178–192. <https://doi.org/10.1093/bib/bbs017>.
57. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550. <https://doi.org/10.1186/s13059-014-0550-8>.
58. Kondili, M., Fust, A., Preussner, J., Kuenne, C., Braun, T., and Looso, M. (2017). UROPA: a tool for Universal Robust Peak Annotation. *Sci. Rep.* 7, 2593. <https://doi.org/10.1038/s41598-017-02464-y>.
59. Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* 38, 576–589. <https://doi.org/10.1016/j.molcel.2010.05.004>.
60. Tremblay, B.J. (2023). Universalmotif: Import, Modify, and Export Motifs with R. *Bioconductor*. <https://github.com/bjmt/universalmotif>.
61. Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842. <https://doi.org/10.1093/bioinformatics/btq033>.
62. Bailey, T.L., Johnson, J., Grant, C.E., and Noble, W.S. (2015). The MEME suite. *Nucleic Acids Res.* 43, W39–W49. <https://doi.org/10.1093/nar/gkv416>.
63. Du, L., Liu, Q., Fan, Z., Tang, J., Zhang, X., Price, M., Yue, B., and Zhao, K. (2021). Pyfastx: a robust Python package for fast random access to sequences from plain and gzipped FASTA/Q files. *Brief. Bioinform.* 22, bbaa368. <https://doi.org/10.1093/bib/bbaa368>.
64. Lee, T., Nobori, T., Illouz-Eliaz, N., Xu, J., Jow, B., Nery, J., and Ecker, J. (2023). A single-nucleus atlas of seed-to-seed development in *Arabidopsis*. Preprint at bioRxiv. <https://doi.org/10.1101/2023.03.23.533992>.
65. 10x Genomics (2019). Chromium Single Cell V(3.1) Chromium Next GEM single cell 3’Reagent Kits V3.1 Document Number CG000204 Rev D.
66. Zhang, X., Henriques, R., Lin, S.-S., Niu, Q.-W., and Chua, N.-H. (2006). Agrobacterium-mediated transformation of *Arabidopsis thaliana* using the floral dip method. *Nat. Protoc.* 1, 641–646. <https://doi.org/10.1038/nprot.2006.97>.
67. Naito, Y., Hino, K., Bono, H., and Ui-Tei, K. (2015). CRISPRdirect: software for designing CRISPR/Cas guide RNA with reduced off-target sites. *Bioinformatics* 31, 1120–1123. <https://doi.org/10.1093/bioinformatics/btu743>.
68. Zheng, G.X.Y., Terry, J.M., Belgrader, P., Ryvkin, P., Bent, Z.W., Wilson, R., Ziraldo, S.B., Wheeler, T.D., McDermott, G.P., Zhu, J., et al. (2017). Massively parallel digital transcriptional profiling of single cells. *Nat. Commun.* 8, 14049. <https://doi.org/10.1038/ncomms14049>.
69. Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W.M., Zheng, S., Butler, A., Lee, M.J., Wilk, A.J., Darby, C., Zager, M., et al. (2021). Integrated analysis of multimodal single-cell data. *Cell* 184, 3573–3587.e29. <https://doi.org/10.1016/j.cell.2021.04.048>.
70. Bais, A.S., and Kostka, D. (2019). scds: computational Annotation of Doublets in Single Cell RNA Sequencing Data. Preprint at bioRxiv. <https://doi.org/10.1101/564021>.
71. Qiu, X., Mao, Q., Tang, Y., Wang, L., Chawla, R., Pliner, H.A., and Trapnell, C. (2017). Reversed graph embedding resolves complex single-cell trajectories. *Nat. Methods* 14, 979–982. <https://doi.org/10.1038/nmeth.4402>.
72. Gu, Z. (2022). Complex heatmap visualization. *Imeta* 1, e43. <https://doi.org/10.1002/imt2.43>.
73. Gu, Z., Eils, R., and Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* 32, 2847–2849. <https://doi.org/10.1093/bioinformatics/btw313>.
74. Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* 9, 671–675. <https://doi.org/10.1038/nmeth.2089>.
75. Lamesch, P., Berardini, T.Z., Li, D., Swarbreck, D., Wilks, C., Sasidharan, R., Muller, R., Dreher, K., Alexander, D.L., Garcia-Hernandez, M., et al. (2012). The *Arabidopsis* Information Resource (TAIR): improved gene annotation and new tools. *Nucleic Acids Res.* 40, D1202–D1210. <https://doi.org/10.1093/nar/gkr1090>.
76. Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based analysis of ChIP-seq (MACS). *Genome Biol.* 9, R137. <https://doi.org/10.1186/gb-2008-9-9-r137>.
77. Cheng, C.Y., Krishnakumar, V., Chan, A.P., Thibaud-Nissen, F., Schobel, S., and Town, C.D. (2017). Araport11: a complete reannotation of the *Arabidopsis thaliana* reference genome. *Plant J.* 89, 789–804. <https://doi.org/10.1111/tpj.13415>.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains		
<i>Agrobacterium tumefaciens</i>	N/A	N/A
<i>E.coli</i>	N/A	N/A
Biological samples (all col-0 background)		
<i>Arabidopsis thaliana</i> (Columbia-0)	N/A	N/A
<i>pGPAT5:NLS-GFP-GUS</i>	Wunderling et al. ¹⁰ ; Naseer et al. ²¹	N/A
<i>pDOT1:NLS-mCitrine</i>	this manuscript	N/A
<i>pSGNH:NLS-mScarlet</i>	this manuscript	N/A
<i>pBGAL8:NLS-mCitrine</i>	this manuscript	N/A
<i>pTLL1:NLS-mTurquoise</i>	this manuscript	N/A
<i>pDOT1:NLS-mCitrine pSGNH:NLS-mScarlet</i>	this manuscript	N/A
<i>Arabidopsis: myb67-1 T-DNA line</i>	Kleinboelting et al. ⁴³	ABRC stock number GK-765A08
<i>Arabidopsis: myb67-2 CRISPR line</i>	this manuscript	N/A
Recombinant DNA		
<i>pDGE332</i>	Stuttmann et al. ⁵⁰	ADDGENE#153241
<i>pDGE333</i>	Stuttmann et al. ⁵⁰	ADDGENE#153242
<i>pDGE335</i>	Stuttmann et al. ⁵⁰	ADDGENE#153244
<i>pDGE337</i>	Stuttmann et al. ⁵⁰	ADDGENE#153246
<i>pDGE651</i>	Stuttmann et al. ⁵⁰	ADDGENE#153229
<i>pAGM1263</i>	Engler et al. ⁵¹	ADDGENE #47985
<i>pICH41258</i>	Engler et al. ⁵¹	ADDGENE#47987
<i>pICH41276</i>	Engler et al. ⁵¹	ADDGENE#47994
<i>pICH41421</i>	Engler et al. ⁵¹	ADDGENE#50339
<i>pICH47732</i>	Engler et al. ⁵¹	ADDGENE#48000
<i>pICSL70002</i>	Engler et al. ⁵¹	ADDGENE#50333
<i>pICH47822</i>	Engler et al. ⁵¹	ADDGENE#48009
<i>pICH54044</i>	Engler et al. ⁵¹	ADDGENE#48068
<i>pICH47772</i>	Engler et al. ⁵¹	ADDGENE#48004
<i>pICH41800</i>	Engler et al. ⁵¹	ADDGENE#48020
<i>pAGM4723</i>	Engler et al. ⁵¹	ADDGENE#48015
Chemicals, peptides, and recombinant proteins		
<i>SUPERASE RNase inhibitor</i>	Invitrogen	CAT# AM2696
<i>spermidine</i>	Sigma-Aldrich	CAS#:124-20-9
<i>Spermine</i>	Sigma-Aldrich	CAS#:71-44-3
<i>cOmplete™, Mini, EDTA-free Protease Inhibitor Cocktail</i>	Sigma-Aldrich	CAS#:4693159001
<i>Optiprep</i>	Sigma-Aldrich	CAS # 92339-11-2
<i>Hoechst</i>	Invitrogen	Cat# H21491
<i>Nile Red</i>	Invitrogen	CAT# N1142
<i>Calcofluor White</i>	Sigma Aldrich	CAS #: 18909
<i>xylitol powder</i>	Sigma Aldrich	CAS#: 87-99-0
<i>sodium deoxycholate</i>	Sigma Aldrich	CAS#: 302-95-4
<i>Urea</i>	Sigma Aldrich	CAS#: 57-13-6
<i>Paraformaldehyde</i>	Sigma Aldrich	CAS#: 30525-89-4
<i>Fluorol Yellow-88</i>	Santa Cruz Biotechnology	CAS #: 81-37-8

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
D,L-lactic acid	Sigma Aldrich	CAS#:79-33-4
Aniline Blue	Sigma Aldrich	CAS 3: 62-53-3
Citifluor PVP-1+Antifadent	Electron Microscopy Sciences	CAT#17970-25
Klenow Fragment (3'-exo)	New England Biolabs	CAT# M0212S
Sera-Mag beads	Sigma Aldrich	PN:GE65152105050250
SPRIselect	Beckman Coulter	PN-B23317
Critical commercial assays/kits		
End-It DNA End-Repair Kit	Biosearch Technologies	PN-NEG772014MC
pENTR/D-TOPO Cloning Kit	Promega	PN-G8090
TNT SP6 Coupled Wheat Germ Extract System	Promega	PN-PRL4130
HS ds DNA Assay Kit	Invitrogen	CAT# Q32854
RNeasy Plant Mini Kit	Qiagen	CAT# 74904
Illumina TruSeq Stranded Total RNA with Ribo-Zero Plant kit	Illumina	CAT#RS-122-2401
DNA High Sensitivity Bioanalyzer Kit	Agilent	Part# 5067-4626
Magne HaloTag Beads	Promega	CAT# G7281
Chromium Next GEM Chip G Single Cell Kit, 48 rxns	10X Genomics	PN-1000120
Chromium Next GEM Single Cell 3' Kit v3.1, 16 rxns	10X Genomics	PN-1000121
Dual Index Kit TT set A, 96 rxns	10X Genomics	PN-1000213
Deposited data		
10X Genomics sequencing reads for snRNA-seq	This study	GEO:GSE261441
snRNA-seq gene counts and nucleus barcodes	This study	GEO:GSE261441
snRNA-seq serialized Seurat analysis object	This study	GEO:GSE261441
Illumina sequencing reads for bulk mRNA-seq	This study	GEO:GSE261437
Bulk mRNA-seq gene counts	This study	GEO:GSE261437
Bulk mRNA-seq differential expression estimates	This study	GEO:GSE261437
Illumina sequencing reads for DAP-seq	This study	GEO:GSE261431
DAP-seq MYB67 DNA binding genomic coordinates	This study	GEO:GSE261431
Predicted MYB67 DNA binding motif position weight matrix	This study	GEO:GSE261431
Predicted MYB67 DNA binding motifs genomic coordinates	This study	GEO:GSE261431
Software and algorithms		
FastQC	Andrews ⁵²	RRID:SCR_014583
FastQ Screen	Wignet and Andrews ⁵³	RRID:SCR_000141
Trimmomatic	Bolger et al. ⁵⁴	RRID:SCR_011848
STAR	Dobin et al. ⁵⁵	RRID:SCR_004463
Integrative Genomics Viewer	Thorvaldsdottir et al. ⁵⁶	RRID:SCR_011793
DESeq2	Love et al. ⁵⁷	RRID:SCR_015687
UROPA	Kondili et al. ⁵⁸	https://github.com/loosolab/UROPA/tree/master
HOMER	Heinz et al. ⁵⁹	RRID:SCR_010881
UniversalMotif	Tremblay ⁶⁰	https://doi.org/10.18129/B9.bioc.universalMotif
BEDtools	Quinlan and Hall ⁶¹	RRID:SCR_006646
The MEME Suite	Bailey et al. ⁶²	RRID:SCR_001783
pyfastx	Du et al. ⁶³	https://github.com/lmdu/pyfastx

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Seeds of *A. thaliana* (ecotype Col-0) were sterilized for 1.5h using chlorine gas (250ml 6% bleach with 3.5ml HCL) in an airtight chamber. Seeds were stratified at 4°C for three days before being sown on half-strength Murashige and Skoog (1/2 MS) media without sucrose at pH5.7. The plants were grown on vertically oriented plates at 21°C under 16/8 day-night cycle conditions. For the seedlings grown for single-nuclei sequencing, seeds were densely sown onto 1/2 MS plates with a mesh layer. This facilitated efficient root harvesting with minimal agarose carry over. Roots for single-nuclei experiments were harvested at 6, 7, 8, 9, 10, 11, 12 and 14 days after germination (DAG) into liquid nitrogen and stored at -80°C until further processing. For each timepoint, 30 plates were harvested. For all other experiments, *Arabidopsis* seedlings were grown under the same conditions but without mesh and with just 8 seedlings per plate.

METHOD DETAILS

Nuclei extraction and FACS

Nuclei extraction was performed as described by Lee et al.⁶⁴ For this, roots were ground to a fine powder in liquid nitrogen and then homogenised in Nuclei Extraction Buffer (NEB) (20mM MOPS pH7, 40mM NaCl, 90mM KCL, 2mM EDTA, 0.5 mM EGTA, 1% SUPERASE inhibitor (Invitrogen), 0.5mM spermidine, 0.2mM spermine, 1 x Roche Complete Protease inhibitor tablet) before being filtered through a 70µm filter. Nuclei were then further filtered through a 40µm filter and then centrifuged at 700 rcf for 5min at 4°C. Supernatant was discarded and the pellet resuspended in NEB supplemented with 0.1% triton. Samples were centrifuged again to wash. This was repeated for a total of two washes and the samples resuspended in 2ml NEB (without triton). Two solutions were made for the density gradients - a sucrose solution (400mM sucrose, 25mM KCL, 5mM MgCl₂, 10mM Tris-Cl (pH8)) and a diluent buffer (120mM Tris-Cl (pH8), 150mM KCL, 30mM MgCl₂). Prior to preparing the two density fractions, 1 volume of the diluent buffer was mixed with 5 volumes of Optiprep (Sigma). Density gradients were prepared in 15ml tubes with a 2ml 45% layer (1.8ml sucrose solution and 0.2ml dilutant buffer) and a 1ml 15% layer (0.3ml sucrose solution per 0.7ml dilutant buffer). The 2ml sample was added carefully to the top of the gradient and centrifuged at 1500 rcf for 5min without breaks. Nuclei were collected at the 45% interface and resuspended in NEB and centrifuged at 1000 rcf to collect nuclei. Samples were submitted to cell sorting in fresh NEB. Nuclei were stained using Hoechst. Stringent gating was used to remove debris and chloroplast signals were also detected and omitted from the sorted nuclei portion. Following the sorting of nuclei, samples were centrifuged at 1000 rcf and resuspended in the appropriate volume of 1xPBS for the 10X Genomics single-nuclei sequencing workflow.⁶⁵ Samples were sequenced using Novaseq with 1050 million paired end reads.

Generation and selection of transgenic lines

A subset of genes identified in the whole data DEG analysis were selected for use in our transcriptional reporter cluster validation experiments. For this subset of promising DEGs we cloned 1.5kb of their promoter sequence. All promoter sequences cloned were synthesised by IDT (Integrated DNA Technologies, Inc.). See [Data S1](#).

Golden Gate reporter lines were created according to the protocol published by Engler et al.⁵¹ Synthesized promoters, fluorescent protein (mCitrine, mScarlet, and mTurquoise) coding sequences, NLS tag, BASTA resistance cassette, and terminator sequence were each cloned into the level-zero (L0) acceptor, pAGM1263. L0 plasmids were then assembled into level-one (L1) constructs to produce complete transcriptional units with the following components: Synthesized promoter – fluorophore – NLS – terminator. L0 promoters were cloned into separate L1acceptor-fluorophore pairings (pICH47732+mCitrine, pICH47822+mTurquoise, pICH47772+mScarlet) to provide the option for future multigene/multifluorescent constructs to be produced. The L0 BASTA construct was cloned into L1 acceptor, pICH47742. Complete L1 gene constructs were each assembled into the Level 2 (L2) acceptor vector pAGM4723, along with a L1 BASTA cassette, to produce binary vectors.

A one-step digestion/ligation reaction was used for all Golden Gate assembly stages. Reagents used in all reactions were 1.5µl 10xbuffer for T4 DNA ligase, 3µl 5xUltrapure Bovine Serum Albumin, and 0.2µl T4 DNA Ligase (New England Biolabs). Acceptor vectors and DNA fragment/plasmid inserts were added in a 1:2 ratio [acceptor: insert(s)]. For L0 and L2 reactions, 0.25µl *Bpi*1-HF (New England Biolabs) was added as the chosen restriction enzyme. For L1 reactions, 0.25µl *Bsa*-HF (New England Biolabs) was used as the restriction enzyme. All reactions were made up to a final volume of 20µl using nuclease-free, filter and steam sterilized water. Reactions were completed in an Eppendorf Mastercycler Nexus X2 Thermal Cycler using the following settings: 20 seconds 37 °C, (3 min 37 °C, 4 min 16 °C) x 26cycles, 5 min 50 °C, 5 min 80 °C, 5 min 16 °C.

After each Golden Gate assembly reaction, 5µl of the digested/ligated product was added to 50µl of Top10 chemically competent *E. coli* cells (Thermofisher) and heat-shocked for 1 min at 42°C. 500µl of 2.5% Miller's LB (Research Products International) was immediately added to each mixture and subsequently incubated at 37°C for one hour. The transformed cells were spread on 1% Miller's LB-agarose (U.S. Biological) plates containing a 0.1% (v/v) antibiotic concentration corresponding to construct level (L0-spectinomycin, L1-carbenicillin, L2-kanamycin) and incubated overnight at 37°C. Multiple colonies of each construct were picked and cultivated in 5 ml of 2.5% Miller's LB (Research Products International), containing 0.1% (v/v) of their construct level's corresponding antibiotic, at 37°C overnight. Plasmid purification from all overnight cultures was completed using the NucleoSpin Plasmid Mini kit mid DNA according to manufacturers' Guidelines (Machery-Nagel). All final constructs were sequenced fully using Primordium sequencing.

MYB67 reporter lines were generated using Gateway cloning. Primer oligos were synthesized (Eton Bioscience) to amplify two genomic fragments: a 3.4kb section of MYB67's upstream promotor region (PRO), and the same promoter sequence connected to MYB67 coding region(PRO+CODING). Fragments were amplified from Col-0 gDNA using the Phusion High-Fidelity PCR Kit (New England Biolabs) following manufacturers guidelines and processed with the NucleoSpin PCR cleanup kit (Machery-Nagel) according to the provided instructions. Amplified MYB67 PRO and PRO+CODING fragments were put into pENTR/D-TOPO vector (kanamycin selection) (Thermofisher) following the manufacturer's reaction instructions. TOPO products were transformed into Top10 *E.coli* cells (Thermofisher) and processed as previously described in the Golden Gate cloning methods. Successful MYB67 PRO and PRO+CDS TOPO constructs were then cloned into the vector, pGWB504 (spectinomycin selection), with Gateway LR Clonase II Enzyme Mix (Thermofisher), following the manufacturer's instruction. LR reaction products were again transformed into Top10 cells and processed using the previously described steps to produce MYB67:eGFP and MYB67:MYB67-eGFP binary vectors for plant transformation.

Transgenic lines of *Arabidopsis* were made using the floral dip method.⁶⁶ For all plant transformation experiments, 1ul of binary vector plasmid was mixed with 50ul of *Agrobacterium tumefaciens* GV3101 competent cells and electroporated at 1.8kW for 3 seconds. 500uL of 2.5% Miller's LB (Research Products International) was immediately added to the cell mixture and incubated on a 28°C shaker for 1.5 hrs. The transformed cells were then spread on 1% Miller's LB agarose (U.S. Biological) plates containing a 0.1% (v/v) antibiotic concentration of gentamicin, rifampicin, and kanamycin (GoldenGate) or spectinomycin (MYB67-Gateway). Single colonies were picked from selection plates, added to 5ml of 2.5% Miller's LB (Research Products International) containing 0.1% (v/v) of each respective selection antibiotic, and incubated at 28°C overnight. These 5ml cultures were used to inoculate larger 300ml of the same solution and left again at 28°C overnight. The cultures were spun down and resuspended in 300mL of 5% sucrose solution containing 67.5μl of Silwet L-77 (PlantMedia). Newly developed Col-0 flower buds were submerged in the solution for 1min and subsequently kept in the growth chamber for 4-5 weeks for senescence to complete. Successful transformants were selected on 1% agarose (Phytotech), 2.22g/1L Murashige & Skoog Basal Salts (pH5.7-5.8) (Phytotech) media containing selection herbicide corresponding to the transformed vector. Subsequent generations were screened using the same selection system until a homozygous population was identified.

Generation of CRISPR/Cas9 k.o. lines for MYB67

Gene editing was used to generate additional *Arabidopsis myb67* mutant alleles. To delete MYB67, gRNAs flanking the gene were designed, where two gRNAs target the 5' end and two target the 3' end. gRNA spacers were designed using CRISPRdirect.⁶⁷ See Figure S3.

The MYB67 CRISPR binary vector was constructed using a multiplexing system from Stuttmann et al.⁵⁰ gRNA oligos were synthesized (IDT), annealed, and subsequently Golden Gate assembled into shuttle vectors (pDGE332, pDGE333, pDGE335, and pDGE337) using the previously described level-one reaction mixture and colony selection system. The four gRNA-containing shuttle vectors were then assembled into the recipient vector pDGE651, to create the final binary vector pAP169, using the level-two reaction mixture and colony selection system. All levels of CRISPR assembly reactions were completed with the following thermocycler settings: (2 min 37°C, 5 min 16°C) x 30 cycles, 10min 50°C, 10min 80°C.

A.tumefaciens [GV3101] cells were transformed with pAP169 and used for floral dip of Col-0 plants as previously described. Successfully transformed T1 seedlings were detected via hygromycin selection plates and transferred to soil. Transferred plants were screened for MYB67 deletions by PCR using primers outside the genomic region containing gRNAs. T2 seeds from plants showing MYB67 edits were screened for Cas9 presence using FAST fluorescent marker selection with a Zeiss AXIO Zoom V16 microscope. Cas9 negative seeds from lines showing a single insertion ratio of fluorescent signal [\sim 25% negative; \sim 75% positive] were transferred to soil and genotyped to identify *myb67* mutants. Absence of the transgene in T2/T3 plants was reconfirmed via PCR (shown in Figure S3). PCR products that indicated deletions of Myb67 were sequenced in the selected T2/T3 lines to confirm exact edits. The primers used in genotyping were as follows:

Upstream Flanking(UF): TTGGTTGAATCTCACACCCAT; Downstream Flanking(DF): TCGTGCTCCCTACTTTACCTC and Mid-Gene(MG): TGATTCCCTCACTGGTCAGATGT.

Nile Red and Calcofluor White staining and confocal imaging

To assess the development of periderm under our growth conditions, a confocal time course experiment was performed using a GPAT5:NLS-mCitrine line¹⁰ grown for 6, 7, 8, 9, 10, 11, 12, and 14 DAG. For this, seedlings were cut just below the hypocotyl and subsequently fixed by vacuum infiltrating the samples for 1.5h in a 4% Paraformaldehyde (PFA) in 1x PBS solution. Roots were cleared for a minimum of 24h at RT in ClearSee solution (10% (w/v) xylitol powder, 15% (w/v) sodium deoxycholate, and 25% (w/v) urea). For detection of suberin-containing mature periderm, as well as suberised endodermis cells, roots were stained in a 0.1% Nile RedTM solution (Invitrogen) (in ClearSee) overnight. The samples were subsequently washed 3x 20min in ClearSee, stained with Calcofluor White (Thermo Scientific) counterstain for 30min and again washed in ClearSee for another 30min. Root samples were imaged at 20x or 40x on a Zeiss 710 confocal. Nile red was excited at 561nm, mCitrine at 514nm and Calcofluor White at 488nm. For transgenic lines developed for cluster validation, mCitrine was visualised at 514nm, mScarlet at 561nm and mTurquoise at 488nm. Images made using this method can be found in Figure 1B.

Fluorol Yellow Staining, Imaging, and Analysis

The primary root of each seedling was cut just below the hypocotyl and all visible lateral roots were removed. Trimmed primary roots were placed in small mesh baskets within an 8-well rectangular dish and submerged in a 0.01% Fluorol Yellow-88 (Chem Cruz), 88% w/w D,L-lactic acid (Sigma-Aldrich) solution. The dish was incubated in a 70°C water bath for 20min in the dark. The root-containing baskets were then removed from the Fluorol Yellow solution, washed for 1min in deionized (DI) H₂O, and then resubmerged in a 0.5% Aniline Blue, 0.4% Acetic Acid (Sigma-Aldrich) solution for 20min in the dark. The baskets were then removed and washed in DI-H₂O for another 10min in the dark. Stained roots were mounted on slides (25 x 75 x 10mm) with Citifluor PVP-1+Antifadent (Electron Microscopy Sciences). The stained roots were imaged at 4x magnification in GFP Channel on BZ-X810 Keyence microscopes using Keyence BZ-X800 Viewer software.

RNA-seq

Col-0 and *myb67* mutant seeds were sown together on 1% agarose (Phytotech) media containing 2.22g/1L Murashige & Skoog Basal Salts (pH5.7-5.8) (Phytotech) in square (120 x 120 x 15mm) growth-plates (Greiner) for 14 days. Starting at the base of the hypocotyl, 1.5cm of mature primary root tissue, with lateral roots removed, was collected in 1.5ml microcentrifuge tubes (Eppendorf) (~40 seedlings per biological replicate) and flash frozen in liquid nitrogen. Frozen tissue was ground by hand into a fine powder using plastic pestles and RNA was isolated using RNeasy Plant Mini Kit (Qiagen) following the manufacturer's instructions. Col-0 and *myb67* purified RNA samples were submitted to the Salk Institute Next Generation Sequencing Core for cDNA library construction and sequencing using the Illumina TruSeq Stranded Total RNA with Ribo-Zero Plant kit and run on Illumina's Nextseq2000 (P2 flowcell, single-end, 100-cycles).

DAP-seq, cloning, immunoprecipitation, library construction and sequencing

Genomic DNA (gDNA) libraries for *Arabidopsis* were created following protocols in Bartlett et al.⁴⁶ and O'Malley et al.⁴⁷ with minor modifications. 10µg of *Arabidopsis* gDNA extracted from leaf tissue was fragmented to ~200bp in a Covaris S2 sonicator. The fragmented gDNA was end-repaired using the End-It DNA End-Repair Kit (Biosearch Technologies), A-tailed using Klenow Fragment (3'-5'exo) (New England Biolabs), and the adaptor was ligated using the truncated Illumina TruSeq Y-adaptor. The adapter-ligated library was confirmed with RTqPCR. The MYB67 cDNA was TOPO cloned into a pIX-Halo plasmid (ABRC, stock # CD3-1742) per manufacturer's instructions (ThermoFisher). This construct, along with an empty plasmid containing only the Halo tag, were incubated with the TNT SP6 Coupled Wheat Germ Extract System to produce Halo-MYB67 or Halo only proteins, respectively (Promega). 40µl of the resulting protein extracts were mixed with 10µl Magne HaloTag Beads (Promega) and 30ul of wash buffer (WB1; PBS + 0.005% NP40) on a rotator for 1h at room temperature. The beads with protein were washed with WB1, then incubated with 100ng of the adapter-ligated gDNA library in WB1 for 1hr. The beads were washed again with WB1 to remove unbound DNA fragments, then incubated at 98°C for 10min to denature the protein and release the bound DNA fragments. To amplify the released DNA, PCR reactions were prepared using Phusion DNA Polymerase (New England Biolabs), a universal primer, and a unique Illumina index primer, to allow multiple samples to be pooled in one sequencing run. PCR products were then size-selected using Sera-Mag beads (Sigma-Aldrich) at a 1.0x ratio. The purified DNA libraries were measured by the Qubit HS ds DNA Assay Kit (Thermofisher) and sequenced on Illumina's Nextseq2000 with 100 cycles and 50bp sequence length.

QUANTIFICATION AND STATISTICAL ANALYSIS

Single-nuclei data bioinformatic work-flow

For the single nuclei RNA-seq libraries, CellRanger (v.6.0.1)⁶⁸ was used to perform sample-demultiplexing, barcode processing and single-nuclei gene-UMI counting (a summary of the cell ranger output for each sample can be found in [Data S1A](#)). An expression matrix was obtained for each experiment by aligning to the *A. thaliana* genome (TAIR 10) using CellRanger with default parameters. For initial quality-control filtering, aligned nuclei and transcript counts from each experiment were processed by Seurat (Version 4.3)⁶⁹ by removing outlier nuclei in terms of the distribution of number of detected genes (nFeature) and transcripts (nCount) (Cutoffs: nFeature>4k & nCount>10k for data from Day 6&7; nFeature>3k & nCount>5k for data from Day8 to 12; nFeature>2k & nCount>5k for data from Day12). Low quality nuclei containing a high abundance of mitochondrial reads (>5% of total transcripts) and chloroplast reads (>5% of total transcripts) were also filtered. After initial filtering, 22,445 out of 23,747 cells were retained.

Expression data of nuclei passing these thresholds were log normalized with NormalizeData() function and the top 2K variable genes were identified with the FindVariableFeatures() function. Next, data from all experiments were integrated using Seurat's Reciprocal Principal component analysis (RPCA) and FindIntegrationAnchors() functions, to identify integration features and correct for potential batch effects. The integrated data was then scaled with ScaleData() function. PCA was carried out with the RunPCA() function and the top 30 principal components (PCs) were retained. Clusters were identified with the FindClusters() function by use of the shared nearest neighbour modularity optimization with a clustering resolution set to 0.5. Clusters with only one nucleus were removed. This method resulted in 26 initial clusters. To further check the quality of remaining nuclei, doublets were detected by applying the hybrid estimation scores from R package scds.⁷⁰ After doublet removal, the remaining nuclei were re-clustered with the top 30 PCs and a clustering resolution at 0.5. This resulted 19,755 nuclei distributed in 24 clusters (see [Figure 1C](#)). Pseudotime analysis was performed using Monocle 3.⁷¹ This was performed including just periderm-related clusters and using cluster 3, comprising pericycle-related nuclei as the progenitor cluster. Following the identification of the pseudotime trajectory, the nuclei

were divided into pseudobins. For this we divided our nuclei into 11 pseudobins of equal size. The division of the UMAP in this way very closely recapitulated the clustering achieved by high resolution (0.8) UMAP clustering, indicating that it appropriately summarises the variation in gene expression across the trajectory. See [Figure 4A](#).

DEG analysis

To identify the most effective marker genes to use for cluster validation, a Differential Gene Expression (DEG) analysis was performed using the Seurat “FindAllMarkers” function. For this we set the minimum percentage of nuclei expressing the DEG and a minimum difference relative to other clusters as 25%. The most significant DEGs were further queried visually using both Feature plots and Dot plot functions in Seurat. A list of all DEGs identified across clusters is provided in [Data S1C](#). A further DEG analysis was later performed on a subset of clusters confirmed to be related to the periderm. These periderm-related DEGs are summarised in [Data S2A](#). [Data S2B](#) shows the results of Gene Ontology (GO) analysis which was performed on significant DEGs (adjusted pValue <0.05) using agriGO v2.0²⁹ and PANTHER²⁸. The top 200 DEGs for each cluster were also plotted in form of a heatmap ([Figure 4D](#)) using the Heatmap() function in the ComplexHeatmap package (V2.16.0).^{72,73}

The transcription factors identified through DEG analysis were identified by providing a list of all DEGs to the PlantTFDB tool.⁴² These TFs are listed in [S2C](#) and a subset labelled on [Figure 4D](#).

Mutant phenotype quantification

Col-0 and *myb67* mutant Keyence images were processed manually using ImageJ.⁷⁴ Image files were randomized using an in-house R script (https://osf.io/uaxbg/?view_only=57ebd24cd6464ac889c816c313325da6) to reduce bias during whole-root, periderm, and phellem cell length measurements. The script duplicated the image files, matched the file names with random numbers generated by the function sample(), and renamed the files with their corresponding random numbers. The renamed, numbered image files were then used to acquire length measurements in ImageJ. After processing, the resulting measurements were matched back together with the file names for analysis. The length of each primary root was measured from the base of the hypocotyl to the root tip. Seedling periderm length was determined by the presence of stained, suberin containing, phellem cells starting at the base of the hypocotyl and measured in the direction of the root tip. Each root’s periderm region was then marked at 10% intervals of the previously determined length. Phellem cells intersected by marker-lines were measured at their longest point. Two-tailed student t-tests were used to compare whole-root, periderm, and phellem cell lengths between Col-0 and *myb67* mutant seedlings (all raw data related to root phenotyping can be found in [Data S3](#)). The mean values with standard error bars are presented in [Figures 5C and 5D](#), where “n” refers to the number of roots used in the quantifications. For phellem cell measurements at least 100 cells were measured for each root sample.

Bulk RNA-seq data processing workflow

Sequence reads for both experiments were downloaded from the Salk Sequencing Core and read quality was judged to be acceptable using FASTQC⁵² and FastqScreen.⁵³ A summary of the bulk RNA-seq sequencing runs can be found in [Data S4A](#) and [S4B](#)). Low quality and adapter sequences were removed using Trimmomatic v0.39.⁵⁴ The parameters used were ILLUMINA_CLIP:TruSeq-fasta:2:30:10 and SLIDINGWINDOW:4:15, with adapter sequences for paired-end reads provided with the Trimmomatic release. Trimmed reads were aligned using STAR v2.7.10a⁵⁵ to a suffix array representation of the *Arabidopsis thaliana* TAIR10 assembly⁷⁵ created using STAR –runMode genomeGenerate with default parameters and –genomeSAindexNbases 12. STAR –runMode alignReads was executed using default parameters and –outFilterMultimapNmax 1 to report only alignments from uniquely mapping reads. Visualizations of alignment coverage were generated using Integrative Genomics Viewer.⁵⁶ First, SAM-format alignment data were converted to BedGraph format using makeUCSCfile from the HOMER⁵⁷ collection then converted to TDF format using igvtools⁵⁶ and loaded into IGV. Expression was quantified based on transcripts from the TAIR10 genome annotation using analyzeREpeats.pl from the HOMER suite. DEGs were determined using DESeq2 v1.40.1⁵⁷ with the design formula ~ batch + group where libraries from respective experiment were classified as batch and *myb67* mutant and WT libraries were classified as group. False discovery rates (FDRs) were estimated using the Benjamini-Hochberg method and genes with a fold change magnitude of at least 1.5 and FDR less than 0.05 were reported ([Data S4C-S4F](#)). Periderm-specific DEG were determined to be those with expression in at least 5% of the nuclei in clusters 9 or 15 and less than 5% mean percentage of expression across all other clusters. RNA-seq tracks for a subset of genes are shown in [Figures 5H and S5](#).

DAP-seq data processing workflow

Sequence reads were downloaded from the Salk Sequencing Core and read quality was judged to be acceptable using FASTQC⁵² and FastqScreen.⁵³ A summary of the DAP-seq sequencing runs can be found in [Data S5A](#)). Low quality and adapter sequences were removed using Trimmomatic v0.39.⁵⁴ The parameters used were ILLUMINA_CLIP:TruSeq-PE.fasta:2:30:10 and SLIDINGWINDOW:4:15, with adapter sequences for paired-end reads provided with the Trimmomatic release. Trimmed reads were aligned to the *Arabidopsis thaliana* TAIR10 assembly using STAR v2.7.10a.⁵⁵ The STAR index was created using default parameters and –genomeSAindexNbases 12, and STAR was run with default parameters and –outFilterMultimapNmax 1 to report only alignments from uniquely mapping reads. Duplicate alignments were removed using the Picard Tools v2.27.1 program MarkDuplicates ([Broad Institute.Picard Tools. Broad Institute, GitHub repository. <http://broadinstitute.github.io/picard/>]). Visualizations of alignment coverage were generated using Integrative Genomics Viewer.⁵⁶ First, SAM-format alignment data were converted to BedGraph

format using makeUCSCfile from the HOMER collection using the parameter -fragLength 150 to smooth the peaks by extending alignment tags to 150bp, then converted to TDF format using igvtools and loaded into IGV.⁵⁶ Regions with enriched read alignments (peaks) were identified using macs3 v3.0.0a7 (q value = <0.05) with replicate libraries for both the MYB67 and the empty vector samples used as inputs.⁷⁶ Genes from the Araport11 gene annotation⁷⁷ were associated with peaks using UROPA v4.0.2⁵⁸ following the rule: if a peak falls within 2000bp upstream or 500bp downstream of the TSS or within the gene body, then the gene is considered to be a putative target of the MYB67, with ties going to the closest gene. A summary of the DAP-seq peaks identified are listed in Data S5B and S5C. A subset of DAP-seq peaks are also presented in Figure 5I.

Predicting MYB67 binding motifs

De novo motif discovery was carried out using findMotifsGenome.pl⁵⁹ of the HOMER package. Peak sequences were searched for overrepresented 5-13 bp sub-sequences. The null distribution consisted of GC-normalized random draws from the entire genome sequence, enrichment was determined using the binomial test, and enriched motifs were compared to remove redundancy. One motif was highly prevalent among peaks and was visualized using the R universalmotif package.⁶⁰ The positions of this motif within the peak loci were determined using fimo⁸ as follows. Peak loci sequences were obtained using BedTools⁶¹ the HOMER-style motif position weight matrix was converted to meme format using chen2meme from the MEME suite⁶² and a log-likelihood ratio score was computed by fimo for each peak subsequence, determining a p-value based on a null model of randomly generated sequences using the background base frequencies calculated from the entire genome using pyfastx.⁶³ Bootstrapping was used to estimate false discovery rates (FDR, q-value) and motifs with an FDR of less than 0.05 were retained. The sequence of this motif is given in Figure 5H.