Single-cell and tissue transcriptomes unveil factors regulating vein initiation and development in grass leaf primordia

Juan Yi[1,3,8], Yonghe Chen[1,3,8], Shilong Tian[1,8], Daniela Vlad[2], Jianzhao Yang[1,3], Olga Sedelnikova[2,7], Caiyao Zhao[4], Hong Su[1,3], Wentao Dong[5], Yijing Zhang[6], Xin-Guang Zhu[1], Jane A. Langdale[2], Jia-Wei Wang[1], Peng Wang[1]*

[1] Key Laboratory of Plant Carbon Capture, CAS Center for Excellence in Molecular Plant Sciences, Institute of Plant Physiology and Ecology, Chinese Academy of Sciences, Shanghai, China
[2] Department of Biology, University of Oxford, South Parks Rd, Oxford, OX1 3RB, UK
[3] University of Chinese Academy of Sciences, Beijing, China
[4] BGI Research, Wuhan, China
[5] State Key Laboratory for Crop Genetics & Germplasm Enhancement and Utilization, Nanjing Agricultural University, Nanjing, China
[6] State Key Laboratory of Genetic Engineering, Collaborative Innovation Center of Genetics and Development, Department of Biochemistry, Institute of Plant Biology, School of Life Sciences, Fudan University, Shanghai, China
[7] Current address: Syngenta Jealott's Hill International Research Centre, Bracknell RG42 6EY, UK
[8] These authors contributed equally to this work

* Author for correspondence: Peng Wang
Email: wangpeng@cemps.ac.cn

Short title: Single-cell transcriptomes unveil factors regulating vein development

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (https://academic.oup.com/plcell/pages/General-Instructions) is: Peng Wang (wangpeng@cemps.ac.cn).

Summary

Compared to the C_3 plant rice, veins in the leaves of the C_4 plant maize are more densely arranged, but the regulatory mechanisms controlling their initiation and differentiation remain unclear. This study systematically investigated the developmental patterning of vein formation in leaf primordia of maize and rice, capturing the stages of vascular initiation and extension through transcriptome analyses. The single-nucleus transcriptomic atlas of maize primordia revealed high cellular heterogeneity, with distinct middle ground tissue and procambium cell populations identified and a role for auxin in vein initiation inferred. Pseudo-time developmental trajectory analysis further facilitated the identification of marker genes in ground tissue and procambial cell types including *ETHYLENE-RESPONSIVE ELEMENT BINDING 114* (*ZmEREB114*), *AUXIN IMPORT CARRIER 3* (*ZmAIC3*), *ZmEREB161*, and *AUXIN AMIDO SYNTHETASE 2* (*ZmAAS2*). To test the role of auxin in grass leaf vein formation, an experimental system was established in which the interplay between auxin and the expression of *SHORTROOT1* (*ZmSHR1*) was used to suppress or restore vein formation in rice leaf primordia. Notably, expression of the rice genes *OsGH3.8* (homologous to *ZmAAS2*) and *OsAUX1* (homologous to *ZmAIC3*) was suppressed by *ZmSHR1* but promoted by auxin, conditions in which vein formation was repectively suppressed or rescued within the primordium. Collectively these results provide single-cell resolved resources for early stages of grass leaf development, an experimental system for manipulating vein initation in rice, and a model to explain how the interplay between SHR function and auxin response dynamics regulates vein patterning in maize and rice leaves.

Introduction

The vascular tissue in plants provides mechanical support and functions to transport water, nutrients, hormones, and other signalling molecules throughout the plant (Canny, 1998; Hoad, 1995; Lucas et al., 2013). In leaves, the veins

play a crucial role in transporting sugars out of photosynthesizing (source) tissues into non-photosynthesizing (sink) tissues and organs. In flowering plants, distinct patterns of veins are seen in eudicotyledon versus monocotyledon leaves, with eudicot leaves displaying reticulate leaf vein patterning and monocot leaves exhibiting parallel leaf venation (Nelson and Dengler, 1997; Perico et al., 2022). To date, investigations of leaf venation patterning have been primarily focused on the development of reticulate networks in the model eudicot *Arabidopsis thaliana* (Scarpella et al., 2006; Sawchuk et al., 2007; Donner et al., 2009; Biedron and Banasiak, 2018), and as such our knowledge of how vein formation is regulated in the leaves of major monocot crop plants such as maize and rice is limited.

Maize and rice are both grasses and as such, some aspects of leaf development are shared between the two species. In the context of vein formation for example, the midvein develops first, extending from the base of the primordium to the tip, the major rank veins develop next, again from the base to the tip, and finally the minor rank veins develop from the tip of the primordium to the base (Sharman, 1942; Nelson and Dengler, 1997; Sedelnikova et al., 2018). Other aspects differ between the two species, however, one of which is the density at which veins form. This difference is associated with the fact that rice photosynthesizes using the ancestral and more common C_3 pathway whereas maize utilizes the C_4 pathway, a pathway that has evolved independently over 60 times (Sage et al., 2012). Closely spaced leaf veins have been identified as a critical pre-adaptive trait for C_4 evolution across diverse phylogenetic contexts (Christin et al., 2013), and the leaves of C_4 plants are characterized by the formation of a greater number of minor veins between each pair of major veins. Understanding the determinants of closely spaced vein patterning in C_4 plants would not only facilitate efforts to introduce C_4 traits into C_3 plants but could also contribute to the optimization of

traits in C_3 plants (Sedelnikova et al., 2018; Kumar and Kellogg, 2019; Perico et al., 2022).

Our previous study comparing transcriptional profiles of maize foliar and husk leaf primordia (with and without close vein spacing) identified potential regulators of vein spacing that included the SHORTROOT (SHR) / SCARECROW (SCR) regulatory module (Wang et al., 2013; Fouracre et al., 2014). Cell type-specific transcriptomes of early maize leaf development further identified genes encoding auxin transporters (e.g. PINFORMED (PIN)1a, PIN1d, and LIKE AUXIN RESISTANT2 (LAX2)) as well as AINTEGUMENTA（ANT）family transcription factors (ZmANT1, ZmANT2, ZmANT3, ZmANT3-1) (Liu et al., 2022; Perico et al., 2024). In terms of gene function, recent work indicated that a SHR-INDETERMINATE DOMAIN (IDD) regulatory circuit mediates auxin transport by negatively regulating PIN expression to modulate minor vein patterning in leaves of both C_3 and C_4 grasses (Liu et al., 2023); while the combined action of SCR and NAKED ENDOSPERM (NKD) IDD controls the number of mesophyll cells specified between veins in the leaves of C_4 but not C_3 grasses (Hughes et al., 2023). Furthermore, in the C_4 plant *Setaria viridis*, loss-of-function mutations in the *ANT1* gene lead to reduced leaf size and abnormal internal anatomy (Liu et al., 2020). Transcription factors such as ANT, SCR, and SHR, along with the auxin pathway thus play crucial roles in leaf vein development in grasses but the specific mechanisms of action and interactions between components of the regulatory network remain to be elucidated.

Here, we have investigated the early stages of leaf development in maize and rice at tissue and/or single cell levels. We first identified and separated the regions of leaf primordia in which veins are being initiated, and then performed comparative transcriptomic analysis with corresponding rice tissues. We then

used single-nucleus RNA sequencing (snRNA-seq) to investigate the critical developmental period within maize leaf primordia. These approaches together identified the gene expression signatures of both the ground meristem layer and the procambial initial cells that arise from it to form veins, and enabled the construction of a co-expression network of factors that may regulate procambial specification and vein initation. Through the development of an experimental system to manipulate vein formation in rice, we further demonstrated that the combined activity of SHR and auxin influences whether these candidate regulatory factors are expressed, providing crtitical insight into how C_4 venation patterns might be engineered in rice.

Results

Analysis of venation patterns in maize and rice leaf primordia and spatiotemporal quantification of vein numbers

Vein formation in maize leaves initiates early during the development of leaf primordia from the shoot apical meristem (SAM) (Sharman, 1942; Johnston et al., 2015). The time interval between the formation of successive primordia from the SAM is referred to as a plastochron (P), and the stage P1 primordium is the youngest/closest to the SAM. Stages P3 to P5 represent a phase of active proliferation of leaf veins (Robil and McSteen, 2023). We focused on these stages of leaf development and utilized maize *pZmPIN1a::ZmPIN1a:YFP* and rice *DR5::VENUS* fluorescent reporter lines to visualize the distribution of veins across the medio-lateral and proximo-distal leaf axes **(Figure 1A, B; Supplemental Figure 1)**. In combination with serial transverse sections from paraffin-embedded samples which show cellular arrangements in the adaxial (upper)-abaxial (lower) leaf axis, we first quantified the number of leaf veins. In 2-mm-long maize leaf primordia, the maximum number of veins across the medio-lateral leaf axis reaches 20 at the widest point **(Figure 1A)**, whereas in similarly sized rice leaf primordia, the maximum count is 13 **(Figure 1B)**. At the

tip of maize leaf primordia an initial ‘Kranz’-like arrangement is observed, whereby veins (V) are surrounded by a ring of bundle sheath (BS) cells which are themselves surrounded by a layer of mesophyll (M) cells (Brown et al., 1975; Fouracre et al., 2014) **(Figure 1C)**. By contrast, sections from the middle of the proximo-distal leaf axis display very few intermediate veins between lateral veins, indicative of an active vein formation stage **(Figure 1D),** and sections from the basal region have yet to develop any intermediate veins between lateral veins **(Figure 1E)**. Serial sections of rice leaf primordia demonstrate a similar arrangement of midvein, lateral and intermediate veins along the proximo-distal leaf axis but the developing intermediate veins are visble closer to the leaf base than in maize **(Figure 1F–H)**. This comparison indicates a more prolonged period of vein formation in maize compared to rice.

To examine changes in vein number as the leaf primordium elongates and widens, maize leaves were sampled from the primordium to expanded leaf stage **(Figure 1I).** The number of veins increased from approximately 27 in a 1.5-mm-long primordium (with a similar width of about 1.5 mm) to around 213 in an unfolded leaf blade with a width of approximately 2 cm. The growth curve in **Figure 1J** indicates that the number of veins increases most rapidly when the leaf primordium reaches 3–5 mm in length, corresponding to the active phase of intermediate vein development. In contrast, changes in vein numbers in rice are less prominent **(Figure 1F)**. As such, maize leaf primordia exhibit a higher frequency of procambium initiation and a more pronounced gradient of change than rice leaf primordia, and thus maize provides a more accessible system for capturing the dynamic processes of procambial cell initiation and differentiation.

Transcriptomic profiling of segmented leaf primordia reveals differentially enriched pathways between maize and rice

To identify gene expression changes associated with the initiation and development of leaf veins, we performed sub-tissue segmentation of leaf primordia and transcriptomic analysis. Micro-dissected 5 mm long maize and rice leaf primordia were divided into upper (tip), middle (middle) and lower (base) segments **(Figure 2A,B)**. The differentially expressed genes between tip, middle, and basal segments of maize leaf primordia were classified into eight co-expression modules **(Figure 2C)**. Modules C1, C6, and C8 comprised genes that were highly expressed specifically in the tip region. GO enrichment analysis indicated that these modules were significantly associated with photosynthesis and chlorophyll biosynthesis. Module C3 genes showed higher expression in both the tip and middle segments compared to the base and were enriched for processes including epidermal cell development, cell differentiation, and gibberellin response. By contrast, genes in module C2 were relatively highly expressed in the base and enriched in terms such as *cis*-regulatory region sequence-specific DNA binding and nucleobase transmembrane transporter activity. Module C5 genes exhibited high expression in both the base and middle segments and were primarily enriched in the auxin-activated signaling pathway and cellular response to auxin stimulus. This module included representative transcription factors and auxin-related genes, such as *ZmEREBs*, *ZmIAAs*, *ZmPINs*, *ZmARF35*, *ZmC2H2-46*, and *ZmDOF43*. Differentially expressed genes between the tip, middle, and basal segments of 5-mm rice leaf primordia were also grouped into eight co-expression modules **(Figure 2D)**. A notable difference between the maize and rice leaf modules was that no significant enrichment of auxin-related genes was observed in the middle segment of rice leaf primordia. There is thus a correlation between the expression of auxin-related genes and the extent of procambium initiation, with higher levels of both observed in maize than rice.

Construction of a single-nucleus transcriptome atlas of maize leaf

primordia

Given that procambium initiation occurs in individual cells of the middle ground meristem layer in the adaxial-abaxial leaf axis, we investigated the transcriptional features of this process at single-cell resolution. To this end, we performed single-nucleus RNA sequencing on 3–4 mm leaf primordia, a stage characterized by active procambium initiation. The cellular anatomy of leaf primordia at this stage is such that five cell layers are identifiable across the adaxial-abaxial leaf axis: the upper and lower epidermis and three internal cell layers. The veins, bundle sheath cells and mesophyll cells between veins all differentiate from the middle ground meristem (mGM) (Esau, 1943; Langdale et al., 1989; Bosabalidis et al., 1994), which locates among the middle layer cells **(Figure 3A)**. Cells adjacent to the upper and lower epidermis resemble palisade mesophyll precursors, while parenchyma cells with typical characteristics are present surrounding the midvein in the basal region of the leaf primordium **(Figure 3A)**. We obtained 7,476 effective cells with an average of 2,312 genes expressed in each nucleus. Principal component analysis and unsupervised clustering categorized 14 distinct cell clusters. Dimensionality reduction was performed using the Uniform Manifold Approximation and Projection (UMAP) algorithm **(Figure 3B)**. By analysing DEGs between the clusters, we identified a series of cluster-enriched or cluster-specific genes. Dot plots of the top 10 marker genes from each of the different cell clusters were generated **(Supplemental Figure 2)** and a subset is shown in **Figure 3C**.

Using previously reported marker genes, cell clusters 0 and 6 were identified as epidermal cells, Cluster 11 as primary phloem, Cluster 12 as primary xylem, and Cluster 13 as undefined vascular cells **(Figure 3C, D; Supplemental Figure 3)**. Cluster 8 was identified as parenchyma and/or middle layer mesophyll cells, and Cluster 10 was as palisade mesophyll cells. These annotations were verified by selecting the enriched marker genes for *in situ*

hybridization **(Figure 3C, D; Supplemental Figure 4)**. *ZmWIP1* (encoding a Bowman-Birk type wound-induced proteinase inhibitor) (Rohrmeier and Lehle, 1993) which was enriched in Cluster 8 (also present in Cluster 0) was distributed in the swollen parenchyma cells (also detected in epidermal cells) close to the midrib and in the adaxial epidermis **(Supplemental Figure 4)**. *GLU3* which was enriched in Cluster 10 and encodes a β-glucosidase (Gómez-Anduro et al., 2011), was specifically expressed in mesophyll cells adjacent to the abaxial epidermis, or in mesophyll cells both adaxially and abaxially close to the lateral veins **(Figure 3C, D; Figure 4A, B)**.

Both Clusters 1 and 5 showed enrichment of the auxin transporter gene *ZmAIC3* (auxin import carrier 3), and *in situ* hybridization revealed transcript localization in middle layer cells **(Figure 3C, D)**. Because cells in Clusters 1 and 10 were enriched with *ZmDEF4* (defensin-like protein 4), and *ZmDEF4* transcripts were detected in mesophyll cells surrounding the veins (consisting of the middle layer and palisade mesophyll cells) **(Figure 3C, D; Figure 4A, B)**, Cluster 1 was preliminarily annotated as middle ground tissue. The vascular and procambium marker gene *ZmPIN1* was enriched in Clusters 5 and 12 **(Figure 3C)**, and both *ZmYAB3* and *ZmGLK6*, which were enriched in Clusters 5 and 13, were verified via *in situ* hybridization to be expressed in procambium. As such, Cluster 5 was annotated as procambial cells **(Figure 3C, D; Supplemental Figure 5)**. Cluster 9, remains to be further characterized but may contain a group of cells at the developmental junction between mGM and mesophyll/parenchymal cells. Cluster 4 cannot be annotated but is enriched with ribosomal protein-related genes. *In situ* hybridization of selected genes from Cluster 4 detected transcripts in veins, in cells around veins, and in the epidermis **(Supplemental Figure 6)**. For the remaining clusters, the presence of cell division markers suggested that Cluster 3 was likely to represent proliferative cells in S phase, and Clusters 2 and 7 proliferating cells in G2/M

phase **(Figure 3C; Supplemental Figure 6)**.

Molecular signatures of the three inner cell layers in maize leaf primordia

Due to the pronounced developmental gradient along the proximo-distal axis of maize leaf primordia cells are more differentiated at the leaf tip than at the base. To identify markers of cells in the different leaf layers of the adaxial-abaxial axis, we examined transcript accumulation patterns at two points along this proximo-distal gradient in P4 primordia. Analysis of *ZmGLU3* expression patterns via *in situ* hybridization revealed that the gene was expressed in abaxial mesophyll precursors in the mid-to-lower regions of the leaf but in more differentiated regions expression was detected in both abaxial and adaxial cells around developing veins **(Figure 4A**, left**)**. *ZmGLU3* is thus a marker of cells in two of the three internal leaf layers. By contrast, *ZmDEF4* transcripts were detected in a ring of mesophyll cells surrounding the veins that included all three (adaxial, abaxial and middle) layers of mesophyll cell precursors **(Figure 4A**, central**)**. Of particular interest was expression of the maize homolog of *Arabidopsis SCPL46* (serine carboxypeptidase-like 46). In the lower middle region of the leaf primordium, transcripts were detected in the middle layer of cells between veins but also weakly in adaxial mesophyll precursors. Towards the tip, however, where the number of intervening cells between veins narrows to two, the expression of the *AtSCPL46* homolog was restricted to these intervening cells **(Figure 4A**, right**)**. Together, therefore, the combined accumulation patterns of *ZmGLU3*, *ZmDEF4*, and the homolog of *AtSCPL46* distinguish the abaxial and adaxial cell layers around veins (*ZmGLU3*, *ZmDEF4)* from the middle layer between veins (*ZmDEF4*, homolog of *AtSCPL46)* and from the ground meristem cells that give rise to procambial initials (homolog of *AtSCPL46)* **(Figure 4B)**.

Having identified markers of all three internal leaf layers, we next sought to

identify expression patterns in the middle leaf layer that were associated with the initiation of procambium. Enrichment in Clusters 1 and 5 of the *AtANT* homologs *ZmEREB114* and *ZmEREB161,* the previously reported vein development-related gene *TOO MANY LATERALS* (*ZmTML1*) (Vlad et al., 2024), the auxin transporter *ZmPIN1* and the GRAS transcription factor *ZmSHR1* were consistent with roles for these genes in vein development. UMAP analyses (**Figure 4C, D**) further showed the relative distribution of these genes and the *AtSCPL46* homolog across Clusters 1 and 5. Notably *ZmEREB114* and *ZmEREB161* were enriched in Clusters 1 and 5, and also in Cluster 3, suggesting a possible role in general cell proliferation in developing veins. In contrast, *ZmPIN1* and *ZmTML1* were preferentially localized in Cluster 5 suggesting a role in early procambium specification whereas *ZmSHR1* was localized in procambium but also in phloem and xylem, suggesting a role in vascular tissue differentiation.

Expression of *ZmEREBs* and *ZmSHR1* in the middle cell layer of the leaf during procambium initiation

To characterize patterns of gene expression in the middle leaf layer during the initiation of procambium, *in situ* hybridization was carried out with *ZmEREB114*, *ZmEREB161* and *ZmSHR1*. Whereas *ZmEREB114* and *ZmEREB161* transcripts were localized to developing veins towards the tip of the primordium, both were detected more uniformly in the lower middle region. *ZmEREB114* expression was not only detected in procambial cells (white triangle indicated), but also observed in the three or four-contiguous cells (orange triangle indicated) between existing procambial centres **(Supplemental Figure 7A-E; Figure 5A, B)**. *ZmEREB161* exhibited a more restricted distribution within the middle layer compared to *ZmEREB114*. In addition to its expression in procambial cells, it was also detected in individual or pairs of adjacent cells within the middle layer **(Figure 5A, C-F)**. This

expression pattern aligned with the observation that *ZmEREB114* was distributed across both Clusters 1 and 5, whereas *ZmEREB161* showed a stronger enrichment in Cluster 5 **(Figures 3C and 4D)**. In contrast to both, *ZmSHR1* was consistently detected in procambium cells in both the upper middle and lower middle primordium regions **(Figure 5A)**.

Using *ZmSHR1* as a comparitor, we next examined the expression of *ZmEREB161* in the context of cell division and arrangement patterns. *In situ* hybridization showed that *ZmEREB161* transcripts accumulated not only at different stages of procambium development (white triangle indicated), but also in mGM cells that had yet to undergo the periclinal divisions required for specification of procambial initials (orange triangle indicated, may be equivalent to one of the three-contiguous cells, at 3C stage, in Liu et al., 2022) **(Figure 5C-F,K)**. The periclinally divided cell in the middle of the three adjacent mGM cells produced two cells, with one of them expressing the *ZmSHR1* gene **(Figure 5G,L)**. With the second periclinal division, *ZmSHR1* gene expression was detected in the middle of the three progeny cells **(Figure 5H,L)**. During subsequent cell division and differentiation, *ZmSHR1* was confined to procambial cells within veins **(Figure 5I,J,L)**. Therefore, the spatiotemporally staggered expression of *ZmEREB114*, *ZmEREB161*, and *ZmSHR1* reflected the highly dynamic state of proliferation and differentiation in middle layer cells.

Pseudo-time analysis reveals dynamic differentiation stages during procambium initiation in the middle cell layer of the primordium

Not all cells within the middle layer are actively differentiating middle ground meristem (mGM) cells. Due to their poorly distinct differentiation states, these cells may not be fully separable by single-cell clustering alone. Therefore, we utilized pseudo-temporal developmental trajectory analysis to further resolve the heterogeneous states among middle layer cells. First, we extracted cells

from the middle ground tissue (Cluster 1 in **Figure 3B**) expressing *ZmEREB114* (with expression level > 2.8) and the homolog of *AtSCPL46* (expression level > 3), along with procambium and vascular cells (specifically clusters 5, 11, 12, and 13), for sub-clustering analysis. By leveraging previously reported marker genes and *in situ* hybridization-validated genes from this study, we annotated the subclustered cell populations as follows: sub-clusters 0 and 1 as procambium, sub-cluster 3 as “middle ground tissue*” (*indicated a subset of cells obtained from the middle ground tissue as described above), sub-cluster 2 as xylem, sub-cluster 4 as phloem, and sub-cluster 5 as vascular bundle cells **(Figure 6A)**.

The sub-clustered cells were subjected to unsupervised pseudo-temporal trajectory analysis using Monocle2 and Monocle3 **(Figure 6B; Supplemental Figure 8)**. The trajectory constructed by Monocle2 revealed a branch point (Branch Point 1) where cell fates diverged toward vascular tissue and procambium cells. At Branch Point 2, vascular bundle cells further differentiated into phloem and xylem lineages. Along the pseudo-temporal axis, all cells could be categorized into five distinct differentiation states. mGM cells should be predominantly concentrated in the early pseudo-time (State I). Procambium cells spanned almost the entire pseudo-temporal trajectory **(Figure 6B, C)**. There was also a special differentiation state composed of some of the procambial cells corresponding to State V, named as “procambium (State V)” **(Figure 6D)**.

Based on the pseudo-temporal trajectory, the dynamic expression patterns of key genes were inferred. At Branch Point 2, xylem-related genes including *ZmTMAAT*, *ZmbHLH143* and *ZmWAT1* initially exhibited low expression but gradually increased along the xylem differentiation branch. Similarly, phloem marker genes such as *ZmSMXL3*, the homolog of *AtNEN3*,

and *ZmZNF30* showed low initial expression that elevated along the phloem branch over pseudo-time **(Supplemental Figure 9A)**. These expression patterns supported the biological validity of our pseudo-time developmental trajectory. Whereas *ZmEREB114* was highly expressed from the initial stage and maintained a comparable level along the "procambium (State V)" branch, *ZmEREB161* started at a relatively lower expression level and showed a continuous increase **(Figure 6E)**. The expression pattern of *ZmAAS2* (auxin amido synthetase 2), which was identified in State V, differed notably: it began at low levels, increased and stabilized, then increased again specifically along the "procambium (State V)" branch **(Figure 6F)**. *In situ* hybridization results showed that *ZmAAS2* signals were detected in individual cells, pairs of adjacent cells, and derivatives of periclinal divisions of procambial initial cells **(Figure 6G)**. As expected, *ZmPIN1* expression began at very low levels and gradually increased along both branched pseudo-temporal trajectories **(Figure 6F)**.

We proposed that cells in State V may represent a developmental status close to mGM, and that genes enriched in State V cells warrant further investigation **(Supplemental Figure 9B)**. Branch Expression Analysis Modeling (BEAM) identified 335 genes with significantly differential expression between the pseudo-temporal branches. These genes were clustered based on their expression patterns, and those similar to *ZmEREB161* and *ZmAAS2* were selected for comparative analysis with cell type-enriched genes from the study by laser capture microdissection (LCM)(Liu et al., 2022) on maize leaf primordia. Most of these genes were expressed in mGM and 3C cells as described by Liu et al.. Notably, *Zm00001d049399* was expressed in bundle sheath cells and veins (BS + V), whereas *Zm00001d034126* was detected in middle layer mesophyll cells (M and 2M) **(Supplemental Figure 10A, B)**. The "2M" samples of the LCM data further enabled us to determine different types of mesophyll cells for the single-cell clusters **(Supplemental Figure 10C-F)**. Accordingly, our

Cluster 8 was identified to be related with middle layer mesophyll in addition to parenchyma cells **(Figure 3B)**.

Co-expression regulatory network of middle leaf layer associated genes

To explore the gene regulatory networks involved in the transition from middle ground tissue to procambium initial cells, we conducted a co-expression correlation analysis of transcription factors (21), protein kinases (11), and auxin-related genes (10) enriched among the “middle ground tissue*” and procambium cells. The top four transcription factors with the highest gene co-expression connectivity (number of correlated genes) were *ZmEREB114*, *ZmEREB41*, *AUXIN RESPONSE FACTOR TF25* (*ZmARFTF25*), and *INDOLE-3-ACETIC ACID 23* (*ZmIAA23*). These genes were all expressed during the early pseudo-time stage **(Figure 6H, I)**. *ZmEREB114* exhibited the highest connectivity with other genes (indicated by brown connecting lines), which might be related to its broad and early expression range in the middle layer cells. In contrast, *ZmEREB161* showed relatively lower connectivity with other genes (indicated by cyan connecting lines). It was associated with genes such as *ZmEREB41*, *ZmEREB114*, *ZmEREB184*, *ZmAIC3*, and *ZmAAS2*, and was expressed during the late pseudo-time stage, coinciding with the progression toward procambium formation. These findings suggested that *ZmEREB114*, *ZmEREB41*, *ZmEREB161*, and *ZmEREB184* might spatiotemporally co-regulate the initiation of procambium in maize alongside auxin-related genes such as *ZmARFTF25*, *ZmIAA23*, *ZmAIC3*, and *ZmAAS2*.

***ZmEREB114* and *ZmEREB161* play a general role in leaf development rather than a specific role in vein patterning**

To test the function of genes expressed in the middle cell layer of the leaf, we first characterized mutant lines for *ZmEREB114* and *ZmEREB161*. Two translation-termination mutant lines, *zmereb114-16* and *zmereb114-1c,* were

obtained and homozygous mutant seeds were successfully generated for *zmereb114-1c*. Compared to B73, both *zmereb114-16* and *zmereb114-1c* exhibited a significant reduction in plant height **(Supplemental Figure 11A, B)**. Whereas no significant change in leaf width was observed in *zmereb114-16*, *zmereb114-1c* showed a noticeable decrease in leaf width **(Supplemental Figure 11B)**. No noticeable changes in cellular organization were observed in either *zmereb114-16* or *zmereb114-1c* mutants and vein density did not differ significantly from that of B73 **(Supplemental Figure 11B, C)**, although both lines exhibited a reduction in leaf vein number. Similarly, although the *zmereb161* mutant influenced vein number in association with altered leaf width, there was no significant effect on vein patterning **(Supplemental Figure 11D-I)**. Collectively these results suggested that *ZmEREB114* and *ZmEREB161* play a general role in leaf development rather than a specific role in vein patterning, a suggestion supported by the observation that constitutive expression of *ZmEREB161* or *ZmEREB114* in rice resulted in perturbations to growth (dwarfing and narrow leaves respectively) rather than vein patterning **(Supplemental Figure 12)**.

SHR suppresses procambium initiation by regulating auxin-related processes in the middle cell layer of the leaf

Although significant enrichment of auxin-related genes was observed both in Cluster 5 (procambium cells) of our single-cell clustering atlas and in transcriptome data from the middle segment (with active vein formation) of maize leaf primordia **(Figures 2B and 3A)**, exogenous auxin application during rice seedling development did not induce a noticeable increase in vein formation **(Figure 7A, B)**. Previous reports noted that overexpression of *ZmSHR1* in rice leads to an increased number of mesophyll cells between veins (Liu et al., 2023; Dong et al., 2023) and that rice seedlings undergoing induced expression of *35S:ZmSHR1-GR* are sensitised to exogenous

application of the auxin analog 2,4-D (Dong et al., 2023). The regulatory role of auxin in vein development and cellular patterning in rice seedlings may thus be more pronounced under conditions where leaf anatomy is disrupted, such as when *SHR* is overexpressed. To investigate whether the expression of middle layer and procambium-associated genes contributes to the conditional rearrangement of rice leaf anatomy after auxin treatment, we generated rice lines in which the expression of *ZmSHR1* could be induced by dexamethasone (*ZmUBI:ZmSHR1-GR*). Unlike previous reports that initiated induction prior to seed germination (Dong et al., 2023), we began induction of *ZmSHR1* (alone or in combination with 2,4-D application) in one- to two-week-old rice seedlings. In the absence of 2,4-D, vein formation was consistently suppressed such that leaves sampled three weeks after treatment initiation exhibited wider vein spacing **(Figure 7A, B)**. However, in the presence of of 2,4-D, suppression by SHR1 was not manifest. Because exogenous application of the auxin analog 2,4-D may have a different effect from that of endogenous auxin, we also used *ZmYUCCA9*, a key gene in the auxin biosynthesis pathway, to generate a transgenic line containing a dual-inducible system combining *ZmUBI:ZmSHR1-GR* with *ZmUBI:XVE:ZmYUCCA9*. Following dual induction with dexamethasone and estradiol, we observed that endogenous auxin was similarly able to rescue the suppressed vein phenotype caused by SHR induction **(Figure 7C–7E)**. To further analyse when this SHR/auxin coordinated rearrangement of vein patterning occurs, we dissected and sectioned rice P3 leaf primordia following *ZmSHR1* induction (alone or in combination with auxin application). Treatments started at germination when the SAM was surrounded by the P1-P3 primordia that were initated during embryogenesis (leaves 1-3), and lasted for 2-3 weeks by which time the expanded blade of leaves 1 (P7) to 4 (P4) were visible and leaves 5 (P3) to 7 (P1) had been initiated. The results confirmed that the SHR-induced phenotype of increased interveinal cell counts and reduced vein number was present at

the P3 stage, and that this phenotype could be partially counteracted by auxin treatment **(Figure 7F)**. The critical stage at which ectopic SHR activity is suppressing vein formation is thus between P0 and P3.

To identify genes that might be moderating SHR/auxin effects on vein formation, we performed transcriptome analysis on rice leaf primordia subjected to the SHR/auxin induction and treatment conditions. Analysis of auxin-related gene expression profiles revealed a cohort that were: i) upregulated to varying degrees in response to auxin treatment alone, ii) significantly downregulated following *ZmSHR1* induction and iii) less downregulated when auxin was applied simultaneously with *ZmSHR1* induction **(Figure 8A)**. These expression patterns correlate with the observed changes in vein number. We next examined the differential expression of rice genes homologous to those in clusters associated with vein formation and vascular tissues in maize. Maize cell Clusters 5, 11, and 12 (procambium, xylem, and phloem) corresponded to a substantial number of rice homologs that were markedly upregulated by auxin treatment but downregulated upon *ZmSHR1* induction **(Figure 8C)**. Although it is unclear whether the rice middle leaf layer is equivalent to the maize middle leaf layer, these results confirmed that the procambium and vascular tissues located in the middle layer serve as the primary sites for the conditional rearrangement of leaf anatomy in response to coordinated SHR and auxin action. Therefore, these tissues represent important sources for identifying relevant regulatory genes.

Further analysis of the rice homologs of middle ground tissue-associated genes revealed that *GRETCHEN HAGEN 3.8 (OsGH3.8*, homolog of *ZmAAS2*) and *AUXIN RESISTANT 1* (*OsAUX1*, homolog of *ZmAIC3*) expression patterns were suppressed by *ZmSHR1* induction and promoted by auxin treatment. In contrast, the expression of *ANT* homologs (not promoted by auxin treatment)

and *PIN* family genes (promoted by auxin but not suppressed by *ZmSHR1* induction) did not conform to this pattern **(Figure 8D; Supplemental Figure 13)**. To further investigate the the role of *GH3* and *AUX* genes, we conducted a 6-hour transient induction of *ZmUBI:ZmSHR1-GR* lines and collected samples for qRT-PCR, targeting *OsGH3.8*, *OsAUX1*, *OsAUX2* and *OsAUX4*. Results showed that these genes were significantly suppressed by *ZmSHR1* even within this short timeframe. Similarly, we performed qRT-PCR assays on the dual-inducible *ZmUBI:ZmSHR1-GR* × *ZmUBI:XVE:ZmYUCCA9* lines and showed that the expression trends of *OsGH3.8*, *OsAUX1*, *OsAUX2*, and *OsAUX4* were also consistent with the observed phenotypic recovery **(Figure 8E)**. Collectively these results showed that auxin-related genes are downregulated in rice in response to *ZmSHR1* induction but that down-regulation can be overcome by concurrent auxin treatment. Importantly, these dynamics are associated with the phenotypic suppression and recovery of vein formation in the leaf, suggesting a causal link.

Discussion

Vein initiation and development differ in maize and rice leaf primordia

Combining fluorescent labelling and sequential sectioning, we were able to analyse vein development patterns in maize and rice leaf primordia in proximo-distal (PD), medio-lateral (ML), and adaxial-abaxial contexts **(Figure 1; Supplemental Figure 1)**. By quantifying the number of veins from primordia to expanded seedling leaf stages, we confirmed that as the maize leaf widens, the number of veins increases rapidly from P3–P5 (approximately 3 mm to 5 mm) and then there is a reduced rate of increase during leaf expansion **(Figure 1I, J)**. Maize leaf primordia thus allowed us to capture the active zones where intermediate veins are being initiatied and are extending. We obtained gene expression signatures for sub-sectioned leaf primordia of maize and rice, contributing data to the early leaf primordium development database at a sub-

tissue level. These sub-tissue transcriptomic profiles and gradients shed light on early photosynthetic development in grass leaves, and represent a significant addition to maize and rice leaf transcriptome databases (Wang et al., 2013; Wang et al., 2014; van Campen et al., 2016).

Cell heterogeneity and clusters associated with intermediate vein formation in maize leaf primordia

Our single-cell resolution transcription map revealed high cell heterogeneity in early maize leaf primordia, showing continuity of the cellular developmental trajectory during vein formation. Because there are few known cell type-specific marker genes in young grass leaf primordia we relied on *in situ* hybridization for subsequent annotation and validation. In this way, we retrieved 7 broad populations of cells with 14 transcriptionally distinct clusters. Importantly, we identified the cell groups (clusters 1/5 in **Figure 3B**) that give rise to intermediate veins and pseudo-time trajectory analysis showed that these clusters include transitional cells (State V in **Figure 6D**) which potentially link the undifferentiated middle layer ground tissue and the procambial initials.

Regulatory impacts of SHR and auxin on vein patterning

The effects of auxin and auxin transport on vein formation and patterning were reported many years ago (Mattsson et al., 1999; Sieburth, 1999; Scarpella et al., 2006) but understanding the spatial and temporal patterns of auxin activity within and between tissues or cells is required to elucidate the role of auxin in this process (Perico et al., 2022). In our analysis, there was a notable enrichment of auxin-related gene expression during vein formation including genes encoding the auxin transporters *ZmAIC3* (homolog to AUX1), *ZmPIN1*, *ZmPIN4*, and auxin signaling pathway genes, such as *ZmARF25*, *ZmIAA23* and *ZmAAS2*. Intriguingly, *ZmAAS2* exhibits an expression pattern similar to that of the ANT family transcription factor gene *ZmEREB161*, being expressed in

individual or pairs of cells within the middle layer tissue as well as in procambium cells. The homolog of *ZmAAS2*, *OsGH3.8*, belongs to subgroup II of the *GRETCHEN HAGEN 3* (*GH3*) gene family and encodes an acyl-amino acid synthetase that catalyzes the conjugation of IAA with amino acids, modulating auxin activity (Li et al., 2016). Overexpression of *OsGH3.8* in rice leads to a significant increase in tiller number and a reduction in plant height (Dai et al., 2018). Collectively the pattern of expression of auxin transporters, of ZmEREB161 and of the conjugating enzyme ZmAAS/OsGH3 led us to hypothesise that a regulatory cascade involving transcription factors and auxin signalling is likely responsible for vein patterning in the middle layer of grass leaf primordia.

To test the hypothesis that auxin plays a conditional role in the regulation of grass leaf vein formation, we established an experimental system to manipulate expression of the transcription factor ZmSHR1, which had previously been shown to suppress vein formation in rice, in the presence or absence of increased auxin. Constitutive expression of *ZmSHR1* in rice primordia resulted in increased interveinal mesophyll cell number and reduced intermediate vein number, a phenotype that was restored to normal by simultaneous auxin treatment **(Figure 7)**. However, treatment with auxin alone did not result in significant venation changes. On the basis of these observations, we hypothesized that *ZmSHR1* may directly or indirectly alter gene expression in cells of the middle leaf layer, leading to changes in cell fate. Transcriptome analysis of rice leaf primordia subjected to *ZmSHR1* induction in the presence or absence of auxin confirmed that, homologs of genes expressed in the vein-initiating middle cell layer of maize leaf primordia (such as *ZmAAS2* and *ZmAIC3*) exhibit response patterns consistent with the changes in vein density observed in rice leaves **(Figure 8)**. That is, SHR1 may function to make cells of the rice leaf competent to respond to auxin signalling

components that can induce vein formation. Although manipulation of *ZmEREB114* and *ZmEREB161* did not alter vein patterning, a role in maintaining proliferative capacity of middle layer cells may influence where and when veins can be initiated. Collectively these functional analyses provide evidence that some of the candidate regulators identified by our single-cell and tissue transcriptomes play important roles in the regulation of vein patterning in grass leaves.

A model of vein development in grass leaf primordia

Using a combination of single-cell, developmental and genetic approaches, this study provided both theoretical and practical platforms for the identification of regulators of leaf vein development. On the basis of the data obtained we propose that the expression of genes such as *ZmEREB114* and *ZmAIC3* in cells of the middle layer of ground tissue (Cluster 1) and expression of genes such as *ZmSHR1*, *ZmAAS2* and *ZmEREB161* in procambial cells (Cluster 5) leads to distinct developmental trajectories **(Figure 9)**. The specification of procambial initials (Cluster 5) from within the middle ground tissue (Cluster 1), the differentiation of middle ground tissue (Cluster 1) into middle layer mesophyll cells (Cluster 8 or 9), and further development of procambial cells (Cluster 5) into phloem (Cluster 11) or xylem (Cluster 12) cells, proceeds under the spatial and temporal regulation of the aforementioned transcription factors and auxin signalling related genes. In maize leaf primordia this sequential process eventually contributes to the formation of Kranz anatomy with veins separated by just two mesophyll cells. Aspects of this process must be conserved in rice because when vein formation is experimentally suppressed by induction of *ZmSHR1* expression, homologs of middle ground tissue or procambium expressed genes in maize, such as *OsAUX1* (homolog to *ZmAIC3*) and *OsGH3.8* (homolog to *ZmAAS2*), are downregulated. Together these results indicate that the combined action of SHR and auxin-related pathways

controls grass leaf vein patterning. Further systematic investigation of genes that act in the middle cell layer and procambium cells of maize leaf primordia is needed to decode the requirement for increasing vein density in rice, which will enhance our understanding of cellular patterning mechanisms and facilitate the engineering of crop leaf anatomy.

Materials and methods

Plant material and growth conditions

In Shanghai, *Zea mays* (L.) ssp. Mays cv. B73 and *pZmPIN1a::ZmPIN1a:YFP* transgenic line (Yang et al., 2015), as well as *Oryza sativa* (L.) ssp. Japonica cv. Nipponbare and *DR5::VENUS* line, were grown in a phytotron, 27 °C in day and 25 °C at night, with 600 μmol photons m^{-2} s^{-1} light intensity, 16 h light and 8 h dark photoperiod. Tissue was harvested from 2-, 3- or 4-week-old maize plants and 2-week-old rice plants.

Zmereb161 and *zmereb114* mutants of maize by EMS (Ethyl methanesulfonate) chemical mutation were obtained from http://www.elabcaas.cn/memd/index.php. The M3-generation was used and genotyping was conducted by genomic PCR with primers listed in **Supplemental Dataset 8**. PCR products were sequenced to confirm the mutant of *zmereb161*, in which the peptide chain coding is terminated prematurely from CAG to TAG. For the mutation of *ZmEREB114*, we obtained two termination mutant lines, namely *zmereb114*-16 and *zmereb114*-1c. The mutation site in the *zmereb114*-16 line is located at Chr 7: 3332556 (RefGen V4.32), where a G-to-A substitution occurred, resulting in a change from TGG to TGA in the coding sequence. The mutation site in the *zmereb114*-1c line is located at Chr 7: 3332836 (RefGen V4.32), where a C-to-T substitution occurred, resulting in a change from CAG to TAG in the coding sequence.

To generate *ZmUBIpro::ZmEREB161* and *ZmUBIpro::ZmEREB114* transgenic plants in rice, coding sequences were PCR amplified from maize

cDNA that had been generated using RNA isolated from P1-5 leaf primordia. The amplified sequence was subcloned into Gateway® donor vector, sequenced, and then cloned into destination vector pSC310, downstream of the maize ubiquitin promoter. The construct was transformed into japonica rice cultivar Nipponbare, and after T1 genotyping and expression analysis, two lines of *ZmUBIpro::ZmEREB161* and three lines of *ZmUBIpro::ZmEREB114* seedlings were used for phenotypic analysis. Primers used are listed in **Supplemental Dataset 8**

ClearSee and confocal imaging

ClearSee assays were performed as previously described (Kurihara et al., 2015). In brief, leaf primordium were immersed in fixation buffer (4% w/v paraformaldehyde, Cat. No. AR-0211) under vacuum at 25 mbar for 1h, washed once with 1× PBS, and then immersed in ClearSee reagent (10% xylitol, 15% sodium deoxycholate, and 25% urea) for 3 days. The primordium were flattened with the adaxial side facing up in ClearSee reagent, and observed under a ZEISS LSM880 confocal microscope using a 514-nm laser excitation and 520- to 560-nm emission for detection of the YFP.

Histology

Leaf primordium samples were fixed overnight in ethanol/acetic acid (3:1), and embedded in Paraplast Plus (Sigma, Cat. No. P3683-1KG) using a modular automated tissue processor (KD-TS3A and KD-BM/BL). Paraffin-embedded samples were sectioned (5 μm) with a Rotary microtome (Lecia, RM2125). Sections not dewaxed were viewed using an Olympus CX23 microscope.

***In situ* hybridization and imaging**

For probe synthesis, templates of RNA probes were amplified from cDNAs using gene-specific primers containing T7 promoter sequences at the 5' end. *In*

vitro transcription was performed with T7 RNA polymerase (Roche, Cat. No. 10881767001; Thermo, Cat. No. EP0111) and DIG RNA labeling mix (Roche, Cat. No. 11277073910). The primers used to generate the probes are listed in **Supplemental dataset 8**.

Tissue embedding and RNA *in situ* hybridization were performed as described with modifications (Langdale 1994; Zeng et al, 2021; Weigel et al., 2002). Briefly, leaf primordium were fixed with FAA (3.7% formaldehyde, 5% glacial acetic acid, and 50% ethanol), embedded in paraffin and sectioned (8 μm) as described above. The sections were dewaxed, digested with proteinase K (Sigma, Cat. No. P6556), dehydrated with gradient ethanol, and hybridized with RNA probes. After washing, the sections were incubated with anti-digoxigenin-AP Fab fragments (Roche, Cat. No. 11093274910). The signals were developed with the NBT/BCIP stock solution (Roche, Cat. No. 11681451001), and the sections were imaged using an Olympus CX23/BX3-CBH microscope. Where necessary, the sections were also stained with 10 mg/mL Calcofluor white for enhancing UV excited cell wall autofluorescence, and imaged with a fluorescence microscope (DM6000B, Leica).

Sample collection for bulk RNA-seq

Maize leaf primordium of about 5 mm were partitioned under dissection microscope into three parts: M3tip≈1.5 mm, M3middle≈2.5 mm, and M3base≈1 mm. About 200 leaf primordium were obtained, partitioned, and pooled into different samples with three biological replicates. The rice leaf primordium of about 5 mm were also partitioned into R3tip≈1.5 mm, R3middle≈2.5 mm, and R3base≈1 mm. Similarly, about 400 leaf primordium were processed into different samples with three biological replicates. Total RNA was extracted using the GeneJET Plant RNA Purification Mini Kit (Thermo Scientific, Cat. No. K0801) according to the manufacturer's instructions.

RNA-seq

Three types of maize tissues and 3 types of rice tissues were sampled as described above for bulk RNA sequencing. Paired-end RNA-seq was performed using the Illumina NovaSeq 6000 platform. Reads were mapped to the genome version of *Zea mays* V4. Differentially expressed genes (DEGs) were determined by DESeq2 (v1.44.0) (Love et al., 2014). The p-values were adjusted using the Benjamini-Hochberg (BH) method to generate False Discovery Rate (FDR). A threshold of FDR < 0.05 and fold change $\geqslant$ 2 was used as the screening criteria.Three biological replicates were used in the data analyses.

Gene Ontology (GO) analysis

For maize, GO enrichment analysis was performed using the R package "clusterProfiler (4.12.6)" with the parameter settings: "p-value Cutoff = 0.05, q-value Cutoff = 0.05". For rice, GO enrichment analysis was analyzed at https://geneontology.org. R package "ClusterGVis (0.1.2)" was used to visualize the gene expression pattern and GO enrichment results. The values have been normalized using the z-score method. The clustering method employed is MFuzz, which is an improved algorithm based on Fuzzy C-Means (FCM). MFuzz defaults to using the weighted average as the cluster centers.

Sample collection and preparation of nuclei

According to our pilot experiments, ~200 maize leaf primordium could yield enough high quality nuclei, using the methods previously described with modifications (Thibivilliers et al., 2020; Conde et al., 2021). To isolate nuclei, freshly collected leaf primordium of 3-4 mm (with group efforts to ensure enough amount within 1 h) were placed on Petri dishes and vigorously chopped with two razor blades for 5 minutes in ~500 µL Nuclei Isolation Buffer [NIB, 10

mM MES-KOH pH 5.7, 10 mM NaCl, 10 mM KCl, 2.5 mM EDTA, 250 mM sucrose, 0.1% protease inhibitor cocktail (Roche), and 0.1% BSA (YESEN), finally adjust pH to 5.7]. Homogenized tissue was first filtered with 40 μm cell strainers (Falcon, Cat. No. 352340), then filtered with 30 μm cell strainers (pluriSelect, Cat. No. 43-50030-01). Nuclei were spun down in a swinging-bucket centrifuge (5 min, 500 rcf.) and re-suspended in 100 μL washing buffer (15 mM Tris-HCl pH 7.5, 160 mM KCl, 40 mM NaCl, finally adjust pH to 7.0). Nuclei quantification was done by staining with trypan blue (final concentration of ~0.1%), and counting on a hemocytometer for a total of 200,000 nuclei. Nuclei suspensions were then spun down (5 min, 500 rcf.) and re-suspended in diluted nuclei buffer (10x Genomics) to a final concentration of 2,000~2,500 nuclei per μL, and used as input for snRNA-seq library preparation (~16,000 nuclei in total). Samples were kept on ice for all intermittent steps.

SnRNA-seq library construction

Approximately 16,000 counted nuclei were loaded on Single Cell A Chip. The libraries were constructed using a Chromium Controller and Chromium Single Cell 30 Reagent Kits V3.1. Qualitative analysis of DNA library was performed by an Agilent 2100 Bioanalyzer. Libraries were sequenced by an Illumina NovaSeq 6000.

Processing of snRNA-seq data

The raw single-nucleus RNA sequencing data were initially processed using Cell Ranger 7.0.0 (10x Genomics). The maize reference genome (Version 4, V4) and GTF annotation file (Zea_mays.B73_RefGen_v4.50) were downloaded from the EnsemblPlants database (https://plants.ensembl.org/Zea_mays/). The "cellranger mkref" command was executed with the parameters "--genome", "--fasta", and "--genes" to construct the reference genome. The "cellranger count" command was run using the parameters "--id", "--transcriptome", "--

fastqs", "--sample", and "--force-cells=8000" to generate single-cell gene counts. The gene-cell matrices (named “filtered_feature_bc_matrix” and “raw_feature_bc_matrix” by 10x Genomics) were corrected for ambient RNA expression using SoupX (v1.5.0) (Young et al., 2020). The filtered gene-cell matrices were served as processed raw data for further analyses.

Data filtering, clustering, and annotation

We used scDblFinder (v1.2.0) software for detection and handling of doublets/multiplets in the snRNA-seq data (Hao et al., 2021), a total of 373 doublets were identified and removed. Downstream analyses were primarily conducted using Seurat (v4.2.1). The workflow included quality control (calculating mitochondrial and chloroplast gene percentages and filtering low-quality cells and genes), data normalization (using the SCTransform function with default parameters), PCA (RunPCA with variable genes and 100 principal components), SNN graph construction and Louvain-based clustering (FindNeighbors and FindClusters), and visualization through nonlinear dimensionality reduction (RunTSNE and RunUMAP).

For quality control, low-quality cells and genes were filtered using the following criteria: cells with fewer than 200 expressed genes were removed, and genes detected in fewer than three cells were excluded. Further cleaning steps were performed using log10GenesPerUMI, and set the log10 fold change more than 0.7 and mitochondria percentage below 5%. A total of 7,473 cells with average of 2,424 genes and an average of 3,050 UMI counts per nucleus.

After filtering, the data were used for downstream analysis. The SCTransform function with the default parameters was used for data normalization, scaling and transformation. The dimensions of the expression matrix were then reduced by the RunPCA function, and the top 30 dimensions were used for

FindNeighbors and UMAP analysis. The cell clusters were identified by the FindClusters function with a resolution of 1.

Cell types were manually annotated for each cluster based on known marker genes **(Supplementary Dataset 7)**. Cluster-enriched marker genes were identified using the FindAllMarkers function in Seurat with the following criteria: Wilcoxon Rank Sum test, $\log_2$ fold change threshold of 0.58 (equivalent to 1.5-fold difference), and a minimum expressing fraction of 0.25 in either group.

Subclustering

To investigate the cell heterogeneity of middle layer ground tissue, we extracted cells expressing *ZmEREB114* (gene_expr_matrix > 3) and homolog of *AtSCPL46* (gene_expr_matrix > 2.8), as well as cells from procambial and vascular tissues (cluster5, cluster11, cluster12, cluster13). The raw count data were imported into Seurat to create a new Seurat object. The analytical pipeline was executed as described above. Specifically, parameters including the total number of principal components (npcs), dimensions (dims), number of nearest neighbors (n.neighbors), resolution, and minimum distance (min.dist) were adjusted to optimally approximate the latent topological structure of the subclusters (npcs = 50, n.neighbors = 30, resolution = 0.8, min.dist = 0.3). Following subclustering, the "FindAllMarkers" function was employed to identify cluster-enriched genes using the following parameters: Wilcoxon rank-sum test, a minimum fold-change threshold of 1.68 (logfc.threshold = 0.75) for differential expression between groups of cells, and a minimum detection proportion of 0.25 for genes (min.pct = 0.25).

Pseudo-time trajectory analysis

Pseudo-temporal developmental trajectory analysis was performed using Monocle 2 (v2.32.0) (Qiu et al., 2017). The Seurat object (containing middle layer ground tissue cells, procambial cells, and vascular cells) was converted

into a CellDataSet object using as.CellDataSet. First, dispersionTable was run to calculate the intercellular expression variance for each gene. Variable genes defining the developmental progression were selected based on mean expression levels (i.e., the mean_expression parameter). Second, the data were reduced to two principal components (max_components = 2, method = 'DDRTree'). Using the expression data in the low-dimensional space, orderCells was applied to describe the transition of cells from one state to another. Cells were ordered along pseudo-time by invoking orderCells. The cellular trajectory was visualized using plot_cell_trajectory in Monocle 2. To specify a prior "starting point" for the tree-shaped trajectory, orderCells was run again with the root_state parameter set accordingly. Branch points were selected to analyze bifurcation events within the differentiation trajectory. The 1139 significant genes were used as the ordering gene set. Gene expression trends were plotted using the plot_genes_in_pseudotime function, and a hierarchical tree of cell lineages was constructed using plot_complex_cell_trajectory.

Construction of co-expression network

Functional gene module analysis was performed using the Python package "Hotspot (v1.1.2)". The input single-cell data for this study consisted of subclustered cell populations from the middle layer ground tissue, procambium, and vascular tissues. Modules were identified via the "create_modules" function with parameters set to min_genes = 100 and an FDR threshold of 0.05. Module scores were calculated using the "module_scores" function to quantify the expression level of genes within each cell. Gene expression was projected onto UMAP-reduced dimensional space and correlated with cell clusters. Target gene correlations and module attributes were extracted using a correlation filter value of 5. A co-expression network was constructed based on autocorrelation z-scores and visualized in Cytoscape (v3.10.3).

Construction of *ZmSHR1* induction vector for rice transformation

Using cDNA reverse-transcribed from mRNA of maize leaf primordia, maize genomic DNA (gDNA), and a plasmid containing the glucocorticoid receptor (GR) gene fragment (All-in-one CRE/LOX system plasmid) as templates, the coding sequence (CDS) of the *ZmSHR1* gene, the *ZmUBI* promoter fragment, and the GR fragment were amplified separately using Phanta Max Super-Fidelity DNA Polymerase (Vazyme). These fragments were then homologously recombined into the Golden Gate cloning system's Level 0 vectors (SC1, PU, and C2) using the ClonExpress II One Step Cloning Kit (Vazyme). Subsequently, these vectors, along with a Level 0 vector containing the *tNOS* terminator, were incubated with T4 DNA ligase, BsaI-HFv2, and an empty Level 1 backbone vector of the Golden Gate system to assemble a Level 1 transcriptional unit. Under the action of BpiI restriction enzyme, this unit was ligated with other Level 1 transcriptional units (*ZmUBI:XVE:ZmYUCCA9*, constructed similarly using Golden Gate cloning system; *35S::HYG*, Golden Gate adapted selection marker) according to demand to form Level 2 vectors. The constructed Level 2 vectors were then introduced into *Agrobacterium tumefaciens* strain EHA105 for rice (Zhonghua 11) transformation.

Dexamethasone (Dex) / β-Estradiol induction and auxin treatment of transgenic rice plants

Rice seeds were soaked in water until germination (visible radicle emergence), then transferred to growth chamber for hydroponic cultivation (photoperiod: 14-hour light/10-hour dark, light intensity: 400–600 μmol m^{-2} s^{-1}, temperature: ~27°C, relative humidity: 65–70%). The seedlings were grown in liquid ½ Murashige and Skoog (½ MS) medium for one or two weeks to establish robust root systems. Subsequently, 10 μM dexamethasone (DEX) was added to the medium, and the plants were cultured for an additional three weeks. The induction with β-Estradiol (10 μM) or treatment with 2,4-D (0.1 mg/L) was performed the same way as above. The most recently fully expanded leaves were then collected for histological sectioning and observation.

Transcript abundance analysis

RNA extraction, cDNA synthesis, and qRT–PCR were performed as previously described (Zhang et al., 2023). Leaf samples were flash-frozen in liquid nitrogen and ground with a tissue grinder. Total RNA was extracted using RNAiso Plus reagent (Takara, #9108) according to the manufacturer's instructions. For cDNA synthesis, 2 μg of total RNA was reverse transcribed using a first-strand cDNA synthesis kit (YEASEN, #11141ES60). qRT–PCR was carried out with Hieff UNICON® Universal Blue qPCR SYBR Master Mix (YEASEN, #11184ES08) in a 20 μL reaction volume. Transcript levels were normalized to *ACTIN* as the internal reference gene and calculated using the ΔΔCt method. The RT–PCR primers used are listed in **Supplemental Dataset 8**.

Quantification and statistical analysis

Two-way or one-way ANOVA with Tukey's HSD test, or Student's two-sided *t* test was used to determine the statistical significance among different samples. *P*-value < 0.05 was considered as statistical significance, and different letters on the bar graphs indicate statistically significant differences. Statistic results and graphs were generated in GraphPad Prism 8 (www.graphpad.com). The numbers of samples and types of statistical analyses are given in figure legends and the quantitative details can be found in Microsoft Excel spreadsheets of the **Supplemental Dataset 10**.

Accession numbers

Accession numbers for the genes analyzed are listed in **Supplemental dataset 7 and 8**.

Data availability

Single-cell RNA-seq data of maize leaf primordium, bulk RNA-seq data of segmented maize/rice leaf primordia, and transcriptome data of SHR1

induction/auxin treated rice leaf primordia have been deposited at NCBI's Sequence Read Archive (SRA) (https://www.ncbi.nlm.nih.gov/sra) under the accession numbers: PRJNA1051680, PRJNA1051184, PRJNA1051294 and PRJNA1329017.

Supplemental data

Supplemental Figure 1. The growth pattern of maize leaf primordia.

Supplemental Figure 2. Dot map showing the top 10 marker genes from each of the different cell clusters.

Supplemental Figure 3. *In situ* expression patterns of representative marker genes from xylem and phloem related cell clusters.

Supplemental Figure 4. *In situ* expression patterns of representative marker genes from mesophyll and parenchymal cell related clusters.

Supplemental Figure 5. The expression patterns of *ZmYAB3* and *ZmGLK6*.

Supplemental Figure 6. In situ expression patterns of representative marker genes related with clusters 1 and 4.

Supplemental Figure 7. Localization of *ZmEREB114* and *ZmEREB41* transcripts in the middle cell layer of maize leaf primordium, and hybridization results with sense RNA probes.

Supplemental Figure 8. Differentiation trajectory in a P4 maize primodium.

Supplemental Figure 9. Expression trends of xylem, phloem, and procambium marker genes along the pseudo-temporal trajectory.

Supplemental Figure 10. Cell fate and differential gene expression analysis along pseudo-time, and cross-reference from the LCM data.

Supplemental Figure 11. Characterization of maize *zmereb114* and *zmereb161* mutants.

Supplemental Figure 12. Characterization of *ZmUBIpro::ZmEREB161* and *ZmUBIpro::ZmEREB114* transgenic rice plants.

Supplemental Figure 13. The impact of induced *ZmSHR1* expression and auxin

treatment on the expression of auxin-related or IDD genes in rice leaf primordia.

Supplemental Dataset 1. Maize-All_gene_TPM

Supplemental Dataset 2. Rice-All_gene_TPM

Supplemental Dataset 3. All-snRNA-seq data

Supplemental Dataset 4. Clustering analysis of the snRNA-seq data

Supplemental Dataset 5. Rice SHR-Auxin All_gene_TPM

Supplemental Dataset 6. 335 DEGs involved in pseudo-time analysis

Supplemental Dataset 7. List of marker genes

Supplemental Dataset 8. Primers used for *in situ* hybridization and qRT-PCR

Supplemental Dataset 9. DEGs from SHR-Auxin rice transcriptome

Supplemental Dataset 10. Statistical analysis

Author contributions

JY and PW conceived the project; JY, YC, and PW performed the experiments; DV, WD, OS and JAL worked on the *ZmSHR1* inducible lines and *ZmEREB161* overexpression lines in rice; JY, YC, ST, HS, CZ, JY and PW analyzed the data and produced the figures; JY, YZ, XZ, JAL, JW and PW interpreted the results, wrote, and revised the paper.

Acknowledgments

We thank Langdale lab members and Prof. Lin Xu for fruitful discussions. We thank Rui Zhang, Xiyu Zeng, Jiachun Wu, and Qiqi Zhang for the team work of dissecting and collection of maize and rice leaf primordia. We are grateful to Yutong Li and Luhuan Ye for helping with mutant analysis. We thank Prof. Fang Yang for providing the maize *pZmPIN1a::ZmPIN1a:YFP* transgenic line, and the late Prof. Dabing Zhang for providing the rice *DR5::VENUS* line. We are grateful to Dr. Hua Wang, Dr. Shuining Yin, Dr. Wenjuan Cai, Fang Yuan and Qing Gao for technical support with *in situ* hybridization, confocal microscopy,

and plant growth in the phytotron.

Funding

This research was supported by the Biological Breeding-National Science and Technology Major Project (2023ZD04072) awarded to Peng Wang, the Bill and Melinda Gates Foundation C_4 Rice grant awarded to the University of Oxford (2015-2019) (OPP1129902), and the supporting grant from the Key Laboratory of Plant Carbon Capture, CEMPS (CNTD001).

References

Biedron M,Banasiak A. 2018. Auxin-mediated regulation of vascular patterning in Arabidopsis thaliana leaves. Plant Cell Rep, 37(9), 1215-1229.

Bosabalidis AM, Evert RF,Russin WA. 1994. Ontogeny of the vascular bundles and contiguous tissues in the maize leaf blade. American Journal of Botany, 81, 745-752.

Bosch M, Mayer CD, Cookson A,Donnison IS. 2011. Identification of genes involved in cell wall biogenesis in grasses by differential gene expression profiling of elongating and non-elongating maize internodes. J Exp Bot, 62(10), 3545-3561.

Brown WV. 1975. VARIATIONS IN ANATOMY, ASSOCIATIONS, AND ORIGINS OF KRANZ TISSUE. American Journal of Botany, 62(4), 395-402.

Canny MJ. 1998. Transporting Water in Plants. American Scientis, 86(2), 152-159.

Christin PA, Osborne CP, Chatelet DS, Columbus JT, Besnard G, Hodkinson TR, Garrison LM, Vorontsova MS,Edwards EJ. 2013. Anatomical enablers and the evolution of C4 photosynthesis in grasses. Proc Natl Acad Sci U S A, 110(4), 1381-1386.

Conde D, Triozzi PM, Balmant KM, Doty AL, Miranda M, Boullosa A, Schmidt HW, Pereira WJ, Dervinis C,Kirst M. 2021. A robust method of nuclei isolation for single-cell RNA sequencing of solid tissues from the plant genus Populus.

Dai Z, Wang J, Yang X, Lu H, Miao X,Shi Z. 2018. Modulation of plant architecture by the miR156f-OsSPL7-OsGH3.8 pathway in rice. J Exp Bot, 69(21), 5117-5130.

Dong W, Chang T, Dai H, Yang W, Su Y, Chao D, Zhu XG, Wang P, Yu N,Wang E. 2023. Creating a C(4)-like vein pattern in rice by manipulating SHORT ROOT and auxin levels. Sci Bull (Beijing), 68(24), 3133-3136.

Donner TJ, Sherr I,Scarpella E. 2009. Regulation of preprocambial cell state

acquisition by auxin signaling in Arabidopsis leaves. Development, 136(19), 3235-3246.
Esau K. (1943). Ontogeny of the vascular bundle in Zea Mays.
Fouracre JP, Ando S,Langdale JA. 2014. Cracking the Kranz enigma with systems biology. J Exp Bot, 65(13), 3327-3339.
Gómez-Anduro G, Ceniceros-Ojeda EA, Casados-Vázquez LE, Bencivenni C, Sierra-Beltran A, Murillo-Amador B,Tiessen A. 2011. Genome-wide analysis of the beta-glucosidase gene family in maize (Zea mays L. var B73). Plant Mol Biol, 77(1-2), 159-183.
Hao Y, Hao S, Andersen-Nissen E, Mauck WM, 3rd, Zheng S, Butler A, Lee MJ, Wilk AJ, Darby C, Zager M, Hoffman P, Stoeckius M, Papalexi E, Mimitou EP, Jain J, Srivastava A, Stuart T, Fleming LM, Yeung B, Rogers AJ, McElrath JM, Blish CA, Gottardo R, Smibert P,Satija R. 2021. Integrated analysis of multimodal single-cell data. Cell, 184(13), 3573-3587 e3529.
Hoad GV. 1995. Transport of hormones in the phloem of higher plants. Plant Growth Regulation, 16(2), 173-182.
Hughes TE, Sedelnikova O, Thomas M,Langdale JA. 2023. Mutations in NAKED-ENDOSPERM IDD genes reveal functional interactions with SCARECROW during leaf patterning in C4 grasses. PLoS Genet, 19(4), e1010715.
Johnston R, Leiboff S,Scanlon MJ. 2015. Ontogeny of the sheathing leaf base in maize (Zea mays). New Phytologist, 205(1), 306-315.
Kumar D,Kellogg EA. 2019. Getting closer: vein density in C(4) leaves. New Phytol, 221(3), 1260-1267.
Kurihara D, Mizuta Y, Sato Y,Higashiyama T. 2015. ClearSee: a rapid optical clearing reagent for whole-plant fluorescence imaging. Development, 142(23), 4168-4179.
Langdale J,Kidner C. 1994. Bundle sheath defective, a mutation that disrupts cellular differentiation in maize leaves. Development, 120.
Langdale JA, Lane B, Freeling M,Nelson T. 1989. Cell lineage analysis of maize bundle sheath and mesophyll cells. Dev Biol, 133(1), 128-139.
Li X, Yang DL, Sun L, Li Q, Mao B,He Z. 2016. The Systemic Acquired Resistance Regulator OsNPR1 Attenuates Growth by Repressing Auxin Signaling through Promoting IAA-Amido Synthase Expression. Plant Physiol, 172(1), 546-558.
Liu Q, Teng S, Deng C, Wu S, Li H, Wang Y, Wu J, Cui X, Zhang Z, Quick WP, Brutnell TP, Sun X,Lu T. 2023. SHORT ROOT and INDETERMINATE DOMAIN family members govern PIN-FORMED expression to regulate minor vein differentiation in rice. Plant Cell.
Liu WY, Lin HH, Yu CP, Chang CK, Chen HJ, Lin JJ, Lu MJ, Tu SL, Shiu SH, Wu SH, Ku MSB,Li WH. 2020. Maize ANT1 modulates vascular development, chloroplast development, photosynthesis, and plant growth. Proc Natl Acad Sci U S A, 117(35), 21747-21756.

Liu WY, Yu CP, Chang CK, Chen HJ, Li MY, Chen YH, Shiu SH, Ku MSB, Tu SL, Lu MJ,Li WH. 2022. Regulators of early maize leaf development inferred from transcriptomes of laser capture microdissection (LCM)-isolated embryonic leaf cells. Proc Natl Acad Sci U S A, 119(35), e2208795119.

Ljun K, Hul AK, Kowalczyk M, Marchant A, Celenza J, Cohen JD,Sandberg G. 2002. Biosynthesis, conjugation, catabolism and homeostasis of indole-3-acetic acid in Arabidopsis thaliana. Plant Mol Biol, 50(2), 309-332.

Love MI, Huber W,Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol, 15(12), 550.

Lucas WJ, Groover A, Lichtenberger R, Furuta K, Yadav SR, Helariutta Y, He XQ, Fukuda H, Kang J, Brady SM, Patrick JW, Sperry J, Yoshida A, Lopez-Millan AF, Grusak MA,Kachroo P. 2013. The plant vascular system: evolution, development and functions. J Integr Plant Biol, 55(4), 294-388.

Mattsson J, Sung ZR,Berleth T. 1999. Responses of plant vascular systems to auxin transport inhibition. Development, 126(13), 2979-2991.

Nazipova A, Gorshkov O, Eneyskaya E, Petrova N, Kulminskaya A, Gorshkova T,Kozlova L. 2021. Forgotten Actors: Glycoside Hydrolases During Elongation Growth of Maize Primary Root. Front Plant Sci, 12, 802424.

Nelson T,Dengler NG. 1997. Leaf Vascular Pattern Formation. The Plant Cell, 9, 1121 - 1135.

Perico C, Tan S,Langdale JA. 2022. Developmental regulation of leaf venation patterns: monocot versus eudicots and the role of auxin. New Phytol, 234(3), 783-803.

Perico C, Zaidem M, Sedelnikova O, Bhattacharya S, Korfhage C,Langdale JA. 2024. Multiplexed in situ hybridization reveals distinct lineage identities for major and minor vein initiation during maize leaf development. Proceedings of the National Academy of Sciences, 121(28), e2402514121.

Qiu X, Mao Q, Tang Y, Wang L, Chawla R, Pliner HA,Trapnell C. 2017. Reversed graph embedding resolves complex single-cell trajectories. Nat Methods, 14(10), 979-982.

Robil JM,McSteen P. 2023. Hormonal control of medial-lateral growth and vein formation in the maize leaf. New Phytol, 238(1), 125-141.

Rohrmeier T,Lehle L. 1993. WIP1, a wound-inducible gene from maize with homology to Bowman-Birk proteinase inhibitors. Plant Mol Biol, 22(5), 783-792.

Sage RF, Sage TL,Kocacinar F. 2012. Photorespiration and the evolution of C4 photosynthesis. Annu Rev Plant Biol, 63, 19-47.

Sawchuk MG, Head P, Donner TJ,Scarpella E. 2007. Time-lapse imaging of Arabidopsis leaf development shows dynamic patterns of procambium formation. New Phytol, 176(3), 560-571.

Scarpella E, Marcos D, Friml J,Berleth T. 2006. Control of leaf vascular patterning by polar auxin transport. Genes Dev, 20(8), 1015-1027.

Sedelnikova OV, Hughes TE, Langdale JA. 2018. Understanding the Genetic

Basis of C4 Kranz Anatomy with a View to Engineering C3 Crops. Annu Rev Gene, 52, 249-270.
Seyfferth C, Renema J, Wendrich JR, Eekhout T, Seurinck R, Vandamme N, Blob B, Saeys Y, Helariutta Y, Birnbaum KD,De Rybel B. 2021. Advances and Opportunities in Single-Cell Transcriptomics for Plant Research. Annu Rev Plant Biol, 72, 847-866.
Sharman BC. 1942. Developmental Anatomy of the Shoot of Zea mays L. Annals of Botany, 6, 245-282.
Sieburth LE. 1999. Auxin is required for leaf vein pattern in Arabidopsis. Plant Physiol, 121(4), 1179-1190.
Thibivilliers S, Anderson D,Libault M. 2020. Isolation of Plant Root Nuclei for Single Cell RNA Sequencing. Curr Protoc Plant Biol, 5(4), e20120.
van Campen JC, Yaapar MN, Narawatthana S, Lehmeier C, Wanchana S, Thakur V, Chater C, Kelly S, Rolfe SA, Quick WP,Fleming AJ. 2016. Combined Chlorophyll Fluorescence and Transcriptomic Analysis Identifies the P3/P4 Transition as a Key Stage in Rice Leaf Photosynthetic Development. Plant Physiol, 170(3), 1655-1674.
Vlad D, Zaidem M, Perico C, Sedelnikova O, Bhattacharya S,Langdale JA. 2024. The WIP6 transcription factor TOO MANY LATERALS specifies vein type in C(4) and C(3) grass leaves. Curr Biol, 34(8), 1670-1686.e1610.
Wang L, Czedik-Eysenberg A, Mertz RA, Si Y, Tohge T, Nunes-Nesi A, Arrivault S, Dedow LK, Bryant DW, Zhou W, Xu J, Weissmann S, Studer A, Li P, Zhang C, LaRue T, Shao Y, Ding Z, Sun Q, Patel RV, Turgeon R, Zhu X, Provart NJ, Mockler TC, Fernie AR, Stitt M, Liu P,Brutnell TP. 2014. Comparative analyses of C_4 and C_3 photosynthesis in developing leaves of maize and rice. Nat Biotechnol, 32(11), 1158-1165.
Wang P, Kelly S, Fouracre JP,Langdale JA. 2013. Genome-wide transcript analysis of early maize leaf development reveals gene cohorts associated with the differentiation of C4 Kranz anatomy. Plant J, 75(4), 656-670.
Xue HC, Xu ZG, Liu YJ, Wang L, Ming X, Wu ZY, Lian H, Han YW, Xu J, Zhang ZD, Shao QL, Liu K, Wang FX, Wang AH, Zhao J, Zhang J, Chen J, Mao Y,Wang JW. 2025. A unified cell atlas of vascular plants reveals cell-type foundational genes and accelerates gene discovery. Cell.
Yang F, Bui HT, Pautler M, Llaca V, Johnston R, Lee BH, Kolbe A, Sakai H,Jackson D. 2015. A maize glutaredoxin gene, abphyl2, regulates shoot meristem size and phyllotaxy. Plant Cell, 27(1), 121-131.
Young MD,Behjati S. 2020. SoupX removes ambient RNA contamination from droplet-based single-cell RNA sequencing data. Gigascience, 9(12).
Zeng J, Li X, Ge Q, Dong Z, Luo L, Tian Z,Zhao Z. 2021. Endogenous stress-related signal directs shoot stem cell fate in Arabidopsis thaliana. Nat Plants, 7(9), 1276-1287.
Zhang Q, Tian S, Chen G, Tang Q, Zhang Y, Fleming AJ, Zhu X-G,Wang P. 2024. Regulatory NADH dehydrogenase-like complex optimizes C4 photosynthetic carbon flow and cellular redox in maize. New Phytologist,

241(1), 82-101.