

Single-cell RNA sequencing reveals a key regulator *ZmEREB14* affecting shoot apex development and yield formation in maize

Langlang Ma^{1,†} , Na Zhang^{1,†}, Peng Liu^{1,2}, Yuru Liang¹, Ran Li¹, Guangsheng Yuan¹, Chaoying Zou¹, Zhong Chen¹, Thomas Lübbertedt³, Guangtang Pan¹ and Yaou Shen^{1,*} 

¹State Key Laboratory of Crop Gene Exploration and Utilization in Southwest China, Maize Research Institute, Sichuan Agricultural University, Chengdu, China

²College of Life Science & Biotechnology, Mianyang Teachers' College, Mianyang, China

³Department of Agronomy, Iowa State University, Ames, USA

Received 24 July 2024;
revised 22 October 2024;
accepted 16 November 2024.

*Correspondence (Tel 028-86290916; fax 028-86290916; email shenyaou@sicau.edu.cn)

†Langlang Ma and Na Zhang contributed equally to this work.

Keywords: maize, shoot apical meristem, single-cell RNA sequencing, *ZmEREB14*, yield formation.

Summary

Shoot apical meristem (SAM) is the origin of aerial structure formation in the plant life cycle. However, the mechanisms underlying the maize SAM development are still obscure. Here, approximately 12 700 cells were captured from the 5-day-old shoot apex of maize using a high-throughput single-cell transcriptome sequencing. According to the gene expression patterns, we partitioned the cells into 8 cell types with 13 transcriptionally distinct cell clusters and traced the developmental trajectory of shoot apex. Regulatory network analysis of transcription factors (TFs) showed that three core TFs, AP2-EREBP-transcription factor 14 (*ZmEREB14*, *Zm00001d052087*), MYB histone 4 (*ZmMYB4*) and HSF-transcription factor 8 (*ZmHSF8*) potentially regulated the SAM development. Functional validation revealed that *ZmEREB14* affected the SAM development and thereby regulated the maize yield formation. Our results characterised the inherent heterogeneity of SAM at single-cell resolution and provided new insights into the mechanisms of SAM development.

Introduction

Plant organogenesis spans the whole life by maintaining the pluripotency of stem-cell populations residing in the shoot apical meristem (SAM) after embryogenesis. All the above-ground organs of a mature plant besides hypocotyls and cotyledons are generated from SAM (Zhang *et al.*, 2021a). For the model plant *Arabidopsis*, SAM can be distinguished into three layers, namely, L1, L2 and L3 (Satina *et al.*, 1940; Zhang *et al.*, 2021a). Most of the cells derived from L1 layer form the epidermis of shoots, flowers and leaves, whereas L2 and L3 layers provide cells for giving rise to the central tissues of the leaves and stems (Satina *et al.*, 1940; Zhang *et al.*, 2021a). In maize, SAM includes only two layers L1 and L2 (Figure 1a). According to the staining patterns and cell division arrangement, SAM can be typically divided into three distinct zones, namely, central zone (CZ), peripheral zone (PZ) and rib zone (RZ) (Kerstetter and Hake, 1997). CZ is composed of apical cells of SAM, which divide less frequently than that of PZ cells. Divisions of CZ cells contribute to maintenance of the indeterminate cell populations and supplementation of the cells for leaf primordia and stems (Kerstetter and Hake, 1997). PZ includes rapidly dividing cells and inconspicuous vacuoles (Kerstetter and Hake, 1997). The PZ cells contribute to the organ initiations (Kerstetter and Hake, 1997; Zhang *et al.*, 2021a). RZ consists of the cells at the bottom of SAM, which are responsible for forming stem tissues (Kerstetter and Hake, 1997).

To date, several transcription factors (TFs) have been shown to participate in SAM function. Class I *KNOTTED1-like homeobox* (*KNOX1*) genes are essential for the SAM maintenance and for the

identification of organ boundaries (Laufs *et al.*, 1998; Shani *et al.*, 2006). Three NAC family transcription factors, *CUP SHAPED COTYLEDON1* (*CUC1*), *CUC2* and *CUC3* are required for SAM establishment and for the formation of organ–organ and meristem–organ boundaries (Aida and Tasaka, 2006; Shani *et al.*, 2006). In *Arabidopsis*, a canonical *WUSCHEL* (*WUS*)–*CLAVATA* (*CLV*) negative feedback loop explains the maintenance of stem cell homeostasis and the alteration of shoot meristem size. In this signalling pathway, *CLV3* expression negatively regulates the *WUS* expression via the *CLV1*–*CLV2* receptor complex, which controls the shoot meristem size (Satterlee *et al.*, 2020; Shani *et al.*, 2006). Recently, as the co-receptors of *CLV1*, *CLV2/CORYNE* and *RECEPTOR-LIKE PROTEIN KINASE 2*, a group of leucine-rich repeat receptor-like protein kinases are found to mediate *CLV3* signalling pathway through phosphorylation (Hu *et al.*, 2018). In addition, hormones cytokinin, gibberellin, auxin and ethylene are involved in meristem maintenance by regulating expressions of the above transcription factors (Shani *et al.*, 2006).

Although some studies have been carried out on SAM, the meristematic functions and differentiation trajectories of SAM cells are still ambiguous due to the difficulties in meristem separation. According to the bulk RNA sequencing data, previous studies used a marker-biased method to profile the gene expressions of SAM stem cells in *Arabidopsis* (Tian *et al.*, 2019; Yadav *et al.*, 2009). Compared with the bulk RNA sequencing, the single-cell RNA sequencing (scRNA-seq) developed in recent years enables to achieve the expression patterns of genes in organs at single-cell resolution. In *Arabidopsis*, rice, maize, etc., scRNA-seq has been applied to dissect the spatial distribution and

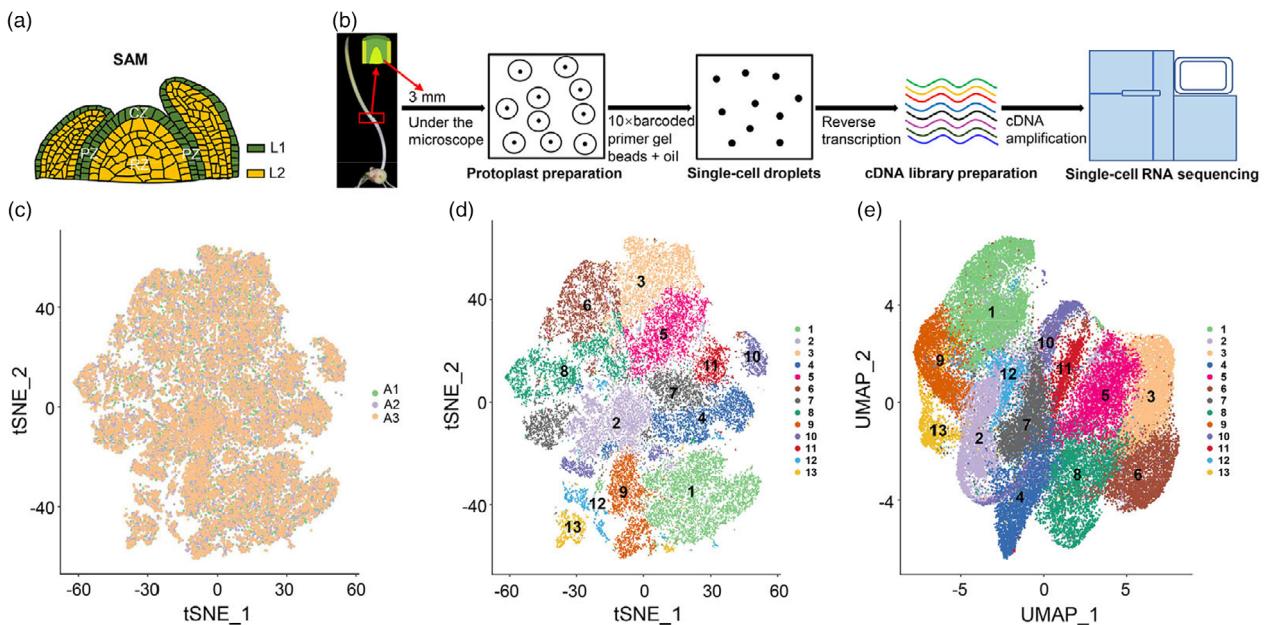


Figure 1 Cell heterogeneity of shoot apex in maize. (a) Schematic of anatomy of shoot apex in maize. CZ, central zone; PZ, peripheral zone; RZ, rib zone. (b) Workflow of single-cell RNA sequencing. (c) t-distributed Stochastic Neighbour Embedding (tSNE) projection plot of cells from three replicates. A1, replicate 1; A2, replicate 2; A3, replicate 3. (d) tSNE projection plot of 13 cell clusters. Each colour represents a cell type. (e) Uniform Manifold Approximation and Projection (UMAP) projection plot of 13 cell clusters. Each colour represents a cell type.

temporal ordering of distinct organ cells (Bezrutczyk *et al.*, 2021; Li *et al.*, 2022; Liu *et al.*, 2022; Ryu *et al.*, 2019; Satterlee *et al.*, 2020; Zhang *et al.*, 2021a, 2021b). Herein, we performed a scRNA-seq to dissect the inherent heterogeneity of SAM at single-cell level and excavate the core regulators involved in SAM development.

Results

scRNA-seq characterizing the gene expression profiles of shoot apex

According to the flow path of scRNA-seq (Figure 1b), we individually profiled 13 672, 13 723 and 13 364 single cells from three biological replicates (Table S1). After quality control of raw data, a total of 12 748, 12 705 and 12 723 single cells from each of three replicates were used for further analysis (Figure S1a-c; Table S1). The scRNA-seq revealed approximately 2500 genes and 6800 unique molecular identifiers (UMIs) in per cell (Table S1). The proportion of mitochondrial genes was equal to 1.58% (Figure S2; Table S1) and was lower than the threshold of 5% reported in the previous study (Osorio and Cai, 2021), indicating that the sequencing data were reliable. Cells from each of the three replicates were largely overlapped and had a high degree of reproducibility according to t-distributed stochastic neighbour embedding (t-SNE) projection (Figure 1c). To assess the robustness of the scRNA-seq and effect of protoplasting, we performed a bulk RNA-seq on shoot apexes of five-day-old maize seedlings without protoplasting. A strong correlation was observed between scRNA-seq and bulk RNA-seq expression data (Pearson's correlation coefficient = 0.84, $P < 2.2\text{e-}16$) (Figure S1d). All captured cells were classified into 13 major clusters by t-SNE and uniform manifold approximation and projection (UMAP) analyses (Figure 1d,e). Most of the top 10 marker genes were uniquely

enriched in each corresponding cluster (Figure S3a). The expression patterns of representative marker genes are displayed in Figure S3b.

Identification of cell types and functional enrichment of differentially expressed genes (DEGs)

To assign the 13 clusters to known cell types, the gene(s) that was among the top 10 zone-specific genes in each cluster and was reported to specifically express in a zone of shoot apex was selected as the marker gene(s) for the zone. Among the top 10 zone-specific genes in cluster 1, *long chain acyl-CoA synthetase2* (*ZmLacs2*), *eceriferum5* (*ZmCer5*) and *outer cell layer5a* (*ZmOcl5a*) were determined as the marker genes for epidermis cell (EC) based on the previous study (Figure S4; Table S2; Satterlee et al., 2020). Among these, *ZmOcl5a* is a homologue of *ZmOcl4* that has been identified as a marker gene for EC (Satterlee et al., 2020; Vernoud et al., 2009). In cluster 9, *ZmOcl4*, *hydroxycinnamoyltransferase12* (*ZmHct12*), *phospholipid transfer protein33* (*ZmPlt33*), *NOD26-like membrane intrinsic protein1* (*ZmNip1a*) and *semi-rolled leaf5* (*ZmSrl5*) were distinguished as marker genes for EC (Figure S4; Table S2; Satterlee et al., 2020). In clusters 3, 6 and 8, two (*blue copper protein*, *ZmBcp*; *ferulic acid 5-hydroxylase2*, *ZmFah2*), three (*cortical cell-delineating protein*, *ZmCcdp*; *Zm00001d004231*; *bHLH-transcription factor40*, *ZmBhlh40*) and four (*triacylglycerol lipase*, *ZmTrl*; *heat shock protein18c*, *ZmHsp18c*; *lifeguard2*, *ZmLfg2*; *Zm00001d029747*) genes were considered as marker genes for vasculature cell (VC) (Figure S4; Table S2; Knauer et al., 2019), respectively. In cluster 4, the three genes that respectively encode photosystem I (Psa)/photosystem II (Psb)-related protein, cytochrome (Pet) and NADH dehydrogenase subunit (Ndh) were identified as the marker genes for mesophyll cell (MLC), because they are closely correlated to photosynthesis in MLC (Figure S4; Table S2; Vinayak et al., 2021). According to the

previous studies (Elhaddad *et al.*, 2014; Juarez *et al.*, 2004; Satterlee *et al.*, 2020; Zhang *et al.*, 2018), we identified xyloglucan endo-transglycosylase/hydrolase5 (*ZmXth5*, in cluster 5), growth-regulating-factor-interacting factor1 (*ZmGif1*, in cluster 7), axial regulator *yabby2* (*ZmYabby2*, in cluster 10) and protein kinase *apk1b* (*ZmApk1b*, in cluster 11) as the marker genes for meristem cell (MMC), meristem determinacy cell (MDC), transit-amplifying cell (TAC) and guard cell (GC), respectively (Figure S4; Table S2). Most of the top 10 genes in cluster 12 were annotated as histone, which has been proven to be a crucial factor in cell proliferation (Figure S4; Table S2; Zhu *et al.*, 2011; Gutierrez *et al.*, 2016), cluster 12 was thereby identified as proliferating cells (PCs). As a previously reported marker gene for PC, *cyclin3* (*ZmCyc3*) was among the top 10 zone-specific genes in cluster 13 (Figure S4; Table S2; Weingartner *et al.*, 2003). Taken together, these evidences indicated that the cell clusters 1 and 9 belong to EC; clusters 3, 6 and 8 belong to VC; clusters 4, 5, 7, 10 and 11 belong to MLC, MMC, MDC, TAC and GC, respectively; clusters 12 and 13 belong to PC. We did not identify known marker genes in cluster 2; therefore, this cluster was termed as ‘unknown cell’. We then randomly selected four genes (*ZmpsaB*, *ZmGif1*, *ZmXth5* and *ZmOcl4*) to conduct RNA *in situ* hybridisation for validating the reliability of cell type classification in the present study. Finally, *ZmpsaB* (belonging to MLC), *ZmGif1* (belonging to MDC), *ZmXth5* (belonging to MMC) and *ZmOcl4* (belonging to EC) were highly expressed in leaf primordia and leaves, centre zones and leaf primordia, meristem and epidermis and leaf primordia, respectively (Figure S3c–j). All these findings suggest that the cell-type annotations for the clusters were accurate.

In addition, we compared the results of the present study and previous results of SAM-related studies. The representative marker genes identified in maize SAM using scRNA-seq (Satterlee *et al.*, 2020) were expressed in the corresponding clusters divided in the present study (Figure S5a–d). The expressions of the common DEGs in each cell type between the results of the present study and previous study (Satterlee *et al.*, 2020) were used to calculate the correlations based on the ‘Spearman Correlation Analysis’. Significant correlations in gene expression levels were observed between our study and the previous study for the cluster pairs VC vs. VC ($r_s = 0.24$, $P = 2.0E-03$), MLC vs. Leaf primordia ($r_s = 0.51$, $P = 2.4E-16$), MDC vs. Indeterminate cell ($r_s = 0.53$, $P = 1.1E-24$) and EC vs. EC ($r_s = 0.47$, $P = 3.1E-14$) (Figure S5e–h). Similarly, we analysed the expression correlations between our study and a gene expression study on distinct regions of maize meristem using LCM-RNAseq (Knauer *et al.*, 2019). Significant correlations were found between EC and L1 ($r_s = 0.67$, $P < 2.2E-16$), MMC and MMC ($r_s = 0.59$, $P < 2.2E-16$), MLC and P2 ($r_s = 0.66$, $P < 2.2E-16$) and VC and VC ($r_s = 0.66$, $P < 2.2E-16$; Figure S6). Collectively, these findings support the reliability of the cell type characterisation in our study. Moreover, the marker genes that were not identified in the previous studies (Table S3) can be considered as new marker genes for distinct cell types. We then investigated expression levels of the representative new marker genes in various cell types reported by Knauer *et al.* (2019). Ultimately, these marker genes were highly expressed in their corresponding cell types (Figure S7), verifying the credibility of the new markers.

In addition, we identified three additional cell types (TAC, GC and PC) in comparison to the previous study (Satterlee *et al.*, 2020). However, the S-phase and G2/M-phase cell types classified in the previous study were undefined in the present study. To verify whether TAC, GC and PC distinguished from the

S-phase and G2/M-phase cell types, we analysed the expressions of the top 10 marker genes of S-phase and G2/M-phase cells in our expression data using Seurat (v3.1.1). As results, the top 10 genes of S-phase and G2/M-phase cells were specifically highly expressed in clusters 12 (PC) and 13 (PC), respectively (Figure S8). Therefore, PC belonged to S-phase and G2/M-phase cells, however, TAC and GC were the new cell types.

To understand the biological functions of the DEGs in each cell type, we performed a Gene Ontology (GO) enrichment analysis (Table S4). The functions of DEGs in different cell types are displayed in Figure S9a. Notably, the DEGs in MMC only shared auxin binding (GO: 0010011) and IAA-Ala conjugate hydrolase activity (GO: 0010179) with other cell types, indicating that the MMC probably perform a relatively unique biological function (Figure S9a). GO enrichment analysis showed that the DEGs in MMC were associated with peptide metabolic/biosynthetic process (GO: 0006518/GO: 0043043), translation (GO: 0006412), cellular amide metabolic process (GO: 0043603), etc. (Figure S9b). In addition, we analysed the expression patterns of the classic SAM-regulated genes, auxin metabolism-related genes and auxin-related genes identified in MMC. Among the SAM-regulated genes, *ZmKN1*, *ZmTD1* and *ZmFEA2* showed high expressions in comparison to *ZmWUS1*, *ZmWUS2*, *ZmCUC1*, *ZmCUC2* and *ZmCUC3* (Figure S10a). Notably, *KN1* was highly expressed in the clusters 5 (MMC) and 8 (VC) (Figure S10a). In addition, *ZmWUS1* was slightly expressed in cluster 5, 6 and 7, whereas it was not detected in maize SAM in the previous study (Satterlee *et al.*, 2020). The phenomenon was likely due to the difference in genetic background between the A188 and B73 lines. An alternative explanation was the different sampling time for shoot apex scRNA-seq between the two studies. For the eight auxin-related genes in MMC, *ZmIAA23* displayed high expressions in the clusters 5 (MMC), 3 (VC) and 11 (GC) (Figure S10b); *ZmIAA32* was highly expressed in all the clusters, especially in the clusters 5 (MMC), 1 (EC), 7 (MDC) and 9 (EC) (Figure S10b). These findings suggest that *ZmIAA23* and *ZmIAA32* probably play critical roles in regulating distinct cells of SAM. In addition, auxin biosynthesis-related gene *ZmYUC6* and auxin transport-related gene *ZmPILS6/PIN14* showed higher expression levels in comparison to the other genes (Figure S10c), supporting the important roles of auxin metabolism in SAM development.

Trajectory tracking of shoot apex development

To know the temporal and spatial patterns of the marker genes during the shoot apex developmental progress, we constructed continuous differentiation trajectories using the Monocle 2 (Qiu *et al.*, 2017). The result showed that the pseudo-time trajectory contained two branch points (Figure 2a). As well known, meristem is the determinant factor affecting shoot apex development, which mainly functions in the prophase of plant development. Here, the MMCs were mainly expressed at early and mid-term developmental stages as expected (Figure S11). In addition, we performed an RNA velocity analysis to confirm the pseudo-temporal model using velocyto (La Manno *et al.*, 2018). RNA velocity is an effective approach for analysing the dynamic changes for transcript expressions and revealing the cell fate direction (La Manno *et al.*, 2018). Herein, the transcriptional profiles in MMCs and MDCs displayed little changes (short or no arrows), whereas they were largely changed (long arrows) in VCs, PCs and ECs (Figure 2b). These findings were consistent with the pseudo-time trajectory results. Moreover, the developmental

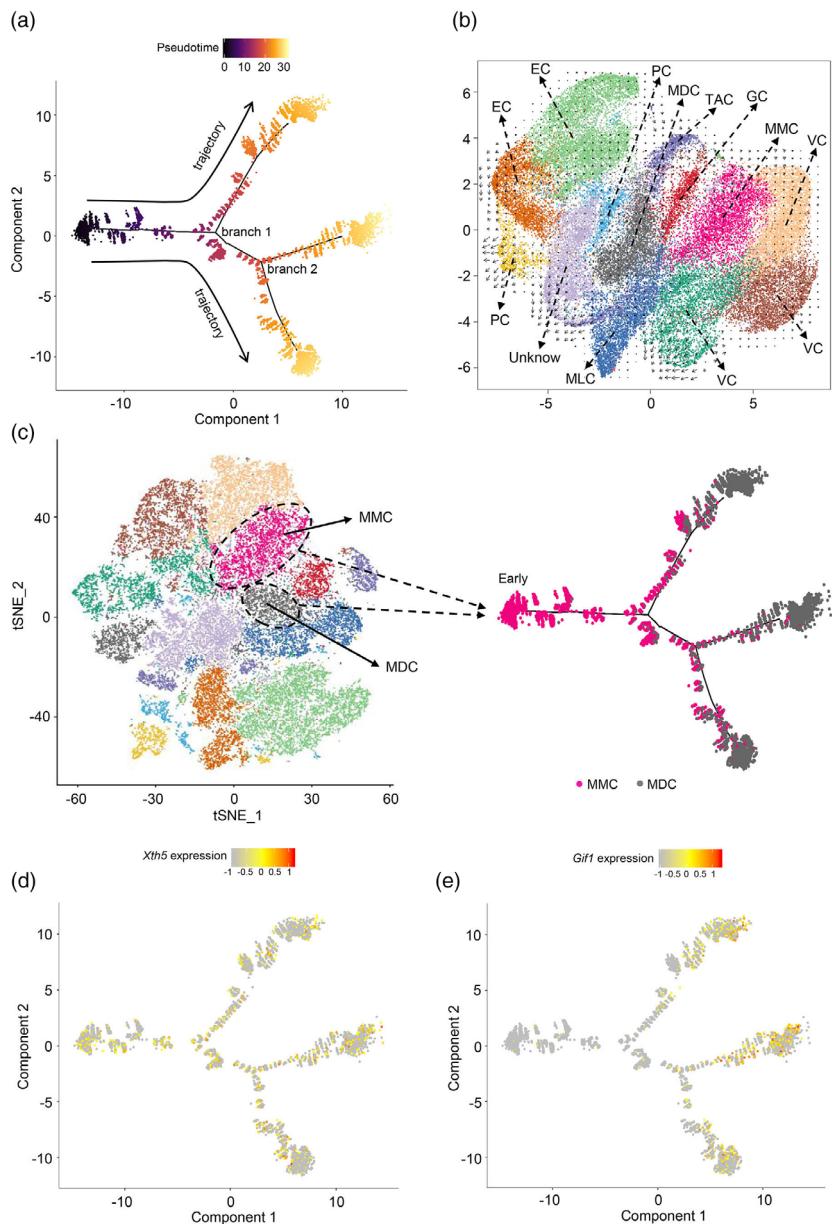


Figure 2 Developmental trajectory of maize shoot apex. (a) Developmental trajectory achieved by pseudo-time analysis. (b) RNA dynamics of different cell types achieved by RNA velocity analysis. EC, epidermis cell; GC, guard cell; MDC, meristem determinacy cell; MLC, mesophyll cell; MMC, meristem cell; PC, proliferating cell; TAC, transit-amplifying cell; VC, vasculature cell. Arrow direction represents the state transition direction of RNA dynamics. Arrow length represents the change extent of RNA dynamics. (c) Developmental trajectory of meristem cell (MMC) and meristem determinacy cell (MDC). (d–e) Expression pattern of marker genes in meristem cell (*ZmXth5*) (d) and meristem determinacy cell (*ZmGif1*) (e).

direction was from MMCs to the other cell types (Figure 2b), suggesting the MMCs were original elements of shoot apex development. For the cell types MMC and MDC, developmental trajectory showed that the MDCs were derived from MMCs (Figure 2c). The marker genes *ZmXth5* (cell type: MMC) and *ZmGif1* (cell type: MDC) were expressed as expected (Figure 2d, e). We applied GeneSwitches (Cao *et al.*, 2020) to uncover the ordering of critical gene expression during SAM development, which can identify the status (on/off) of switching genes along pseudotime. As a result, the SAM-regulated gene *ZmKNO1* (belonging to branch 1) was deactivated early, whereas *ZmTD1* (belonging to branch 2) and *ZmWOX2b* (belonging to branch 2)

were activated mid-late (Figure S12a). Among the top 20 switching genes, *ZmRPS13* (40S ribosomal protein S13) and *ZmIHT2* (Lysine histidine transporter 2) showed the highest quality of fit determined by McFadden's Pseudo R² (Figure S12a,b), suggesting the two genes probably act as critical switching genes in regulating the SAM development. Notably, we uncovered several 40S/60S ribosomal proteins (*ZmRPS13*, *ZmRPS5*, *ZmRPS10*, *ZmRPL28*, *ZmRPL34* and *ZmRPL19*) in branch 1 (Figure S12a), which have been validated to mediate plant SAM development (Jirásková, 2022; Luo *et al.*, 2023). We also analysed the genes expressed in MMC using GeneSwitches tool. Five TFs were identified as switching genes, of which *ZmWRKY64* and

ZmEREB14 showed the highest R^2 (Figure S12c,d). Among the top 20 switching genes in MMC, *ZmCNR9* (Cell number regulator 9) and *ZmPGIP3* (Polygalacturonase-inhibiting protein 3) were two lead candidates (Figure S12c,d), which have been confirmed to affect cell number and leaf primordium growth, respectively (Guo et al., 2010; Yang et al., 2021). The above genes can be considered as candidate markers to identify differentiation process of SAM cell. We then conducted GO analyses based on the highly variable genes. For branch point 1, the highly variable genes were clustered into three modules, which were separately enriched in two (ion transport; hormone metabolism), two (biosynthesis and molecular complex assembly; gene expression regulation) and one (single cellular localisation) biological processes (Figure S13). For branch point 2, two (ion transport;

metabolism of lignin, phenylpropanoid and phosphorus), two (molecular metabolic process; molecular complex assembly) and one (establishment of cellular localisation) GO terms were obtained for the three modules, respectively (Figure S14).

Identification of core TFs of shoot apex development

The development of maize SAM is controlled by a complex transcription regulatory network. Several TFs (including *KNOX1*, *WUS* and *NAC*) regulating the SAM development have been identified in previous studies (Aida and Tasaka, 2006; Laufs et al., 1998; Satterlee et al., 2020; Shani et al., 2006). As such, we investigated the TFs expressed in MMC and MDC for identifying the causal regulators of SAM development. A total of 10 and 5 TFs were identified to regulate 187 and 99 target

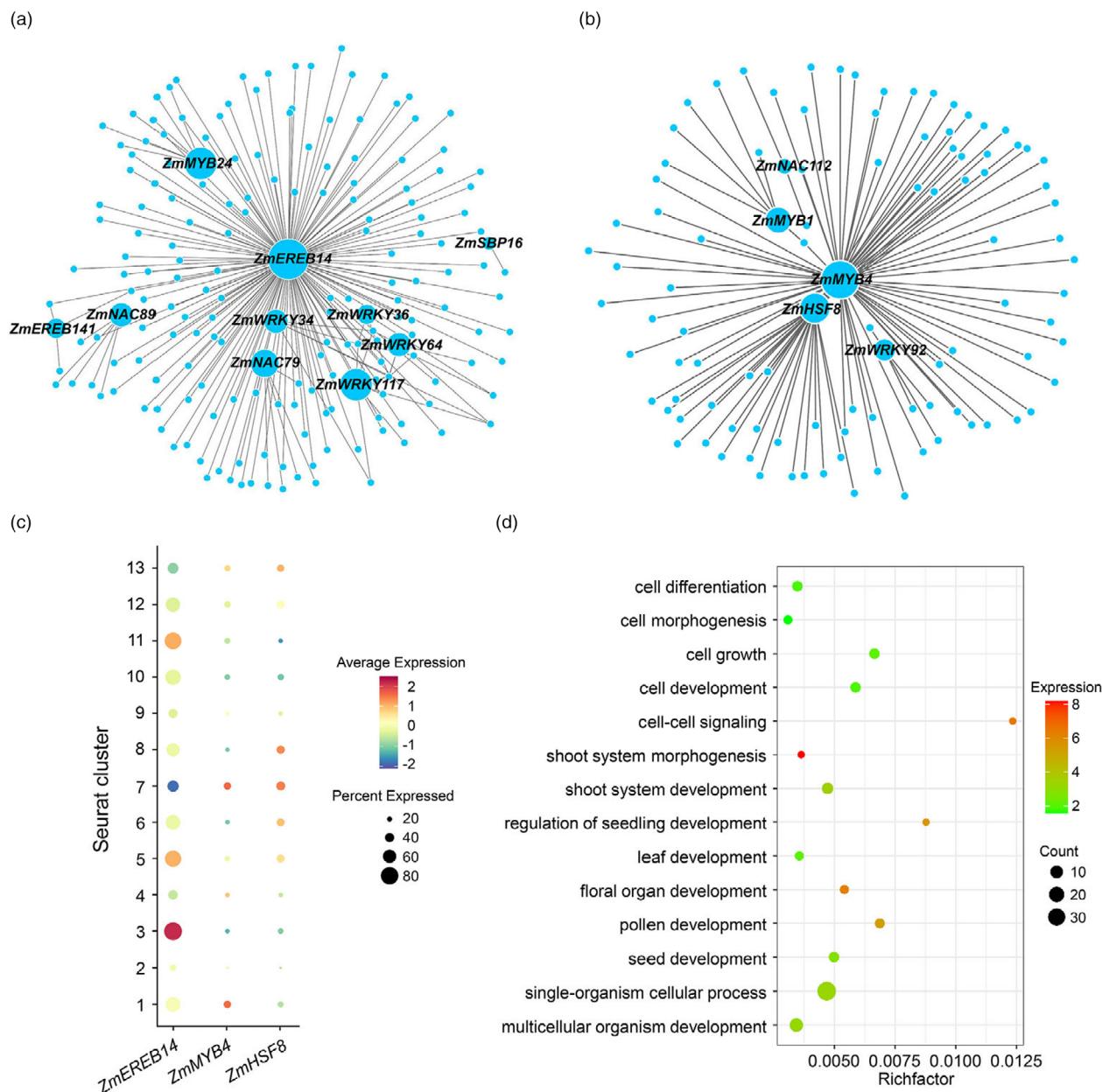


Figure 3 Identification of transcription factors (TFs) involved in maize shoot apex. (a,b) Regulatory networks of TFs and their target genes in meristem cell (MMC) (a) and meristem determinacy cell (MDC) (b). (c) Expression patterns of three core TFs detected in MMC and MDC in 13 clusters. (d) Gene Ontology (GO) enrichment of 145 target genes of *ZmEREB14*.

genes in MMC and MDC, respectively, with the threshold P -values $<1 \times 10^{-5}$ (Figure 3a,b). As a core TF in MMC, AP2-EREBP-transcription factor 14 (*ZmEREB14*) regulated 145 target genes (Figure 3a). In MDC, 56 and 31 target genes were separately controlled by the core TFs single MYB histone 4 (*ZmMYB4*) and HSF-transcription factor 8 (*ZmHSF8*) (Figure 3b). These core TFs were highly expressed in the corresponding cell types (Figure 3c). Previous studies had revealed that the TFs AP2-EREBP, MYB and HSF were involved in somatic embryogenesis, shoot developments or shoot meristem maintenance in *Arabidopsis*, rice, oil palm and *Liriodendron chinense* (Hammoudi et al., 2021; Khadem et al., 2023; Lin et al., 2007; Morcillo et al., 2007; Yang et al., 2002; Zong et al., 2021). Interestingly, the 145 target genes of *ZmEREB14* were enriched in cell differentiation, shoot system development, regulation of seedling development, leaf development, multicellular organism development and other categories, suggesting that *ZmEREB14* was a multiple-effect factor on organ development (Figure 3d). Moreover, the targets of *ZmMYB4* were mainly enriched in translation, regulation of cell cycle phase transition (G1/S and G2/M) and cellular component assembly categories (Table S5); the targets of *ZmHSF8* were mainly enriched in chaperone-mediated protein complex assembly, cellular response to calcium ion and positive regulation of translational initiation categories (Table S6). Among these three TFs, *ZmEREB14* was identified as a critical switching gene in MMC using the GeneSwitches analysis. Thus, *ZmEREB14*

was considered as a causal gene controlling shoot apex development and was selected for further functional validation.

ZmEREB14 regulates the SAM development and yield formation in maize

Previous studies have shown that some EREB proteins were involved in the determination of floral organ identity, manipulation of leaf epidermal cell identity, specification of organ and meristem identity, seed development and response to various biotic and abiotic stresses (Riechmann and Meyerowitz, 1998). Phylogenetic analysis showed that *ZmEREB14* was closely to SORBI_3004G321300 from sorghum and Zm00001d018305 from maize among 47 putative homologues (Figure S15). Up to now, *ZmEREB14* and the two homologues have not been functionally characterised in SAM development. To understand the expression pattern of *ZmEREB14*, we performed real-time quantitative PCR (qRT-PCR) in seven tissues, including leaves, roots, stems, tassels, ears, kernels and shoot tips with the primers displayed in Table S7. In results, *ZmEREB14* showed the highest expression in shoot tips and the lowest expression in tassels, respectively (Figure 4a). RNA in situ hybridisation revealed that *ZmEREB14* was highly expressed in the vasculature, meristem and leaf primordium of the shoot apex (Figure S16a,b), which was consistent with the scRNA-seq data. In the developing ear, the *ZmEREB14* transcripts were mainly enriched in the inflorescence meristem (IM), spikelet pair meristem (SPM) and spikelet meristem

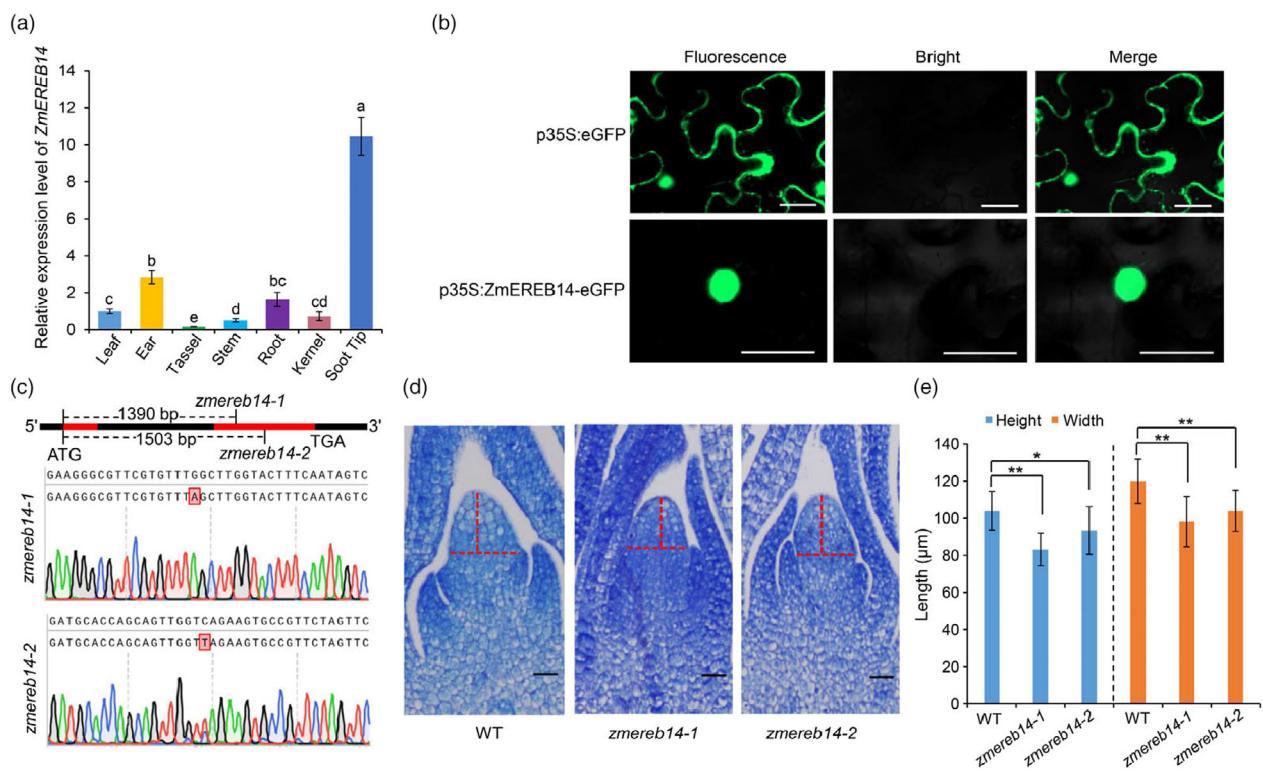


Figure 4 *ZmEREB14* regulating the shoot apical meristem (SAM) size in maize. (a) Quantitative real-time polymerase chain reaction (qRT-PCR) analysis of *ZmEREB14* in distinct tissues. Different small letters above the columns represent significantly different at $P \leq 0.05$ level based on Student's *t*-test. (b) Subcellular localisation of *ZmEREB14* protein. Bar = 20 μm . (c) Characteristics of the two mutants *zmereb14-1* and *zmereb14-2*. (d,e) Morphology examination (d) and statistics (e) on SAM height and width of wild-type (WT) line and two mutants. The sample numbers of wild-type (WT), *zmereb14-1* and *zmereb14-2* were 10, 9 and 9, respectively. Bar = 40 μm . ** and * indicate significantly different at $P \leq 0.01$ and $P \leq 0.05$ based on Student's *t*-test, respectively.

(SM) of 2–4 mm developing ear, in the SM and floral meristem (FM) of 6–8 mm developing ear and in the FM of 10–12 mm developing ear (Figure S16c–h). This was supported by the development process of maize ear. Subcellular localisation analysis using tobacco leaves indicated that the ZmEREB14-GFP fusion protein was specifically targeted to the nucleus (Figure 4b), supporting that ZmEREB14 acts as a TF.

To clarify the role of ZmEREB14, we used two ethyl methane sulphonate (EMS) mutants *zmereb14-1* and *zmereb14-2*, in which two stop-gain mutations individually occurred on the second exons and led to knockout of ZmEREB14 in the mutants (Figure 4c). Homozygous mutants were screened using ZmEREB14-based RCR with the primers (Table S7), which were then subjected to phenotypic investigation. The SAM size (height and width) of *zmereb14-1* and *zmereb14-2* were significantly ($P \leq 0.05$) reduced when compared to that of the wild-type B73, with the reduction proportions of 20.02% on height and 18.19% on width for *zmereb14-1* and 10.13% on height and 13.32% on width for *zmereb14-2* on SAM (Figure 4d,e). To verify whether

ZmEREB14 affected the axillary meristem (AM), we measured the AM size of these lines. The height and width were both significantly ($P \leq 0.05$) decreased in the mutants relative to those in B73 (Figure S17). Moreover, we examined the vegetative phenotypes at the maturation stage. The mutants *zmereb14-1* and *zmereb14-2* showed significantly ($P \leq 0.05$) decreased plant height, primary tassel length, ear height, flowering time, aboveground biomass, average length of internodes below ear, average length of internodes above ear and internode number, in comparison with those of the wild-type B73 (Figure 5a–j). Since the average length of internodes below ear and average length of internodes above ear were both reduced in *zmereb14-1* and *zmereb14-2*, we investigated the cell size of internodes below/above ear using cytological analysis. The results displayed that the cell length and width of B73 were both greater than those of *zmereb14-1* and *zmereb14-2* (Figure 5k–m). The yield-related traits including ear length, ear weight, kernel weight per ear, hundred kernel weight, kernel number per ear, kernel length and kernel width, were all significantly ($P \leq 0.05$) lower in

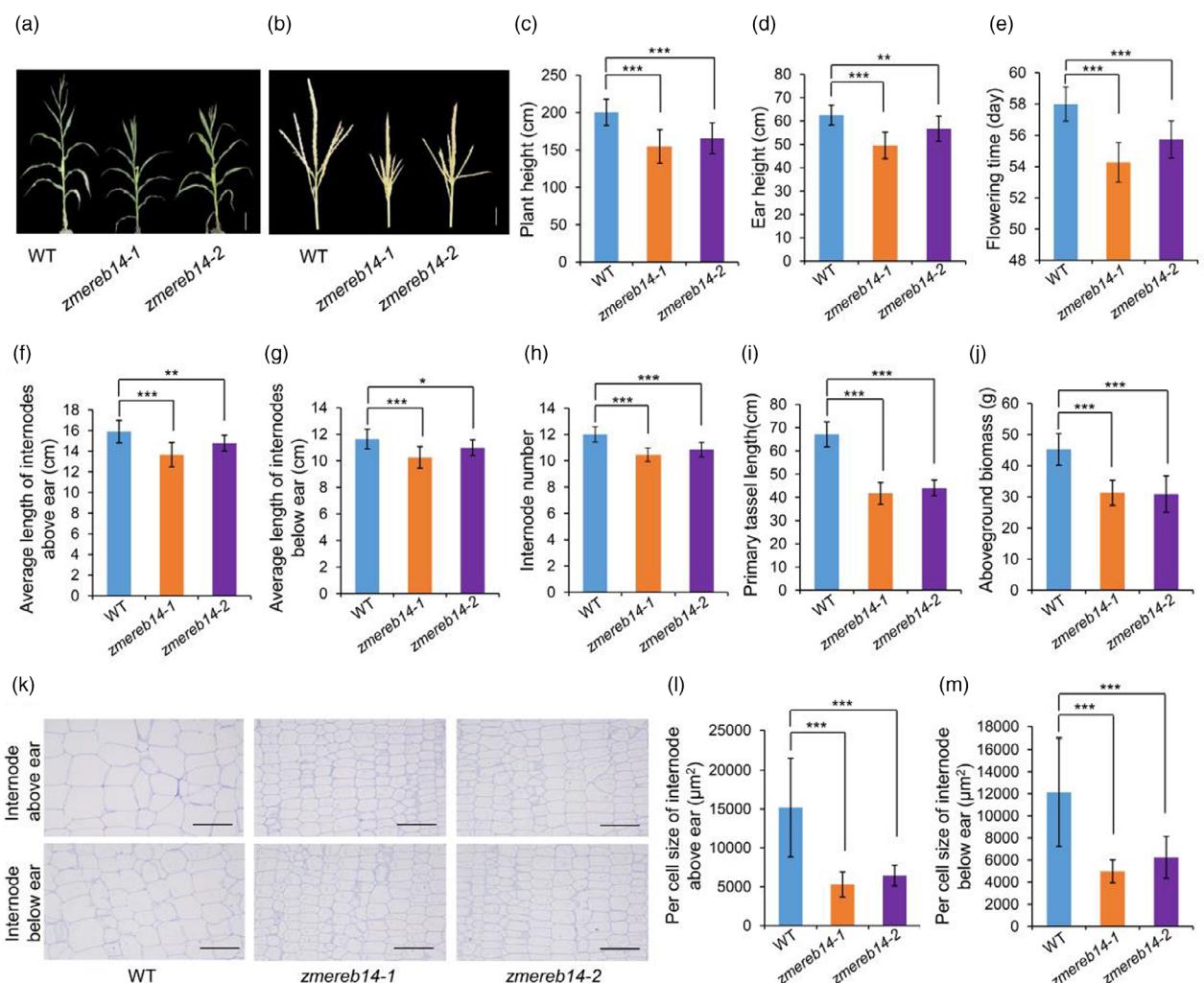


Figure 5 ZmEREB14 regulating the vegetative traits in maize. (a,b) Plant (a) and tassel (b) characteristics of the wild-type (WT) line and two mutants *zmereb14-1* and *zmereb14-2*. Bar = 20 cm and Bar = 5 cm, respectively. (c–j) Plant height (c), ear height (d), flowering time (e), length of internodes above ear (f), length of internodes below ear (g), internode number (h), primary tassel length (i) and aboveground biomass (j) of WT line and two mutants *zmereb14-1* and *zmereb14-2*. (k–m) Cell morphology examination (k) and statistics on internodes above (l) and below (m) the ear of WT line and two mutants *zmereb14-1* and *zmereb14-2*. Bar = 200 μm . ***, ** and * indicate significant difference at $P \leq 0.001$, $P \leq 0.01$ and $P \leq 0.05$ levels based on Student's *t*-test, respectively.

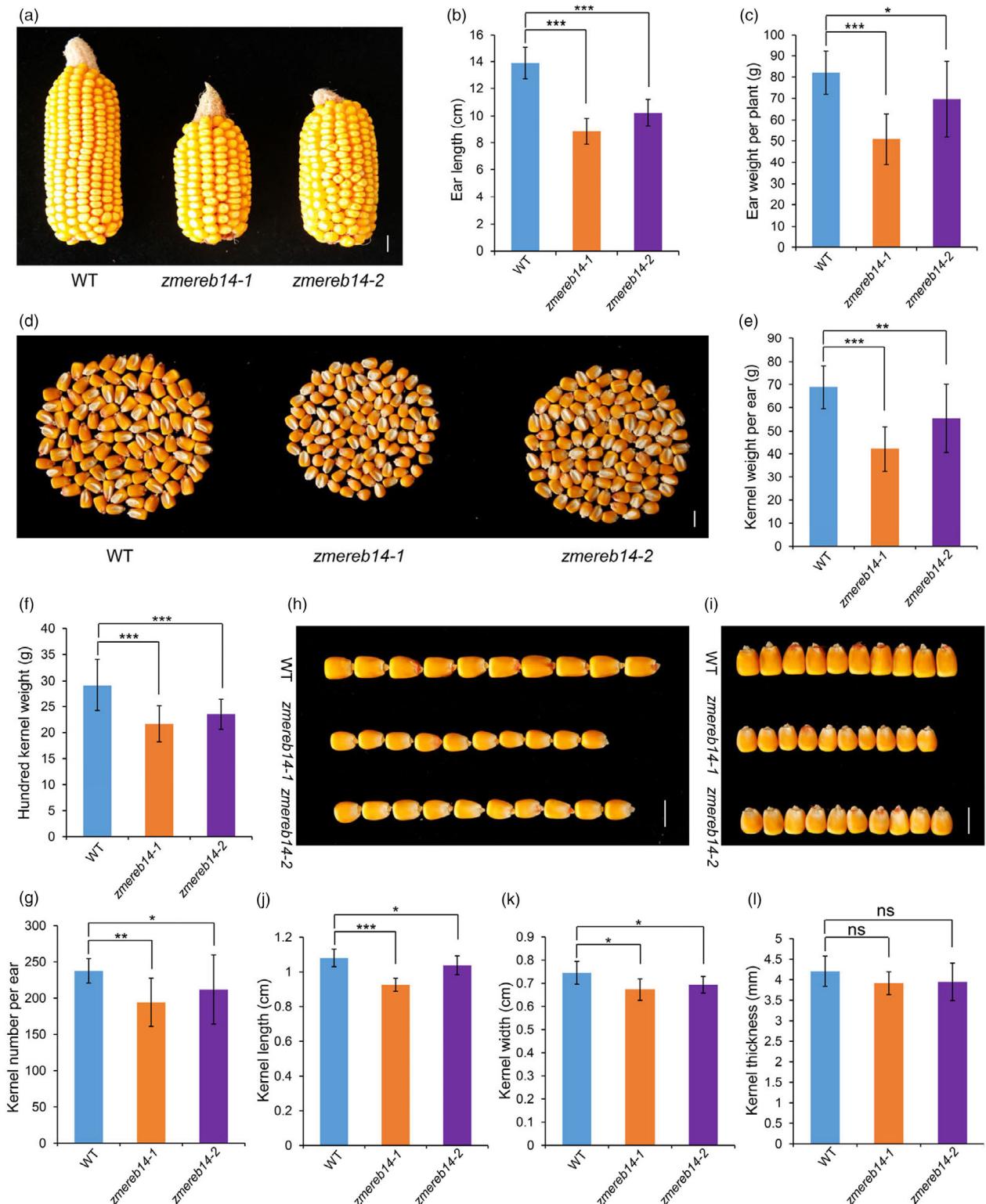


Figure 6 ZmEREB14 affecting maize yield. (a, d) Ear (a) and kernel (d) characteristics of the wild-type (WT) line and two mutants zmereb14-1 and zmereb14-2. Bar = 1 cm. (b, c, e, f, g) Statistics of ear length (b), ear weight (c), kernel weight per ear (e), hundred kernel weight (f) and kernel number per ear (g) of WT line and two mutants zmereb14-1 and zmereb14-2. (h–l) Kernel length (h, j), width (i, k) and thickness (l) of WT line and two mutants zmereb14-1 and zmereb14-2. Bar = 1 cm. ***, ** and * indicate significant difference at $P \leq 0.001$, $P \leq 0.01$ and $P \leq 0.05$ levels based on Student's *t*-test, respectively. ns indicates no significance.

zmereb14-1 and *zmereb14-2* than those in B73 (Figure 6a–k), except for kernel thickness (Figure 6l). All these findings indicate that *ZmEREB14* affects the maize SAM development and yield formation.

To investigate the regulatory network of *ZmEREB14* in SAM development, we compared the transcriptomes of the shoot apices from wild-type B73 and *zmereb14-1* mutant. According to the high-quality data obtained from RNA-seq (Figure S18a; Table S8), a total of 1929 (upregulated genes: 971; downregulated genes: 958) DEGs were identified in *zmereb14-1* vs. B73 (Figure S18b). GO enrichment analysis revealed that the upregulated DEGs were significantly ($P \leq 0.05$) enriched in ion import, ion homeostasis, guard cell differentiation, G1/S transition of mitotic cell cycle and etc. (Figure S18c). Notably, the downregulated DEGs were mainly enriched in response to cytokinin, regulation of meristem growth, regulation of seed growth, cytokinin-activated signalling pathway and meristem maintenance (Figure S18d). We analysed the expressions of SAM-regulated genes and auxin metabolism-related genes in shoot apices. Finally, *ZmCLV2* (*ZmFEA2*) was significantly ($P \leq 0.05$) increased, whereas the *ZmKN1*, *ZmCUC2* and *ZmCUC3* were significantly ($P \leq 0.05$) reduced in the mutant compared to those in B73 (Figure S18e). These implied that *ZmFEA2* was the negatively regulatory factor, and *ZmKN1*, *ZmCUC2* and *ZmCUC3* were the positively regulatory factors on SAM development. This finding was supported by the previous studies (Liu et al., 2021; Takada et al., 2001; Vollbrecht et al., 2000). We did not detect the expression values of *ZmCUC1*, *ZmWUS1* and *ZmWUS2* in both B73 and *zmereb14-1*. In addition, the expressions of auxin biosynthesis- (*ZmYUC6*), transport- (*ZmPIN1a* and *ZmPIN1b*) and signal transduction-related (*ZmTIR1*, *ZmBIF1/AUX22* and *ZmBIF4/IAA16*) genes were significantly ($P \leq 0.05$) lower in the mutant than those in B73 (Figure S19). Furthermore, the IAA content in B73 was significantly higher than that in the mutants (Figure S20). All these findings suggest that *ZmEREB14* probably participates in auxin biosynthesis, transport and signal transduction.

Discussion

scRNA-seq is efficient for cell-type identification of maize shoot apex

In recent years, scRNA-seq has become increasingly popular in studying various plants, which enables to unravel the heterogeneity and expression level of individual cells from target organs (Kim et al., 2021; Liu et al., 2022; Ryu et al., 2019; Satterlee et al., 2020; Zhang et al., 2021a, 2021b). In the present study, we captured ~12 700 cells of shoot apex at maize seedling stage using scRNA-seq and identified 13 major clusters that were assigned to eight cell types. *In situ* hybridisation experiments of representative genes confirmed the accuracy of cell-type annotation (Figure S3c–j). Currently, the studies on SAM mainly focus on *Arabidopsis*, which cause the rarity of marker genes of SAM in other plant species. The scRNA-seq of shoot apices of 14-day-old maize seedlings identified six cell categories in a recent research report (Satterlee et al., 2020). Therefore, we could not assign all clusters to the corresponding cell types. For example, none of the top 10 genes (or their homologous genes) in cluster 2 were previously reported in SAM studies, and no marker genes were thus identified in this cluster. Thereby, the cluster 2 was annotated as 'unknown cell' in the present study. However, the cell types EC, VC and PC all included multiple clusters (Figure 2b).

In the present study, we identified two additional cell types (TAC and GC) in comparison to the previous study (Satterlee et al., 2020). Correlations of gene expressions in the same cluster identified in the present study and the two previous studies were ranged from 0.24 to 0.67 (Figures S5 and S6). The relatively limited coefficients were probably caused by two reasons. Firstly, the cell type annotation is still a challenge due to the deprivation of marker genes in maize. Several cell types including companion cell, shoot endodermic cell and etc., which have been annotated in *Arabidopsis* (Zhang et al., 2021a), were undefined in the three studies (LCM-RNAseq, Knauer et al., 2019; scRNA-seq, Satterlee et al., 2020; scRNA-seq, the present study) on maize SAM. With the deepening of researches on SAM-regulated functional genes, the more sophisticated cell types of maize SAM would be classified in future. Secondly, the material background and sampling time were different between the present and previous studies (the present study: A188, five-day-old seedling; the previous studies: B73, 14-day-old seedling). In addition, GO terms enriched in MMC only share two overlap categories (auxin binding and IAA-Ala conjugate hydrolase activity) with those of other cell types (Figure S9a), indicating that the hormone signalling plays an important role in the entire process of SAM development. This is consistent with the results in the previous studies (Shani et al., 2006; Zeng et al., 2021) and is confirmed by the terms enriched with the highly variable genes in the developmental trajectory (Figure S13).

Developmental trajectory reconstruction of maize shoot apex

As a powerful tool, pseudo-time analysis has been widely used to determine the dynamics of gene expressions in scRNA-seq researches. The developmental process of SAM at different stages is affected by striking changes of transcription, biosynthesis and metabolism and cell differentiation and localisation. In this study, developmental trajectory analysis indicated that the cell types were originated from MMC and the pseudo-time trajectory contained two branch points (Figure 2a,b), which was consistent with the previous study (Satterlee et al., 2020). As expected, the representative marker genes, *ZmXth5* in MMC and *ZmGif1* in MDC were mainly expressed at early and mid-term, and late developmental stages, respectively (Figure 2c–e). GO analysis of the highly variable genes from the corresponding branches showed different enriched categories (Figures S13 and S14), supporting that these genes play distinct roles at different stages of SAM development. In addition, RNA velocity analysis showed that MMC and MDC had low RNA metabolism activity (Figure 2b). In contrast, RNA metabolism activity in VC, PC and EC was relatively high, suggesting that cells from these cell types were highly differentiated in the developmental process of SAM (Figure 2b). In addition, we also identified several novel switching genes (including *ZmRPS13*, *ZmRPS5*, *ZmRPS10*, *ZmRPL28*, *ZmRPL34*, *ZmRPL19*, *ZmlHT2*, *ZmWRKY64*, *ZmEREB14*, *ZmCNR9* and *ZmPGIP3*), which probably play important roles in regulating the SAM development. All these findings indicated that it is practicable to reconstruct the developmental trajectory of maize shoot apex using the expression abundance of each cell.

Identification of core TFs regulating the development of maize SAM

Currently, several TFs controlling SAM development have been characterised, such as *KNOXI*, *NAC* (*CUC1*, *CUC2* and *CUC3*) and *WUS* (Aida and Tasaka, 2006; Laufs et al., 1998; Satterlee

et al., 2020; Shani et al., 2006). Here, we screened the TFs regulating the target genes in MMC and MDC. Finally, we identified ten (*ZmEREB14*, *ZmEREB141*, *ZmWRKY64*, *ZmWRKY36*, *ZmWRKY34*, *ZmWRKY117*, *ZmNAC79*, *ZmNAC89*, *ZmMYB24* and *ZmSBP16*) and five (*ZmMYB1*, *ZmMYB4*, *ZmHSF8*, *ZmWRKY92* and *ZmNAC112*) TFs in MMC and MDC, respectively (Figure 3a,b). Among these TFs, *ZmEREB14*, *ZmMYB4* and *ZmHSF8* were identified as the core TFs, which separately regulated 145, 56 and 32 target genes (Figure 3a,b). As an EREB family member, *EgAP2-1* affects the embryo proliferation in oil palm (Morcillo et al., 2007). Moreover, several EREB family genes were found to regulate shoot development in *Liriodendron chinense* (Zong et al., 2021). In rice, an EREB gene *OsEBP-89* was confirmed to control shoot development (Yang et al., 2002). Khadem et al. (2023) revealed the regulation of HSFs on the core-enriched genes during somatic embryogenesis in *Arabidopsis thaliana* by using a gene set enrichment analysis (GSEA) (Khadem et al., 2023). As an HSF member, *SUMO* is critical for maintaining *Arabidopsis* shoot meristem (Hammoudi et al., 2021). In *Arabidopsis thaliana*, *AtCDC5* (belonging to MYB TF family) is essential for the G2/M phase transition and regulates the SAM function by controlling the expression of *STM* and *WUS* (Lin et al., 2007).

We investigated the *ZmEREB14* expression profile in various maize tissues by referring to several previous studies (Shen et al., 2023; Stelpflug et al., 2016; Yi et al., 2019). Ultimately, *ZmEREB14* was expressed in all the investigated tissues (Figure S21), which explained why *ZmEREB14* loss-of-function led to the changes on multiple vegetative and reproductive phenotypes. In leaves, the *ZmEREB14* expression showed peaks at V1 and VT stages (Figure S21g; Stelpflug et al., 2016), suggesting that it participated in leaf formation and transition from vegetative to reproductive growth. In the early developing kernels, embryos and endosperms, *ZmEREB14* indicated the highest expressions at 8–12 h after pollination, the critical stage for double fertilisation (Figure S21a–c; Yi et al., 2019). In the 2–24 DAP (days after pollination) kernels, embryos and endosperms, the expression peak occurred at 12–18 DAP, the crucial period for kernel formation (Figure S21d–f; Stelpflug et al., 2016). These implied that *ZmEREB14* played a role in regulation of fertilisation and kernel formation. In the immature ears, *ZmEREB14* maintained a stable expression level throughout the early development of ears and showed a slight upregulation at the stage of 3–5 cm ears, the decisive stage for silking (Figure S21h; Shen et al., 2023). Collectively, the core TF *ZmEREB14* probably played complex roles in meristem development, plant growth and yield formation in maize.

Previous studies have shown that the above-ground organs of mature plants and ears were originated from SAM and AM, respectively (Sun et al., 2020; Zhang et al., 2021a). In the present study, the SAM and AM were both reduced in the two mutants in comparison with those in B73, which explained the significant changes in most of the plant traits and ear traits (Figures 5 and 6). However, no significant changes were observed on tassel branch number and kernel row number per ear (Figure S22). We then examined the expression levels of several classical tassel branch-regulated and kernel row number-regulated genes in B73 and *zmereb14-1* according the transcriptome data. The expressions of *ZmRA2*, *ZmRA3*, *ZmBIF2* and *ZmKRN4* did not show significant difference between B73 and *zmereb14-1* (Figure S23), suggesting that these genes were not regulated by *ZmEREB14*. This could be one reason that

knockout of *ZmEREB14* did not cause the changes on tassel branch number and kernel row number.

Collectively, we characterised the cell types and reconstructed the developmental trajectory of maize shoot apex. Subsequently, we identified the core TFs involved in SAM size and kernel yield of maize. Our results provide new information about shoot apex development and will be helpful for perfecting the mechanisms of shoot apex development.

Materials and methods

Plant materials

Maize inbred line A188 was used for scRNA-seq of shoot apex in this study. Seeds with healthy and consistent growth were soaked with 6% sodium hypochlorite for 15 min and rinsed five times with distilled water. Thirty sterilised seeds were subjected to a germination experiment in a growth chamber with the conditions of a 16 h light at 28°C/8 h darkness at 22°C, and a 65% relative humidity for 5 days. For each seedling, approximately a 3 mm segment was excised above the node of the shoot apex. Most of the leaves wrapping the shoot apex were then stripped out under the microscope. The shoot apices with the 2–3 most recently initiated leaf primordia were collected for preparing protoplasts (Figure S24a). To ensure the reliability of the scRNA-seq data, we carried out three independent biological repetitions.

Two EMS mutants of *ZmEREB14* under the B73 background were ordered from MEMD (https://elabcaas.cn/memd/public/index.html#). Each EMS mutant was backcrossed with B73 (recurrent parent) to generate the BC₂F₂ population. In the populations, sibling plants of positive homozygous mutation at the target sites (*zmereb14-1* and *zmereb14-2*) were used to investigate the trait phenotypes. The non-mutant sibling plants were considered the negative control (wild type).

Protoplast isolation

The shoot apices (~3 mm) were cut into pieces of approximately 1 mm and submerged in protoplasting solution (1.5% cellulase R10, 0.5% macerozyme R10, 0.5 M mannitol, 0.03 M 4-Morpholineethanesulfonic acid hydrate [PH = 5.7], 0.03 M KCl, 0.015 M CaCl₂, 7.5 mM β-mercaptoethanol and 0.15% bovine serum albumin [BSA]) by vacuum infiltration (at 55 kPa) for 10 min. To facilitate enzyme solubilisation, the protoplasting solution was heated to 55°C for 10 min, prior to the supplement of CaCl₂, β-mercaptoethanol and BSA. Then, the samples were incubated at 25°C for 2 h at 50 rpm/min, and the mixture (Figure S24b) was filtered by a 40-μm nylon strainer and centrifuged at 1000 rpm/min for 2 min. After removing the supernatant, the retained protoplasts were resuspended in the wash solution (protoplasting solution without cellulase R10 and macerozyme R10) and rinsed three times using the same centrifugation conditions (Figure S24c). The trypan blue staining was used to test the protoplast activity. The concentration of protoplasts was determined using a haemocytometer. Herein, a sample with the proportion of living cells >85% and the number of living cells >1 × 10⁵ was subjected to scRNA-seq.

scRNA-seq

The flow path of the scRNA-seq is shown in Figure 1b. Briefly, the protoplast suspension was subjected to a 10x Chromium Controller (10x Genomics) to form single-cell droplets. Then, the RNA within the cells of barcoded droplet was reversely transcribed as cDNA, and the cDNA libraries were constructed

based on the manufacturer's instructions of the Chromium Single Cell 3' Reagent Kit v3. The libraries were sequenced with an Illumina NovaSeq 6000 platform (Illumina Inc., San Diego, CA, USA).

Bulk RNA-seq

Shoot apices of five-day-old seedlings were collected from A188 for extracting total RNA using the HiPure Plant RNA Maxi Kit (Guangzhou Magen Biotechnology Co., Ltd., Guangzhou, China). RNA libraries were constructed using NEBNext Ultra RNA Library Prep Kit (NEB, Ipswich, MA, USA) following the manufacturer's recommendations and were then sequenced on the Illumina sequencing platform with three biological repetitions. After removing the low-quality reads, the clean reads were aligned to the maize reference genome (B73 RefGen_V4) via Hisat2. The Pearson's correlation coefficient between scRNA-seq and bulked RNA-seq was calculated using the \log_2 -transformed read counts from scRNA-seq and bulked RNA-seq.

Analysis of scRNA-seq data

Raw reads were subjected to preliminary quality control using the 10 \times Genomics Cell Ranger software pipeline (v3.1.0) (Butler *et al.*, 2018), and then were mapped to the B73 reference genome (B73 RefGen_V4) using the STAR aligner (Dobin *et al.*, 2013). In further quality control, two criteria were applied to remove the low-quality cells and likely multiplet captures, that is, cells with UMI/gene numbers exceeding the mean value \pm twofold of standard deviations and cells where $>10\%$ of the counts belonging to mitochondrial genes were both discarded (Liu *et al.*, 2022). The 'NormalizeData function' in Seurat (v3.1.1) was used to normalise the gene expression levels for each cell (Butler *et al.*, 2018). The normalised expression values of genes were log-transformed for the following analysis.

According to the method described in previous study (Macosko *et al.*, 2015), we identified the top variable genes across single cells using the FindVariableGenes function (mean.function = ExpMean, dispersion.function = LogVMR) in Seurat (v3.1.1) (Butler *et al.*, 2018). We applied the principal component analysis to reduce the dimensionality of the top variable genes. Cells were clustered using the FindClusters function according to their gene expression levels and visualised using the t-SNE algorithm and the UMAP analysis (Becht *et al.*, 2019; Butler *et al.*, 2018).

Cell type annotation was performed according to the cell type-specific marker genes. The FindAllMarkers function (test.use = bimod, logfc.thresold = 0, min.pct = 0.25) in Seurat (v3.1.1) was used to identify the DEGs between clusters (Butler *et al.*, 2018). In each cell cluster, top 10 DEGs were considered as the zone-specific marker genes. Among the top 10 marker genes in each cluster, those (or their homologous gene) zones identified in previous studies were selected for cell-type annotation. Pseudo-time trajectory analysis was performed using the Monocle 2 (Qiu *et al.*, 2017). RNA velocity analysis was conducted using the velocyto algorithm (La Manno *et al.*, 2018). GO enrichment analysis of DEGs was implemented in GENE DENOVO platform (<https://www.omicshare.com>).

Core TFs identification

For each cell type, the core TFs were identified according to 'Binding Site Prediction' function of the PlantTFDB database with the threshold P -value of 1×10^{-5} (<http://planttfdb.gao-lab.org/>). Cytoscape software was used to plot the regulation networks for the TFs and their target genes (Shannon *et al.*, 2003). Here, we

focused on a core transcription factor *ZmERE14* in MMC type for functional validation. The neighbour-joining phylogenetic tree was constructed using the MEGA11 software (<https://www.megasoftware.net/>) with 1000 replicates for bootstrap analysis.

Expression analysis of *ZmERE14*

Total RNA was separately extracted from leaves (ear leaf), roots, stems, tassels, ears, kernels and shoot tips using TRIzol reagent (Invitrogen, California, USA). Herein, the leaves, roots, stems, tassels and ears were collected at flowering period. The kernels and the shoot tips were collected on 7 DAP and 5 days after sowing, respectively. The PrimeScriptTMRT reagent kit (TaKaRa, Japan) was used to synthesise cDNA. An ABI 7500 real-time PCR System (Torrance, CA, USA) was used to examine the relative expression level of *ZmERE14* in each tissue with three biological replicates. *ZmActin 1* (*Zm00001d010159*) was used as the reference gene and the relative expression level of *ZmERE14* was calculated by the $2^{-\Delta\Delta CT}$ method.

Subcellular localisation of *ZmERE14*

The subcellular localisation of *ZmERE14* protein was conducted as follows: First, we amplified the coding sequence (CDS) of *ZmERE14* without the stop codon from the maize line B73. Next, the cloned CDS was inserted into the pCAMBIA2300-35S-eGFP vector to generate the recombinant plasmid p35S:*ZmERE14*-eGFP. Then, the fusion construct was introduced into tobacco leaves using *Agrobacterium* injection. Finally, we examined the eGFP fluorescence at 48 h of *Agrobacterium* injection under a confocal microscope (ZEISS LSM800, Germany).

Cell morphology examination

For the wild-type B73, *zmereb14-1* and *zmereb14-2*, shoot apices (for SAM morphology examination), shoot apices (for AM morphology examination) and internodes (the middle section of the first internode above/below the ear) were sampled from five-day-old seedlings, 15-day-old seedlings and plants 15 day after pollination, respectively. These samples were fixed in FAA solution (formalin:acetic acid:70% ethanol [1:1:18, v/v]) at 4°C for 24 h, dehydrated through an ethanol series, waxed for 3 h, embedded in paraffin at 65°C (JB-P5, JB-L5, Wuhan Junjie Electronics Co., LTD, Wuhan, China) and then sliced into 8 µm-thick sections (RM2016, Shanghai Leica Instrument Co., LTD, Shanghai, China). Slices were stained by toluidine blue and then observed under an optical microscope (Nikon Eclipse E100, Japan).

In situ hybridisation

In situ hybridisation was performed as described by Satterlee *et al.* (2020), with some modifications. Briefly, RNA was extracted using TRIzol reagent (Invitrogen, California, USA) from the shoot apex of five-day-old seedlings. cDNA was prepared using PrimeScriptTMRT reagent kit (TaKaRa Bio Inc., Shiga, Japan) and was used as a template to amplify the coding sequences of target gene. The amplification products were then inserted into pGEM-T vector and linearised with *Hind*III and *Eco*RI. Antisense and sense RNA probes were synthesised using SP6 and T7 RNA polymerases. Slices were deparaffinised with transparent dewaxing liquid (Servicebio Technology Co., LTD, Wuhan, China), rehydrated with an ethanol series and digested with Proteinase K (Servicebio Technology Co., LTD, Wuhan, China) at 37°C for 15 min. Dehydrated sections were then hybridised with probe at 40°C overnight, incubated with anti-DIG-AP (Jackson ImmunoResearch Laboratories, Inc., PA, USA) at 37°C for 40 min and treated with

BCIP/NBT chromogenic agent (Boster Biological Technology Co., LTD, Wuhan, China). Images were collected using an optical microscope (Nikon Eclipse E100, Japan).

RNA-seq analysis of maize mutant *zmereb14-1* and B73 lines

For RNA-seq analysis, shoot apices of five-day-old seedlings were respectively collected from wild-type B73 and *zmereb14-1* mutant. RNA isolation, library construction and sequencing were carried out as previously described (Zhang *et al.*, 2021c), with three biological replicates. Clean reads were then mapped to the maize reference genome (B73 RefGen_V4). Transcripts per million (TPM) was calculated to normalise the gene expression values. Genes with $|log_2$ fold-change (FC) ≥ 1 and false discovery rate (FDR) < 0.05 were determined as DEGs. GO enrichment analysis of DEGs was conducted on the GENE DENOV0 platform (<https://www.omicshare.com>).

IAA concentration measurement

Approximately 3 mm segments above the nodes of the shoot apices were collected from five-day-old seedlings of B73 and two mutants (*zmereb14-1* and *zmereb14-2*), respectively. Then, the samples were used for measuring IAA concentration using an IAA ELISA kit (Shanghai Renjie Biotechnology Company, Shanghai, China, <http://www.elisakit.cc>). The detailed experimental procedures were described in the instruction manual. For each sample, three biological replicates were included.

Acknowledgements

This work was supported by the National Key Research and Development Program of China (2021YFF1000303) and National Nature Science Foundation of China (32101777 and 32072073). The authors thank OE Biotech Co., Ltd (Shanghai, China) for providing single-cell RNA-seq, and Lanying Wang and Hongyan Liu for assistance with bioinformatics analysis.

Conflict of interest

The authors declare no competing interests.

Author contributions

Y.S. and L.M. conceived and supervised the project; L.M., N.Z. and P.L. sequenced and processed the raw data; N.Z., Y.L., R.L., C.Z. and Z.C. conducted most of the experiments. L.M. and Y.S. wrote the manuscript with the contributions from G.Y., T.L. and G.P. All authors have read and approved the manuscript.

Data availability statement

All single-cell transcriptome data and bulk RNA sequencing data generated for this study have been deposited Genome Sequence Archive (GSA) in National Genomics Data Center (NGDC) database with the accession numbers CRA012259 and CRA014263.

References

- Aida, M. and Tasaka, M. (2006) Genetic control of shoot organ boundaries. *Curr. Opin. Plant Biol.* **9**, 72–77.
- Becht, E., McInnes, L., Healy, J., Dutertre, C.A., Kwok, I.W.H., Ng, L.G., Ginhoux, F. *et al.* (2019) Dimensionality reduction for visualizing single-cell data using UMAP. *Nat. Biotechnol.* **37**, 38–44.
- Bezrutczyk, M., Zöllner, N.R., Kruse, C.P.S., Hartwig, T., Lautwein, T., Köhrer, K., Frommer, W.B. *et al.* (2021) Evidence for phloem loading via the abaxial bundle sheath cells in maize leaves. *Plant Cell* **33**, 531–547.
- Butler, A., Hoffman, P., Smibert, P., Papalexi, E. and Satija, R. (2018) Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**, 411–420.
- Cao, E.Y., Ouyang, J.F. and Rackham, O.J.L. (2020) GeneSwitches: ordering gene expression and functional events in single-cell experiments. *Bioinformatics* **36**, 3273–3275.
- Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P. *et al.* (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21.
- Elhaddad, N.S., Hunt, L., Sloan, J. and Gray, J.E. (2014) Light-induced stomatal opening is affected by the guard cell protein kinase APK1b. *PLoS One* **9**, e97161.
- Guo, M., Rupe, M.A., Dieter, J.A., Zou, J., Spielbauer, D., Duncan, K.E., Howard, R.J. *et al.* (2010) Cell Number Regulator1 affects plant and organ size in maize: implications for crop yield enhancement and heterosis. *Plant Cell* **22**, 1057–1073.
- Gutierrez, C., Desvoyes, B., Vergara, Z., Otero, S. and Sequeira-Mendes, J. (2016) Links of genome replication, transcriptional silencing and chromatin dynamics. *Curr. Opin. Plant Biol.* **34**, 92–99.
- Hammoudi, V., Beeren, B., Jonker, M.J., Helderman, T.A., Vlachakis, G., Giesbers, M., Kwaaitaal, M. *et al.* (2021) The protein modifier SUMO is critical for *Arabidopsis* shoot meristem maintenance at warmer ambient temperatures. *bioRxiv*. <https://doi.org/10.1101/2021.02.04.429700>
- Hu, C., Zhu, Y., Cui, Y., Cheng, K., Liang, W., Wei, Z., Zhu, M. *et al.* (2018) A group of receptor kinases are essential for CLAVATA signalling to maintain stem cell homeostasis. *Nat. Plants* **4**, 205–211.
- Jirásková, V. (2022) *The role of ribosomal proteins in the plant development*. Bachelor thesis, Charles University <http://hdl.handle.net/20.500/11956/173111>
- Juarez, M.T., Twigg, R.W. and Timmermans, M.C.P. (2004) Specification of adaxial cell fate during maize leaf development. *Development* **131**, 4533–4544.
- Kerstetter, R.A. and Hake, S. (1997) Shoot meristem formation in vegetative development. *Plant Cell* **9**, 1001–1010.
- Khadem, A., Moshtaghi, N. and Bagheri, A. (2023) Regulatory networks of hormone-involved transcription factors and their downstream pathways during somatic embryogenesis of *Arabidopsis thaliana*. *3 Biotech* **13**, 1–11.
- Kim, J.Y., Symeonidi, E., Pang, T.Y., Denyer, T., Weidauer, D., Bezrutczyk, M., Miras, M. *et al.* (2021) Distinct identities of leaf phloem cells revealed by single cell transcriptomics. *Plant Cell* **33**, 511–530.
- Knauer, S., Javelle, M., Li, L., Li, X., Ma, X., Wimalanathan, K., Kumari, S. *et al.* (2019) A high-resolution gene expression atlas links dedicated meristem genes to key architectural traits. *Genome Res.* **29**, 1962–1973.
- La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., Lidschreiber, K. *et al.* (2018) RNA velocity of single cells. *Nature* **560**, 494–498.
- Laufs, P., Grandjean, O., Jonak, C., Kiêu, K. and Traas, J. (1998) Cellular parameters of the shoot apical meristem in *Arabidopsis*. *Plant Cell* **10**, 1375–1389.
- Li, X., Zhang, X., Gao, S., Cui, F., Chen, W., Fan, L. and Qi, Y. (2022) Single-cell RNA sequencing reveals the landscape of maize root tips and assists in identification of cell type-specific nitrate-response genes. *Crop J.* **10**, 1589–1600.
- Lin, Z., Yin, K., Zhu, D., Chen, Z., Gu, H. and Qu, L.J. (2007) AtCDC5 regulates the G2 to M transition of the cell cycle and is critical for the function of *Arabidopsis* shoot apical meristem. *Cell Res.* **17**, 815–828.
- Liu, L., Gallagher, J., Arevalo, E.D., Chen, R., Skopelitis, T., Wu, Q., Bartlett, M. *et al.* (2021) Enhancing grain-yield-related traits by CRISPR–Cas9 promoter editing of maize CLE genes. *Nat. Plants* **7**, 287–294.
- Liu, Z., Wang, J., Zhou, Y., Zhang, Y., Qin, A., Yu, X., Zhao, Z. *et al.* (2022) Identification of novel regulators required for early development of vein pattern in the cotyledons by single-cell RNA-sequencing. *Plant J.* **110**, 7–22.

- Luo, J., Zhao, H., Chen, L. and Liu, M. (2023) Multifaceted functions of RPS27a: An unconventional ribosomal protein. *J. Cell. Physiol.* **238**, 485–497.
- Macosko, E.Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I. et al. (2015) Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**, 1202–1214.
- Morcillo, F., Gallard, A., Pillot, M., Jouannic, S., Aberlenc-Bertossi, F., Collin, M., Verdeil, J.L. et al. (2007) EgAP2-1, an ANTIGUMENTA-like (AIL) gene expressed in meristematic and proliferating tissues of embryos in oil palm. *Planta* **226**, 1353–1362.
- Osorio, D. and Cai, J.J. (2021) Systematic determination of the mitochondrial proportion in human and mice tissues for single-cell RNA-sequencing data quality control. *Bioinformatics* **37**, 963–967.
- Qiu, X., Hill, A., Packer, J., Lin, D., Ma, Y.A. and Trapnell, C. (2017) Single-cell mRNA quantification and differential analysis with Census. *Nat. Methods* **14**, 309–315.
- Riechmann, J.L. and Meyerowitz, E.M. (1998) The AP2/EREBP family of plant transcription factors. *Biol. Chem.* **379**, 633–646.
- Ryu, K.H., Huang, L., Kang, H.M. and Schiefelbein, J. (2019) Single-cell RNA sequencing resolves molecular relationships among individual plant cells. *Plant Physiol.* **179**, 1444–1456.
- Satina, S., Blakeslee, A.F. and Avery, A.G. (1940) Demonstration of the three germ layers in the shoot apex of datura by means of induced polyploidy in periclinal chimeras. *Am. J. Bot.* **27**, 895–905.
- Satterlee, J.W., Strable, J. and Scanlon, M.J. (2020) Plant stem-cell organization and differentiation at single-cell resolution. *Proc. Natl. Acad. Sci. USA* **117**, 33689–33699.
- Shani, E., Yanai, O. and Ori, N. (2006) The role of hormones in shoot apical meristem function. *Curr. Opin. Plant Biol.* **9**, 484–489.
- Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N. et al. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504.
- Shen, X., Xiao, B., Kaderbek, T., Lin, Z., Tan, K., Wu, Q., Yuan, L. et al. (2023) Dynamic transcriptome landscape of developing maize ear. *Plant J.* **116**, 1856–1870.
- Stelpflug, S.C., Sekhon, R.S., Vaillancourt, B., Hirsch, C.N., Buell, C.R., de Leon, N. and Kaeplinger, S.M. (2016) An expanded maize gene expression atlas based on RNA sequencing and its use to explore root development. *Plant Genome* **9**, 1–16.
- Sun, Y., Dong, L., Zhang, Y., Lin, D., Xu, W., Ke, C., Han, L. et al. (2020) 3D genome architecture coordinates trans and cis regulation of differentially expressed ear and tassel genes in maize. *Genome Biol.* **21**, 1–25.
- Takada, S., Hibara, K., Ishida, T. and Tasaka, M. (2001) The CUP-SHAPED COTYLEDON1 gene of *Arabidopsis* regulates shoot apical meristem formation. *Development* **128**, 1127–1135.
- Tian, C., Wang, Y., Yu, H., He, J., Wang, J., Shi, B., Du, Q. et al. (2019) A gene expression map of shoot domains reveals regulatory mechanisms. *Nat. Commun.* **10**, 141.
- Vernoud, V., Laigle, G., Rozier, F., Meeley, R.B., Perez, P. and Rogowsky, P.M. (2009) The HD-ZIP IV transcription factor OCL4 is necessary for trichome patterning and anther development in maize. *Plant J.* **59**, 883–894.
- Vinayak, V., Khan, M.J., Jha, A.N. and Harish (2021) Photosystem I P700 chlorophyll a apoprotein A1 as PCR marker to identify diatoms and their associated lineage. *J. Eukaryot. Microbiol.* **68**, e12866.
- Vollbrecht, E., Reiser, L. and Hake, S. (2000) Shoot meristem size is dependent on inbred background and presence of the maize homeobox gene, knotted1. *Development* **127**, 3161–3172.
- Weingartner, M., Pelayo, H.R., Binarova, P., Zwerger, K., Melikant, B., de la Torre, C., Heberle-Bors, E. et al. (2003) A plant cyclin B2 is degraded early in mitosis and its ectopic expression shortens G2-phase and alleviates the DNA-damage checkpoint. *J. Cell Sci.* **116**, 487–498.
- Yadav, R.K., Girke, T., Pasala, S. and Reddy, G.V. (2009) Gene expression map of the *Arabidopsis* shoot apical meristem stem cell niche. *Proc. Natl. Acad. Sci. USA* **106**, 4941–4946.
- Yang, H.J., Shen, H., Chen, L., Xing, Y.Y., Wang, Z.Y., Zhang, J.L. and Hong, M.M. (2002) The OsEBP-89 gene of rice encodes a putative EREBP transcription factor and is temporally expressed in developing endosperm and intercalary meristem. *Plant Mol. Biol.* **50**, 379–391.
- Yang, Y., Anderson, C.T. and Cao, J. (2021) Polygalacturonase45 cleaves pectin and links cell proliferation and morphogenesis to leaf curvature in *Arabidopsis thaliana*. *Plant J.* **106**, 1493–1508.
- Yi, F., Gu, W., Chen, J., Song, N., Gao, X., Zhang, X., Zhou, Y. et al. (2019) High temporal-resolution transcriptome landscape of early maize seed development. *Plant Cell* **31**, 974–992.
- Zeng, J., Li, X., Ge, Q., Dong, Z., Luo, L., Tian, Z. and Zhao, Z. (2021) Endogenous stress-related signal directs shoot stem cell fate in *Arabidopsis thaliana*. *Nat. Plants* **7**, 1276–1287.
- Zhang, D., Sun, W., Singh, R., Zheng, Y., Cao, Z., Li, M., Lunde, C. et al. (2018) GRF-interacting factor1 regulates shoot architecture and meristem determinacy in maize. *Plant Cell* **30**, 360–374.
- Zhang, T.Q., Chen, Y. and Wang, J.W. (2021a) A single-cell analysis of the *Arabidopsis* vegetative shoot apex. *Dev. Cell* **56**, 1056–1074.
- Zhang, T.Q., Chen, Y., Liu, Y., Lin, W.H. and Wang, J.W. (2021b) Single-cell transcriptome atlas and chromatin accessibility landscape reveal differentiation trajectories in the rice root. *Nat. Commun.* **12**, 2053.
- Zhang, X., Liu, P., Qing, C., Yang, C., Shen, Y. and Ma, L. (2021c) Comparative transcriptome analyses of maize seedling root responses to salt stress. *PeerJ* **9**, e10765.
- Zhu, Y., Weng, M., Yang, Y., Zhang, C., Li, Z., Shen, W.H. and Dong, A. (2011) *Arabidopsis* homologues of the histone chaperone ASF1 are crucial for chromatin replication and cell proliferation in plant development. *Plant J.* **66**, 443–455.
- Zong, Y., Hao, Z., Tu, Z., Shen, Y., Zhang, C., Wen, S., Yang, L. et al. (2021) Genome-wide survey and identification of AP2/ERF genes involved in shoot and leaf development in *Liriodendron chinense*. *BMC Genomics* **22**, 1–17.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Figure S1 Quality control analysis of the single-cell RNA sequencing raw data.

Figure S2 Single-cell RNA sequencing data of maize shoot apex before and after quality control.

Figure S3 Expression patterns of representative genes.

Figure S4 Expression patterns of all markers used for cell type annotation.

Figure S5 Correlations of gene expressions between the results of the present study and previous study (Satterlee et al., 2020) on maize shoot apex.

Figure S6 Gene expression correlations between the results of the present study and previous study (Knauer et al., 2019).

Figure S7 Expression patterns of the representative new marker genes in maize shoot apex reported by the previous study (Knauer et al., 2019).

Figure S8 Expression patterns of top 10 marker genes of S-phase and G2/M-phase cells (Satterlee et al., 2020) in each of our identified clusters.

Figure S9 Gene Ontology (GO) enrichment of the differentially expressed genes (DEGs) in each cell type.

Figure S10 Expression patterns of related genes in distinct clusters.

Figure S11 Distribution of cells in each cell type on the pseudo-time trajectory.

Figure S12 GeneSwitches analyses of scRNA-seq data.

Figure S13 Expression, clustering, and GO enrichment of the branch point 1-dependent genes over pseudo-time.

Figure S14 Expression, clustering, and GO enrichment of the branch point 2-dependent genes over pseudo-time.

Figure S15 Phylogenetic analysis of ZmEREB14.

Figure S16 In situ hybridisation of *ZmEREB14* in shoot apex and developing maize ear.

Figure S17 Axillary meristem size of B73 and two mutants.

Figure S18 Transcriptomic analysis of shoot apices from B73 line and *zmereb14-1* mutant.

Figure S19 Expression levels of auxin metabolism-related genes in B73 and *zmereb14-1* mutant.

Figure S20 IAA concentrations in shoot apices of B73 and two mutants.

Figure S21 Expression patterns of *ZmEREB14* in various maize tissues.

Figure S22 Comparison of tassel branch number and kernel row number between B73 and two mutants.

Figure S23 Expression levels of tassel branch-regulated and kernel row number-regulated genes in B73 and *zmereb14-1* mutant.

Figure S24 Images of shoot apex sampling and protoplast isolation.

Table S1 Single-cell transcriptome data before and after quality control.

Table S2 Marker genes used for cell type annotation.

Table S3 New marker genes for distinct cell types.

Table S4 Differentially expressed genes in each cell type.

Table S5 Representative GO enrichment categories for target genes of *ZmMYB4*.

Table S6 Representative GO enrichment categories for target genes of *ZmHSF8*.

Table S7 Primers used in this study.

Table S8 Summary of transcriptome data from B73 and *zmereb14-1* mutant.