

Construction of a single-cell transcriptome atlas for *Pogostemon cablin* embryoids reveals PcNAC048 as a dual regulator coordinating lateral root morphogenesis and patchouli alcohol biosynthesis

Received: 28 April 2025

Accepted: 15 December 2025

Cite this article as: Wang, X., Ge, X., Zhong, W. *et al.* Construction of a single-cell transcriptome atlas for *Pogostemon cablin* embryoids reveals PcNAC048 as a dual regulator coordinating lateral root morphogenesis and patchouli alcohol biosynthesis. *Commun Biol* (2025). <https://doi.org/10.1038/s42003-025-09434-5>

Xiaobing Wang, Xiaojin Ge, Wenxuan Zhong, Daidi Wu, Zhiye Yang, Ruoting Zhan, Likai Chen & Hua Li

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

Construction of a single-cell transcriptome atlas for *Pogostemon cablin* embryoids reveals PcNAC048 as a dual regulator coordinating lateral root morphogenesis and patchouli alcohol biosynthesis

Xiaobing Wang^{1,2}, Xiaojin Ge¹, Wenxuan Zhong¹, Daidi Wu¹, Zhiye Yang^{2,3}, Ruoting Zhan¹, Likai Chen^{1*}, Hua Li^{2,3,4*}

¹ Research Center of Chinese Herbal Resource Science and Engineering, Guangzhou University of Chinese Medicine; Key Laboratory of Chinese Medicinal Resource from Lingnan (Guangzhou University of Chinese Medicine), Ministry of Education; Joint Laboratory of National Engineering Research Center for the Pharmaceuticals of Traditional Chinese Medicines, Guangzhou, 510006, China

² Guangdong Institute for Drug Control, Guangzhou, 510663, China

³ NMPA Key Laboratory of Rapid Drug Inspection Technology, Guangzhou, 510663, China

⁴ Key Laboratory of Traditional Chinese Medicine Standards Research in the Guangdong Hong Kong Macao Greater Bay Area of Guangdong Medical Products Administration, Guangzhou, 510663, China

* Correspondence author:

Likai Chen: chenlk@gzucm.edu.cn

Hua Li: gdidc_lihua@gd.gov.cn

Abstract

Pogostemon cablin Benth. (*P. cablin*) is an annual aromatic medicinal plant. In this study, anthers were cultured in vitro during the microspore development stage of *P. cablin*, inducing somatic embryogenesis. Globular somatic embryos (GSE), heart-shaped somatic embryos, torpedo-shaped somatic embryos, and cotyledonary somatic embryos (CSE) were observed and isolated. Single-cell RNA sequencing was then employed to generate single-cell maps for GSE and CSE. Using reported marker genes, a total of eight cell types were identified. Pseudo-temporal analysis reconstructed the continuous differentiation trajectory of apical meristem cells and epidermal cells. Further investigation identified PcNAC048 as a putative transcription factor that regulates embryonic cell differentiation. Gene expression analysis showed *PcNAC048* is expressed in different tissues of *P. cablin* and responds to abiotic stress. Transgenic studies in *Arabidopsis*

thaliana showed that *PcNAC048* can promote lateral root development. Results from yeast one-hybrid and dual luciferase assays showed that *PcNAC048* can interact with the promoter of the patchouli alcohol synthase gene (*PcPTS*) and inhibit its activity. Transient overexpression and virus-induced gene silencing (VIGS) analysis further confirmed that *PcNAC048* can negatively regulate the biosynthesis of patchouli alcohol. Overall, this study provides theoretical support for germplasm development and regulation of medicinal compounds in *P. cablin*.

Introduction

Plants develop organs, including flowers, roots, and leaves. These organs are associated with a complete embryonic structure essential for producing offspring. Anthers, constituting a part of the male reproductive organs in flowers, serve as one of the explants utilized in plant tissue culture¹. It is now understood that in the presence of growth regulators such as hormones, anthers can differentiate into embryogenic calli, potentially leading to the formation of somatic embryos. In the subsequent stages of embryonic development, these tissues undergo further differentiation to develop into unique structures, such as the phloem², xylem³, and epidermal layer, until they become complete plants.

Somatic embryogenesis has been successfully established in various plant species, including *Centaurium erythraea* Rafn⁴, *longan*⁵, and *Lilium pumilum* DC. Fisch.⁶, serving as a model system for studying embryonic development and plant regeneration *in vitro*. Somatic embryos have strong division ability and often have multiple different cell types. Identifying the cell types of somatic embryos is beneficial for more accurate research on the functions of different types of cells during development. The differences in cell types are usually determined by gene expression. In recent years, high-throughput single-cell transcriptomics has emerged as a fundamental and widely utilized method in plant research. It can isolate single cells from somatic embryonic tissue for high-throughput sequencing, detect gene expression, genetic variation, and other information of individual cells, thereby revealing heterogeneity in cell function, state, and developmental stages, and can reconstruct the developmental trajectory of cells, elucidating how cells gradually differentiate from undifferentiated states into mature cells with specific functions. ScRNA-seq is of great significance in studying plant cell pluripotency, molecular mechanisms of plant

regeneration, plant genetics and breeding, and germplasm resource preservation. Single-cell transcriptome profiles have been reported for species such as *Prunus mume*⁷, tea leaves⁸, cotton⁹, and *Hylocereus undatus*¹⁰. In addition, the spatio-temporal developmental trajectory of early somatic embryos in the renowned traditional Chinese medicine, longan, has been elucidated through single-cell RNA sequencing (scRNA-seq)⁵.

The complex transcriptional regulatory network plays an important role in the development of plant cells or tissues, as well as the synthesis of secondary metabolites. For example, ethylene responsive factor 6 (ERF6), selected from single cell sequencing of early somatic embryos in *Dimocarpus longan* Lour., is a thermosensitive factor that can negatively regulate longan somatic embryogenesis under high temperature stress conditions⁵. The novel gene *NAC TRANSCRIPTON FACTOR-LIKE 9* (*GbiNTL9*) discovered in single-cell sequencing of cotton cotyledons can affect the formation of cotton pigment glands¹¹. Similarly, scRNA-seq of maize leaves revealed that TFs such as NAC and WRKY are the main participants in the early development of leaf mesophyll cells¹², and another study showed that the ZmNAC1 can promote lateral root development in maize (*Zea mays* L.)¹³. In addition, studies have shown that NAC TFs are abundantly present in plants and play crucial roles in the response to various abiotic stresses¹⁴, such as cold, heat, salt, and drought.

Pogostemon cablin (*P. cablin*) is an important essential oil plant native to the Philippines and other Southeast Asian countries. It is currently widely cultivated in Guangdong Province and other regions of China for medicinal and essential oil extraction¹⁵. Anther is an important component of *P. cablin* reproductive organs, containing various cell types with different functions and developmental stages, such as pollen mother cells, microspores, etc. These cells are involved in a complex gene expression regulatory network during development, and scRNA-seq analysis can comprehensively reveal the dynamic expression patterns of these genes, which helps us to gain a deeper understanding of the reproductive and developmental mechanisms of *P. cablin*. In addition, although the flowering of *P. cablin* cultivated in China is rare¹⁶ and somatic embryos have not been successfully established, cells in anthers have high differentiation potential under specific conditions, providing a potential model for studying gene regulation related to somatic embryogenesis.

P. cablin is rich in patchouli alcohol, which is the core component of *P. cablin* essential oil. It exhibits antiviral, antibacterial, and antitumor activities¹⁷⁻¹⁹, which makes it of great scientific

value and application prospect to further study its biosynthesis mechanism. Therefore, identifying the key genes regulating the biosynthesis of patchouli alcohol has become the central goal of current research. Patchouli alcohol is mainly produced in the leaves, which serve as the main organs for photosynthesis and gas exchange in plants. The epidermal cells of the leaves play a key role in the synthesis and accumulation of secondary metabolites. Therefore, when studying the biosynthesis of patchouli alcohol, the epidermal cells in the leaves should be the focus. In addition, the lateral roots of *P. cablin* are well-developed and play crucial roles in stabilizing the plant body, absorbing water, and nutrients^{20,21}, and improving soil quality. However, the regulatory genes behind the development of lateral roots are still unknown. To address these two issues, single-cell transcriptome data of *P. cablin* embryos can be analyzed to capture transcriptional information related to leaves and roots.

This study presents high-resolution scRNA-seq expression profiles of *P. cablin* globular somatic embryos (GSE) and cotyledonary somatic embryos (CSE), capturing detailed spatiotemporal dynamics, identifying major cell types, and revealing key regulatory factors. We explored the complex developmental trajectories associated with the differentiation of apical meristem cells into epidermal cells. Furthermore, we compiled a comprehensive list of genes involved in the patchouli alcohol biosynthesis pathway and assessed the highly interconnected transcription factor PcNAC048, investigating its role in lateral root development and its functions in patchouli alcohol biosynthesis. Our investigation not only provides a unique spatio-temporal view of somatic embryo differentiation in *P. cablin* with unprecedented resolution but also establishes PcNAC048 as a dual regulatory factor that coordinates lateral root morphogenesis and patchouli alcohol biosynthesis.

Results

Identification of microspore development stage and somatic embryo induction in *P. cablin*

The flowering period of *P. cablin* typically occurs from mid to late March annually, though only a minimal fraction of the plants produces flower buds. In the early phases of this study, we successfully observed and documented this event. As illustrated in Fig. 1A, the inflorescence of this species consists of 10 to numerous individual flowers, arranged in a spike-like formation that is densely adorned with fine hairs. The corolla, when in bloom, displays a distinct purple coloration.

To determine the optimal sampling period for *in vitro* culture of patchouli anthers, green flower buds were selected for dissection. The results showed that the anthers were full and tender

at the early stage of flower bud growth. After dissociation with 1 mol/L HCl and staining with acetic acid magenta staining solution, mononuclear microspores were observed under microscopy (Fig. 1B).

During the development of *P. cablin* callus tissue, four stages of embryoid differentiation were observed (Fig. 1C): globular somatic embryos (GSE), heart-shaped somatic embryos, torpedo-shaped somatic embryos, and cotyledonary somatic embryos (CSE). Given that GSE and CSE exhibited the highest abundance among these stages, they were selected for single-cell RNA transcriptome sequencing in subsequent research. This approach aimed to elucidate the histological differences between the two embryonic bodies at different developmental stages, as well as to conduct functional analyses of embryonic body development and secondary metabolite biosynthesis involving differentially expressed genes.

Eight cell cluster were identified by scRNA-seq in *P. cablin* GSE and CSE

Protoplasts of GSE and CSE in *P. cablin* were identified by enzymatic hydrolysis and visualized using trypan blue staining. The results were as follows: GSE protoplast concentration was 635 cells/ μ L, with a total cell count of 63,000 and a viability of 86%. The rate of large fragmentation was 25%. Based on concentration and relevant experience, 15,000 cells were loaded onto the automated cell counter. CSE protoplast concentration was 1,000 cells/ μ L, with a total cell count of 100,000 and a viability of 82%. The rate of large fragmentation was 23%. Based on concentration and relevant experience, 15,000 cells were loaded onto the automated cell counter. Subsequently, single-cell transcriptome sequencing was performed according to the procedure shown in Fig. 2A.

Utilizing the 10x Genomics Cell Ranger and Seurat software, 14,202 high-quality cells were identified in the GSE group, while 14,444 were identified in the CSE group. In addition, the cell data in each sample after filtration showed that the unique molecular identifier (UMI) of GSE was 34698383, while the UMI of CSE was 32812489, with an average UMI of 2443.2 and 2271.7 per cell, respectively. In the GSE and CSE samples, the median number of genes per cell after filtration was 1298 and 1164, respectively, and the median number of UMI per cell after filtration was 1589 and 1488, respectively. As detailed in Supplementary Fig. S1 and Table S1, the proportion of high-quality cells was relatively high.

To identify clusters of distinct cell types within *P. cablin* GSE and CSE from the single-cell transcriptome data, unsupervised clustering analysis was performed using Seurat software based

on the expression patterns of differentially expressed genes. Twelve distinct clusters were identified, ranging from cluster 0 to cluster 11. In GSE, each cluster comprised 19 to 6716 cells, with cluster 10 containing the fewest cells and cluster 0 containing the most. In CSE, each cluster comprised 25 to 6676 cells, with cluster 11 containing the fewest cells and cluster 0 containing the most (Supplementary Table S1).

T-distributed Stochastic Neighbor Embedding (t-SNE) nonlinear dimensionality reduction was used to visualize the results of single-cell cluster classification (Fig. 2B). In the t-SNE classification plot, single cells were divided into twelve distinct clusters, numbered 0 to 11. These clusters were color-coded for visual differentiation, with each point representing a single cell.

To map the 12 cell clusters to cell subtypes, it was necessary to identify cell clusters with the expression of marker genes within various cellular clusters. The bubble plot was used to visualize the expression levels of known marker genes within each cluster (Fig. 2C), while Table S2 was used to display the expression levels and distribution proportions of marker genes in different clusters, both of which contribute to the identification of cell types.

In this research, these cell clusters were characterized using genes known to be specifically expressed. Among them, Enriched expression of photosynthesis-related genes Pat_A17G021800 (*LHCB4.1*) and Pat_B06G080900 (*RBCS1*) identified clusters 0 and 4 as mesophyll cells (MC). High expression of genes Pat_B02G038000 (*BDGI*), Pat_A05G156600 (*PDF1*), and Pat_A18G041800 (*LTPI*) served as markers classifying cluster 1 as epidermal cells (EC). Enriched gene expression of Pat_A16G084900 (*SPCH*) and Pat_A30G046000 (*FAMA*) categorized cluster 9 as guard cells (GC). Enriched gene expression of Pat_A32G007400 (*HIS2A*), Pat_A23G029900 (*HIS2B*), and Pat_A01G224300 (*HTR2*) characterized cluster 3 as apical meristem tissue cells (AMC). Enriched gene expression of Pat_A21G069200 (At3g51280), Pat_A22G022400 (*TUBA*), and Pat_A22G058500 (At1g18250) identified cluster 5 as meristematic tissue cells (MeC). Enriched gene expression of Pat_A17G044500 (At5g03610), Pat_B29G077600 (*STP13*), Pat_A01G075800 (*CSA*), Pat_B18G005200 (*NHL10*), and Pat_B28G085500 (*PUMP5*) categorized clusters 2 and 8 as root endothelial cells (RC). Enriched gene expression of Pat_A28G048000 (*RTM2*) and Pat_B29G018300 (*SEOB*) characterized cluster 10 as phloem cells (PC). Enriched gene expression of Pat_A14G006600 (*IRX9*), Pat_A03G058000 (*RAC9*), and Pat_A10G149200 (*XCPI*) categorized cluster 11 as xylem cells (XC). In addition, clusters 6 and 7 were identified as unknown cell clusters Uk_1 and Uk_2, respectively, as no corresponding

marker genes were found for these clusters. To further verify the accuracy of cluster identity, T-SNE maps were used to show the expression of representative cluster specific marker genes in GSE and CSE, respectively (Supplementary Fig. S2). The results of cell cluster identification were mapped onto the classification t-SNE plot shown in Fig. 2D. Twelve cell clusters were identified, comprising eight known cell types and two unknown cell types.

Potential differentiation existed between apical meristematic cells and epidermal cells

Since scRNA-seq captured cells in transitional states, it was feasible to explore the successive differentiation trajectories of cells during development. The cellular differentiation trajectories of GSE and CSE in *P. cablin*, and the differentially expressed genes along these differentiation trajectories, were analyzed using Monocle 2 software. Patchouli alcohol is a key bioactive compound in *P. cablin*, mainly synthesized in leaves, especially in the glandular scales formed by the differentiation of epidermal cells²². Therefore, genes upregulated during epidermal cell differentiation are reasonable candidates for actively regulating the biosynthesis of patchouli alcohol. In the pseudotime analysis, cluster 3 (apical meristematic cells) and 1 (epidermal cells) were selected to construct a continuous differentiation trajectory of cells within *P. cablin* embryoids. Fig. 3A(1) showed that cells of cluster 3 were positioned at the beginning of pseudo time, while the cells of cluster 1 were located at the endpoint of the pseudo time axis, separated from cells of cluster 3. Fig. 3A(2) presents information on cell cluster, including apical meristematic tissue cells and epidermal cells. Fig. 3A(3) depicts the differentiation state information graph, revealing three bifurcation points in the pseudo time course, thereby dividing all cells into seven states. Fig. 3A (4) represents the sample information graph, illustrating two distinct samples of GSE and CSE. The above results indicate that over time, apical meristematic cells can differentiate into epidermal cells in the late stage of development.

Epidermal cells play an important role in plant secondary metabolism and response to environmental stress. As the first barrier between plants and the external environment, epidermal cells are directly exposed to various abiotic stress factors, and their gene expression changes can sensitively reflect the plant's response to stress, thereby affecting the synthesis of secondary metabolites. In pseudotime analysis, there is a differentiation relationship between apical meristematic cells and epidermal cells. Therefore, it is necessary to investigate whether the key genes involved in the differentiation of apical meristematic cells into epidermal cells play a regulatory role in the response to abiotic stress and secondary metabolism processes of *P. cablin*.

Differential gene analysis showed that the total number of pseudo timeline differential genes, differentiation status differential genes, and differentiation fate differential genes are 1928, 3272 and 2060, respectively. In the analysis of cell differentiation trajectories, the focus was on differentiation fate differential genes. During differentiation from the apical meristem to epidermal cells, three nodes were identified. Branch Node ① contained 665 differentiation fate-related genes, including 33 TFs that promoted cell development toward the 5,6 apical meristem lineages, a direction not relevant to this study. Branch Node ② comprised 64 differentiation fate-related genes (no TFs). Branch Node ③ included 1331 differentiation fate-related genes, including 137 TFs. In this study, the 13 TFs with the strongest significance among the top 40 genes were selected for heatmap visualization in ascending order of *p*-value (Fig. 3B). The pseudo time axis extends on both sides of the heatmap, and the middle starting branch of the heatmap is the pseudo time axis starting point, which includes two states: state 5, 4, 2, 1, 7 and state 5, 4, 3. These 13 TFs are all highly expressed in the 5, 4, 3 direction, driving apical meristematic cells to differentiate into epidermal cells in the 5, 4, 3 directions. In contrast, their expression in the 5, 4, 2, 1, 7 direction is relatively weak, although this direction also tends towards epidermal cell differentiation. Bubble plots (Fig. 3C) illustrate the expression levels of the 13 TFs across clusters 0-11, revealing varied expression patterns. Notably, *NAC081* and *NAC048* were predominantly expressed in clusters 2 and 8 (root endodermal cells).

Candidate TF PcNAC048 identified by TF target gene regulatory network analysis

To explore the interrelationships between the 13 TFs involved in the differential expression of genes at branch node ③ and the key genes of the mevalonate (MVA) and methylerythritol 4-phosphate (MEP) pathways of patchouli alcohol biosynthesis during the differentiation of apical meristematic tissue cells into epidermal cells in *P. cablin* embryoids, the expression regulatory networks of the 13 TFs and their target genes within these two pathways were constructed in this study.

The predicted TF-target gene regulatory network revealed correlations between 13 TFs and genes involved in the MVA and MEP metabolic pathways. Notably, Pat_A22G120500 (*NAC081*) exhibited interactions with several other transcription factors. Pat_B04G150300 (*WRKY11*) and Pat_A02G122400 (*ERF109*) were predicted to regulate the expression of Pat_A16G063700 (*TPS9*). Pat_B19G076600 (*ZAT10*) exhibited close relationships with Pat_A02G122400 (*ERF109*), Pat_B01G105100 (*ERF109*), and Pat_B02G094700 (*ERF109*). Furthermore, Pat_B16G144500

(*NAC048*), predicted to function as a transcription factor, could regulate the expression of key genes involved in the biosynthesis of patchouli alcohol such as Pat_B04G175400 (*HMGR2*), Pat_B03G170100 (*HMGR2*), Pat_A03G179400 (*HMGR2*), Pat_A03G023200 (*TPS11*), and Pat_A03G034000 (*TPS7*).

In supplementary Fig. S3A and Table S3, the larger the degree value, the more neighbors the node has directly connected to in the network, indicating that it has a higher direct influence in the network. The betweenness value is an indicator used in network analysis to measure the importance of a node in the network. The larger its value, the more significant the node's role as a "bridge" or "key hub" in the network. From supplementary Table S3, PcNAC081 (Pat_A22G120500\Pat_B22G113900) and PcNAC048 (Pat_B16G144500) have the highest degree value of 14, while PcNAC048 (Pat_B16G144500) has the highest betweenness value of 81.7394319131161, which is much higher than other TFs. The combined degree and betweenness values of PcNAC048 (Pat_B16G144500) indicate that the *PcNAC048* gene is prioritized in the gene regulatory network (GRN). To further predict the function of *PcNAC048*, the tSNE plots were used to display the expression patterns of *PcNAC048* and terpenoid synthase genes *TPS11*/*TPS7* in epidermal cells, respectively. The results indicate that the expression pattern of *PcNAC048* (Pat_B16G144500) is like *TPS* (Supplementary Fig. S3B), and we predict that it may play an important role in the biosynthesis of patchouli alcohol.

Based on the predictive results, scatter plots were generated to visualize the gene expression trends of *PcNAC048*. As depicted in Fig. 3D, Pat_B16G144500 (*PcNAC048*) exhibited gene expression patterns aligned with the solid line trajectory 5,4,3, indicating high expression abundance along the pathway of epidermal cell differentiation. Therefore, this study aims to preliminarily observe the potential role of *PcNAC048* in lateral root development and secondary metabolism.

PcNAC048 bioinformatics and expression profile analysis

The *PcNAC048* gene's open reading frame (ORF) was 840 bp (Supplementary Fig. S4) and encoded 279 amino acids. ProtParam analysis revealed that the protein encoded by the *PcNAC048* gene possessed a theoretical isoelectric point of 6.24 and a molecular weight of 31,335.61 Da. The instability index (II) was 43.67, suggesting a potentially unstable protein. The aliphatic index was 70.68, while the grand average of hydropathicity (GRAVY) was -0.448. Analysis using the ProtScale online software further indicated that PcNAC048 exhibited predominantly hydrophilic

behavior (Supplementary Fig. S5A). SOPMA software was employed to predict the secondary structure of the PcNAC048 protein. The results showed that the amino acid sequence contained four types of secondary structures, with the random coil being the most abundant, represented by the purple region, constituting 69.53% of the structure (194 amino acids). The alpha helix, depicted in blue, comprised 13.98% of the sequence (39 amino acids). The extended strand was illustrated in red, accounting for 12.90% of the structure (36 amino acids). The beta turn was depicted in green, comprising 3.58% of the sequence (10 amino acids) (Supplementary Fig. S5B). The conserved region of NAC was present within the amino acid sequence of PcNAC048, and it was classified as a member of the NAM supergene family (Supplementary Fig. S5C).

At the N-terminal end of a typical NAC protein, the highly conserved NAC domain comprising five conserved subdomains (A-E, depicted in pink). Substructural domains C and D were responsible for DNA binding (Fig. 4A). Multiple sequence alignment of the amino acid sequence encoded by the *PcNAC048* gene with the NAC48 amino acid sequences from millet, maize, *Acorus calamus*, *Apostasia shenzhenica*, *Oryza sativa*, and *Cocos nucifera* revealed that the amino acid sequence of PcNAC048 showed high similarity within the conserved NAC domain to the NAC48 amino acid sequences of several other plant species (Fig. 4B). Phylogenetic tree analysis revealed that PcNAC048 clustered with AcNAC48 from *Acorus calamus*, indicating high homology (Supplementary Fig. S5D).

Expression profile analysis of *PcNAC048* in various *P. cablin* tissues showed that *PcNAC048* exhibited the highest expression in leaves, with 57.95% lower expression in roots compared to leaves (Fig. 4C(1)). Following treatment at 4 °C for 3 and 6 hours, the expression level of *PcNAC048* increased in *P. cablin*, although the difference was not significant. However, at 12 hours, a significant increase in expression level was observed (Fig. 4C(2)). Analysis of *P. cablin* under 200 mM NaCl salt stress showed that *PcNAC048* responded to salt stress. The *PcNAC048* expression level significantly increased after 3 hours of treatment with 200 mM NaCl, reaching a peak at 6 hours, followed by a subsequent decrease (Fig. 4C(3)). Analysis of *P. cablin* under 20% PEG6000 drought stress revealed that *PcNAC048* responded to drought stress. *PcNAC048* expression significantly increased after 3 hours of treatment with 20% PEG6000, followed by a decrease at 6 hours and reaching a peak expression at 12 hours (Fig. 4C(4)).

PcNAC048 localized to the nucleus and cytoplasm of *Arabidopsis* protoplasts

The subcellular localization of PcNAC048 was predicted to be nuclear using the online tool

WoLF PSORT. To experimentally determine its localization, the ORF of *PcNAC048*, excluding the stop codon, was fused with the enhanced green fluorescent protein (*EGFP*) gene to construct the fusion vector *PcNAC048-EGFP* (Fig. 5A). This construct was then transformed into *Arabidopsis* protoplasts using PEG 4000. Green fluorescence expression was subsequently observed using laser scanning confocal microscopy. The results showed that *EGFP_{control}* exhibited diffuse green fluorescence throughout the cell, while *PcNAC048-EGFP* showed green fluorescence primarily in the nucleus with a few distinct green fluorescent particles visible in the cytoplasm, indicating that the *PcNAC048* protein was localized in the nucleus with a minor cytoplasmic presence (Fig. 5B).

Overexpression of *PcNAC048* led to an increase in the lateral root count in transgenic *Arabidopsis thaliana* plants.

During the development of *P. cablin* somatic embryos, *PcNAC048* is highly expressed in the root endothelial cells. To investigate the function of *PcNAC048* in plant root development, we conducted transgenic research in the model plant *Arabidopsis*, as the transgenic system of *P. cablin* is not yet developed. The structure of the genetic transformation vector for *PcNAC048* in *Arabidopsis thaliana* is shown in Fig. 5C. A comparison of the growth of transgenic *PcNAC048 A. thaliana* and wild type (WT) is shown in Fig. 5D. The results demonstrated that #048-1, #048-2 and #048-3 exhibited an increased lateral root count compared to the WT, and the plants displayed larger overall size than the WT. Gene expression analysis showed that *PcNAC048* expression in pBWA(V)BS-*PcNAC048-2* was 146.61-fold higher than that of the WT, while *PcNAC048* expression in pBWA(V)BS-*PcNAC048-3* was 185.05-fold higher than that of the WT (Fig. 5E). Taproot length analysis did not reveal any significant differences between WT and pBWA(V)BS-*PcNAC048-1*, pBWA(V)BS-*PcNAC048-2*, and pBWA(V)BS-*PcNAC048-3* (Fig. 5F). Examination of the quantity of lateral roots revealed that stable overexpression of *PcNAC048* in pBWA(V)BS-*PcNAC048* lines significantly increased the lateral root count compared to the WT. Specifically, the lateral root count in pBWA(V)BS-*PcNAC048-1*, pBWA(V)BS-*PcNAC048-2*, and pBWA(V)BS-*PcNAC048-3* was 4-fold, 3.67-fold, and 4.67-fold higher than that of the WT, respectively (Fig. 5G).

PcNAC048* is a transcriptional inhibitor of *PcPTSPro

To investigate potential interactions between *PcNAC048* and the promoter of the patchouli alcohol synthase gene (*PcPTS*), a yeast one-hybrid (Y1H) assay was performed. The Y1H assay

results demonstrated that the reporter strain harboring pHis2-PcPTSPro (1865bp) and pGADT7-PcNAC048 exhibited growth on synthetic dextrose (SD) medium lacking L-histidine, L-leucine, and L-tryptophan (SD/-His/-Leu/-Trp) supplemented with 10 mM or 20 mM 3-amino-1,2,4-triazole (3-AT), consistent with the positive control (pHis2-p53+pGAD53m).

In contrast, the negative control strain containing pHis2-PcPTSPro (1865bp), and the empty vector (pGADT7) did not exhibit growth on SD/-His/-Leu/-Trp supplemented with 10 mM or 20 mM 3-AT. These findings suggest that, within the yeast system, the transcription factor PcNAC048 could interact with the 1865 bp promoter region of the *PcPTS* gene (Fig. 6A).

To analyze the regulatory effect of PcNAC048 on PcPTSPro, a dual-luciferase assay was performed in *Nicotiana benthamiana*. The effectors and reporters were constructed as depicted in Fig. 6B. *In vivo* imaging results in *N. benthamiana*, as shown in Fig. 6C, revealed that the region injected with PcPTSPro::LUC+35S as the control group exhibited more intense red fluorescence, whereas the region injected with PcPTSPro::LUC+35S::PcNAC048 as the experimental group displayed minimal red fluorescence, indicating PcNAC048 effectively inhibited the activity of PcPTSPro given the significantly stronger signal intensity in the control group than in the experimental group. As shown in Fig. 6D, PcNAC048 significantly inhibited *PcPTS* promoter activity. The LUC activity in the experimental group was reduced by 78.94% compared to the control group, suggesting that PcNAC048 could effectively inhibit the transcriptional expression of *PcPTS* genes by regulating the activity of the PcPTSPro. These findings were consistent with the *in vivo* imaging results obtained in tobacco.

Transient overexpression of *PcNAC048* reduced patchouli alcohol synthesis

To investigate the role of *PcNAC048* in patchouli alcohol biosynthesis, transient overexpression analysis was employed in this study. This approach was selected due to the inherent challenges associated with establishing a stable transformation system in *P. cablin*. Fig. 7A illustrates the schematic diagram of the pJLTRBO-*PcNAC048* vector.

Following transient overexpression of *PcNAC048* for four days, the relative expression of *PcNAC048* was significantly increased, exhibiting 38.38% upregulation compared to the CK (Fig. 7B). Conversely, the expression of *PcPTS* was significantly downregulated, showing a 31.24% decrease compared to the CK (Fig. 7C). Changes in patchouli alcohol content in leaf tissues of pJLTRBO-*PcNAC048* and CK were determined using gas chromatography-mass spectrometry (GC-MS), and the content was calculated based on a standard curve formula ($Y=6431562.84X-$

18583.01, $R^2=0.9997$). As shown by the results, the trend in patchouli alcohol content mirrored the expression pattern of *PcPTS*, exhibiting a significant reduction. Overexpression of *PcNAC048* resulted in a significant decrease in patchouli alcohol biosynthesis accumulation. The patchouli alcohol content (1.741 mg/g) was reduced by 36.74% in the *PcNAC048*-overexpressing lines compared to the CK (2.752 mg/g) (Fig. 7D). Chromatographic analysis revealed that at a retention time of 14.687 min, the peaks for CK, pJLTRBO-*PcNAC048*, and the patchouli alcohol standard coincided. The peak area of patchouli alcohol in the *PcNAC048*-overexpressing lines was smaller than that in the CK (Fig. 7E), which aligned with the quantitative detection results, indicating that *PcNAC048* could inhibit the biosynthesis of patchouli alcohol.

Virus-induced *PcNAC048* silencing increased the synthesis of patchouli alcohol

To further investigate the role of *PcNAC048* in patchouli alcohol biosynthesis, the vector pTRV2-*PcNAC048* was constructed (Fig. 8A), and the empty vector pTRV1+pTRV2 served as the CK. After three weeks of *PcNAC048* gene silencing, the relative expression level of *PcNAC048* in *P. cablin* leaves showed a significant decrease of 48.67% compared to the CK (Fig. 8B). Conversely, the expression level of *PcPTS* increased significantly, showing 35.02% upregulation compared to the CK (Fig. 8C). Changes in patchouli alcohol content in leaf tissues of VIGS-*PcNAC048* and CK were determined using GC-MS, and a standard curve formula ($Y=6465126.01X-14738.05$, $R^2=0.9992$) was used to calculate the content. As shown by the results, the trend in patchouli alcohol content mirrored the expression pattern of *PcPTS*, exhibiting a significant increase. Gene silencing of *PcNAC048* resulted in a significant increase in patchouli alcohol biosynthesis accumulation. The content of patchouli alcohol (1.134 mg/g) was increased by 59.47% in the VIGS-*PcNAC048* lines compared to the CK (0.711 mg/g) (Fig. 8D). Chromatographic analysis revealed that at a retention time of 14.645 min, the peaks for CK, VIGS-*PcNAC048*, and the patchouli alcohol standard coincided. The peak area of patchouli alcohol in the VIGS-*PcNAC048* lines was larger than that in the CK after *PcNAC048* silencing (Fig. 8E), consistent with the results of the quantitative detection, indicating that *PcNAC048* exerts an inhibitory effect on the biosynthesis of patchouli alcohol.

Discussion

P. cablin exhibits unique characteristics compared to other plants, notably its distinctive aroma²³. In our previous study, transcriptome profiling of various tissues in *P. cablin*, including leaves, stems, and roots, identified 78 genes involved in terpenoid synthesis. Among these genes,

28 were associated with the MVA pathway, 38 with the MEP pathway, and 12 with downstream terpenoid pathways²⁴. Further insights were gained from a previous study on the analysis of full-length transcriptome sequencing of genes involved in patchouli biosynthesis regulation and the profiling of methyl jasmonate (MeJA)-induced gene expression. This study, focusing on phytohormone signaling, suggested that MeJA could stimulate patchouli alcohol synthesis²⁵. In the present study, scRNA-seq analysis was employed to elucidate the cell types present in GSE and CSE of *P. cablin*. A total of 12 cell clusters were identified, including epidermal cells, guard cells, mesophyll cells, apical meristem cells, meristematic cells, endodermal cells, phloem cells, and xylem cells. Each cluster was defined by the presence of marker genes expression. The identified cell types aligned with previous findings, such as those obtained from scRNA-seq analysis of longan somatic embryos, which identified twelve cell clusters encompassing vascular cells, proliferating cells, and epidermal cells⁵. The scRNA-seq data generated in this study on *P. cablin* GSE and CSE provides a crucial foundation for future research into *P. cablin* embryonic development at the cellular level. A highly interconnected regulatory network was established based on the identified TFs and genes involved in patchouli alcohol biosynthesis.

Plant scRNA-seq studies have demonstrated the technical advantages of large-scale and high-throughput data generation, exemplified by investigations into cotyledon vein development²⁶ and stomatal lineage trajectories²⁷. Based on the analysis, cell clusters ranging from 0 to 11 were identified as eight distinct cell types. Cell frequency analysis revealed a significant change in the number of cells within cluster 1 of both the GSE and the CSE. Subsequent marker gene analysis identified the cells corresponding to cluster 1 as epidermal cells. The significantly greater number of cells in CSE compared to GSE suggests a dynamic enhancement of epidermal cell differentiation during the transition from the GSE to CSE development. As a non-model plant, *P. cablin* has limited reported marker genes available for cell type identification. Therefore, this study utilized a combination of reported marker genes, including *PROTODERMAL FACTOR 1* (*PDF1*)²⁸ and *RUBISCO BISPHOSPHATE CARBOXYLASE SMALL SUBUNIT* (*RBCS*)²⁹, to identify cell types in *P. cablin*. Furthermore, the gene expression landscape of each cell cluster was characterized by its unique set of marker genes, reflecting the high precision of the scRNA-seq technique. Pseudotime analysis can reveal cell differentiation pathways and dynamic processes in single-cell transcriptome analysis. In this study, pseudotime analysis revealed a potential differentiation relationship between apical meristematic cells and epidermal cells. It is interesting

that the apical meristematic cells mainly differentiate into epidermal cells along the 5-4-2-7-1 and 5-4-3 trajectories. Although trajectory 5-6 also undergo developmental processes, 6 is in a relatively later stage and contains only a small number of epidermal cells, which may indicate that a small number of apical meristematic cells have already differentiated into epidermal cells during this stage. Overall, these results demonstrate the feasibility and effectiveness of the scRNA-seq method for cell type identification in non-model species. Importantly, pseudotime analysis revealed 13 TFs belonging to the ERF, C2H2, NAC, and WRKY families as potential regulators of epidermal cell differentiation. What caught our attention is the TF *PcNAC048* of the NAC family, whose expression increases towards the differentiation trajectory of epidermal cells with increasing pseudotime. Similarly, this gene is expressed in all cell clusters, but has a higher expression level in cluster 2, which is the root endothelial cells. We speculate that *PcNAC048* may play a role in the differentiation of epidermal cells in *P. cablin* somatic embryos and may have a more comprehensive impact on root endothelial cells, but further verification is needed.

Previous research has demonstrated that NAC TFs are expressed in a stage-specific manner during leaf development and play crucial roles in regulating leaf morphological transformations³⁰. It can be inferred that the NAC TFs identified in this study may also function as key regulators of plant development and morphological transformations.

In this study, the expression of *PcNAC048* was analyzed under low-temperature, drought, and salt stress conditions. The results indicated that *PcNAC048* exhibited high sensitivity to temperature fluctuations, high salinity, and drought stress. In *Hibiscus hamabo*, 182 *HhNACs* were identified. RNA-seq expression analysis showed that most *HhNACs* were differentially expressed under salt and drought stress conditions, especially *HhNAC54*³¹. Importantly, the *CaNAC46* gene in pepper demonstrated comparable expression patterns³².

TFs are core components of cellular signaling networks and are typically localized in the nucleus. *PcNAC048*, identified in this study, also exhibited nuclear localization, consistent with previous reports. Interestingly, a small number of fluorescent particles expressing *PcNAC048*-EGFP were also observed in the cytoplasm. It has been reported that approximately 10% of TFs are membrane-bound transcription factors (MTFs), which are latent TFs intrinsically anchored in the cell membrane in a dormant state. Upon exposure to environmental stress, these MTFs are released from the membrane through the ubiquitin-proteasome pathway or by intramembrane proteases. Following their release, they are transported across the cytoplasm and subsequently

localize to the nucleus³³. Based on these observations, we hypothesize that PcNAC048 may function as an MTF, although further investigation is required to confirm this hypothesis.

It has been shown that *NAC* genes play a significant role in regulating lateral root development in plants³⁴. *ZmNAC1* was isolated from two inbred maize lines, Zong3 and 87-1. Transgenic *Arabidopsis thaliana* overexpressing *ZmNAC1* displayed a higher count of lateral roots in comparison to the wild type, suggesting a crucial role for *ZmNAC1* in lateral root development³⁵. The soybean *GmNAC20*, which is also a NAC-type TF, was transferred into the rice genome by *Agrobacterium*-mediated transformation. Overexpression of *GmNAC20* was also found to induce lateral root formation³⁶. *NAC056*, a nitrate-induced NAC family TF, was identified by Xu Pei Pei. The *NAC056* mutant displayed impaired root growth, while overexpression of *NAC056* promoted lateral root development and enhanced tolerance to nitrate deficiency³⁷. However, no studies have hitherto investigated NAC-related transcription factors in *P. cablin*. In this study, *PcNAC048* was stably overexpressed in *A. thaliana*. The results demonstrated that *A. thaliana* overexpressing *PcNAC048* exhibited a significantly higher lateral roots count compared to the WT. The lateral roots count in the overexpression lines was three-fold higher than that of the WT. The results agreed with those of earlier research studies. Importantly, these results suggest that the *PcNAC048* gene can induce lateral root development and formation. This conclusion is also supported by scRNA-seq results, which revealed that *PcNAC048* was highly expressed in cluster 2, namely, root endothelial cells.

The phenotype of lateral roots is excellent, and *PcNAC048* can promote lateral root development while also promoting the differentiation of apical meristematic cells into epidermal cells. The common feature between the initiation and growth of lateral roots and the differentiation of epidermal cells may be cell division, and *PcNAC048* may be involved in both processes simultaneously. However, further research is needed for cellular level analysis.

Patchouli alcohol is an important bioactive component of *P. cablin*, which is mainly synthesized in leaves and widely employed in various applications, including soaps, decorative cosmetics, the perfumery industry, and the medical industry³⁸. Driven by the growing global demand, metabolic engineering approaches have been recognized as an effective way to increase useful secondary metabolites in medicinal plants. It has been reported that several strategies have been used to facilitate the industrialization of the production of patchouli alcohol using brewer's yeast³⁹. In addition, many TFs have proven to play crucial roles in regulating patchouli alcohol

biosynthesis, including *PatDREB*⁴⁰, *PatSWC4*⁴¹, and *PcWRKY44*⁴², etc. The C-terminus of the NAC transcription factor functions as a transcriptional repressor or activator⁴³. In this study, PcNAC048 was found to bind to PcPTSpro by the Y1H, and the dual-luciferase assay confirmed the ability of PcNAC048 to inhibit PcPTSpro activity. The LUC activity in the experimental group was reduced by 78.94% compared to the CK, indicating that PcNAC048 could regulate the transcriptional expression of the *PcPTS* gene by inhibiting *PcPTS* promoter activity. A previous study in *P. cablin* demonstrated that overexpression of *PatJAZ6*, a repressor of patchouli alcohol synthesis, significantly reduced the transcript abundance of *PcPTS*, thereby inhibiting patchouli alcohol production⁴⁴. This finding was consistent with the observed trend in this study. In addition, an efficient VIGS system has been established in *P. cablin*⁴⁵. *PcNAC048* gene silencing resulted in a significant increase in patchouli alcohol accumulation, contrary to the results observed following transient overexpression of *PcNAC048*. These findings in the leaves collectively confirm that *PcNAC048* plays an inhibitory role in patchouli alcohol biosynthesis.

As is well known, leaves are the main site for important physiological processes such as photosynthesis and substance synthesis in plants, and many key biosynthetic pathways and metabolic regulatory networks are actively expressed in leaves. Although somatic embryos and leaves are different plant tissues, they both follow the basic laws of plant cell development and metabolism. Some transcription factors involved in cell division and differentiation play important regulatory roles in the formation of somatic embryos and the growth and development of leaves. For example, the WUSCHEL (WUS)-related homeobox (WOX) transcription factors play important roles in embryogenesis, meristem maintenance, and leaf blade outgrowth in flowering plants⁴⁶. In this study, although scRNA-seq analysis was performed on *P. cablin* somatic embryos, selecting leaves as the gene function research system helped us to comprehensively understand the biological function of the *PcNAC048* gene. In the future, we may be able to develop stable overexpression lines and CRISPR knockout lines of *PcNAC048* in *P. cablin* plants, analyze the metabolic profile of leaf epidermal cells, and further clarify the role of *PcNAC048* in determining embryonic epidermal fate and continuously regulating key metabolite synthesis in later leaves.

Out of interest in the synthesis of patchouli alcohol, Seurat software was also used to score the identified 8 types of cells and the gene sets of MVA/MEP pathways. The results showed that there were significant differences between GSE and CSE samples of root endothelial cells, and the scores were relatively higher in GSE samples, indicating that root endothelial cells may have a

greater contribution to the MVA/MEP pathway (Supplementary Fig. S6), but the specific molecular mechanism still needs further in-depth exploration. Moreover, In the process of single-cell data analysis, we also discovered potential differentiation relationships between meristematic cells and root endothelial cells, providing new clues for further understanding the complexity of root development processes and pointing the way for future research.

Conclusions

The scRNA-seq results in this study provide novel spatiotemporal insights into the differentiation of somatic embryos in *P. cablin*. The identified *PcNAC048* gene contributes to elucidating the processes of lateral root development and the biosynthesis of patchouli alcohol in *P. cablin*. Due to the limited number of identified marker genes related to somatic embryogenesis in *P. cablin*, the identification of other cell types was challenging. However, our scRNA-seq analysis results of somatic embryos in *P. cablin* may offer valuable clues for cell type identification in non-model species. Furthermore, the newly identified dual-function *PcNAC048* gene may provide a foundation for in-depth studies of *P. cablin*.

Methods

Plant materials

Pogostemon cablin plants were collected in Zhaoqing City, Guangdong Province, China. *P. cablin* seedlings were planted in the school's medicinal plant cultivation base and allowed to grow under natural conditions. Plants were watered every 3-4 days, except during periods of rainfall, and were cultivated for approximately one year. In mid-March of the second year, some *P. cablin* plants exhibited the formation of flower buds (Fig. 1A). In addition, some patchouli seedlings were cultivated in an incubator maintained at a constant temperature of 25°C, with a 16-hour light cycle and an 8-hour dark cycle, for gene function analysis. The sterilized seeds of *Arabidopsis thaliana* were sown onto MS solid medium and incubated under a 16-hour light (22 °C)/ 8-hour dark (18 °C) cycle. *Nicotiana benthamiana* seeds were sown in nutrient soil and cultured outdoors in sufficient light. Non-bolting seedlings were used for analysis.

Identification of microspores in *P. cablin*

Upon the emergence of flower buds in *P. cablin*, unopened, tender green buds were collected, and the anthers were peeled off with medical forceps under a dissecting microscope. The peeled anthers were placed on a standard glass slide, dissociated by the dropwise addition of 1 mol/L HCl for 10 minutes, and subsequently stained with acetic acid-carmin staining solution. The stained

anthers were then gently pressed and examined under a microscope using conventional pressing techniques. Images were captured for subsequent analysis.

Somatic embryo induction and protoplast isolation

Flower buds of *P. cablin* with microspores at the mononuclear stage of development were selected. The anthers were carefully separated and then inoculated onto an induction medium containing 0.1 mg/L NAA and 2.0 mg/L 6-BA. The culture plates were incubated in the dark at 25 °C for 20 days to induce the formation of embryogenic callus tissue. Subsequently, the cultures were transferred to a plant culture room maintained at 25 °C under a 16 h light/8-hour dark photoperiod for further growth and development.

Somatic embryos were observed and isolated using a stereomicroscope. Protoplast suspensions were prepared from *P. cablin* GSE and CSE using the enzymatic hydrolysis method. To prepare the protoplast cleaning solution, 20 µL of 500× BSA and 3.43 mL of ddH₂O were added to 6.25 mL of 0.8 M mannitol. The solution was then pre-cooled in a refrigerator at 4°C. To prepare the enzyme solution, cellulose R10 (0.075 g), pectinase (0.025 g), 3.125 mL of 0.8 M mannitol, 500 µL of 0.1 M MES (pH 5.7), 500 µL of 0.2 M KCl, and 870 µL of ddH₂O were weighed or measured and put into a 15 mL centrifuge tube and incubate the tube in a 55°C-water bath for 10 minutes, with vortexing and mixing performed twice during this period. Once the solution had cooled to room temperature, 25 µL of 500× BSA and 50 µL of 1 M CaCl₂ were added. The solution was then mixed thoroughly and filtered through a 0.22 µm filter into another 15 mL centrifuge tube.

The *P. cablin* GSE and CSE were washed twice with protoplast washing solution. Subsequently, they were transferred to centrifuge tubes containing enzyme solution and incubated on a shaking table at 30 °C with 75 rpm in the dark for 0.5-24 h. A 5 µL aliquot of 0.4% trypan blue staining solution was mixed with 5 µL of protoplast suspension. Cell concentration and viability rates were then determined using the Countess[®] II automated cell counter. Protoplasts with greater than 80% viability were used for 10x single-cell library construction within 30 min.

scRNA-seq library construction, sequencing, and raw data quality control

To construct scRNA-seq library, cellular suspensions were loaded on a 10X Genomics GemCode Single-cell instrument that generates single-cell Gel Bead-In-Emulsion (GEMs). Libraries were generated and sequenced from the cDNAs with Chromium Next GEM Single Cell 5' Reagent Kits v3.1. Upon dissolution of the Gel Bead in a GEM, primers containing

(i) an Illumina® R1 sequence (read 1 sequencing primer), (ii) a 10 nt Unique Molecular Identifier (UMI), (iii) a 16 nt 10x Barcode, and (iv) a poly-dT primer sequence were released and mixed with cell lysate and Master Mix. Full-length cDNAs, barcoded were then reverse-transcribed from poly-adenylated mRNA. Silane magnetic beads were used to remove leftover biochemical reagents and primers from the post GEM reaction mixture. Full-length, barcoded cDNAs were then amplified by PCR to generate sufficient mass for library construction. R1 (read 1 primer sequence) were added to the molecules during GEM incubation. P5, P7, a sample index, and R2 (read 2 primer sequence) were added during library construction via End Repair, A-tailing, Adaptor Ligation, and PCR. The final libraries contained the P5 and P7 primers used in Illumina bridge amplification. The Single Cell 5' Protocol produced Illumina-ready sequencing libraries. A Single Cell 5' Library comprised standard Illumina paired-end constructs which begin and end with P5 and P7. The Single Cell 5' 16 bp 10x Barcode and 10 bp UMI were encoded in Read 1, while Read 2 was used to sequence the cDNA fragment. Sample index sequences were incorporated as the i7 index read. Read 1 and Read 2 were standard Illumina® sequencing primer sites used in paired-end sequencing. The Illumina HiSeq2500 sequencing platform (Genedenovo Biotechnology Co., Ltd, Guangzhou, China) was utilized to conduct high-throughput sequencing of the prepared libraries.

10X Genomics Cell Ranger software was used to convert raw BCL files to FASTQ files, alignment and counts quantification. Reads uniquely mapped to the transcriptome and intersecting an exon at least 50% were considered for UMI counting. The cell by gene matrices were produced via UMI counting and cell barcodes calling. Cell Ranger software harnessed the Spliced Transcripts Alignment to a Reference (STAR) tool to align cDNA sequence fragments to the reference genome of *P. cablin*⁴⁷.

Cell clustering analysis and cell type identification

After removing low-quality cells, we employed the Harmony to aggregate all samples. This approach effectively mitigated the influences of batch effects and behavioral variations on clustering by projecting cells into a unified embedding space, where they cluster primarily based on cell type rather than being segregated by dataset-specific factors. Integrated expression matrix was then scaled and performed on principal component analysis (PCA) for dimensional reduction. Seurat implements a graph-based clustering approach. Briefly, Seurat embed cells in a shared-nearest neighbor (SNN) graph, with edges drawn between cells via similar gene expression

patterns. To partition this graph into highly interconnected quasi-cliques or communities, we first constructed the SNN graph based on the Euclidean distance in PCA space and refined the edge weights between any two cells based on the shared overlap in their local neighborhoods (Jaccard distance). To visualize the clusters, we generated t-distributed Stochastic Neighbor Embedding (t-SNE) using the same principal components (PCs). Subsequently, the log-normalized matrices were loaded on SingleR packages for cell type annotation, which based on correlating gene expression of reference cell types with single-cell expression. Firstly, a Spearman coefficient was calculated for single-cell expression with each of the samples in the reference data set. Secondly, multiple correlation coefficients per cell type were aggregated to provide a single value per cell type per single cell. Finally, SingleR reran the correlation analysis, but only for the top cell types from the previous step. The analysis was performed only on variable genes. The cell type corresponding to the top value after the last run was assigned to the single cell.

Identification of cell types using known marker genes. Currently, multiple marker genes have been reported in plant species. *LHCB 4.1*⁴⁸ and *RIBULOSE BISPHTHOSPHATE CARBOXYLASE SMALL SUBUNIT (RBCS)*⁴⁹ were used as marker genes of mesophyll. *PROTODERMAL FACTOR 1 (PDF1)*^{50,51} and *LIPID TRANSFER PROTEIN 1 (LTP1)*⁵² are highly specific for epidermal cells. Basic helix-loop-helix (bHLH) DNA-binding superfamily protein *FAMA*⁵³ was used as marker genes for guard cell. *HIS2A*⁵⁴, *HIS2B* and *HTR2*⁵⁵ were used as marker genes for apical meristem. *STP13*⁵¹, *CSA*⁵⁶, *NHL10*⁵¹ and *PUMP5*⁵⁷ were used as marker genes for root endodermis. *RESTRICTED TEV MOVEMENT 2 (RTM2)* and *SEOB*⁵⁸ were used as marker genes for phloem. *IRREGULAR XYLEM 9 (IRX9)*^{51,59}, *RAC9* and *XYLEM CYSTEINE PROTEASE 1 (XCPI)* was used as marker genes for xylem.

Pseudotime trajectory analysis

Cells were analyzed using the Monocle2 software and arranged along a cellular trajectory based on their temporal changes. Monocle2 reduces high-dimensional data to a two-dimensional space, enabling the visualization of cell differentiation trajectories within this reduced dimensional space. Differentially expressed genes (DEGs) were identified by Monocle2, and their significance was statistically analyzed. This analysis identified essential genes associated with developmental and differentiation processes (false discovery rate (FDR) < 1e-5). Furthermore, genes exhibiting similar expression patterns were clustered, suggesting potential shared biological functions within these populations. Branching Expression Analysis Modeling (BEAM), employing generalized

linear modeling (GLM), was utilized to analyze genes dynamically expressed along pseudo-time-dependent or branching-dependent trajectories. This analysis identified DEGs on both pseudo-time-dependent cellular differentiation trajectories and on branching-dependent trajectories (upregulated, $FDR < 1e-5$).

Analysis of transcription factors and target gene regulatory network

Pearson correlation coefficient calculations were performed on TFs, MEP pathway genes, and MVA pathway genes involved in patchouli alcohol biosynthesis (Supplementary Table S3). Genes were screened based on a threshold of $|\text{correlation coefficient}| > 0.6$. Cytoscape was selected to redraw and visualize the regulatory network diagram of TF-target genes based on the resulting data.

Cloning and multiple sequence alignment analysis of *PcNAC048*

The target gene *PcNAC048* (Pat_S16G144500 gene) was identified within the genome of *P. cablin*⁴⁷. Specific primers were designed using SnapGene software for amplifying the open reading frame (ORF) region. The primers are listed in Supplementary Table S4. The target fragment was amplified and inserted into the pLB vector, followed by sequencing to confirm the accuracy of the *PcNAC048* ORF. Multiple sequence alignment was conducted using DNAMAN, while MEGA7 was utilized for the construction of the phylogenetic tree.

Expression profile analysis of *PcNAC048*

Seedlings of *P. cablin* were cultivated within an artificial climate incubator. Upon reaching the 6-8 leaf stage, the uppermost pair of leaves were designated as young leaves (YL), and the third pair of leaves were designated as mature leaves (L). Stem segments (S) were collected from the bottom to the top of the plant, encompassing 3-4 internodes. Root samples (R) were also collected in sufficient quantities. Immediately freeze all samples in liquid nitrogen and subsequently subjected to RNA extraction for subsequent analysis. The *PcNAC048* expression levels in mature leaves (L) served as the control for the analysis.

For stress response analysis, cultured *P. cablin* plants within an artificial climate incubator were divided into experimental and control groups. To assess the response of *PcNAC048* to 4 °C stress, the experimental group was transferred to an incubator maintained at 4 °C, while the lighting conditions remained unchanged. The control group was cultured under standard growth conditions. For salt stress analysis, the experimental group was treated with 200 mM NaCl, while the control group was cultured under standard growth conditions. For drought stress analysis, the experimental

group was treated with 20% PEG6000, while the control group was cultured under standard growth conditions. Leaf tissues were excised and stored in liquid nitrogen from both groups at 0 h, 3 h, 6 h, and 12 h after the initiation of stress treatments. The expression levels of the target gene at 0 h were used as a control for subsequent analysis.

Subcellular localization analysis of PcNAC048

The ORF of *PcNAC048*, excluding the termination codon, was fused to the N-terminus of the pAN580-GFP vector to generate the PcNAC048-EGFP fusion construct (Fig. 5A). Both the empty vector pAN580 and the PcNAC048-EGFP fusion vector were transiently transformed into *Arabidopsis* protoplasts using PEG4000. Subcellular localization was subsequently observed using a Nikon C2-ER laser confocal microscope (Nikon, Japan).

Transgenic analysis of *PcNAC048* in *Arabidopsis thaliana*

For transgenic analysis in *A. thaliana*, the *PcNAC048* ORF, lacking the termination codon, was fused to the pBWA(V)BS vector to generate the construct pBWA(V)BS-PcNAC048 (Fig. 5C). This construct was subsequently transformed into *Agrobacterium tumefaciens* GV3101, and *A. thaliana* WT plants were transformed using the floral dip method. The harvested seeds were plated on a selective medium containing glyphosate to identify transgenic plants. The FastPure[®] Plant DNA Isolation Mini Kit (Vazyme Biotech Co., Ltd., China, DC104) was used to extract genomic DNA from these plants. To confirm the integration of the bialaphos resistance gene (*bar*) into the plant genome, PCR was then performed, and RT-qPCR was selected to analysis the expression levels of *PcNAC048*.

Yeast One-Hybrid analysis

The *PcNAC048* ORF was fused in the pGADT7 vector digested with *Bam*H I and *Eco*R I, resulting in the construction of the prey vector pGADT7-*PcNAC048*. The bait vector, pHis2-*PcPTSP*ro (1865bp) was obtained from our laboratory. Co-transformation of pHis2-*PcPTSP*ro (1865bp)+pGADT7-*PcNAC048* into Y187 yeast cells served as the experimental group, pHis2-*PTSP*ro (1865bp)+pGADT7 served as the negative control, while pHis2-p53 co-transformed with pGAD53m served as the positive control. Transformants were cultured on SD/-Trp/-Leu. Positive clones were then transferred to SD/-Trp/-His/-Leu medium containing 0 mM 3-AT/10 mM 3-AT/20 mM 3-AT respectively and cultured at 29 °C for 48-96 h. The growth of yeast colonies in the positive, negative, and experimental groups was observed to assess the interaction between PcNAC048 and *PcPTSP*ro (1865 bp) within the yeast system.

Dual luciferase (LUC) analysis

For the LUC analysis, the *PcNAC048* ORF was fused in the pGreenII 62-SK vector digested with *EcoR* I and *BamH* I, resulting in the construction of the effector vector pGreenII 62-SK-*PcNAC048* (35S::*PcNAC048*). The reporter vector PcPTSpro::LUC was obtained from our laboratory (Fig. 6B). The effector, reporter, and the empty vector pGreenII 62-SK were individually transformed into *A. tumefaciens* strain GV3101 (pSoup-p19). A 1:1 mixture of *A. tumefaciens* cultures containing PcPTSpro::LUC and 35S::*PcNAC048* was used as the experimental group. A 1:1 mixture of *A. tumefaciens* cultures containing PcPTSpro::LUC and the pGreenII 62-SK served as the control. Add acetosyringone (AS) to the prepared bacterial solution to a final concentration of 150 $\mu\text{mol/L}$. The solution was then incubated in the dark for 3-5 hours prior to infiltration into *Nicotiana benthamiana*. Dual-luciferase assay kit (Biyuntian, China, RG027) was used to quantify the firefly luciferase and Renilla luciferase after 4 days of infiltration. The ability of *PcNAC048* to regulate transcription was evaluated based on the ratio of LUC to REN.

Transient overexpression analysis of *PcNAC048*

For transient overexpression analysis, the *PcNAC048* ORF was fused in-frame to the *Pac* I and *Not* I restriction sites within the pJLTRBO vector, resulting in the construction of the vector pJLTRBO-*PcNAC048* (Fig. 7A). The competent cells of *A. tumefaciens* GV3101 (pSoup-p19) were transformed with both the empty pJLTRBO vector and the pJLTRBO-*PcNAC048* construct. Subsequently, *P. cablin* leaves were infected with the transformed *Agrobacterium* strains. Leaf samples were harvested four days post-infection for subsequent analysis of target gene expression levels and patchouli alcohol content.

Analysis of virus induced *PcNAC048* silencing in *P. cablin* leaves

For virus induced gene silencing analysis, the *PcNAC048* ORF was fused in-frame to the pTRV2 vector digested with *EcoR* I and *BamH* I to construct the vector pTRV2-*PcNAC048* (Fig. 8A). The competent cells of *A. tumefaciens* GV3101 were separately transformed with the pTRV2-*PcNAC048* construct and the empty pTRV1/pTRV2 vectors. Plants infected with pTRV1+pTRV2 served as the CK, while plants infected with pTRV1+pTRV2-*PcNAC048* served as the experimental group. Following infiltration, an artificial incubator at 24 °C was selected to cultivate those plants. Leaf samples were collected three weeks post-infiltration for subsequent analysis of target gene expression levels and detection of patchouli alcohol content.

Total RNA extraction, synthesis of cDNA, and gene expression analysis

The HiPure Plant RNA Mini Kit (Magen Biotech Co., Ltd., China, R4151-02C) was used to extract total RNA. cDNA for conventional PCR was synthesized using the HiScript III 1st Strand cDNA Synthesis Kit (Vazyme Biotech Co., Ltd., China, R312-01/02). Highly uniform cDNA for RT-qPCR was synthesized using the HiScript III RT SuperMix for qPCR (+gDNA wiper) (Vazyme Biotech Co., Ltd., China, R223-01).

The CFX96™ Real-Time PCR detection system and the ChamQ Universal SYBR qPCR Master Mix Kit (Vazyme Biotech Co., Ltd., China, Q711-02/03) were used for RT-qPCR analysis. The 18S ribosomal RNA gene (*Pc18S*) from *P. cablin* and *AtACT2* from *Arabidopsis* were served as internal reference genes. The relative expression levels of the target genes were calculated by the $2^{-\Delta\Delta C_t}$ method⁶⁰.

Detection of patchouli alcohol content by gas chromatography-mass spectrometry

For patchouli alcohol extraction, samples were retrieved from liquid nitrogen and immediately ground in a pre-cooled mortar. A precise mass of 200 mg of the ground sample was accurately weighed and placed it into a pre-cooled 2mL centrifuge tube. Subsequently, add 1.5 mL of n-hexane into centrifuge tube, and the mixture was subjected to ultrasonic extraction at 60 Hz for 30 minutes. The centrifuge tube was then transferred to a water bath maintained at 56 °C for 1 hour, followed by centrifugation. A 0.22 µm organic filter membrane was selected to filter the supernatant to yield the test solution.

For the reference solution, 10.70 mg of patchouli alcohol reference (Feiyubio, Nantong, China) was dissolved and diluted to the mark in a 10 mL volumetric flask with methanol to obtain the mother liquor. The mother liquor was diluted by a factor of 5, 10, 20, 40, 80, and 100 with methanol respectively. The relationship between concentration and peak area was then used to establish the standard curve for the determination of patchouli alcohol content in transient overexpression of *PcNAC048*. For the experiment of silencing *PcNAC048*, the mother liquor was diluted by a factor of 10, 40, 80, 100, and 200 with methanol. Similarly, the relationship between concentration and peak area was utilized to establish the standard curve for determining the content of patchouli alcohol.

Agilent 7890B gas chromatography mass spectrometer (Agilent, USA) equipped with HP-5MS capillary column (30 m×0.25 mm×0.25 µm) was used for GC-MS analysis. The injection port temperature and ion source temperature were both set at 230 °C, and the quadrupole

temperature was 150 °C. The temperature rise settings were as follows: First, the temperature was held at 50 °C for 2 min. Then, it rose at a rate of 20 °C/min to 130 °C. After that, it rose at a rate of 2 °C/min to 150 °C and was maintained for 5 min. Finally, it rose at a rate of 20 °C/min to 230 °C. The flow rate of the carrier gas (He) was 1 mL/min, and the injection volume was 1 µL.

Statistics and Reproducibility

The significance of the difference was determined by one-way analysis of variance (ANOVA) and unpaired Student's *t*-test. A *p*-value less than 0.05 was statistically significant. Three biological replicates were used for all sample groups in all experiments with the exception of the scRNA-seq experiment, where only single replicates were used.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material. The scRNA-seq data (GSE304316) was deposited in the GEO repository. Uncropped and unedited gel image for Fig. S4 was available as Fig. S7 in the Supplementary Information file. The numerical source data for charts is shown in Supplementary Data 1. Further inquiries can be directed to the corresponding author.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (No. 82373976), the Natural Science Foundation of Guangdong Province (No. 2023A1515030194), the Special Project for Rural Revitalization Strategy in Guangdong Province (No. 2024-NPY-00-041 and 2023-NBH-00-022), and the Guangdong Province Seed Industry Revitalization Action Project (No. 2025-NQD-21-001). We thank Home for Researchers editorial team (www.home-for-researchers.com) for language editing service.

Author contributions

LC, XW, HL and RZ designed the research protocols; XW, XG, WZ, DW and ZY conducted the experiments; XW and DW analyzed the raw data; XW and XG wrote the manuscript. All authors have read the manuscript and agree to submit the final version.

Competing interests

The authors declare no competing interests.

Figure captions/legends:

Fig. 1 Anatomy of *Pogostemon cablin* flowers and observation of somatic embryos. (A) A photograph of the flowering stage of *P. cablin*. (B) Anatomy of a single flower of *P. cablin*. The last image shows the microspores. Scale bar, 500 µm. (C) Morphological display of somatic embryos at different developmental stages. The first and fourth images are globular somatic

embryos (GSE) and cotyledonic somatic embryos (CSE), representing the early and late stages of somatic embryonic development. Scale bar, 500 μ m.

Fig. 2 Single-cell RNA sequencing of globular somatic embryo (GSE) and cotyledon somatic embryo (CSE) of *P. cablin*. (A) Brief flowchart of *P. cablin* somatic embryos scRNA-seq. (B) Visualization of 12 cell clusters identified using tSNE nonlinear clustering method. (C) Bubble chart of marker gene expression distribution. The bubble size reflects the percentage of each cluster expressing the gene, whereas color represents the normalized average expression level of genes in clusters. (D) Visualization of the 10 cell clusters (including two unknown cell clusters) identified by tSNE based on marker gene analysis.

Fig. 3 Pseudotime analysis of apical meristem cells and epidermal cells, and screening of key genes for differentiation fate. (A1-A4) The cell trajectory sequentially displays pseudotime information, cell cluster information, differentiation status information, and sample information. (B) Heat map of 13 transcription factors (TFs) changing with branching expression. (C) Bubble chart of the expression distribution of 13 TFs in the 0-11 cell clusters. (D) Scatter plot of the expression trend of Pat_B16G144500 (*PcNAC048*) gene before and after cell differentiation.

Fig. 4 Bioinformatic analysis and expression profile analysis of *PcNAC048*. (A) The structural diagram of the NAC protein. (B) Multiple sequence alignment of *PcNAC048* with 6 reported NAC48 proteins. AsNAC48 (PKA57102.1), OsNAC48 (sp|Q7F2L3.1), PmNAC48 (RLN03093.1), ZmNAC48 (NP_001136644.1), AcNAC48 (KAK1301065.1), CnNAC48 (KAG1369931.1). (C) RT-qPCR was used to analyze the expression patterns of *PcNAC048*. (1) The expression analysis of *PcNAC048* in *P. cablin* L (leaf), YL (young leaf), S (stem) and R (root). The relative expression level was calculated according to the expression of *PcNAC048* in leaf. (2-4) The expression analysis of *PcNAC048* after treatment at 4 °C, 200 mM NaCl, and 20% PEG6000 for 0 h, 3 h, 6 h, and 12 h, respectively. The relative expression level was calculated according to the expression of *PcNAC048* at 0 h. n=3 biologically independent samples. (Ordinary one-way ANOVA, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$).

Fig. 5 Subcellular localization and transgenic analysis of *PcNAC048*. (A) Schematic diagram of recombinant vector construction for subcellular localization. The *PcNAC048* ORF without terminator fused to enhanced green fluorescent protein (EGFP). (B) Subcellular localization results of *PcNAC048* protein. Green fluorescence indicates the location of EGFP_{control} or *PcNAC048*-EGFP. From left to right, fluorescence channels, chloroplast channels, light field, and superimposed fields are shown. Scale bar, 10 μ m. (C) Schematic diagram of pBWA(V)BS-*PcNAC048* stable overexpression vector construction. (D) Growth of transgenic *Arabidopsis* on plant culture medium. (E) Expression level analysis of *PcNAC048* in transgenic lines. (F) Analysis of root length of transgenic lines. (G) Examination of the quantity of lateral roots in transgenic plants. n=3 biologically independent samples. (Ordinary one-way ANOVA, * $p < 0.05$, ** $p < 0.01$).

Fig. 6 PcNAC048 is a transcriptional inhibitor of *PcPTSP*ro (1865bp). (A) Y1H assays showed that PcNAC048 interacts with *PcPTSP*ro. The interaction between PcNAC048 and *PcPTSP*ro was indicated by the ability of yeast cells to grow on the SD lacking His, Leu, and Trp. pHis2-p53+pGAD53m was used as a positive control, while pHis2-PcPTSPro+pGADT7 was used as a negative control. (B) Schematic diagrams of the reporter and effector used in the dual-LUC (dual-luciferase) assay. REN, Renilla luciferase; PcPTSPro, the promoter of *PcPTS*. (C) Dual-LUC imaging analysis. The *Agrobacterium* liquid containing PcPTSPro::LUC+35S and PcPTSPro::LUC+35S::PcNAC048 was injected into the leaves at the distal axis of *Nicotiana benthamiana*, respectively. The redder the area, the stronger the fluorescence. (D) Relative LUC activity assay in *N. benthamiana* leaves. n=3 biologically independent samples. (Student's *t*-test, **** $p < 0.0001$).

Fig. 7 Transient overexpression analysis of *PcNAC048* in *P. cablin* leaves. (A) Schematic diagram of the pJLTRBO-PcNAC048 vector. (B-C) The relative expression levels of *PcNAC048* and *PcPTS* after transient overexpression for 4 days by RT-qPCR. (D) The patchouli alcohol content detected in CK and pJLTRBO-*PcNAC048* leaves by GC-MS. FW, fresh weight. (E) Chromatographic peaks of each sample. The blue peak represents patchouli alcohol standard, the purple peak represents pJLTRBO-*PcNAC048*, and the black peak represents CK. n=3 biologically independent samples. (Student's *t*-test, * $p < 0.05$).

Fig. 8 Virus induced *PcNAC048* silencing. (A) Schematic of pTRV2-*PcNAC048* vector. (B-C) The relative expression levels of *PcNAC048* and *PcPTS* after 3 weeks of virus induced gene silencing by RT-qPCR. (D) Patchouli alcohol content detected in CK and VIGS-*PcNAC048* leaves by GC-MS. (E) The chromatographic peaks of each sample. The blue peak represents the standard of patchouli alcohol, the purple peak represents VIGS-*PcNAC048*, and the black peak represents CK. n=3 biologically independent samples. (Student's *t*-test, * $p < 0.05$, *** $p < 0.001$).

References

1. Grogg, D. et al. Callus Induction from Diverse Explants and Genotypes Enables Robust Transformation of Perennial Ryegrass (*Lolium perenne* L.). *Plants (Basel)* **11**, (2022).
2. Hardtke, C.S. Phloem development. *New Phytol* **239**, 852-867 (2023).
3. Li, H. et al. Single-cell RNA sequencing reveals a high-resolution cell atlas of xylem in *Populus*. *J Integr Plant Biol* **63**, 1906-1921 (2021).
4. Bogdanović, M.D. et al. Secondary Somatic Embryogenesis in *Centaurea erythraea* Rafn. *Plants (Basel)* **10**, (2021).
5. Zhang, S. et al. Single-cell RNA sequencing analysis of the embryogenic callus clarifies the spatiotemporal developmental trajectories of the early somatic embryo in *Dimocarpus*

- longan. *Plant J* **115**, 1277-1297 (2023).
6. Yan, R. et al. Establishment of Efficient Genetic Transformation Systems and Application of CRISPR/Cas9 Genome Editing Technology in *Lilium pumilum* DC. Fisch. and *Lilium longiflorum* White Heaven. *Int J Mol Sci* **20**, (2019).
 7. Guo, Y. et al. Single-cell RNA sequencing reveals a high-resolution cell atlas of petals in *Prunus mume* at different flowering development stages. *Hortic Res* **11**, uhae189 (2024).
 8. Wang, Q. et al. Single-cell transcriptome atlas reveals developmental trajectories and a novel metabolic pathway of catechin esters in tea leaves. *Plant Biotechnol J* **20**, 2089-2106 (2022).
 9. Long, L. et al. Single-cell transcriptome atlas identified novel regulators for pigment gland morphogenesis in cotton. *Plant Biotechnol J* **21**, 1100-1102 (2023).
 10. Li, X. et al. Single-cell and spatial RNA sequencing reveal the spatiotemporal trajectories of fruit senescence. *Nat Commun* **15**, 3108 (2024).
 11. Sun, Y. et al. Single-cell transcriptomic analysis reveals the developmental trajectory and transcriptional regulatory networks of pigment glands in *Gossypium bickii*. *Mol Plant* **16**, 694-708 (2023).
 12. Tao, S. et al. Single-Cell Transcriptome and Network Analyses Unveil Key Transcription Factors Regulating Mesophyll Cell Development in Maize. *Genes (Basel)* **13**, (2022).
 13. Li, J. et al. miRNA164-directed cleavage of ZmNAC1 confers lateral root development in maize (*Zea mays*L.). *BMC Plant Biology* **12**, 220 (2012).
 14. Puranik, S., Sahu, P.P., Srivastava, P.S. & Prasad, M. NAC proteins: regulation and role in stress tolerance. *Trends Plant Sci* **17**, 369-81 (2012).
 15. Thakur, A. Therapeutic potential of Pogostemon cablin herb: a comprehensive review. *Pharm Pat Anal* **11**, 213-224 (2022).
 16. Zhang, C. et al. An integrated transcriptome and metabolome analysis reveals the gene network regulating flower development in *Pogostemon cablin*. *Front Plant Sci* **14**, 1201486 (2023).
 17. Fan, Y. et al. Inhibitory effects of Patchouli alcohol on the early lifecycle stages of influenza A virus. *Front Microbiol* **13**, 938868 (2022).
 18. Lee, H.S., Lee, J., Smolensky, D. & Lee, S.H. Potential benefits of patchouli alcohol in prevention of human diseases: A mechanistic review. *Int Immunopharmacol* **89**, 107056 (2020).
 19. Lee, J., Lee, H.S. & Lee, S.H. Preventive Activity of Patchouli Alcohol Against Colorectal Cancer and Diabetes. *J Med Food* **26**, 255-261 (2023).
 20. Orman-Ligeza, B. et al. The Xerobranching Response Represses Lateral Root Formation When Roots Are Not in Contact with Water. *Curr Biol* **28**, 3165-3173.e5 (2018).
 21. Péliissier, P.M., Motte, H. & Beeckman, T. Lateral root formation and nutrients: nitrogen in the spotlight. *Plant Physiol* **187**, 1104-1116 (2021).
 22. Peng, X. et al. PcCD2, PcGL2, and PcML1 Promote Patchouli Essential Oil Production by

- Regulating Glandular Trichome Development in *Pogostemon cablin*. *Physiologia Plantarum* **177**, e70466 (2025).
23. Fatima, S. et al. A Comprehensive Review on Pharmacological Activities of Pachypodol: A Bioactive Compound of an Aromatic Medicinal Plant *Pogostemon Cablin* Benth. *Molecules* **28**, (2023).
 24. Tang, Y., Zhong, L., Wang, X., Zheng, H. & Chen, L. Molecular identification and expression of sesquiterpene pathway genes responsible for patchoulol biosynthesis and regulation in *Pogostemon cablin*. *Bot Stud* **60**, 11 (2019).
 25. Chen, X. et al. Full-length transcriptome sequencing and methyl jasmonate-induced expression profile analysis of genes related to patchoulol biosynthesis and regulation in *Pogostemon cablin*. *BMC Plant Biol* **19**, 266 (2019).
 26. Liu, Z. et al. Identification of novel regulators required for early development of vein pattern in the cotyledons by single-cell RNA-sequencing. *Plant J* **110**, 7-22 (2022).
 27. Lopez-Anido, C.B. et al. Single-cell resolution of lineage trajectories in the Arabidopsis stomatal lineage and developing leaf. *Dev Cell* **56**, 1043-1055.e4 (2021).
 28. Sun, X. et al. Single-cell transcriptome reveals dominant subgenome expression and transcriptional response to heat stress in Chinese cabbage. *Genome Biol* **23**, 262 (2022).
 29. Liu, H. et al. Single-cell RNA-seq describes the transcriptome landscape and identifies critical transcription factors in the leaf blade of the allotetraploid peanut (*Arachis hypogaea* L.). *Plant Biotechnol J* **19**, 2261-2276 (2021).
 30. Singh, S., Grover, A. & Nasim, M. Biofuel Potential of Plants Transformed Genetically with NAC Family Genes. *Front Plant Sci* **7**, 22 (2016).
 31. Wang, Z. et al. Genome-Wide Identification and Characterization of NAC Family in *Hibiscus hamabo* Sieb. et Zucc. under Various Abiotic Stresses. *Int J Mol Sci* **23**, (2022).
 32. Ma, J., Wang, L.Y., Dai, J.X., Wang, Y. & Lin, D. The NAC-type transcription factor CaNAC46 regulates the salt and drought tolerance of transgenic *Arabidopsis thaliana*. *BMC Plant Biol* **21**, 11 (2021).
 33. Shu, L., Li, L., Jiang, Y.Q. & Yan, J. Advances in membrane-tethered NAC transcription factors in plants. *Plant Sci* **342**, 112034 (2024).
 34. Singh, S., Koyama, H., Bhati, K.K. & Alok, A. The biotechnological importance of the plant-specific NAC transcription factor family in crop improvement. *J Plant Res* **134**, 475-495 (2021).
 35. Li, J. et al. miRNA164-directed cleavage of ZmNAC1 confers lateral root development in maize (*Zea mays* L.). *BMC Plant Biol* **12**, 220 (2012).
 36. Yarra, R. & Wei, W. The NAC-type transcription factor GmNAC20 improves cold, salinity tolerance, and lateral root formation in transgenic rice plants. *Funct Integr Genomics* **21**, 473-487 (2021).
 37. Xu, P., Ma, W., Hu, J. & Cai, W. The nitrate-inducible NAC transcription factor NAC056 controls nitrate assimilation and promotes lateral root growth in *Arabidopsis thaliana*. *PLoS*

- Genet* **18**, e1010090 (2022).
38. Swamy, M.K. & Sinniah, U.R. A Comprehensive Review on the Phytochemical Constituents and Pharmacological Activities of *Pogostemon cablin* Benth.: An Aromatic Medicinal Plant of Industrial Importance. *Molecules* **20**, 8521-47 (2015).
 39. Ma, B. et al. Significantly Enhanced Production of Patchoulol in Metabolically Engineered *Saccharomyces cerevisiae*. *J Agric Food Chem* **67**, 8590-8598 (2019).
 40. Chen, X. et al. PatDREB Transcription Factor Activates Patchoulol Synthase Gene Promoter and Positively Regulates Jasmonate-Induced Patchoulol Biosynthesis. *J Agric Food Chem* **70**, 7188-7201 (2022).
 41. Chen, X. et al. PatSWC4, a methyl jasmonate-responsive MYB (v-myb avian myeloblastosis viral oncogene homolog)-related transcription factor, positively regulates patchoulol biosynthesis in *Pogostemon cablin*. *Industrial Crops and Products* **154**, 112672 (2020).
 42. Wang, X. et al. Functional analysis of *Pogostemon cablin* farnesyl pyrophosphate synthase gene and its binding transcription factor PcWRKY44 in regulating biosynthesis of patchouli alcohol. *Front Plant Sci* **13**, 946629 (2022).
 43. Han, K., Zhao, Y., Sun, Y. & Li, Y. NACs, generalist in plant life. *Plant Biotechnol J* **21**, 2433-2457 (2023).
 44. Wang, X. et al. PatJAZ6 Acts as a Repressor Regulating JA-Induced Biosynthesis of Patchouli Alcohol in *Pogostemon Cablin*. *Int J Mol Sci* **20**, (2019).
 45. Chen, S., Wang, X., Zheng, H. & Chen, L. Application of virus-induced gene silencing in *Pogostemon cablin* advances our understanding of patchoulol and sesquiterpene metabolism. *The Journal of Horticultural Science and Biotechnology* **98**, 322-333 (2023).
 46. Wang, H. et al. WOX family transcriptional regulators modulate cytokinin homeostasis during leaf blade development in *Medicago truncatula* and *Nicotiana glauca*. *The Plant Cell* **34**, 3737-3753 (2022).
 47. Shen, Y. et al. Chromosome-level and haplotype-resolved genome provides insight into the tetraploid hybrid origin of patchouli. *Nature Communications* **13**, 3511 (2022).
 48. Kim, J.Y. et al. Distinct identities of leaf phloem cells revealed by single cell transcriptomics. *Plant Cell* **33**, 511-530 (2021).
 49. Liu, Z. et al. Global Dynamic Molecular Profiling of Stomatal Lineage Cell Development by Single-Cell RNA Sequencing. *Mol Plant* **13**, 1178-1193 (2020).
 50. Abe, M., Takahashi, T. & Komeda, Y. Identification of a cis-regulatory element for L1 layer-specific gene expression, which is targeted by an L1-specific homeodomain protein. *Plant J* **26**, 487-94 (2001).
 51. Zhang, T.Q., Xu, Z.G., Shang, G.D. & Wang, J.W. A Single-Cell RNA Sequencing Profiles the Developmental Landscape of *Arabidopsis* Root. *Mol Plant* **12**, 648-660 (2019).
 52. Potocka, I., Baldwin, T.C. & Kurczynska, E.U. Distribution of lipid transfer protein 1 (LTP1) epitopes associated with morphogenic events during somatic embryogenesis of

- Arabidopsis thaliana*. *Plant Cell Rep* **31**, 2031-45 (2012).
53. Guo, X. et al. Single-cell transcriptome reveals differentiation between adaxial and abaxial mesophyll cells in *Brassica rapa*. *Plant Biotechnol J* **20**, 2233-2235 (2022).
 54. Satterlee, J.W., Strable, J. & Scanlon, M.J. Plant stem-cell organization and differentiation at single-cell resolution. *Proc Natl Acad Sci U S A* **117**, 33689-33699 (2020).
 55. Denyer, T. et al. Spatiotemporal Developmental Trajectories in the *Arabidopsis* Root Revealed Using High-Throughput Single-Cell RNA Sequencing. *Dev Cell* **48**, 840-852 e5 (2019).
 56. Farmer, A., Thibivilliers, S., Ryu, K.H., Schiefelbein, J. & Libault, M. Single-nucleus RNA and ATAC sequencing reveals the impact of chromatin accessibility on gene expression in *Arabidopsis* roots at the single-cell level. *Mol Plant* **14**, 372-383 (2021).
 57. Shulse, C.N. et al. High-Throughput Single-Cell Transcriptome Profiling of Plant Cell Types. *Cell Rep* **27**, 2241-2247 e4 (2019).
 58. Jean-Baptiste, K. et al. Dynamics of Gene Expression in Single Root Cells of *Arabidopsis thaliana*. *Plant Cell* **31**, 993-1011 (2019).
 59. Peña, M.J. et al. *Arabidopsis* irregular xylem8 and irregular xylem9: Implications for the Complexity of Glucuronoxylan Biosynthesis. *The Plant Cell* **19**, 549-563 (2007).
 60. Livak, K.J. & Schmittgen, T.D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**, 402-8 (2001).
-

Editor's Summary

In patchouli (*Pogostemon cablin*), anther-derived somatic embryos were profiled by single-cell RNA-seq to define eight cell types, identifying PcNAC048 that drives embryonic cell differentiation, promotes lateral root development in Arabidopsis, and negatively regulates patchouli alcohol biosynthesis.

Peer review information:

Communications Biology thanks the anonymous reviewers for their contribution to the peer review of this work. Primary Handling Editors: Shahid Mukhtar and David Favero. A peer review file is available.

ARTICLE IN PRESS

A



Flowering period of
Pogostemon cablin

B

ARTICLE IN PRESS



A single flower of
P. cablin



Anatomy of a single
flower of *P. cablin*

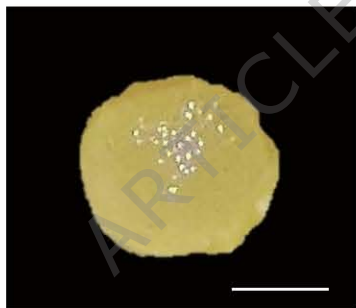


The anthers of
P. cablin



Development period of
microspores in *P. cablin*

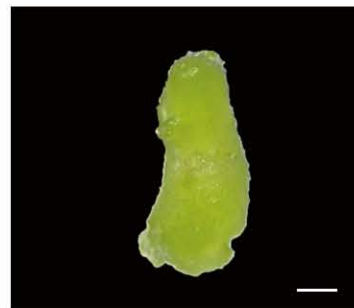
C



Globular somatic
embryo (GSE)



Heart-shaped
somatic embryo

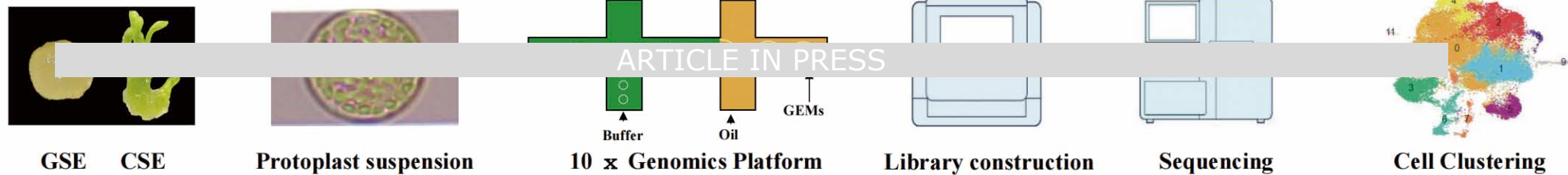


Torpedo-shaped
somatic embryo

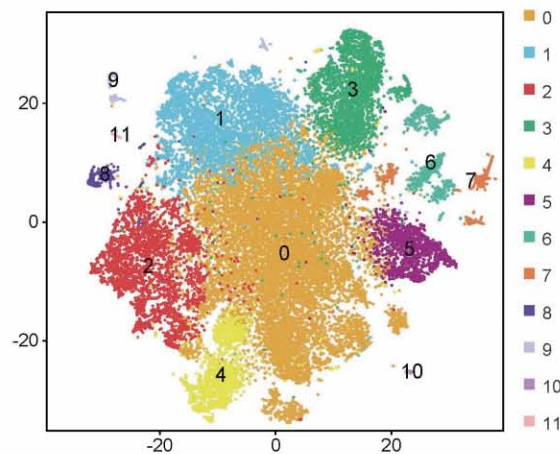


Cotyledonary somatic
embryo (CSE)

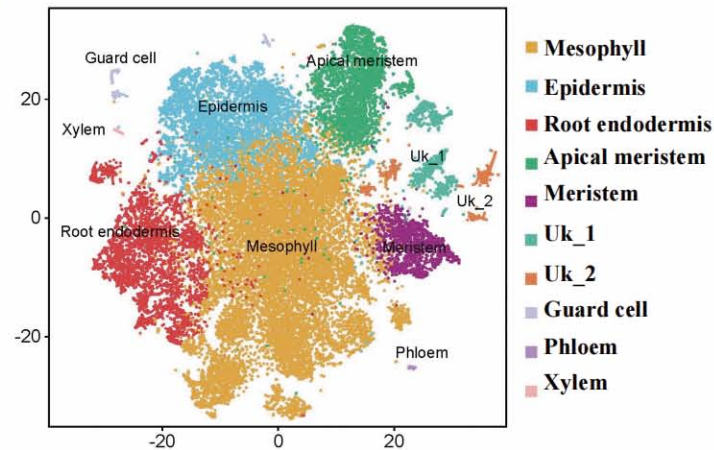
A



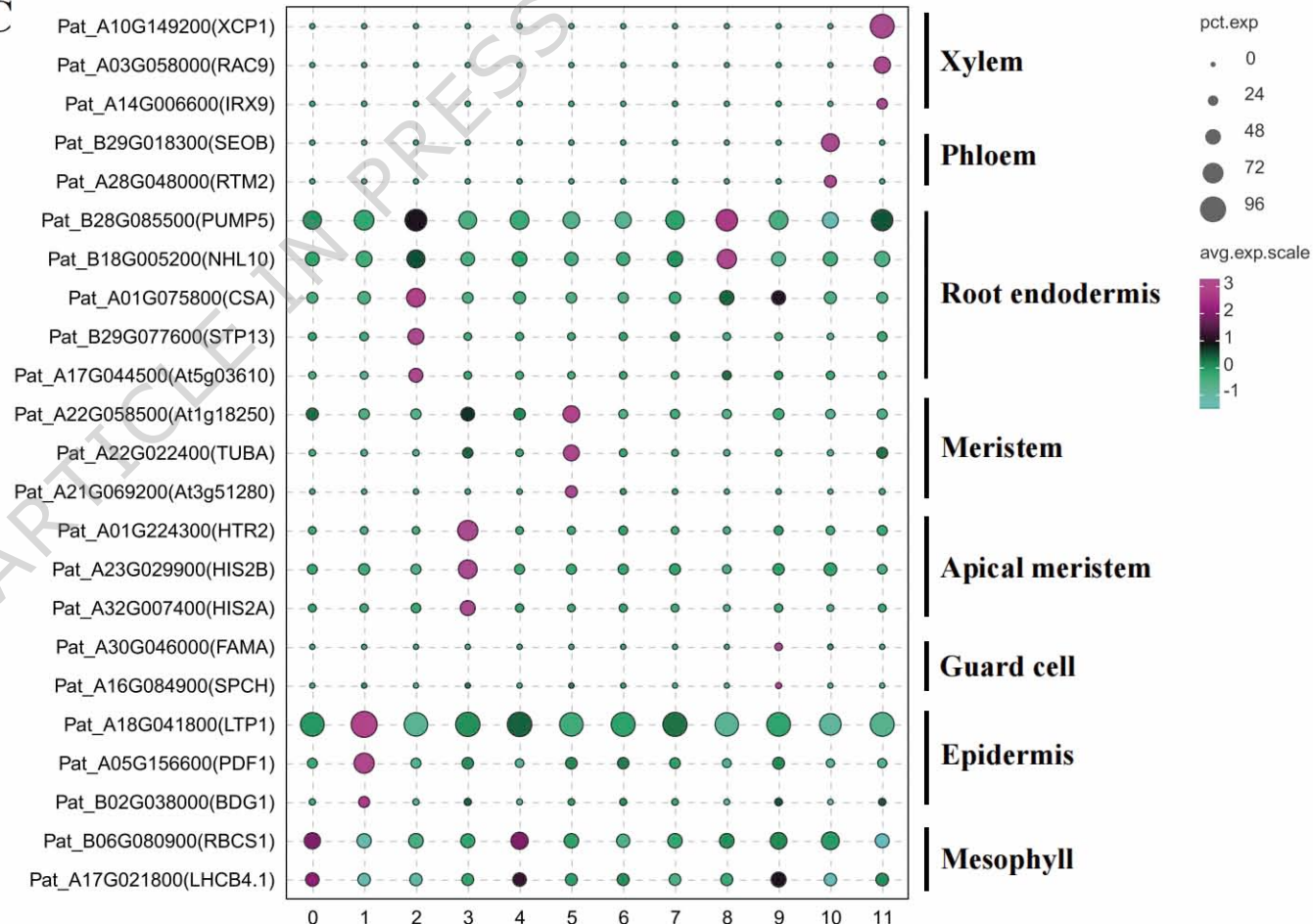
B

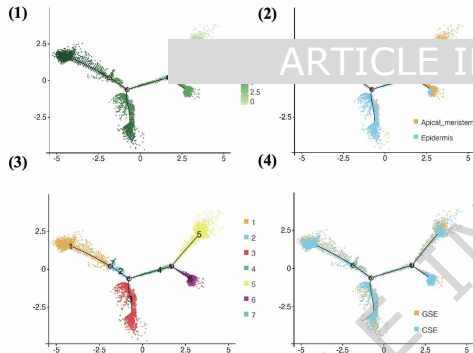
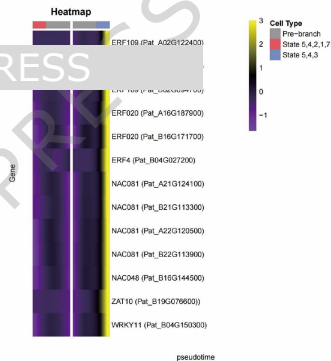
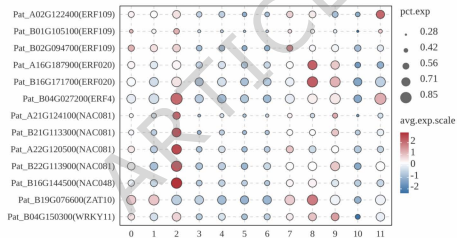
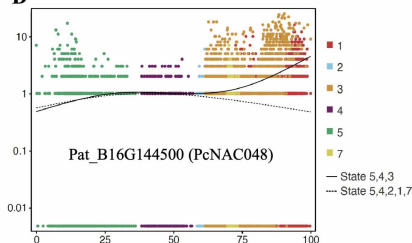


D

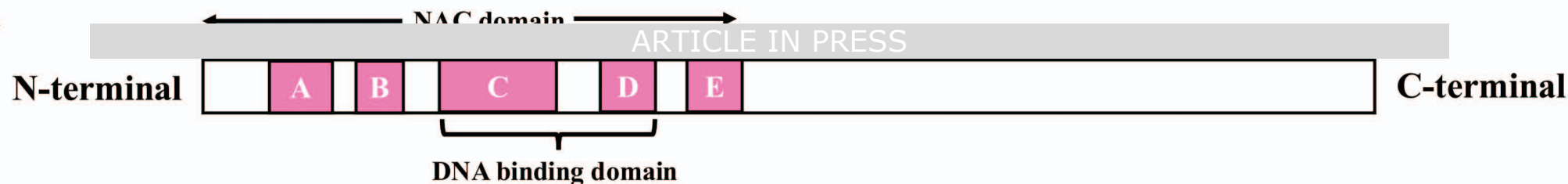


C

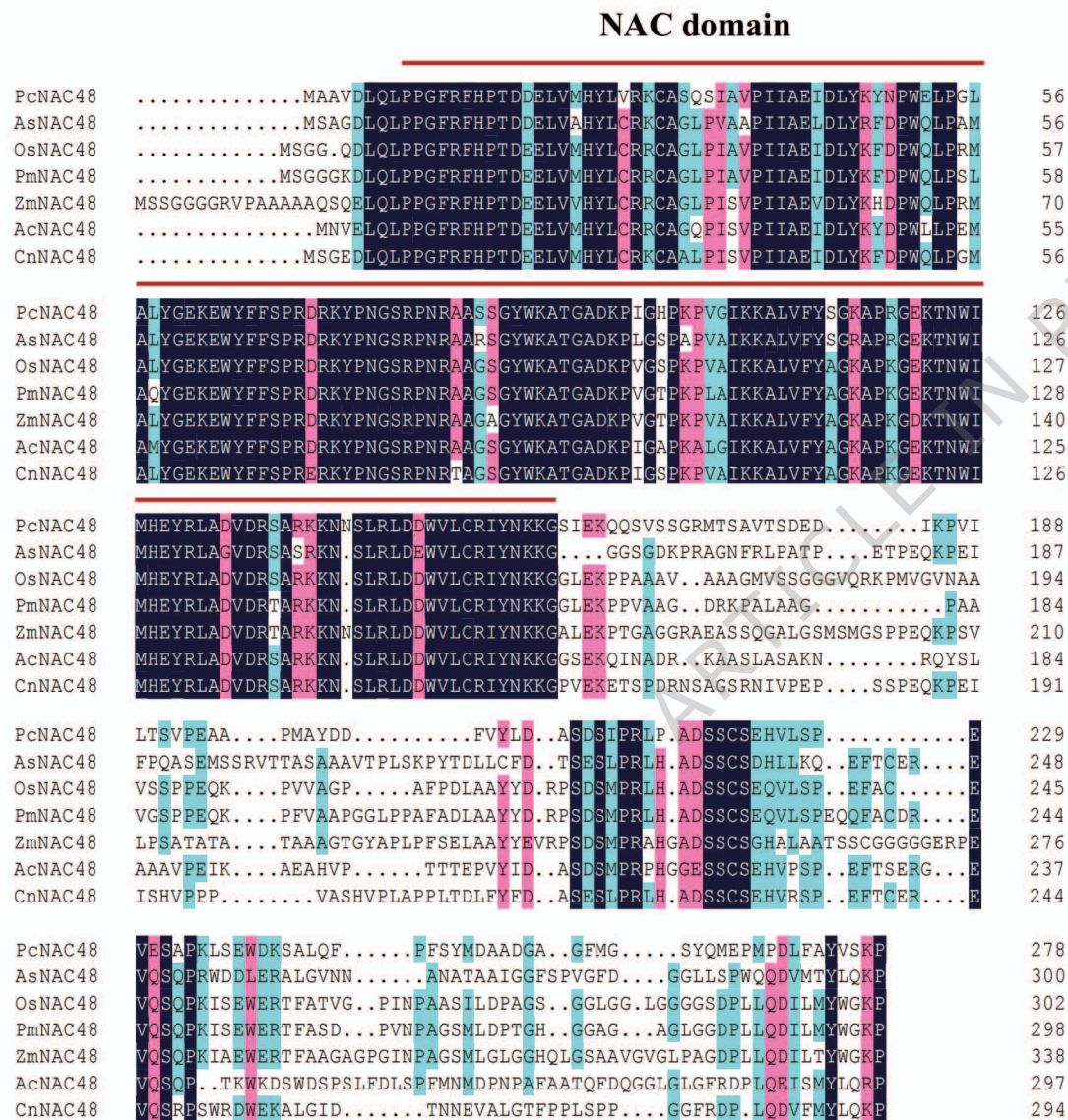


A**B****C****D**

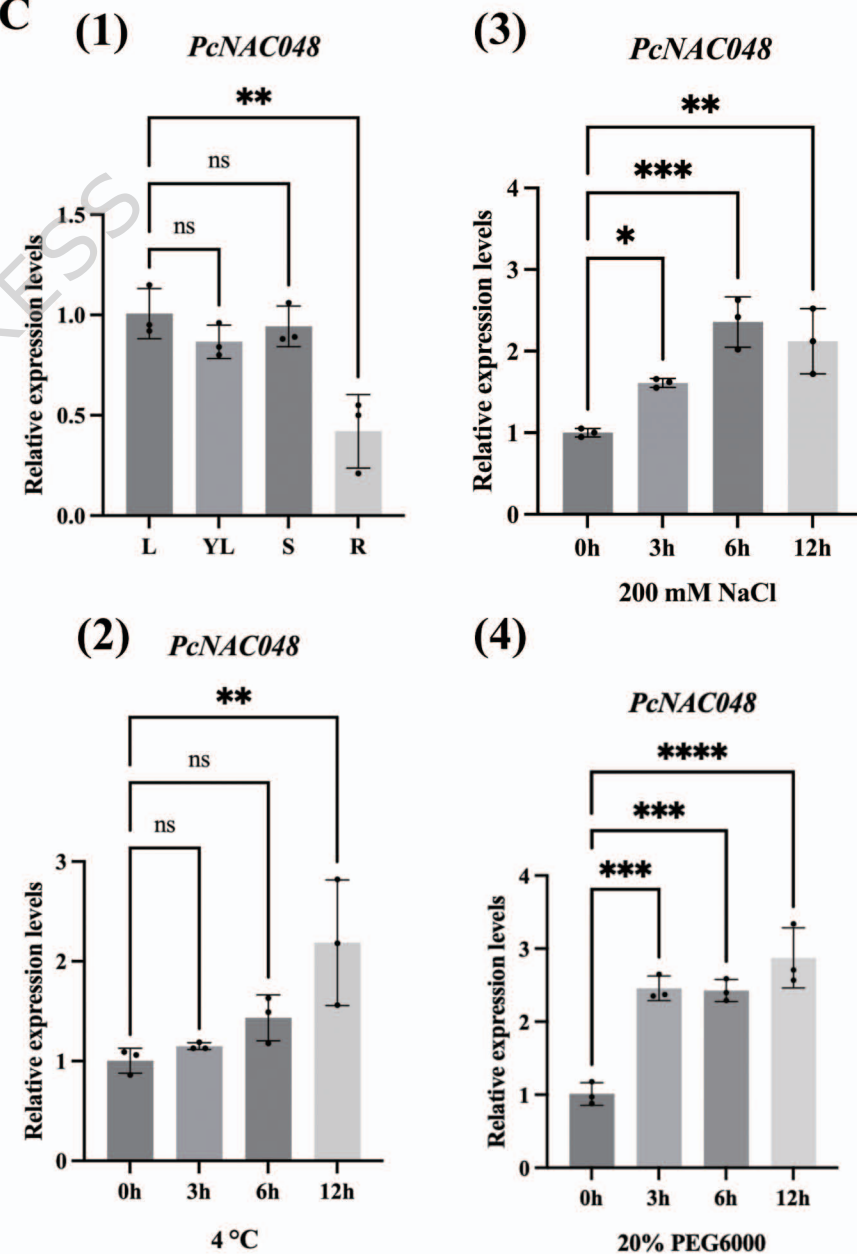
A



B



C



A

PcNAC048-EGFP

u535

ARTICLE IN PRESS

PcNAC048

EGFP

nos

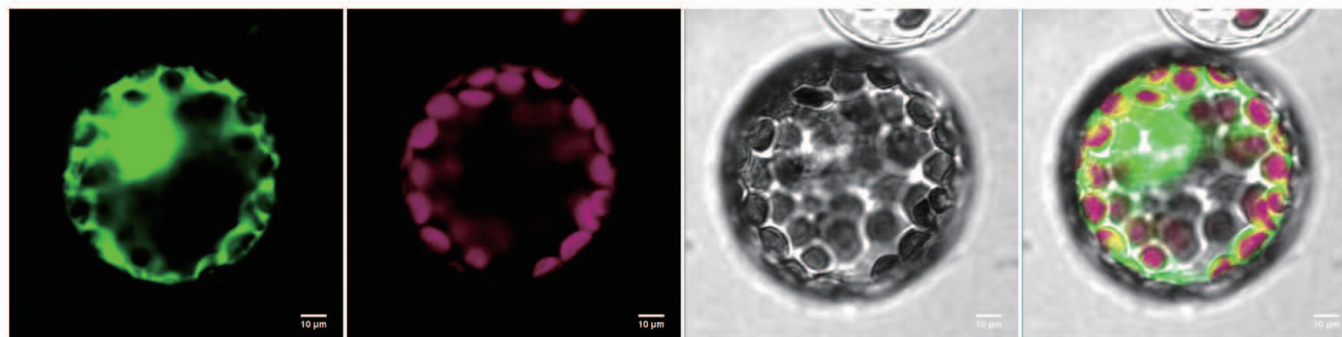
B

GFP

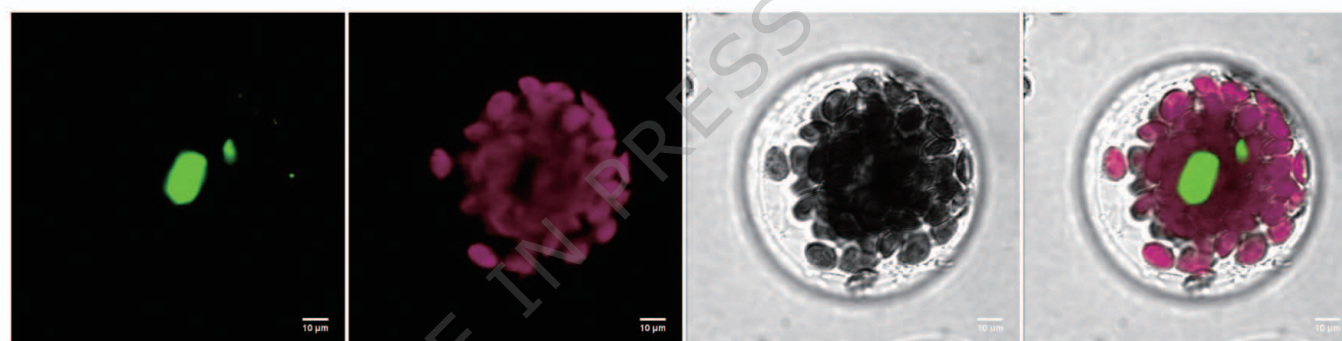
Chlorophyll

Light field

Merged

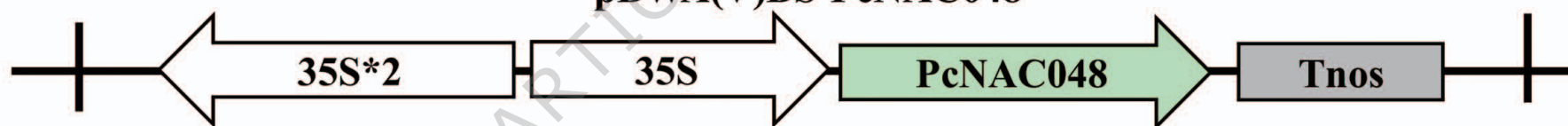
EGFP_{control}

PcNAC048-EGFP



C

pBWA(V)BS-PcNAC048



D

WT

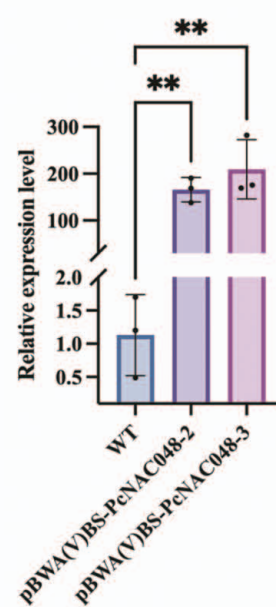
#048-1

#048-2

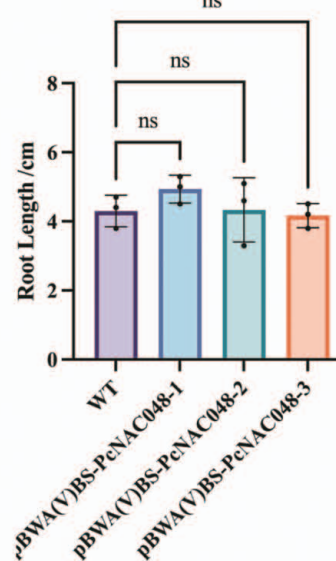
#048-3



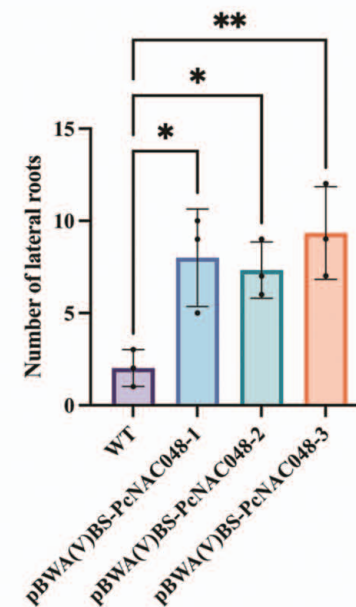
E

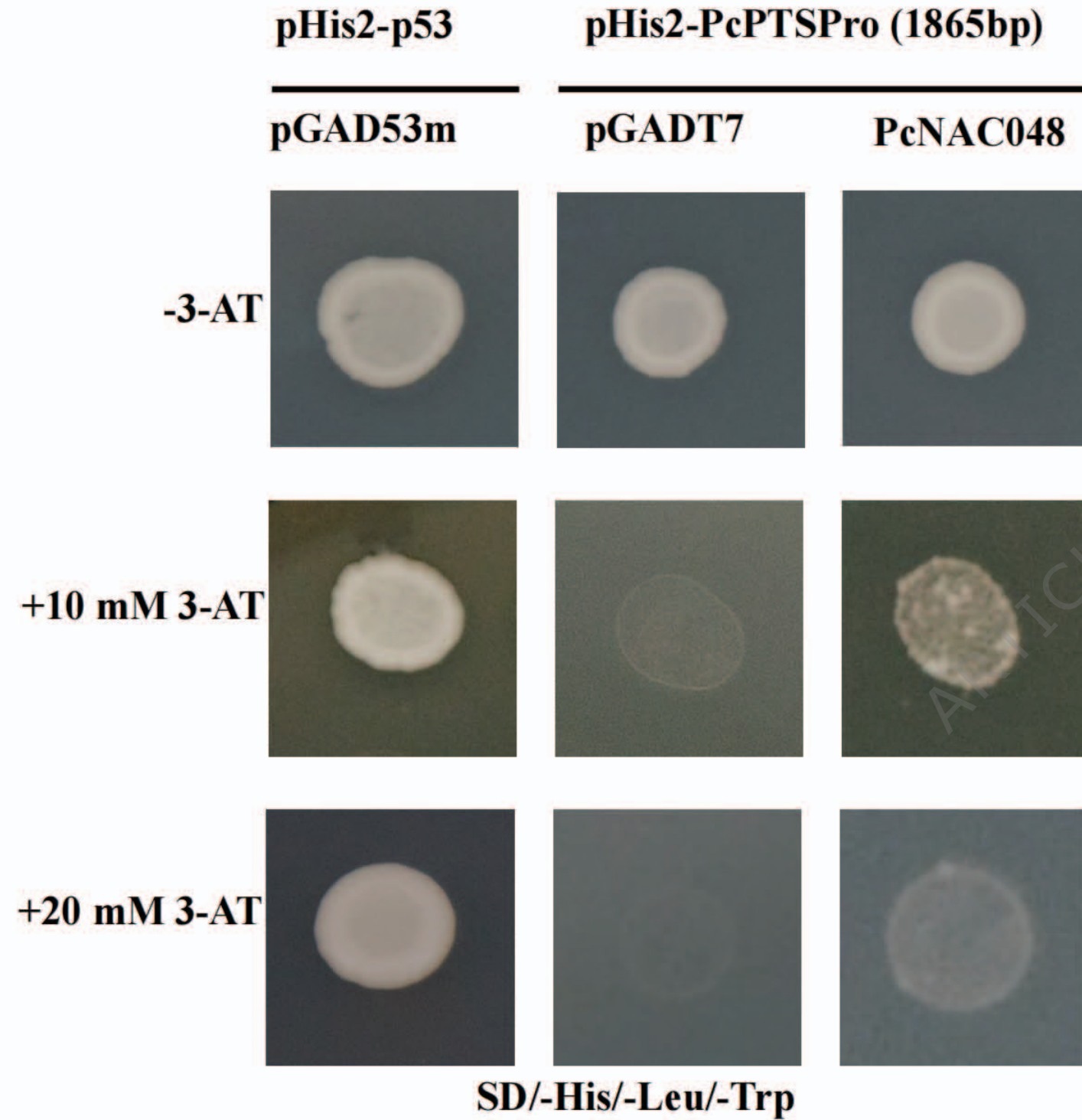
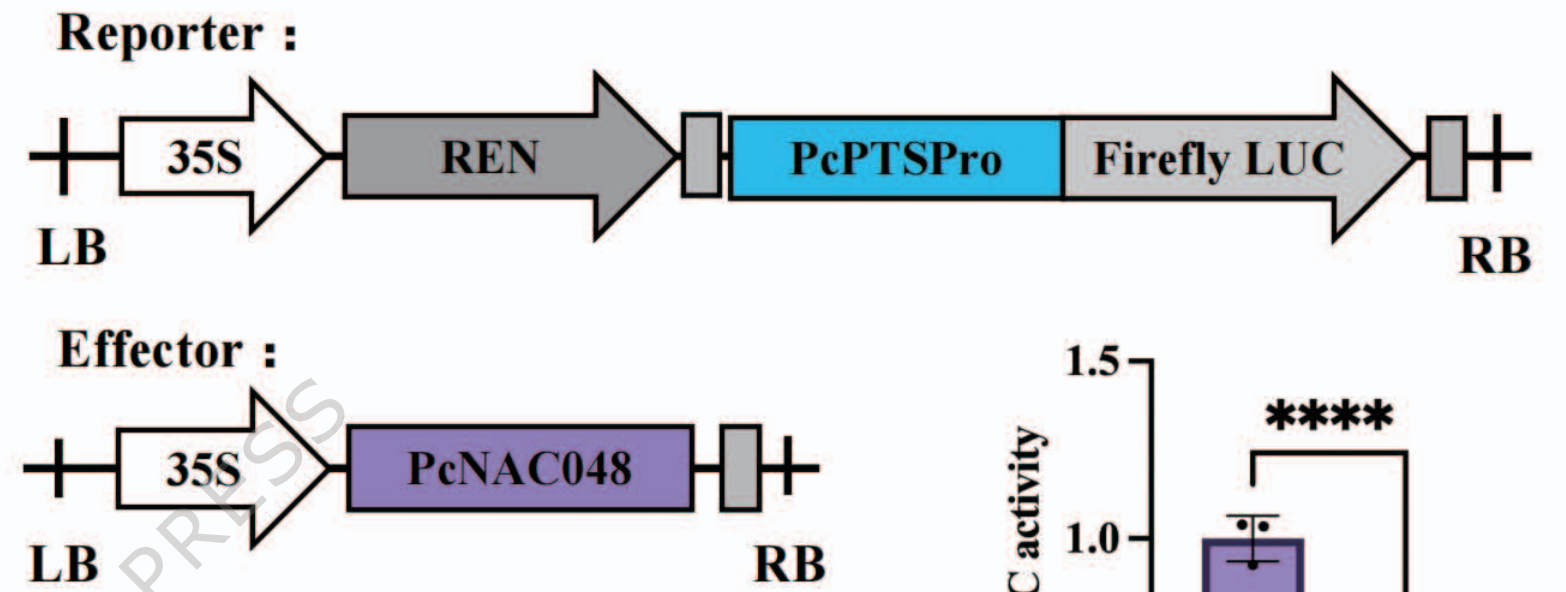
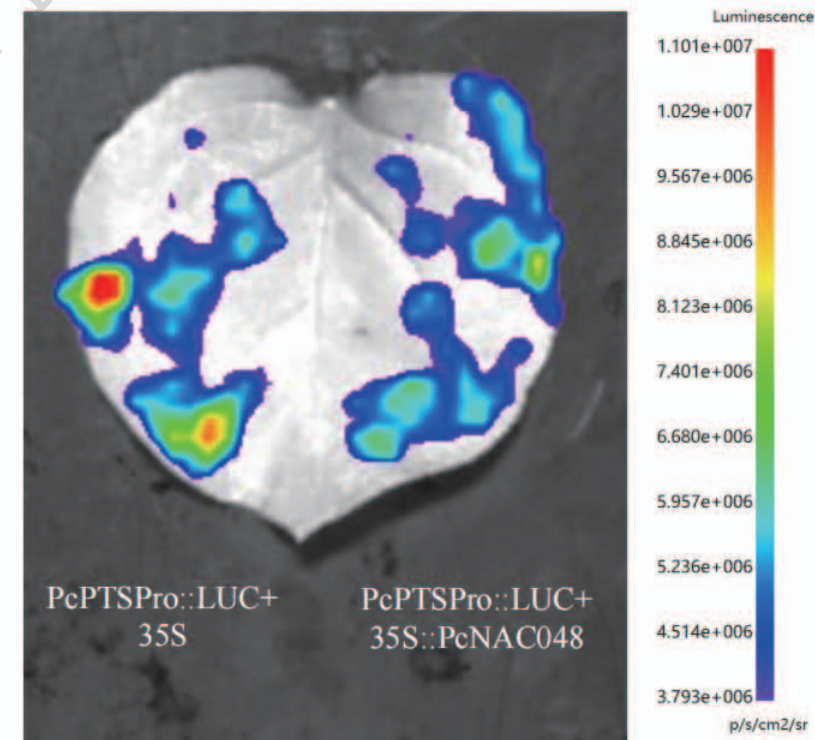
PcNAC048

F



G



A**B****C****D**