ANALYSIS OF EXPRESSION PATTERNS OF BRASSINOSTEROID-RESPONSIVE GENES
IN *ARABIDOPSIS* ROOT USING SINGLE-CELL RNA SEQUENCING DATASETS

Miss Thanaporn Wongkham

A Thesis Submitted in Partial Fulfillment of the Requirements

for the Degree of Master of Science in Bioinformatics and Computational Biology (Interdisciplinary

Program)

Inter-Department of Bioinformatics and Computational Biology

GRADUATE SCHOOL

Chulalongkorn University

Academic Year 2022

การวิเคราะห์รูปแบบการแสดงออกของยีนที่ตอบสนองต่อบราสซิโนสเตียรอยด์ในรากอะราบิดอป
ซิสโดยใช้ชุดข้อมูลลำดับเบสอาร์เอ็นเอระดับเซลล์เดี่ยว

น.ส.ธนภรณ์ วงศ์คำ

Thesis Title   ANALYSIS OF EXPRESSION PATTERNS OF BRASSINOSTEROID-

RESPONSIVE GENES IN *ARABIDOPSIS* ROOT USING SINGLE-

CELL RNA SEQUENCING DATASETS

By   Miss Thanaporn Wongkham

Field of Study   Bioinformatics and Computational Biology (Interdisciplinary

Program)

Thesis Advisor   SIRA SRISWASDI, Ph.D.

Thesis Co Advisor   Assistant Professor JUTHAMAS CHAIWANON, Ph.D.

Accepted by the GRADUATE SCHOOL, Chulalongkorn University in Partial Fulfillment of the Requirement for the Master of Science

                                   Dean of the GRADUATE SCHOOL

(Associate Professor YOOTTHANA CHUPPUNNARAT, Ph.D.)

THESIS COMMITTEE

                                   Chairman

(Professor SUPACHITRA CHADCHAWAN, Ph.D.)

                                   Thesis Advisor

(SIRA SRISWASDI, Ph.D.)

                                   Thesis Co-Advisor

(Assistant Professor JUTHAMAS CHAIWANON, Ph.D.)

                                   Examiner

(Associate Professor TEERAPONG BUABOOCHA, Ph.D.)

                                   External Examiner

(Associate Professor Varodom Charoensawan, Ph.D.)

ธนภรณ์ วงศ์คำ : การวิเคราะห์รูปแบบการแสดงออกของยีนที่ตอบสนองต่อบราสซิโนสเตียรอยด์ในรากอะราบิดอปซิสโดยใช้ชุดข้อมูลลำดับเบสอาร์เอ็นเอระดับเซลล์เดี่ยว. (ANALYSIS OF EXPRESSION PATTERNS OF BRASSINOSTEROID-RESPONSIVE GENES IN *ARABIDOPSIS* ROOT USING SINGLE-CELL RNA SEQUENCING DATASETS) อ.ที่ปรึกษาหลัก : อ. ดร.สิระ ศรีสวัสดิ์, อ.ที่ปรึกษาร่วม : ผศ. ดร.จุฑามาศ ชัยวนนท์

บราสิโนสเตียรอยด์ (BR) มีบทบาทซับซ้อนต่อการเจริญของราก โดยควบคุมการแสดงออกของยีนเป้าหมายผ่านเส้นทางการส่งสัญญาณ ข้อมูลลำดับเบสอาร์เอ็นเอระดับเซลล์เดี่ยวของรากอะราบิดอปซิส เผยให้เห็นถึงบทบาทเฉพาะของบราสิโนสเตียรอยด์ต่อเนื้อเยื่อและระยะการเจริญของราก เอทริโคบลาสต์ซึ่งเป็นเอพิเดอร์มิสที่จะเจริญไปเป็นเซลล์ไม่มีขน มียีนที่ถูกเหนี่ยวนำโดยบราสิโนสเตียรอยด์เป็นสัดส่วนสูงสุดที่บริเวณยืดตามยาว และบริเวณการเจริญของเซลล์ การวิเคราะห์วิถีแสดงให้เห็นว่า การแสดงออกของยีนที่สร้างเอนไซม์ในวิถีสังเคราะห์บราสิโนสเตียรอยด์แตกต่างไปตามประเภทเซลล์ โดยพบว่าการแสดงออกของยีนที่สร้างเอนไซม์ *DET2* และ *ROT3* และยีนในวิถีการส่งสัญญาณของสังเคราะห์บราสิโนสเตียรอยด์ *BRI1* และ *BSK1* นั้น มีการแสดงออกในเอทริโคบลาสต์ที่บริเวณยืดตามยาว และบริเวณการเจริญของเซลล์ สูงกว่าการแสดงออกในไตรโคบลาสต์ จึงสรุปได้ว่า เอทริโครบลาสต์อาจมีปริมาณของฮอร์โมนและกิจกรรมของวิถีการส่งสัญญาณบราสิโนสเตียรอยด์สูงกว่าในไตรโคบลาสต์ จึงทำให้มีสัดส่วนของยีนที่กระตุ้นการแสดงออกด้วยบราสิโนสเตียรอยด์ในเอทริโคบลาสต์สูงกว่าในไตรโคบลาสต์ด้วย การค้นหาหน้าที่และความสัมพันธ์ของกลุ่มยีน เผยให้เห็นถึงความแตกต่างของชุดยีนที่ตอบสนองต่อบราสิโนสเตียรอยด์ ที่เกี่ยวข้องกับกระบวนการของผนังเซลล์ระหว่างเซลล์ทั้งสองประเภท โดยกลุ่มยีน *EXTs* ซึ่งเป็นยีนที่ยับยั้งการแสดงออกด้วยบราสิโนสเตียรอยด์ มีการแสดงออกเฉพาะในไตรโคบลาสต์ที่ระยะการเจริญของเซลล์ ในขณะที่กลุ่มยีน *XTHs* ซึ่งเป็นยีนที่กระตุ้นการแสดงออกด้วยบราสิโนสเตียรอยด์ มีการแสดงออกที่สูงขึ้นในเอทริโคบลาสต์ที่บริเวณยืดตามยาว และบริเวณการเจริญของเซลล์ นอกจากนี้การตรวจสอบความถูกต้องของรูปแบบการแสดงออกของยีน ที่เกี่ยวข้องกับลักษณะที่สนใจด้วยชุดข้อมูลระดับเซลล์เดี่ยวเพิ่มเติม มีความสอดคล้องกับการค้นพบก่อนหน้า และยังเป็นการเน้นถึงความสำคัญของยีนที่เกี่ยวข้องกับผนังเซลล์ที่เฉพาะเจาะจงในการกำหนดชะตาของเอทริโคบลาสต์

สาขาวิชา	ชีวสารสนเทศศาสตร์และชีววิทยาเชิงคอมพิวเตอร์ (สหสาขาวิชา)	ลายมือชื่อนิสิต ....
ปีการศึกษา	2565	ลายมือชื่อ อ.ที่ปรึกษาหลัก ....
		ลายมือชื่อ อ.ที่ปรึกษาร่วม ....

6380142820 : MAJOR BIOINFORMATICS AND COMPUTATIONAL BIOLOGY (INTERDISCIPLINARY PROGRAM)

KEYWORD: Root Development, Epidermis, Brassinosteroid, Single Cell RNA-Seq, Arabidopsis

Thanaporn Wongkham : ANALYSIS OF EXPRESSION PATTERNS OF BRASSINOSTEROID-RESPONSIVE GENES IN *ARABIDOPSIS* ROOT USING SINGLE-CELL RNA SEQUENCING DATASETS. Advisor: SIRA SRISWASDI, Ph.D. Co-advisor: Asst. Prof. JUTHAMAS CHAIWANON, Ph.D.

Brassinosteroid (BR) plays a complex role in root growth and development by regulating the expression of downstream target genes through the BR signaling pathway. In this scRNA-seq study of *Arabidopsis* roots, tissue- and developmental-specific roles of BR were identified. Atrichoblast, an epidermal cell that develops into a non-hair cell, displayed the highest enrichment of BR-induced genes in the elongation and maturation zones among all cell types. Trajectory analysis revealed distinct expression patterns of BR biosynthesis and signaling genes between atrichoblasts and trichoblasts. Atrichoblasts exhibited higher expression levels of BR biosynthesis genes (*DET2* and *ROT3*) and BR signaling genes (*BRI1* and *BSK1*) during elongation and maturation stages compared to trichoblasts. This suggests that the higher levels of BR and BR signaling activity in atrichoblasts likely contribute to their larger percentages of BR-induced gene expression. Gene ontology analysis further revealed differences in BR-responsive gene sets related to cell wall processes between the two cell types. Specifically, *EXTs* that are BR-repressed genes showed specific expression in trichoblasts at the maturation stage, while *XTHs* that are BR-induced genes displayed higher expression in atrichoblasts during the elongation and maturation stages. Additionally, the validation of candidate genes on additional single-cell datasets supported the consistency of the findings and highlighted the significance of specific cell wall-related genes in determining atrichoblast fate.

Field of Study:	Bioinformatics and Computational Biology (Interdisciplinary Program)	Student's Signature ....
Academic Year:	2022	Advisor's Signature ....
		Co-advisor's Signature ....

v

ACKNOWLEDGEMENTS

I would like to express my heartfelt appreciation to my thesis advisors, Dr. Sira Sriswasdi and Assistant Professor Dr. Juthamas Chaiwanon, for their invaluable assistance with my thesis. Their advice, mentorship, and dedicated efforts have significantly influenced the progress of my research and motivated me to strive for excellence.

I am grateful to the members of the thesis committee, consisting of Professor Dr. Supachitra Chadchawan, Associate Professor Dr. Teerapong Buaboocha, and Associate Professor Dr. Varodom Charoensawan, for their contributions and guidance. The feedback and suggestions substantially enhanced the quality of the thesis.

I would like to express my deepest gratitude to lecturers in the Bioinformatics and Computational Biology program for sharing their substantial knowledge over the course. I appreciate my classmates in the program for their assistance in coaching me in programming. The opportunities for collaborative training have been essential to self-improvement.

I would like to convey my sincere gratitude to the Development and Promotion of Science and Technology Talents Project (DPST) scholarship for believing in my potential and providing substantial financial support throughout my education.

Lastly, I am grateful to my family for their encouragement and support as I pursue this master's degree.

Thanaporn Wongkham

TABLE OF CONTENTS

จุฬาลงกรณ์มหาวิทยาลัย

CHULALONGKORN UNIVERSITY

LIST OF FIGURES

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

PART 1

INTRODUCTION

1.1 Background and motivation

Steroids are chemical substances that include various critical metabolism-regulating hormones. They have been discovered in plants, mammals, and fungi, suggesting that their functional roles originated very early in the evolution of life (Ferreira-Guerra et al., 2020) Plant hormones contribute to adaptability of plant growth and development. Primary regulator hormones include auxins, cytokinins (CK), gibberellins (GA), strigolactones (SL), brassinosteroids (BR), and polyamines. Other hormones, such as abscisic acid (ABA), ethylene, salicylate (SA), and jasmonates (JA), have also been associated with stress responses (Munné-Bosch & Müller, 2013). Brassinosteroid (BR) is a major steroid hormone in plants. BR binds to BRASSINOSTEROID-INSENSITIVE (BRI1), a plasma membrane-localized receptor kinase that subsequently initiates a cascade of signaling transductions (Wang et al., 2012). When BR levels increase (Planas-Riverola et al., 2019), the glycogen synthase kinase3 (GSK3)-like kinase BRASSINOSTEROID INSENSITIVE2 (BIN2) is inactivated. This is followed by rapid dephosphorylation of the BRASSINAZOLE RESISTANT 1 (BZR1) family transcription factors by protein phosphatase 2A (PP2A) (He et al., 2005) which results in nuclear accumulation of unphosphorylated BZR1 and modulation of thousands of BR target genes (He et al., 2005).

Studies of *Arabidopsis* thaliana as a model system had led to significant breakthroughs in comprehending how plants perceive and respond to BR signaling (Vert et al., 2005). Previous research suggested that BRs promote root (Müssig et al., 2003) and shoot cell elongation (Yin et al., 2002), cell division in the quiescent center (QC) (González-García et al., 2011), and flowering (Li & He, 2020) in *Arabidopsis*. The roots of *Arabidopsis* are often examined in research problems because they are an easy-to-understand model. Their structure and classification of cell types are straightforward, especially the segmentation of root tip developmental zones. Roots have different developmental zones, including the meristem, elongation, and maturation zones, which

are arranged along the proximal-distal axis of the root. The root meristem zone undergoes repeated cell divisions. At the elongation zone, cellular proliferation ceases and there is a rapid anisotropic elongation of cells. Upon reaching the maturation zone, cells attain their final size and begin differentiation into specialized roles (Cajero Sánchez et al., 2018). Roots are made up of concentric layers of vascular tissue, including xylem and phloem, as well as the epidermis, cortex, and endodermis. The maintenance of the stem cell niche is facilitated by the quiescent center located at the root tip. This serves an essential function in regulating the fate of individual cells and cell types retain their distinct identities throughout their lifespan (Rich-Griffin et al., 2020). The root cap, comprising the columella and lateral root caps, encircles the root meristem and protects the stem cell niche (Huysmans et al., 2018).

Roots respond to BR hormones in gradients along the developmental zone in a tissue/zone-specific manner (Chaiwanon & Wang, 2015; Fridman et al., 2014; Graeff et al., 2020; Hacham et al., 2011; Vragovíć et al., 2015). BR signaling is required for epidermis but not in the inner cell files to modulate the meristem size (Hacham et al., 2011). According to cell-type-specific translatome data (Vragovíć et al., 2015), BR-responsive genes have distinct functions between epidermal and inner vascular cells. The outer cells (epidermis) of the transition zone are highly sensitive to BR and express high levels of BR-induced genes. Meanwhile, inner cells (stele) of apical meristem express BR-repressed genes (Vragovíć et al., 2015). Furthermore, BR signaling has diverse impacts even among differentiated epidermal cell types (trichoblast and atrichoblast). While BRI1 promotes cell elongation in trichoblasts, the expression of BRI1 in atrichoblasts increases ethylene production which prevents cell elongation via crystalline cellulose accumulation (Fridman et al., 2014).

Previous studies have evaluated the involvement of BRs in epidermal cell fate determination (Cheng et al., 2014; Kuppusamy et al., 2009). During epidermal cell differentiation, cells can differentiate into hair or non-hair cells, which are also known as trichoblast and atrichoblast, respectively. BR signaling influences the early stage of

epidermal cell fate by modulating the activity of WEREWOLF (WER), CAPRICE (CPC), and GLABRA 2 (GL2) to promote atrichoblast development (Kuppusamy et al., 2009).

Single-cell RNA sequencing (scRNA-seq) is an advanced sequencing technology that focuses on characterizing the gene expression patterns of individual cells. This approach identifies uncommon cellular types and tracks the development of diverse cellular lineages (Hwang et al., 2018). The scRNA-seq technique effectively reveals the major root cell types of *Arabidopsis* using individual cell transcription signatures (Shahan et al., 2022). Moreover, single-cell data can be employed for pseudotime analysis to determine the developmental process of roots by tracking the alterations in gene expression patterns across temporal and spatial dimensions.

Pseudotime analysis has been performed to reconstruct the spatiotemporal changes in gene expression patterns in root development. Although a comprehensive root atlas containing all major cell types and developmental stages has been proposed by several research groups (Denyer et al., 2019; Shahan et al., 2022; Shulse et al., 2019; Zhang et al., 2019), the mechanisms and genes directly involved in the differentiation of many tissues remain unexplored. Here, Monocle (Qiu et al., 2017) was used to infer single-cell trajectories. This algorithm employs an unsupervised strategy to identify cell subtype-specific differentially expressed genes and reconstruct the trajectory paths using reversed graph embedding (RGE) technique. Monocle then lets the user select the root location of the trajectory and infers pseudotime based on the distance from each cell to the root. The branch points or the cell fate decisions are assigned synchronously with pseudotime computation (Qiu et al., 2017).

Single-cell assay for transposase-accessible chromatin sequencing (scATAC-seq) methods have been developed to decipher the gene regulatory mechanisms underlying gene expression dynamics in and across cells and cell types. The open chromatin landscapes are probed by the Tn5 transposase enzyme, which then inserts sequencing adapters into these regions prior to amplification and sequencing. The integrative analysis of single-cell transcriptomic and epigenomic experiments (Dorrity et al., 2021; Farmer et al., 2021) gives insight into the impact of chromatin accessibility on

gene activity, allowing for predicting critical regulators and regulatory events of *Arabidopsis* cell development.

Moreover, because there are numerous research projects that have collected data from *Arabidopsis* root systems employing a variety of different methodologies, our analyses were validated on additional scRNA-seq and scATAC-seq datasets, such as (Dorrity et al., 2021), to confirm that the findings obtained are not specific to a particular set of data.

In summary, several studies have demonstrated tissue-specific roles of BR in regulating root development and have generated BR-response datasets from bulk root tips or whole seedlings. Moreover, recent studies showed that BRs play a critical role in epidermal cell fate determination by promoting non-hair cell differentiation. However, how BR regulates the onset of cell elongation in the transition zone or how BR elicits differential responses in hair and non-hair cells remains unclear. With the availability of scRNA-seq datasets of *Arabidopsis* root, this research aims to characterize the expression patterns of BR-responsive genes in different tissues and developmental zones and to understand the effects of BRs on epidermal cell differentiation. This study improves our understanding of how *Arabidopsis* root responds to BR and generate hypotheses for further research into plant hormonal responses under diverse environmental conditions.

1.2 Research objective

1.2.1 To investigate the expression patterns of brassinosteroid-responsive genes in each cell type/developmental zone of *Arabidopsis thaliana* roots using available scRNA-seq data.

1.2.2 To infer the roles of relevant genes in epidermal cell development.

1.3 Scope of the research

1.3.1 To assess gene expression data acquired from published scRNA-seq in 12 cell types from all three developmental zones of the root system of *Arabidopsis thaliana* (wild type).

1.3.2 To observe gene expression across the pseudotime trajectory of epidermal cells. We considered genes that have previously been demonstrated to respond to brassinosteroids to explore the hormone response during the development of epidermal cells.

1.4 Expected outcome

The methodology developed in this research can be applied to other single cell datasets from the model plant *Arabidopsis* as well as other crops to uncover genetic regulatory networks underlying root development.

PART 2

LITERATURE REVIEW

2.1 Brassinosteroid hormone

Steroid hormones have existed in vertebrates for a very long time. However, steroids were subsequently identified in invertebrates. It contributes to insect and other arthropod molting. Similarly, steroids were discovered in fungus (Zhabinskii et al., 2015). In 1970, scientists identified a new type of plant hormone known as "Brassins". Brassins are extracted from rape pollen (*Brassica napus L*.). As evidenced by the second-internode bioassay with brassins, these chemicals are hormonal because they are particular translocatable organic molecules that influence stem elongation and cell elongation (Mitchell et al., 1970). Similarly, in 1979, American scientists revealed a novel steroidal lactone with plant hormone characteristics. This substance was termed brassinolide (BL) and was isolated from the pollen of *Brassica napus* L. gathered by bees (Grove et al., 1979). Brassinolide was eventually identified as the most biologically active BR, capable of binding with high affinity to plasma membrane-localized brassinosteroid insensive 1 (BRI1) (Wang et al., 2001). Castasterone was detected in the insect galls of *Castanea crenata spp*. two years later (Yokota et al., 1982). Compounds with similar structures and activities were discovered. These are classified as brassinosteroids (BRs) (Khripach et al., 1998).

Brassinosteroids (BRs) are polyhydroxylated plant steroid hormones that play a key role in several plant mechanisms (Nolan et al., 2019). Some of the growth and development processes that BRs regulate include cell elongation (Wang et al., 2001), photomorphogenesis (Zhang et al., 2021), xylem differentiation (Saito et al., 2018) and seed germination (Zhong & Patra, 2021) in addition to both abiotic and biotic stress responses (Planas-Riverola et al., 2019; Ramirez & Poppenberger, 2020). BRs have the potential to promote physiological and developmental processes by interacting with other hormones such as auxin, jasmonic acid, abscisic acid and other plant signaling pathways (Kour et al., 2021; López-Ruiz et al., 2020).

The biosynthetic pathway of BRs involves the conversion of campesterol to BL, which is the most effective and final product of this pathway. This conversion process entails several reactions involving reduction, hydroxylation, epimerization, and oxidation (Clouse et al., 1996). The biosynthetic pathways for the synthesis of BL have been extensively studied through feeding experiments with labeled intermediates (Noguchi et al., 1999). These studies have revealed a complex network of possible pathways for BL biosynthesis. A sequential biosynthetic pathway and enzyme arrangement were proposed based on the findings of (Ohnishi et al., 2012; Ohnishi et al., 2006). The findings have been studied in mutants involving the specificity of the enzymatic substrate and intermediate content. All enzymes identified in the biosynthetic pathway of brassinosteroids (BRs) are cytochrome P450 (CYP450) mono-oxygenases, except for DE-ETIOLATED2 (DET2), which acts as a steroid 5α-reductase (Clouse et al., 1996).

2.2 Brassinosteroid signaling in plants

BR regulates transcription via binding to the extracellular domain of BRI1 (Clouse et al., 1996) to activate kinase activity, which in turn triggers dimerization with and activation of the BRI1-ASSOCIATED KINASE 1 (BAK1) co-receptor (Nam & Li, 2002). Then, BRI1 phosphorylates BRASSINOSTEROID-SIGNALING KINASE 1 (BSK1) (Tang et al., 2008) and CONSTITUTIVE DIFFERENTIAL GROWTH 1 (CDG1) (Kim et al., 2011). BRI1 also enhances the binding and phosphorylation of BRI1-suppressor1 (BSU1) (Kim et al., 2009) by BSK1/CDG1. When levels of BR are low (Figure 1 (left)), BRI1 KINASE INHIBITOR 1 (BKI1) suppresses BRI1 activation by inhibiting BRI1 interaction with its putative BAK1 coreceptor (Wang & Chory, 2006). This prevents BSU1 from being phosphorylated and allowing the GSK3-like kinase BRASSINOSTEROID INSENSITIVE 2 (BIN2) to remain active and phosphorylate (Kim & Wang, 2010). BIN2 phosphorylates BRASSINAZOLE RESISTANT1 (BZR1) and BZR2 [also termed BRI1-EMS-SUPPRESSOR 1 (BES1)] (He et al., 2002; Yin et al., 2005) to prevent DNA binding and enhance their binding to 14-3-3 proteins, resulting in their cytoplasmic retention and degradation (Gampala et al., 2007). When BR levels are high (Figure 1 (right)), BIN2 is inactivated by

BSU1 (Kim & Wang, 2010) and ubiquitinated by KIB1 (Kink suppressed in *bzr1-1D1*), resulting in its destruction (Zhu et al., 2017). PROTEIN PHOSPHATASE 2A (PP2A) dephosphorylates the homologous transcription factors BZR1 and BZR2 (Tang et al., 2011). Dephosphorylated BZR1 and BZR2 can enter the nucleus and bind to their target genes at the promoters containing transcription factor binding sites (BBRE and/or E-box motifs). This can either activate the target genes or inhibit their expression (Mao & Li, 2020).
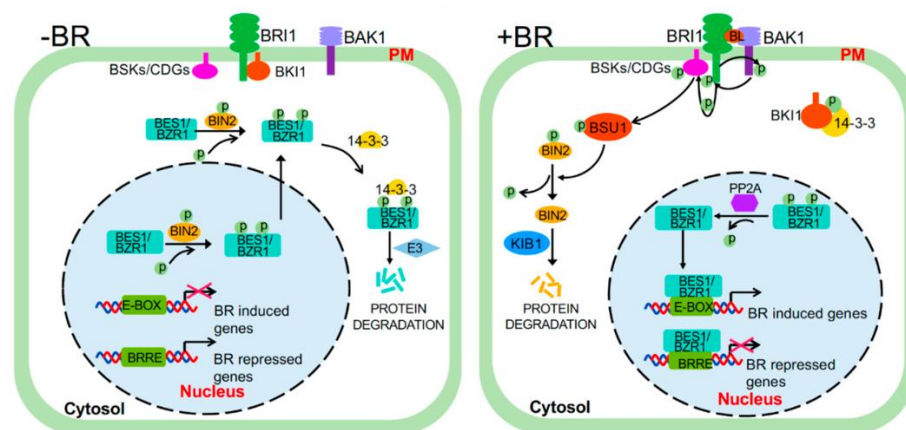
Figure 1 The model of BR signaling in plant
(Mao & Li, 2020)

2.3 Brassinosteroids play important roles in epidermal cell fate specification

The mature root has an outer layer of cells called the epidermis. This structure is composed of atrichoblasts and trichoblasts arranged in alternating layers. Trichoblasts differentiate and create root hairs, whereas atrichoblasts do not. Size and degree of vacuolation distinguish between these two cell types at an early stage of development (Berger et al., 1998). In addition, before the trichoblasts develop into hair cells, these cells possess distinct properties from atrichoblasts, such as a higher rate of cell division (Berger et al., 1998), a shorter cell length and a denser cytoplasm (Duckett et al., 1994). During development, the distribution of both cell types is controlled by position-dependent mechanism (Berger et al., 1998). Epidermal cells contact underlying cortical cells. Cells linked to cortical two cells develop into hair cells, whereas cells attached to

single cortical cell become nonhair cells. Based on this pattern, the positional information eventually determines the cell type specification in the epidermis (Lee & Schiefelbein, 2002).

Due to its simplicity, *Arabidopsis* root is considered as a model for studying the fate of epidermal cells. Several genes have been implicated in the epidermal development process. The transcription factors WEREWOLF (WER), GLABRA 3 (GL3), or ENHANCER OF GLABRA3 (EGL3), and TRANSPARENT TESTA GLABRA1 (TTG1), form a complex which stimulates the expression of GLABRA 2 (GL2) and CAPRICE (CPC). BRs have been identified as hormones that contribute to this process to sustain position-dependent fate specification in roots. BR production and response are necessary for normal expression levels and pattering of WER and GL2 (Kuppusamy et al., 2009). When BRs are absent (Figure 2), both atrichoblast and trichoblast have lower WER expression. In trichoblast, EGL3 is phosphorylated by active BIN2 to increase its cytoplasmic localization, which decreases the assembly of the nuclear WER-GL3/EGL3-TTG1 complex in both epidermal cell types. When this complex is lacking, *GL2* is repressed. In addition, activated BIN2 phosphorylates TTG1 which in turn blocks the activity of WER-GL3/EGL3-TTG1. Furthermore, the lack of BR signaling reduces the expression level of *CPC*, which is a downstream target of WER, resulting in the loss of hair cells in positions where hair cells should exist.

When BR signaling emerges (Figure 2), BR activates WER and blocks BIN2. BIN2 inhibition causes unphosphorylated EGL3 accumulate in the nucleus of the trichoblast. This mechanism also enables both cell types to have normal levels of unphosphorylated TTG1. Thus, WER-EGL3-TTG1 complexes are formed in trichoblasts while WER-GL3-TTG1 complexes are generated in atrichoblasts to enhance *GL2* expression, resulting in an atrichoblast cell fate. Nevertheless, trichoblast cell fate is negatively regulated by GL2. WER in atrichoblast promotes the accumulation of CPC. CPC translocates from atrichoblast to neighboring trichoblast, where it competes with WER for binding to GL3/EGL3, resulting in a CPC-GL3/EGL3-TTG1 complex. After CPC binding, GL2 is blocked and SCRAMBLED (SCM) accumulation is increased. SCM is

essential for cells to correctly comprehend their location inside the epidermal cells of developing root (Kwak et al., 2005). SCM further suppresses WER activity to enhance hair cell identity (Kwak & Schiefelbein, 2007). Along with CPC, the GL3/EGL3-TTG1 complex can also interact with TRY (Schellmann et al., 2002) and ETC1 (Kirik et al., 2004), which migrate from the atrichoblast to the trichoblast. This is indications of a lateral inhibition mechanism (Schiefelbein et al., 2014). Expression of *TRY* and *ETC1* is regulated by HDA6, which modifies the acetylation status of their histones (Li et al., 2015).

Thus, BR differs from other hormones in that it affects cell fate at an early stage by regulating the expression of *WER*, *GL2*, and *CPC* (Kuppusamy et al., 2009).
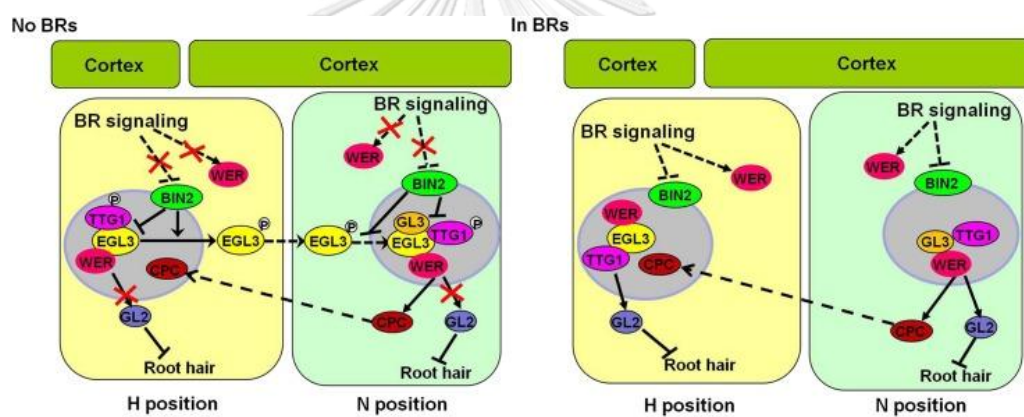
Figure 2 BR signaling regulates root epidermal cell fate.

(Cheng et al., 2014)

2.4 Single-cell RNA sequencing (scRNA-seq) technology

The advance of next-generation sequencing (NGS) technology in recent years has yielded insights into complex biological systems. NGS technologies are used to characterize individual cells in terms of genomes, transcriptomics, and epigenomics. Compared to conventional profiling methods, which examine entire populations, these single-cell analyses enable researchers to get new and possibly unexpected biological findings (Hwang et al., 2018). The emergence of single-cell RNA sequencing (scRNA-seq) offers new possibilities for studying gene expression profiles in single cells. While bulk RNA-seq primarily represents the average of gene expression throughout

thousands of cells, the use of scRNA-seq has increased in popularity for addressing the fundamental biological concern of cell diversity by providing information on the gene expression in individual cells. Moreover, this technique uncovers infrequent and complicated cell populations, follows the development of differentiated cell lineages and elucidates regulatory interactions between genes (Hwang et al., 2018).

In both humans and animals, single-cell genomics research gains popularity while plant research has just begun (Denyer et al., 2019; Jean-Baptiste et al., 2019; Ryu et al., 2019; Shulse et al., 2019; Zhang et al., 2019) in 2019. The basic structure and unpredictable growth of the *Arabidopsis* root makes it an appropriate model for studying the spatiotemporal transcriptional patterns determining developmental trajectories during stem cells to cell differentiation (Shahan et al., 2022).

The droplet-based technique dominates *Arabidopsis* plant single-cell transcriptomics because it is efficient at collecting large numbers of cells and detecting cell subpopulations in complex tissues (Bawa et al., 2022). Figure 3 shows the steps of droplet-based method of scRNA-seq analysis in plants. Sample preparation, which includes single-cell isolation, is the primary step required the collection of plant materials. Due to the strength and rigidity of the plant cell wall, isolating a single cell from plant tissue is challenging. Therefore, cell wall digesting enzyme is required to extract protoplast for subsequent single cell transcriptome analyses (Seyfferth et al., 2021).

After isolating individual cells and obtaining a cell mixture, the cells have to be distributed into tubes, wells, or droplets (Seyfferth et al., 2021) that contain the necessary components for transcript amplification and library preparation. Droplet-based single-cell approaches currently govern the plant single-cell transcriptomic analyses, particularly use in *Arabidopsis* (Denyer et al., 2019; Jean-Baptiste et al., 2019; Ryu et al., 2019; Shulse et al., 2019; Zhang et al., 2019). Using this method, only the 3' end of the transcript is captured and sequenced. This is in contrast to whole transcript scRNA, where this technique typically yields a higher throughput of cells and a reduced sequencing cost per cell.

In the droplet-based method, droplets contain lysis buffer, reverse-transcription mixture, and hydrogel microspheres with barcoded primers. After encapsulating protoplasts with droplets, primers are released, and cDNA is labeled with cell- and transcript-specific barcodes (UMIs) that contain a poly(T) sequence during reverse transcription. To facilitate precise bioinformatic identification of PCR duplicates, a unique molecular identifier (UMI) is appended to each molecule in a droplet as a tag prior to amplification. After the droplets are shattered, the content is linearly amplified and sequenced from all cells (Klein et al., 2015). After sequencing, the sequence files are demultiplexed to reassign the sequences to their original cells using UMIs. Demultiplex is performed in another round to determine the quantity of mRNAs taken from a single cell.
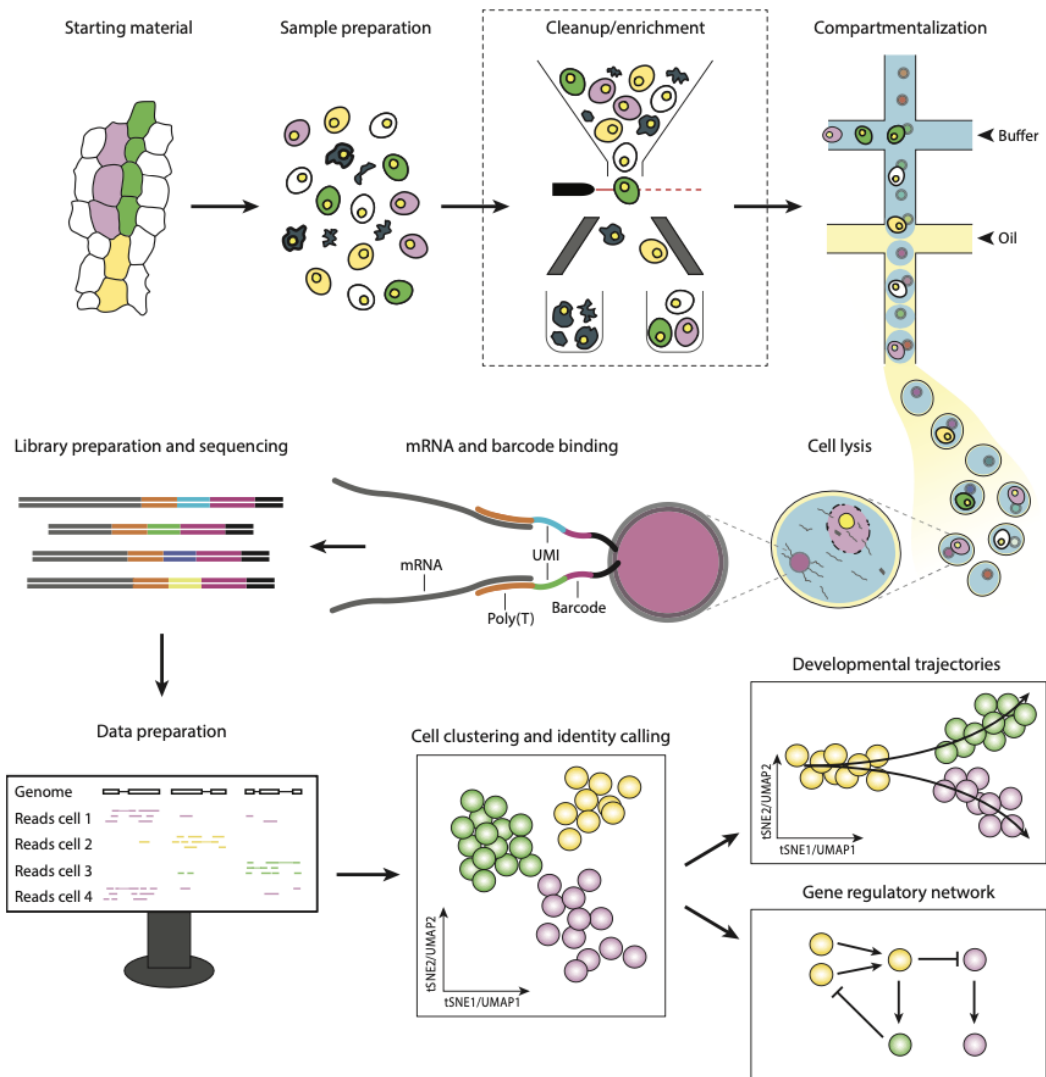
Figure 3 Overview of droplet-based plant scRNA-seq experiment.

(Seyfferth et al., 2021)

2.5 The analysis of scRNA-seq data

Single-cell transcriptomic analysis involves multiple steps that include quality control, normalization, dimensionality reduction, visualization, and cell type identification. There are several software packages, for example, Cell Ranger (Zheng et al., 2017), Seurat (Butler et al., 2018), and COPILOT (Shahan et al., 2022), that are used to perform these steps.

2.5.1   Quality control (QC)

In the case of microfluidic droplet-based approaches, over 90% of droplets do not contain real cells. However, ambient RNA from cell solutions may be apparent in these droplets, requiring algorithms to determine which droplets are likely to contain real cells. Some single droplets may have two or more cells, resulting in a mixture of the gene expression profiles of those two cells. At this point, only droplets containing actual cells are evaluated, and low-quality cells or multiplets should be removed prior to normalization. Thus, valid results from scRNA-seq data analysis depend on comprehensive QC being performed (Hong et al., 2022).

2.5.2   Normalization

This step is crucial for keeping expression counts comparable among genes and/or datasets. Single-cell normalization seeks to preserve actual biological variation while eliminating the impact of technological influences on the underlying molecular counts. Normalized gene variance should mostly be due to biological differences, regardless of gene abundance or sequencing depth. So, highly variable genes (HVGs) should exhibit differential expression between cell types and be used to identify cell clusters, whereas housekeeping genes should display less variation (Hafemeister & Satija, 2019).

2.5.3   Dimensionality reduction and visualization

Selected HVG expression data is high-dimensional since it includes cell-specific data. So, dimensionality reduction simplifies high-dimensional single transcriptomic data to lower-dimensions without losing biological information. The commonly used dimensionality reduction methods include principal component analysis (PCA), t-

distributed stochastic neighbor embedding (t-SNE), and Uniform Manifold Approximation and Projection (UMAP) (Su et al., 2022).

PCA can help identify the most important genes that contribute to the variation in the data. This is accomplished by projecting the data onto a new set of axes called principal components, which are linear combinations of the original features. This method can minimize the dimensionality of the data while conserving the most essential information by retaining only the top principal components that account for most of the variation (Su et al., 2022).

t-SNE and UMAP, both nonlinear dimensionality reduction techniques, are appropriate for visualizing scRNA-seq data and can assist researchers in identifying cell types and subtypes based on gene expression patterns. Both methods demonstrate robust local clustering by grouping similar categories together, while UMAP does it much more clearly in balance with global data structures. Clusters of cells with comparable gene expression profiles are represented in the resulting plot (Su et al., 2022).

2.5.4 Cell type identification

With prior information of gene expression in the cell type, it is possible to detect the cell identity of each cluster by layering the expression of specific marker genes. The *Arabidopsis* root cell types can be characterized using correlation analyses of marker gene expression in clusters (Li et al., 2016) and the calculation of an index of cell identity score (Denyer et al., 2019; Shahan et al., 2022; Shulse et al., 2019). However, clustering cells and determining their identities is an iterative process, with outcomes highly dependent on the cluster parameters selected at the outset (Seyfferth et al., 2021).

The primary outcomes of scRNA-seq analysis are differentially expressed genes (DEGs) indicating genetic variations between cell populations. This enables a high spatial resolution for the interaction of undiscovered regulators and the different responses of multiple cell types to a stimulus, such as environmental stress or phytohormone treatment. Furthermore, the high temporal resolution of scRNA-seq data facilitates the study of transition of cellular state and developmental trajectory. Combining spatial and temporal resolution and a large number of cells, gene regulatory networks

(GRNs) can be constructed to explore the interactions between transcription factors and target genes that regulates cellular function and cell phase transition (Seyfferth et al., 2021).

2.6 The analysis of single-cell trajectory

Trajectory analysis is a relatively recent computational technique used to investigate the dynamics of biological processes, particularly cellular differentiation and development. The analysis enables the assessment of cell fate determination and the processes involved in the transition between cell states. The methods of trajectory analysis comprise inferring a developmental trajectory and subsequently mapping cells to the inferred trajectory. This assumes that cells progress continuously along a developmental trajectory, and that differences in gene expression between cells reflect their progression (Deconinck et al., 2021).

Since existing approaches are unable to capture complex trajectories, a new method is required. This study reconstructs epidermal single-cell trajectories with Monocle2, a new method developed by Qiu et al. (2017). This method employs reversed graph embedding (RGE) (Figure 4), a graph-based approach for reconstructing cell trajectories without assuming a specific trajectory form. Monocle2 identifies the genes that define a specific biological process by using an unsupervised approach called dpFeature. Alternately, we can select genes whose expression differs between clusters of cells. The expression level of these feature genes relates to dimensions (Qiu et al., 2017).

To construct single-cell trajectories, each cell is projected in high-dimensional space. DDRTree is one of the dimensionality reduction methods used in the RGE. It then reduces the dimensionality of the data to embed cells into a low-dimensional space that captures their developmental progression. This low-dimensional space is the input for the RGE algorithm to infer a developmental trajectory. Using k-means clustering, the centroids of the groups of cells are determined, and a spanning tree is then formed. The cells are assigned onto the nearest point on the tree. The newly spanning tree is then

constructed and updated to high-dimensional space. In this algorithm, users can choose a root tip and calculate the distance of each cell to the root, which is interpreted as its pseudotime. In this procedure, the branch, which represents the decision of cells to develop into distinct cell types, is assigned simultaneously (Qiu et al., 2017). Branch expression analysis modeling (BEAM), a new regression model, is then employed to detect genes that differentially expressed after cell fate decision (Qiu et al., 2017) . Thus, Monocle2 contains tools for constructing single-cell trajectories and analyzing significant branch-dependent genes, enabling us to comprehend the genetic architecture driving cell fate determination.
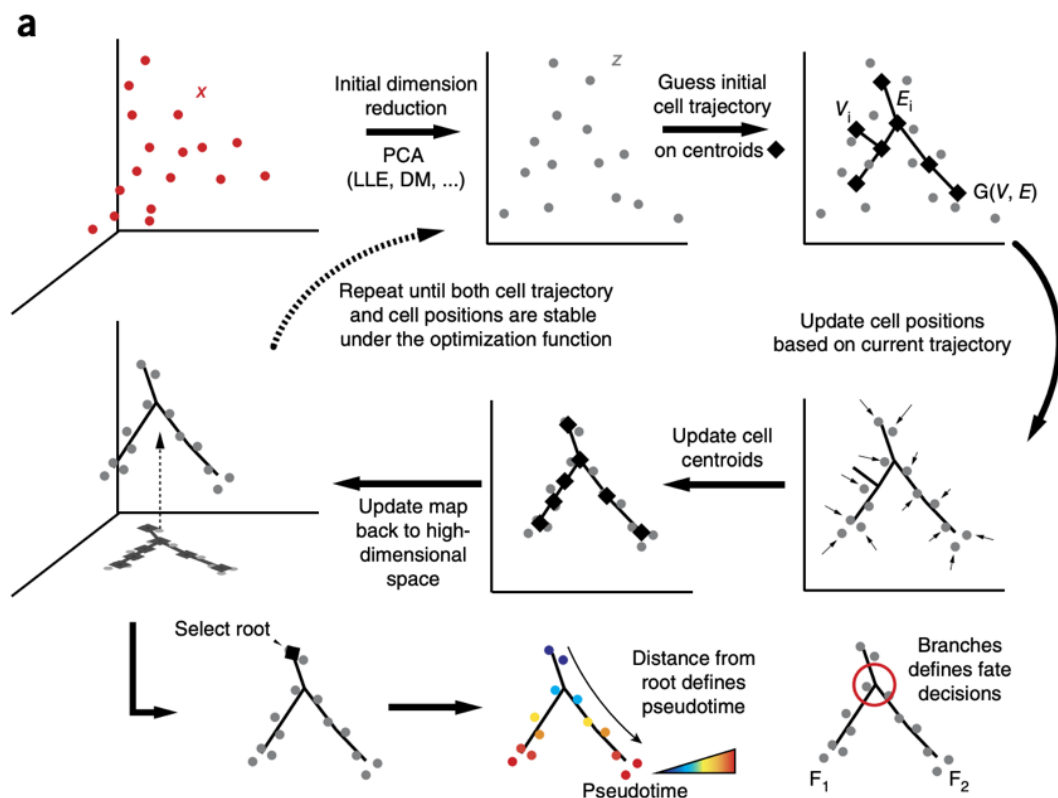
Figure 4 Reversed Graph Embedding (RGE) method implementation in Monocle2 for constructing single-cell trajectory.

(Qiu et al., 2017)

2.7 Single-cell assay for transposase-accessible chromatin sequencing (scATAC-seq) technology

Specific localization of regulatory elements in accessible chromatin facilitates transcriptional regulation. Open chromatin configuration is dependent on transcription factor binding to activate the target gene, while condensed chromatin blocks the access of transcription factors and transcriptional regulators to the promoter and/or enhancer, leading to gene silencing (Figure 5A). ATAC-seq technique is used to investigate chromatin accessibility by employing a Tn5 transposase enzyme to break open chromatin regions. This facilitates the insertion of adaptors for DNA fragment sequencing. The information obtained from the sequenced DNA fragments provides insight into the accessibility of chromatin regions such as transcription factors, enhancers, and insulators (Sun et al., 2019).

Recent advancements in single-cell ATAC-seq are a high-throughput approach for studying the epigenetic landscape of individual cells (Chen et al., 2019). By analyzing the chromatin accessibility of single cells, scATAC-seq can reveal valuable insights into the complexity of gene transcription in a specific cell type throughout cellular differentiation and development (Sun et al., 2019). However, each method of analyzing single cells has its own advantages and disadvantages and can only provide information on a limited subset of biological characteristics. It is essential to obtain data from one dataset to improve the evaluation of another. Consequently, the integrative analysis is designed to address this issue. It anchors single-cell transcriptomic and epigenomic investigations to reveal the effect of chromatin accessibility on gene activity and regulatory networks of *Arabidopsis* cell development (Dorrity et al., 2021; Farmer et al., 2021).
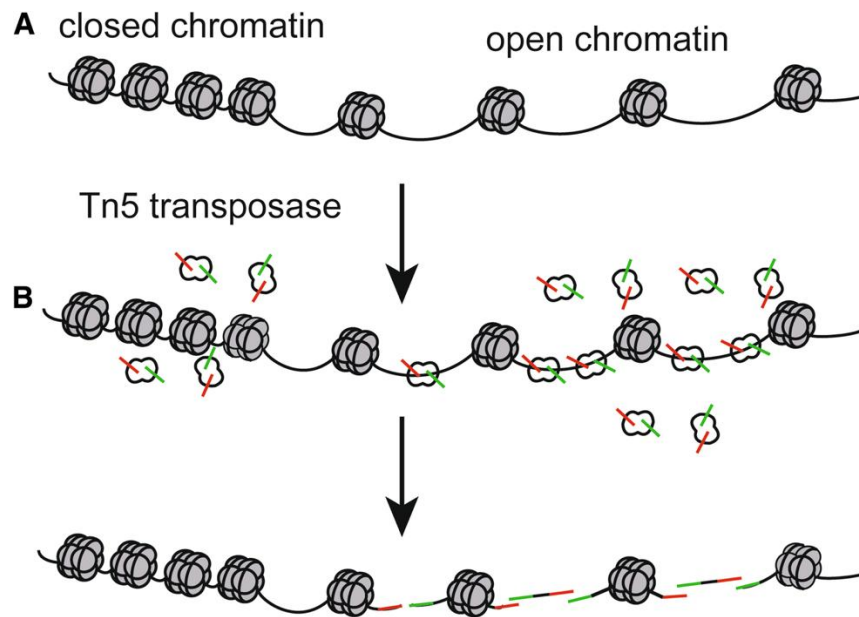
Figure 5 The mechanism of identifying chromatin accessibility using the Tn5

transposase.

(Sun et al., 2019).

2.8 Integrative analyses of single-cell data

While the biological information in the data generated by various techniques may be comparable, some of them might identify distinct cell populations that are currently unidentified (Figure 6A). Furthermore, the classification of cells into subcategories based on data obtained from a single technique poses a challenge (Dorrity et al., 2021; Farmer et al., 2021). Seurat (Version 3) was created with the aim of offering a complete approach for performing reference assembly and transfer learning when examining single-cell data. The goal is to obtain a more extensive comprehension of gene regulation and the diversity of cell types in intricate biological systems. This software has the capability to efficiently process a range of data types, such as transcriptomic, epigenomic, and proteomic data (Stuart et al., 2019).

In the pre-processing of scATAC-seq data, an initial step involves the quantification of reads that map to regions of promoters or enhancers, which enables the construction of a gene activity matrix (Pliner et al., 2018). This matrix serves as a representative for gene activity in synthetic scRNA-seq data and facilitates subsequent

data integration procedures. Furthermore, to ensure data quality, cells with a number of peaks below a specified threshold are excluded from the analysis.
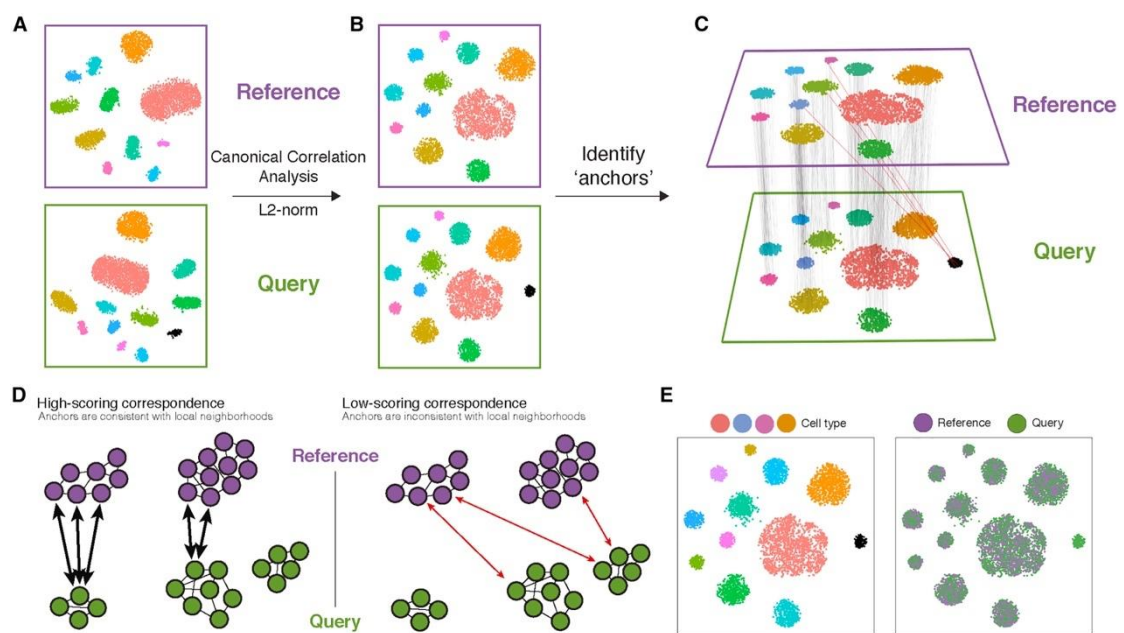
Figure 6 The overview of integration of single-cell datasets using Seurat version 3. (Stuart et al., 2019)

To decrease the dimensionality of the scATAC-seq peak matrix, the application of latent semantic indexing (LSI) is implemented (Cusanovich et al., 2018). Prior to applying LSI, the technique of Term Frequency-Inverse Document Frequency (TF-IDF) is utilized to assign weights to important terms within the scATAC-seq peak matrix. This enables the identification of genomic regions that exhibit high variability across cells and are highly selective for differentiating between various cell types or states.

To calculate the TF (term frequency) for each genomic region in each cell, the total number of reads that are mapped to that region is divided by the total number of reads that are mapped to all regions in that cell. IDF (inverse document frequency) reflects the inverse frequency of each region across all cells by calculating the logarithm of the total number of cells in the dataset divided by the number of cells that contain the region. To avoid computing the logarithm of 0, 1 is added as a pseudocount. Thus, a

weighted score for each accessible region in each cell is obtained by the multiplication between TF and IDF terms (Cusanovich et al., 2018). Highly variable peaks are given higher weights, while peaks that are frequently observed across cells are given lower weights.

A lower-dimensional representation of the TF-IDF matrix is generated by utilizing singular value decomposition (SVD), which returns Latent Semantic Indexing (LSI) components and scaled LSI loadings for each cell, scaled to have a mean of 0 and standard deviation of 1. These steps are used to learn the weighting of anchors in the scATAC-seq dataset and are implemented through the 'RunLSI' function in Seurat (Stuart et al., 2019).

The process of identifying anchors, which are the relationships between individual cells in scATAC-seq (gene activity matrix) and scRNA-seq (gene expression matrix) datasets, is depicted in Figure 6C. This is accomplished by identifying anchor pairs that consist of highly variable features, which exhibit high correlation across both chromatin accessibility and gene expression. Then, a shared low-dimensional space is constructed based on these features using canonical correlation analysis (CCA), as shown in Figure 6B. For each anchor pair, a score is assigned based on the consistency of anchors across the neighborhood structure of each dataset (Figure 6D). This facilitates the efficient transfer of gene expression data from scRNA-seq reference dataset to a low-dimensional space learned from scATAC-seq query dataset. Eventually, we can combine gene expression data from these two modes to create uniform atlases at the single-cell level (Figure 6E) (Stuart et al., 2019).

PART 3

RESEARCH METHODOLOGY

3.1 Obtaining datasets used in this study.

Published scRNA-seq (Shahan et al., 2022) and scATAC-seq (Dorrity et al., 2021) datasets were retrieved from GEO: GSE152766 and GEO: GSE173834, respectively. The first dataset (Shahan et al., 2022) used the 10X Genomics platform to profile 96,000 cells taken from thirteen different sets of wild-type *Arabidopsis thaliana* root tip aged between 5-7 days old. They additionally integrated the data with three published datasets (Denyer et al., 2019; Ryu et al., 2019) to add further depth and evaluate the variability of the data across different labs. Consequently, this dataset contains the gene expression profile of more than 110,000 cells, encompassing 17,513 genes that were conserved in all integrated samples. Cells in this dataset have already accomplished cell type and developmental stage annotation by combining information from the four annotation approaches including (1) annotation based on spatial mapping, (2) marker annotation, (3) correlation annotation and (4) index of cell identity (ICI) calculation. The developmental stage of each cell was determined by comparing its transcriptome to the gene expression profiles of manually dissected root tissue segments. Additionally, this dataset provided high-quality gene markers of cell types and many newly discovered genes prominent in the cell types/developmental zones. This was performed by using differential expression analysis to uncover a total of 7,867 tissue- and stage-specific genes with log 2-fold change $\geq$ 1, the difference in percentage of gene expression between the clusters considered and the number of remaining cells > 0.25 and the area under the curve (AUC) > 0.75. Thus, this dataset provided high-quality gene markers of cell types and many newly discovered genes prominent in the cell types/developmental zones.

This study also assessed the validity of the tools, methods, and results obtained from the analysis on two additional scRNA-seq datasets (Denyer et al., 2019; Ryu et al., 2019) and scATAC-seq dataset (Dorrity et al., 2021). The lists of BR-responsive

(induced/repressed) genes were retrieved from a previous study (Chaiwanon & Wang, 2015). BR-responsive genes were mapped to specific cell subpopulations identified in the scRNA-seq data (Shahan et al., 2022) (in the form of observed marker genes and novel genes distinguished in certain subpopulations).

3.2 Pseudotime analysis

Monocle (version 2.20.0) (Qiu et al., 2017) package implemented in R programming language (RStudio Version 1.4.1103) was used to investigate the process of cell differentiation and epidermal cell fate determination. The epidermal (trichoblast and atrichoblast) cell data was extracted from large dataset (Shahan et al., 2022) and used for downstream analysis. The following methods are explained in the tutorial developed by Qiu et al. (2017).

Here, the data were preprocessed before continuing to the main process of constructing the trajectory. Epidermal single-cell expression data were stored as object of the CellDataSet class using 'newCellDataSet' function. The class comprised three data files including (1) exprs is a matrix of numeric gene expression values, with rows representing genes and columns representing cells. (2) phenoData contains the attributes (such as cell type and developmental stage annotation) of each cell. (3) featureData contains the attributes (such as gene short name) of each gene. The size factors for each cell in the dataset was calculated by 'estimateSizeFactors' function. Size factor was used to normalize the read counts for each cell, considering differences in sequencing depth and library size between cells. The dispersion parameter for each gene measuring the amount of biological variation in gene expression across cells was then estimated by 'estimatDispersions' function.

There were three major processes involved in pseudotime estimation. Firstly, genes that were differentially expressed across cell types and developmental stages were identified using the functions 'differentialGeneTest' and 'FoldChange' with q-value < 0.001 and log 2-fold change > 1, respectively. These genes were set into CellDataSet by 'setOrderingFilter' function to be used for ordering cells in downstream analysis.

Secondly, the dimensionality of the data was reduced through a reversed graph embedding technique by 'reduceDimension' function to learn the underlying manifold structure that captures cell-cell relationships of the data.

Thirdly, the cells were arranged using the 'orderCells' function, and the trajectories were shown in the reduced dimensional space using 'plot_cell_trajectory' function. The occurred branching indicated the separation of two cell types into their lineages. The Monocle package allows the user to specify the cell that marks the initiation of pseudotime. This enables the computation of pseudotime through the implementation of the 'orderCells' function, wherein the cell located prior to the branched point and in the meristematic stage is selected as the starting point for pseudotime.

3.3 Differential expression and enrichment analysis

The statistical significance of the overlap in BR-responsive genes expressed across cell types was assessed with hypergeometric distribution and multiple testing corrections using Benjamini-Hochberg procedure and with a statistically significant p-value < 0.05.

Differentially expressed genes across cell types and developmental stages, which were initially obtained, were used to determine whether they were differently expressed between branches using a statistical test in Monocle called "branched expression analysis modeling" (BEAM) and cutting off significant genes with a q-value < 0.0001. These genes were categorized hierarchically and visualized by 'plot_genes_branched_heatmap' function. This provided insight into the branched developmental processes.

To explore the biological functions of differentially expressed genes in each cluster and branch, functional enrichment analysis was performed using the Gene Ontology (GO). This analysis was conducted using the David Bioinformatics Resources platform (https://david.ncifcrf.gov/home.jsp). To manage the large number of functionally related genes and terms obtained from the analysis, we employed the "DAVID Gene Functional Classification Tool" to group the genes and terms into a more manageable

number of biological modules. The GO term with the least FDR was chosen as the representative of each module.

3.4 Visualization of the expression of candidate genes in the root cell atlas and epidermal cell fate trajectory

For branching trajectory, the function 'plot_genes_branched_pseudotime' was employed to construct the gene expression curves for each lineage. This function generated trend plots for genes of interest that pass through a progenitor state before separating into two different branches. Each branch was stretched to a maturity level between 0 and 100.

Genes involved in epidermal cell development, BR signaling and synthesis pathways were analyzed. Additionally, BR-responsive genes with remarkable expression patterns, which might be associated with those pathways, were explored.

3.5 Integration of single-cell ATAC-seq and single-cell RNA-seq to relate gene expression with chromatin accessibility.

The integration of single-cell datasets derived from different techniques was accomplished using the Signac package (Version 1.9.0) (Stuart et al., 2019). The identification of anchors linking the scRNA-seq (Shahan et al., 2022) and scATAC-seq (Dorrity et al., 2021) datasets was accomplished through using 'FindTransferAnchors' function, with reference data being the gene expression matrix of scRNA-seq data and query data being the gene activity matrix of scATAC-seq data. The cell type labels from the scRNA-seq data were transferred to the scATAC-seq cells using these anchors. This was done by 'TransferData' function with the weight.reduction parameter set to be the LSI components of the scATAC-seq peak data. The LSI components were previously determined and used to ensure that the transferred data was consistent with the scATAC-seq data in the shared low-dimensional space in this function. The predicted cell types of scATAC-seq data were visualized on a UMAP representation. Finally, to examine the two data modalities together, cells from the scRNA-seq and scATAC-seq

experiments were co-embedded on the same low-dimensional space using the identified anchors.

3.6 Validation of findings across multiple datasets

To ensure the generalizability and robustness of the findings, additional scRNA-seq datasets (Denyer et al., 2019; Ryu et al., 2019) as well as scATAC-seq data (Dorrity et al., 2021) were utilized for validation. Expression patterns of candidate genes identified in this study were visualized on other scRNA-seq datasets to confirm that the same conclusions can be obtained. To annotate the enriched peaks detected from the scATAC-seq analysis, the statistical programming package ChIPpeakAnno (Zhu et al., 2010) was employed. This powerful tool facilitated the identification of regulatory elements that play critical roles in gene regulation by exerting control over the expression of nearby genes. In the scATAC-seq and scRNA-seq integration atlas, each scATAC-seq cell was assigned a color based on the peaks associated with its nearest target genes. The color assignment allowed for the visualization and interpretation of chromatin accessibility patterns in the context of gene regulation.

PART 4

RESULTS AND DISCUSSIONS

4.1 BR-induced and -repressed genes are differentially enriched in different cell types and developmental zones.

To explore the specific roles of BR in regulating root development at the cell-type level, we conducted a comparative analysis utilizing lists of BR-responsive genes and cell-type-specific genes derived from previous studies (Chaiwanon & Wang, 2015; Shahan et al., 2022). The BR-responsive gene list was generated through bulk RNA-seq experiments involving BR treatments and BR mutants, enabling the identification of genes induced or repressed by BR in *Arabidopsis* root tips (Chaiwanon & Wang, 2015). The scRNA-seq dataset provided the cell-type-specific gene lists, identifying genes with specific expression patterns in various cell types and developmental zones of *Arabidopsis* root tips (Shahan et al., 2022). To assess the enrichment of BR-responsive genes, the percentages of BR-induced and BR-repressed genes were calculated within each cell type of the meristematic, elongation, or maturation zone. The enrichment of BR-induced or BR-repressed genes in specific cell types was assessed through hypergeometric analysis to determine whether they were significantly over- or under-represented (FDR < 0.05).

Figure 7 illustrates the distribution of BR-induced and BR-repressed genes across different cell types and developmental zones. It reveals a substantial underrepresentation of BR-induced genes in the meristematic zone, while exhibiting an overrepresentation in the elongation and maturation zones across multiple cell types (Figure 7A). In contrast, figure 7B demonstrates a notable overrepresentation of BR-repressed genes specifically within the meristematic cell types, including the stem cell niche, quiescent center, columella, lateral root cap, and phloem. The atrichoblast cell type, which refers to non-hair cells in the epidermis, exhibited the highest enrichment percentages of BR-induced genes in both the elongation and maturation zones (46.18% and 47.84%, respectively). This contrasts with the trichoblast (16.67% and 26.64%,

respectively), which is adjacent to the atrichoblast cell files in the epidermis. These findings imply a significant involvement of BR in the development of atrichoblasts and highlight the distinct roles of BR in regulating epidermal development between these two cell types.

Upon comparing cell types derived from the same stem cell initials, we observed similar enrichment patterns in the cortex and endodermis. In contrast, distinct enrichment patterns emerged in the xylem and phloem of the stele. Specifically, the xylem exhibited a significant enrichment of BR-induced genes, whereas the phloem showed enrichment of BR-repressed genes while being underrepresented in the xylem. Overall, these findings from the enrichment analysis indicate that BR exerts cell-type- and developmental-zone-specific functions.

Prior studies have established that BR signaling is low during the meristematic zone, leading to the upregulation of BR-repressed genes, which serves to sustain stem cell activities and control meristem size. Conversely, BR signaling reaches its peak during the elongation zone, thereby facilitating cell elongation, consistent with the large number of BR-induced genes (Chaiwanon & Wang, 2015). Additionally, earlier research identified a number of BR-repressed genes in the inner tissue (stele) of the meristematic zone (Vragović et al., 2015). Based on our analysis of scRNA-seq data, BR-repressed genes are mostly enriched in the phloem, but not xylem.
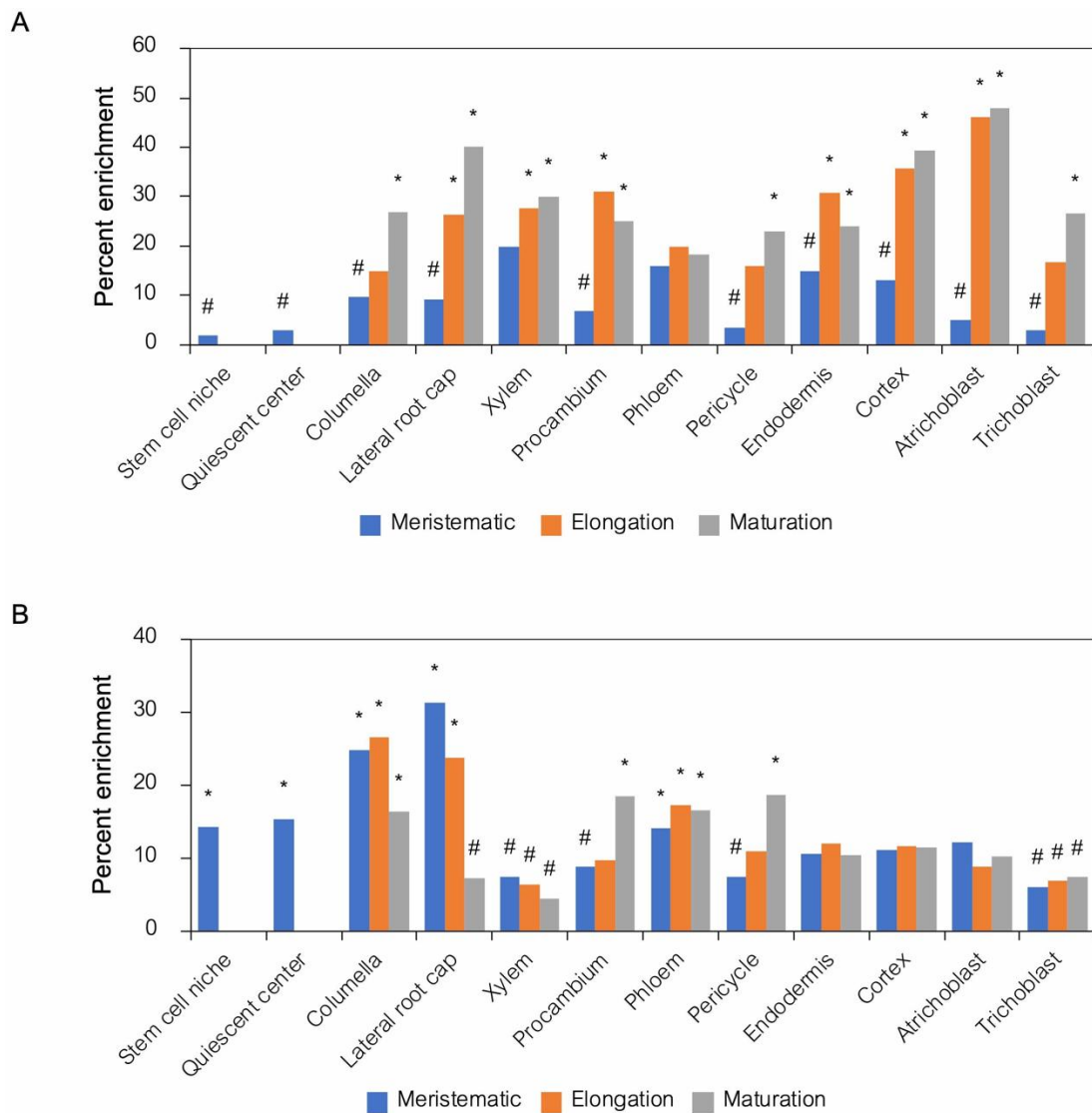
Figure 7 Enrichment analysis of BR-induced and BR-repressed genes across various cell types and developmental zones in the *Arabidopsis* root tips.

Bar graphs visualize percentage enrichment of BR-induced (A) and BR-repressed genes (B) in each tissue. Significant overrepresentation is denoted by an asterisk (*), while significant underrepresentation is indicated by a sharp (#) (FDR < 0.05).

4.2 Reconstruction of the continuous differentiation trajectory of epidermal cells from scRNA-seq dataset

Epidermal meristematic cells undergo differentiation, giving rise to distinct cell types known as trichoblasts (hair cells) and atrichoblasts (non-hair cells). To investigate

the dynamic expression of genes associated with BR biosynthesis and signaling along the developmental axis of the epidermis, it is essential to establish the trajectory of epidermal cell development. Accordingly, the scRNA-seq dataset (Shahan et al., 2022) was utilized to extract cell clusters corresponding to atrichoblasts and trichoblasts.

The Monocle 2 algorithm (version 2.22.0) was employed to reconstruct the trajectory of root cell differentiation and elucidate the determination of epidermal cell fate. In this process, genes exhibiting differential expression across cell types and developmental stages ($\log_2$fold change > 1 and q-value < 0.001) were identified as ordering genes for trajectory construction. A total of 847 ordering genes were selected. Subsequently, each epidermal cell was visualized in a reduced dimensional space using UMAP plots. Cells were represented based on the annotated cell types (atrichoblast or trichoblast; Figure 8A), developmental stages (meristem, elongation or maturation; Figure 8B), or along the pseudotime (Figure 8C).

Our analysis reveals that atrichoblast and trichoblast cells in the meristematic stage exhibit clustering at the initial phase of the pseudotime trajectory, while the two lineages diverge at a branching point corresponding to cells in the elongation stage. Subsequently, the cells progress through the maturation stage towards the end of the pseudotime trajectory. This observation demonstrates a precise ordering of cells based on their developmental stages, consistent with the existing annotations of cell types and developmental stages. In addition, cell-cycle genes can be eliminated by analyzing differentially expressed genes across multiple cell types and developmental stages. This suggests that these genes were not differentially expressed between atrichoblasts and trichoblasts. Cell-cycle genes are recognized as a significant source of biological noise in single-cell data. Their presence can obscure clusters of cell types or lead to the emergence of novel clusters, as well as potentially impacting the accurate reconstruction of single-cell trajectories (Barron & Li, 2016). Consequently, the pseudotime analysis of each epidermal cell type has the potential for studying the expression patterns of genes of interest.
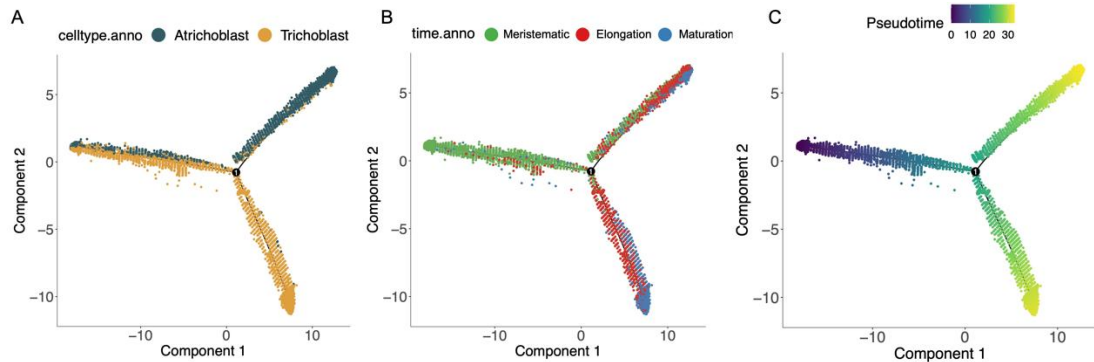
Figure 8 UMAP plots depicting the trajectory of epidermal cells in a reduced dimensional space.

Each dot represents an individual epidermal cell, and the dots are color-coded based on cell types (A), developmental stages (B), and pseudotime (C).

4.3 The analysis of differential gene expression between pseudotime branches and enrichment analysis

Through the input of 847 ordering genes into the branched expression analysis modeling (BEAM) function in Monocle 2 (Qiu et al., 2017), the statistical significance of differentially expressed genes in each branch of the epidermal cell fate trajectory was then determined. We discovered that 825 genes exhibited branch-dependent expression in a statistically significant manner.

The patterns of the significantly branch-dependent genes across pseudo-time were depicted simultaneously in both lineages with a heatmap (Figure 9). The start of pseudotime was in the middle of the heatmap, whereas the columns indicated pseudotime points and the rows represented genes. Both sides of the heatmap displayed pseudo-time terminations. Based on similar lineage-dependent expression patterns, the gene were clustered hierarchically into 5 groups and a gene ontology analysis was then performed on each cluster.

Based on the heatmap (Figure 9), genes in cluster 1 exhibited an initial surge in expression levels during the early stages of pseudotime. This was followed by a subsequent decline in expression levels across both lineages. The cluster displayed enrichment of meristematic cell-specific genes (72.29% and 80.09% in trichoblast and atrichoblast, respectively) (Figure 10A), involved in various early-stage cellular processes, including protein synthesis such as translation and assembly of functional ribosomes (Figure 10C). Cluster 1 also included genes regulating RNA transport and protein quality control, ensuring proper RNA processing and modification before integration into the ribosome.

Cluster 2 exhibited gene expression specific to meristematic tissue (35.90% in trichoblast and 41.03% in atrichoblast) and elongation tissue (20.51% in trichoblast and 3.21% in atrichoblast) (Figure 10A). Interestingly, these genes displayed an up-regulated expression pattern near the branch point and maintained expression levels during the middle of the pseudotime, followed by a decrease towards both ends (Figure 9). The gene signatures within cluster 2 were associated with crucial cellular processes including homeostasis and microtubule organization, indicating their importance in maintaining normal cell function. Additionally, the cluster exhibited genes related to methylation processes, which are involved in post-translational modifications (Figure 10C).
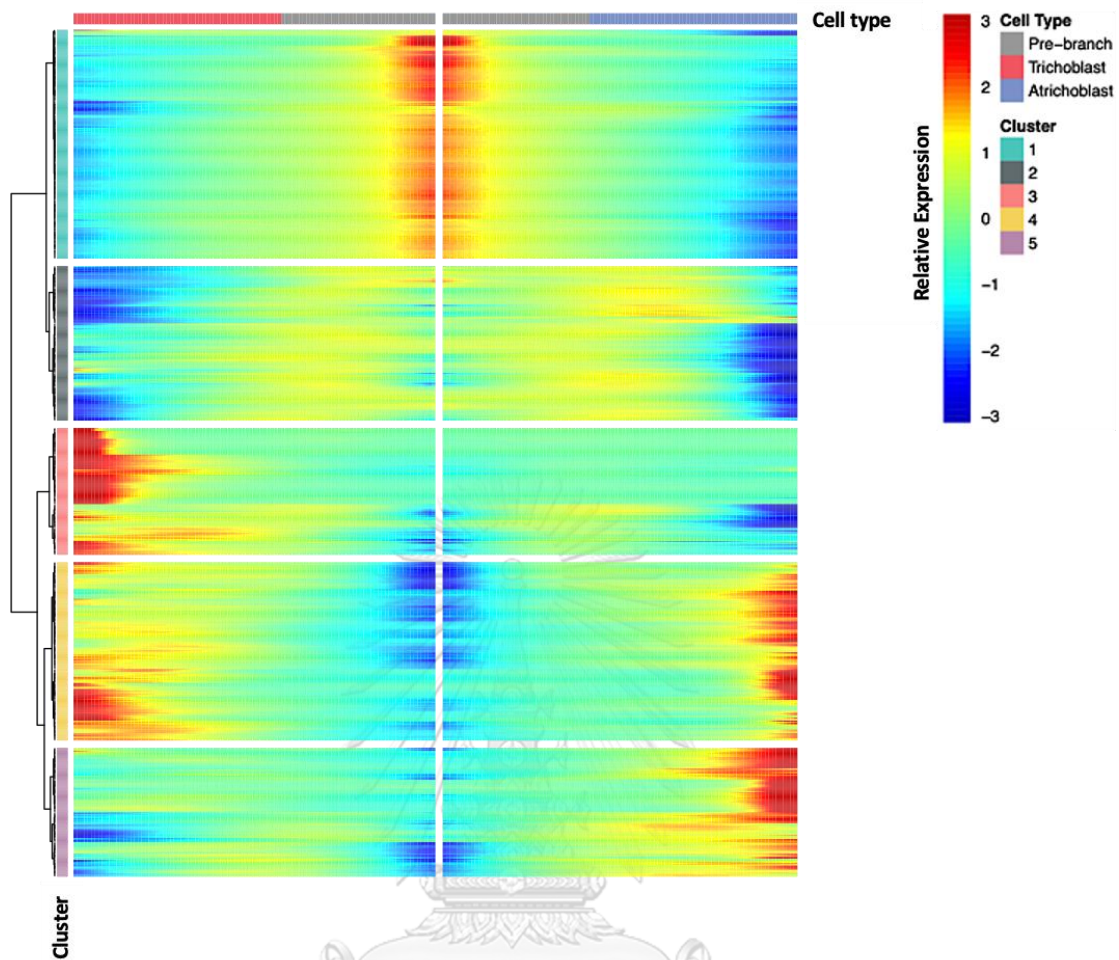
Figure 9 Heatmap illustrating the changes of the significantly branch-dependent genes over pseudotime.

The color bar represents the relative expression levels of the genes.

Genes in cluster 3 exhibited an up-regulated pattern during the midpoint of the pseudotime in the trichoblast lineage (Figure 9). This was consistent with the significant presence of genes specific to elongation and maturation exclusively found in trichoblasts (61.72% and 60.16%, respectively) (Figure 10A). However, expression levels of these genes began to decline towards the end of pseudotime (Figure 9). Cluster 3 genes were primarily related to plant-type cell wall organization and hydrogen peroxide catabolic processes, potentially influencing cell elongation (Figure 10C). Hydrogen peroxide serves as a signaling molecule that promotes root cell elongation (Richards et al., 2014).

Additionally, *COW1* and *XTH14* (Bruex et al., 2012; Shahan et al., 2022), which are known markers for trichoblast cell clusters, were presented in this cluster.

Genes in cluster 4 initially showed an upregulation during the midpoint of the pseudotime, followed by a subsequent increase towards the end, evident on both sides of the heatmap (Figure 9). The signature genes of cluster 4 were involved in root morphogenesis and water transport (Figure 10C). Additionally, this cluster included genes that were specific to the elongation and maturation stages of trichoblasts (20% and 36.11%, respectively) and atrichoblasts (28.89% and 55%, respectively) (Figure 10A). This suggests that these genes likely govern the development of both epidermal cell types.

Cluster 5 demonstrated an up-regulated pattern during the midpoint of the pseudotime, particularly concentrated on the atrichoblast lineage. This aligned with the presence of atrichoblast-specific genes primarily associated with elongation and maturation stages (48.46% and 36.15%, respectively). Cluster 5 mainly consisted of genes involved in cell wall organization, plant epidermis development, and water transport.

The detection of aquaporin genes, a member of the major intrinsic protein family (MIPs) implicated in water transport, within clusters 4 and 5 was consistent with previous research. Kaldenhoff et al. (1995) demonstrated the predominant expression of aquaporins in elongation cells of *Arabidopsis* roots (Kaldenhoff et al., 1995). Aquaporins are responsible for regulating cellular and tissue water content (Yang & Cui, 2009), facilitating cell expansion (Johansson et al., 2000).

Some of the genes utilized in this analysis have been identified as tissue- and developmental-stage-specific genes in previous research (Shahan et al., 2022). However, upon analyzing the expression patterns of genes grouped according to their expression along the pseudotime, we observed distinct intervals of expression for developmental stage-specific genes, particularly in clusters 1 and 2. Cluster 1 comprised genes specific to the meristematic stage, exhibiting upregulation at an early pseudotime point followed by a subsequent decline over time. In contrast, genes in cluster 2

demonstrated a gradual upregulation, peaking at the center of the pseudotime. Interestingly, this method revealed that even genes annotated to the same developmental stage might exhibit differential expression at distinct time points.

Taken together, the findings highlight the ability of scRNA-seq to reconstruct the continuous differentiation trajectory of epidermal cells, offering insight into the expression of signature genes at different developmental stages. Cluster 2 displayed transcriptome profiles representing the intermediate state of cell differentiation. On the other hand, clusters 3, 4, and 5 contained genes involved in the initiation of differentiation and development processes in both cell types.
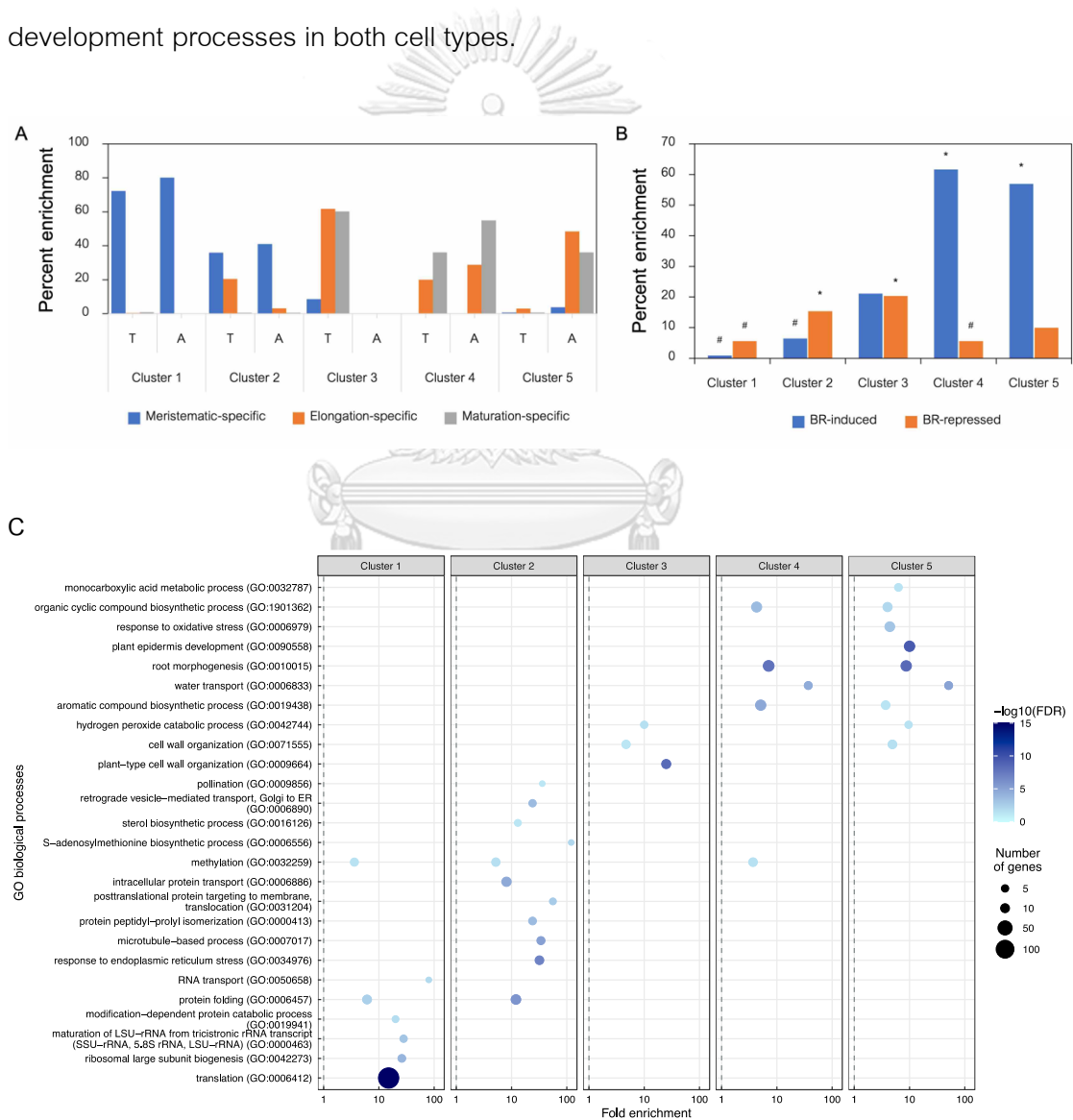

Figure 10 Branch-dependent gene expression analysis

(A) Enrichment percentage of developmental stage-specific genes in each cluster represented by stacked bar plot. T and A are trichoblast and atrichoblast, respectively.

(B) Enrichment percentage of BR-responsive genes in each cluster represented by bar plot. Significant overrepresentation is indicated by an asterisk (*), while significant underrepresentation is indicated by a sharp (#) (p-value < 0.05).

(C) Scatter plot of the most representative GO terms enriched by differentially expressed genes across branches (q-value < 0.0001), cell types and developmental stages (FDR < 0.05).

4.4 Visualization of the expression of candidate genes in the epidermal cell fate trajectory

      4.4.1 Expression of genes involved in BR biosynthesis and signaling pathways.

Differential BR responses could be due to differences in cellular BR levels and/or signaling activity. To test whether genes involved in BR biosynthesis and signaling pathways vary along the developmental stages in atrichoblast and trichoblast cell types, we visualized the expression patterns of selected genes along the pseudotime trajectory of each cell type. The pseudotime was stretched and presented on unified graphs.

Our analysis revealed distinct expression patterns of genes involved in BR biosynthesis. *DWF4*, *DET2*, and *ROT3* were highly expressed, while *CPD*, *CYP90D1*, *BR6OX1*, and *BR6OX2* were expressed at very low levels (Figure 11B-F). *DWF4* and *DET2* were preferentially expressed in trichoblast during the meristematic stage, whereas in atrichoblast, their expression was observed during the transition from the meristematic to elongation stages. The expression of *ROT3* was elevated along the developmental stages in atrichoblast but decreased in the trichoblast during the maturation stage.

Endogenous levels of BR are tightly regulated through a negative feedback loop mediated by the BR signaling pathway (Xuan & Beeckman, 2021). High levels of BR

activate BR signaling and lead to downregulation of BR biosynthesis enzymes and upregulation of BR catabolic enzymes. In our analysis, we observed a similar expression pattern between *BAS1*, encoding a BR degradation enzyme, and *ROT3* (Figure 11J). In addition, *BIA1* expression was increased along the developmental zones in both cell types (Figure 11I). These results suggest that BR levels and signaling activity may be increased along the developmental zones in the epidermis, consistent with the high percentage of BR-induced genes in these cell types.

Regarding BR signaling, the expression levels of *BRI1* and *BSK1* were higher in atrichoblast cells compared to trichoblast cells (Figure 12A and C). These expression levels exhibited a decrease during the maturation stage of the trichoblast. The expression levels of the BR-activated transcription factors *BZR1* and *BES1* also showed a similarly declining trend during the maturation stage in both cell types (Figure 12F-G). Conversely, the expression levels of *BAK1*, *BIN2*, and *PP2A* remained relatively stable and exhibited similar patterns in both cell types (Figure 12B, D-E).

Our results suggest that expression levels of BR biosynthesis genes (*DET2* and *ROT3*) and BR signaling genes (*BRI1* and *BSK1*) in atrichoblast were higher than those in trichoblast, particularly in the elongation and maturation stages. This finding may support that the higher percentages of BR-induced gene expression in the atrichoblast is partly due to the higher BR levels and BR signaling activity (Figure 11A).

Our findings showed that BR biosynthesis gene expression is cell-type-specific. This is consistent with previous studies using promoter reporters, which revealed that not all enzymes in the BR biosynthesis pathway are expressed in each cell type, and thus intermediates have to be transferred between cells to complete BR biosynthesis (Vukašinović et al., 2021). Expression of *DWF4* and *DET2* was reduced in trichoblasts at the elongation stage (Figure 11B and D), while *BZR1* and *BES1* expression was increased (Figure 12F-G). *DWF4* is most abundantly expressed the epidermis (Vukašinović et al., 2021) and is a key factor in determining the synthesis of BR (Kim et

al., 2006). Consistently, a previous study has shown that *DWF4* expression was high before the BR signaling activity, which was observed by the nuclear fluorescence intensity of BES1–GFP, reached its peak (Vukašinovi**ć** et al., 2021). BZR1 and BES1 positively regulate BR signaling pathway and mediate BR biosynthesis feedback regulation (Wang et al., 2002). When endogenous BR levels reach the threshold, BZR1 and BES1 directly block expression of BR biosynthesis genes, including *DWF4* and *DET2*, via negative feedback loop (Zhang & Xu, 2018).

In addition, the inactivation of BRs is a crucial factor in maintaining a steady level of BR. BIA1 and BAS1 are involved in the inactivation of BR (Roh et al., 2012; Youn et al., 2016). Our results indicated that *BIA1* emerged to be expressed during the elongation and maturation stages (Figure 11I). At the elongation stage, BIA1 is promoted by BR signaling to reduce active BR by acyl conjugation (Youn et al., 2016). *BAS1* encodes cytochrome P450 (CYP734A1), which is involved in the conversion of castasterone (CS) and brassinolide (BL) to their inactive metabolites (Roh et al., 2012). *BAS1* gene expression increased upon entering the elongating stage (Figure 11J), which was consistent with the levels of BZR1 and BES1 transcription factors (Figure 12F-G), both of which are responsible for inducing the expression of the BR catabolic gene *BAS1* (Youn et al., 2016). In atrichoblast, however, *BAS1* expression continued to increase until maturation (Figure 11J). This trajectory analysis demonstrated that the inactivation of BR started in the elongation stage, together with the increase in expression of BR biosynthesis genes and BR signaling activities.

BR signaling genes began to be expressed at the beginning of the pseudotime and maintained their expression levels until the onset of maturation (Figure 12). Interestingly, the majority of BR signaling gene expression was downregulated in the trichoblast before the atrichoblast. *BSK1*, which is activated by BRI to block BRI1-suppressor1 (BSU1), was also up-regulated in the atrichoblast until the maturity stage (Figure 12C). This suggests that the higher percentage of BR-induced genes in the

atrichoblast may be due to the higher endogenous BR levels as well as higher BR activity from the receptor BRI1 and the positive regulator BSK1. The other BR signaling components, such as BAK1, BIN2 and PP2A, did not show variation in expression levels along the pseudotime, suggesting that activity of these proteins, which are kinases and phosphatases, are mostly regulated at the post-translational modification level, rather than at the transcriptional level.
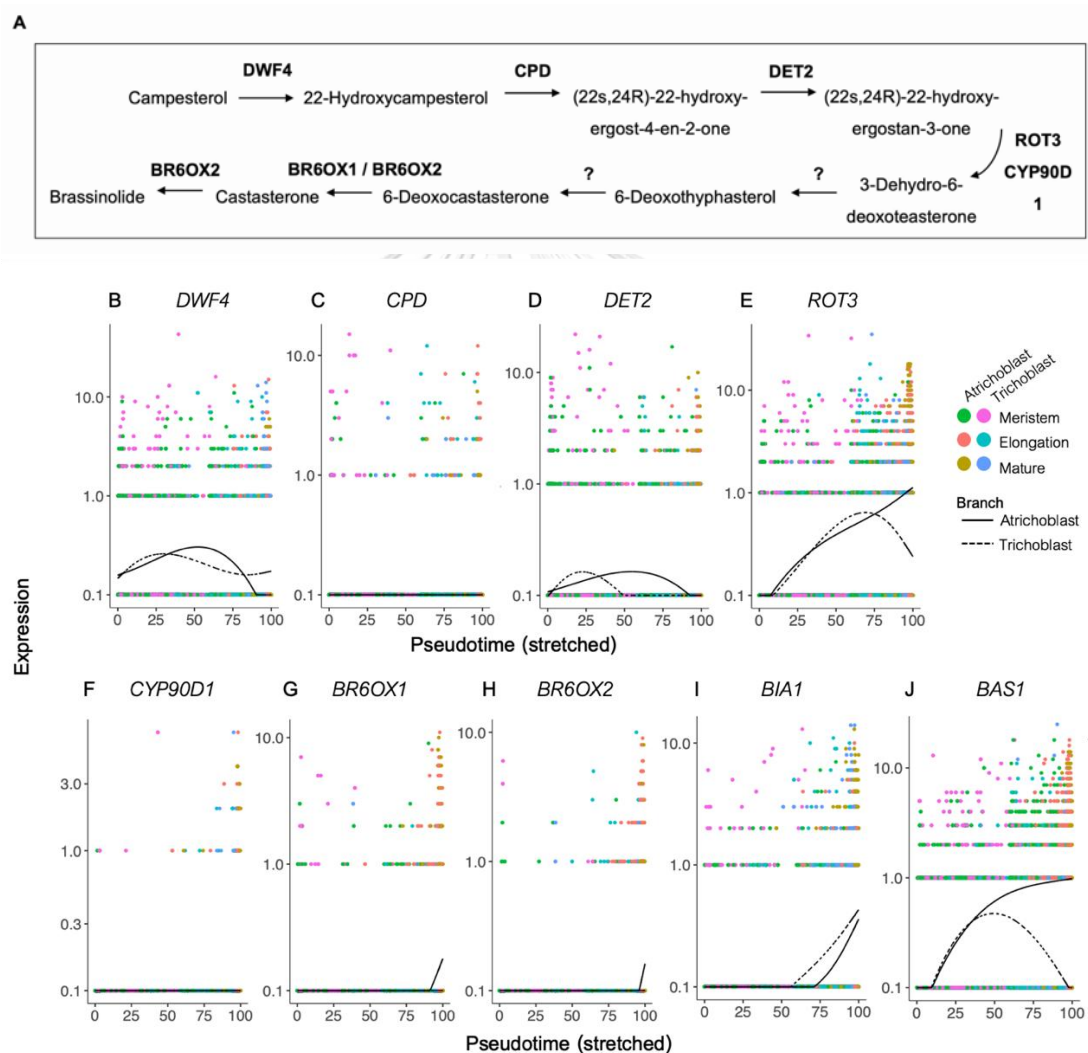
Figure 11 Visualization of the expression of BR biosynthesis and catabolism genes. (A) Diagram of the BR biosynthesis pathway with enzymes indicated in bold. The expression of genes involved in BR biosynthesis (B-H) and catabolism (I-J) across pseudotime trajectory of epidermal development.
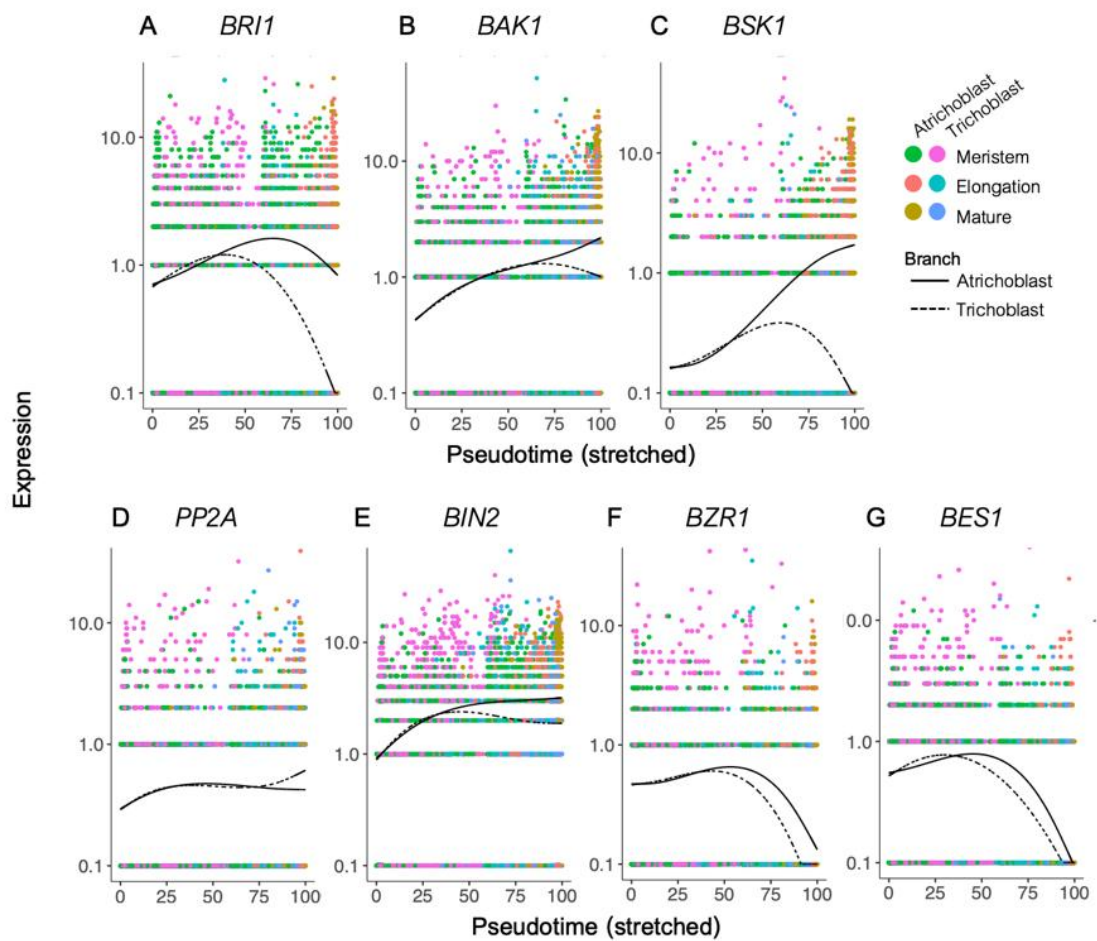
Figure 12 The expression of genes involved in BR signaling across pseudotime trajectory of epidermal development.

4.4.2 Expression of BR-responsive genes in each cluster and cell wall modification genes

Hierarchical clustering of 825 ordering genes yielded five distinct clusters (Figure 9). Subsequently, we investigated the BR-responsive genes enrichment with each cluster (Figure 10B).

Cluster 2 showed a significant enrichment of BR-repressed genes (15.38%) (Figure 10B), indicating that these genes were predominantly active during the early pseudotime corresponding to the meristematic stage (Figure 9). This was consistent with

previous studies that endogenous BR signaling remains low in the meristem zone (Chaiwanon & Wang, 2015), thereby allowing the expression of BR-repressed genes.

Cluster 3 exhibited a significant enrichment of BR-repressed genes, accounting for 20.31% of all genes within the cluster (p-value < 0.05). In contrast, clusters 4 and 5 showed a significant enrichment of BR-induced genes, comprising 61.67% and 56.92% of all genes in clusters 4-5, respectively (Figure 10B). Notably, clusters 3-5 predominantly consisted of genes involved in the differentiation and development of epidermal cells (Figure 10C). These differential patterns of BR-responsive genes among clusters suggest that BR signaling plays different roles in regulating trichoblast and atrichoblast development.

We investigated the GO terms associated with the BR-responsive genes set in cluster 3-5 (FDR < 0.05). The enrichment of BR-repressed gene in cluster 3 was related to plant-type cell wall organization, hydrogen peroxide catalytic process, and developmental growth. Cluster 4 and 5 showed enrichment in similar GO terms. The BR-induced genes in both clusters were associated with biological processes such as water transport, plant epidermis development, and root morphogenesis. Xyloglucan metabolic process was considerably found in cluster 5.

Within cluster 3, BR-repressed genes associated with the "plant-type cell wall organization" term included *EXTENSINs* (*EXTs*) gene family, specifically *EXT6–10* (Figure 13). These genes exhibited a sharp increase in expression towards the end of the pseudotime and displayed specificity for trichoblast cells at the maturation stage (Figure 13). This was in accordance with a previous finding that during the transition to the differentiating stage, these genes encode extensins to form crosslinked networks, leading to the reinforcement and rigidity of the cell wall (Somssich et al., 2016).
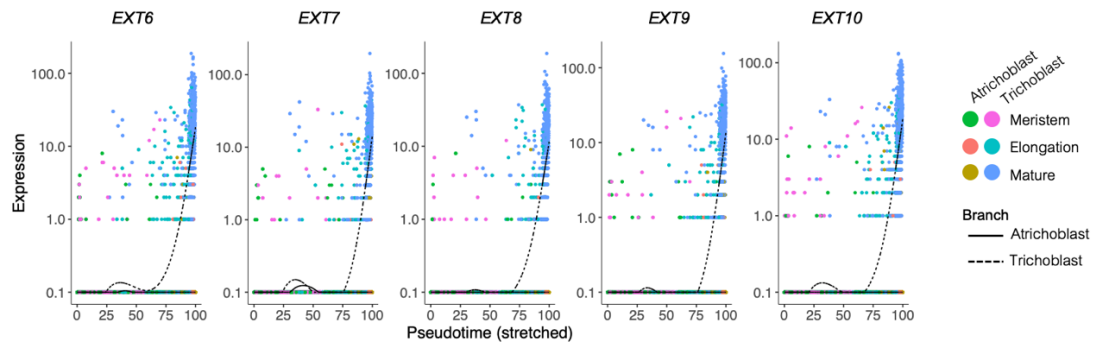
Figure 13 The expression of EXTENSINs (*EXTs*) genes across pseudotime trajectory of epidermal development

Cluster 5, consisting of BR-induced genes associated with the "xyloglucan metabolic process," exhibited the presence of xyloglucan endotransglucosylase/hydrolase (*XTHs*) gene family members, specifically *XTH4* and *XTH17-19* (Figure 14). These genes are known to play a role in enhancing cell wall extensibility (Somssich et al., 2016) and their expression can extend into the differentiating zone (Vissenberg et al., 2005). Initially, the expression of these genes showed consistency in both cell types during the early stages of pseudotime. However, as pseudotime progressed towards the midpoint, atrichoblast cells displayed a gradual increase in gene expression levels, reaching higher levels than trichoblasts (Figure 14).
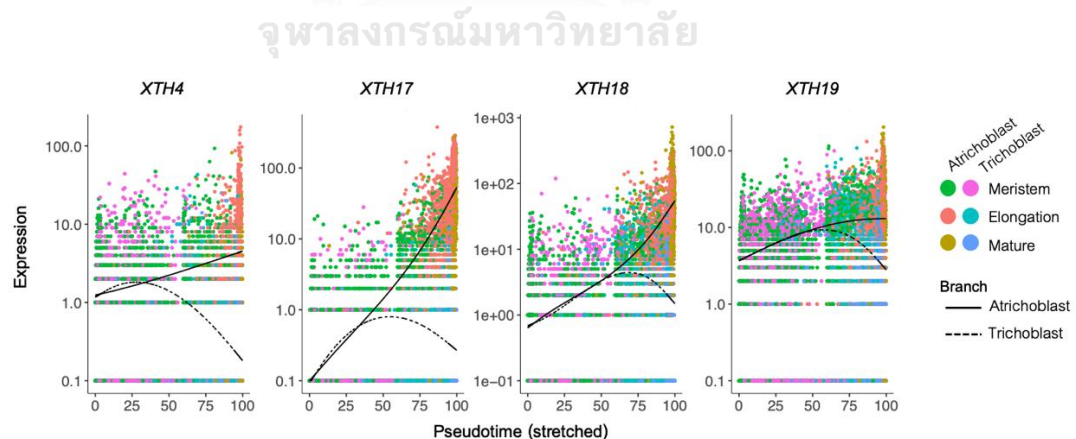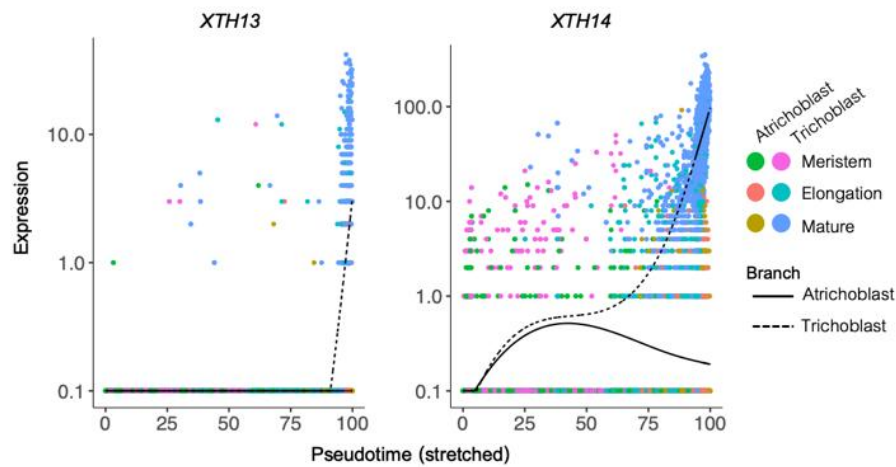

Figure 14 The expression of xyloglucan endotransglucosylase/hydrolase (*XTHs*) genes across pseudotime trajectory of epidermal development

Figure 15 The expression pattern of xyloglucan endotransglucosylase/hydrolase (*XTHs*) genes specifically observed in trichoblasts during the maturation stage.

During the later stages of pseudotime, genes involved in cell wall processes exhibited distinct expression patterns. Specifically, some *XTHs* were predominantly expressed in atrichoblasts, while some *EXTs* were predominantly expressed in trichoblasts.

Previous research has shown that the xyloglucan endotransglucosylase (XET) enzyme activity of *XTH* was distributed throughout the root, with enhanced activity observed in the elongating region and trichoblast cells (Vissenberg et al., 2001). In our analysis, we observed that atrichoblasts had higher expression levels of BR biosynthesis and signaling genes (Figures 11 and 12) during the elongation and maturation stages. This may contribute to the upregulation of *XTHs* expression in atrichoblasts compared to trichoblasts. The upregulation of *XTHs* is important for providing the necessary cell wall extensibility required for atrichoblasts to reach their mature size (Somssich et al., 2016).

The specific expression patterns of *EXTs* in trichoblasts (Figure 13) correlated with the decreased expression of *BZR1* and *BES1* in these cells prior to the end of pseudotime (Figures 12F-G). The reduced availability of these transcription factors leads to a decrease in the regulation of induced genes, allowing for the expression of BR-repressed genes (Mao & Li, 2020). The activation of EXTs in trichoblasts can be mediated by auxin and the oxidation of tyrosine via peroxidase (Schnabelrauch et al.,

1996). In addition to the enrichment findings, we identified *XTH13* and *XTH14* as exhibiting specific expression in trichoblasts during the maturation stage. Both of these genes are BR-repressed, whereas *XTH14* has been previously reported to have an expression pattern independent of hormone pathways (Vissenberg et al., 2001).

These findings suggest that, towards the end of pseudotime, the regulation of cell wall extensibility and assembly in atrichoblasts and trichoblasts is influenced by BR signaling. However, in trichoblasts, alternative processes might also be involved in the regulation of cell wall processes through trichoblast-specific genes.

4.5 Integration of single-cell ATAC-seq and single-cell RNA-seq to relate gene expression with chromatin accessibility.

4.5.1 Integration of single-cell ATAC-seq and single-cell RNA-seq data

We used Signac package (Stuart et al., 2019) to integrate data obtained from scATAC-seq (Dorrity et al., 2021) and scRNA-seq (Shahan et al., 2022) methodologies to investigate variations in chromatin structure within closely related subsets of epidermal cells. We specifically focused on extracting cell clusters that corresponded to the epidermis from the scATAC-seq dataset (Dorrity et al., 2021). In the previous study, scATAC-seq data were utilized to estimate gene transcriptional activity by quantifying ATAC-seq counts within the gene body and 400 bp upstream of the transcription start site (Dorrity et al., 2021). The obtained quantification data was subsequently transformed into a gene activity matrix for further analysis.

Common features, known as anchors, were identified between the gene activity matrix and the gene expression profile obtained from the scATAC-seq and scRNA-seq methods, respectively. Additionally, the underlying correlation structure of these common features was estimated. Leveraging these anchors, the annotation from the scRNA-seq dataset was effectively transferred to the scATAC-seq cells, enabling the prediction of cell types within the scATAC-seq data (Figure 16). Figure 16B depicts the predicted epidermal cell types of the scATAC-seq data, clearly showing that trichoblast and atrichoblast cells were separated on different sides of the UMAP.
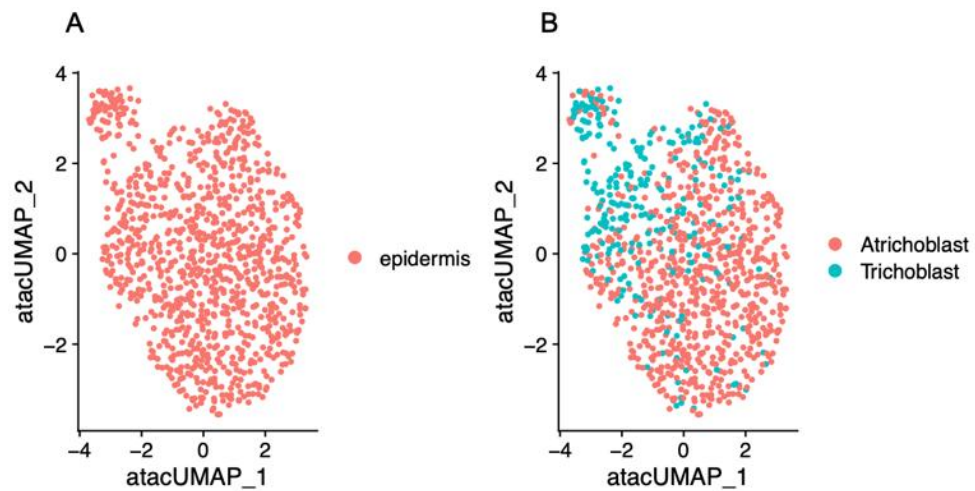
Figure 16 The UMAP visualization displaying the cell type labels of the scATAC-seq data. (A) Before and (B) after the prediction of cell types.

To achieve a comprehensive visualization of cells obtained from both scRNA-seq and scATAC-seq experiments, a shared low-dimensional space was created where the cells from both datasets were co-embedded. The same set of anchors used for transferring cell type annotations was also employed for imputing the expression data of the scATAC-seq cells. Subsequently, the expression data of variable features obtained from both scRNA-seq and scATAC-seq experiments were integrated. The UMAP technique was employed to generate a comprehensive visualization of all cells collectively. The analysis revealed that the cells generated by both techniques exhibited homogeneity (Figure 17A). Moreover, cells with similar characteristics demonstrated co-localization regardless of their data source (Figures 17B-C). This observation can be attributed to the imputation of gene expression data onto scATAC-seq cells, facilitating the grouping and clustering of identical cells from both datasets (Stuart et al., 2019). Additionally, we employed visualizations to examine the distribution of cells based on their developmental stages in the scRNA-seq data (Figure 18). This aimed to compare the positioning of scATAC-seq cells within these developmental stages.

Thus, employing this approach allows us to gain insights into the diverse cell types and their specific attributes within the scATAC-seq data, despite the absence of direct gene expression measurements in those cells.
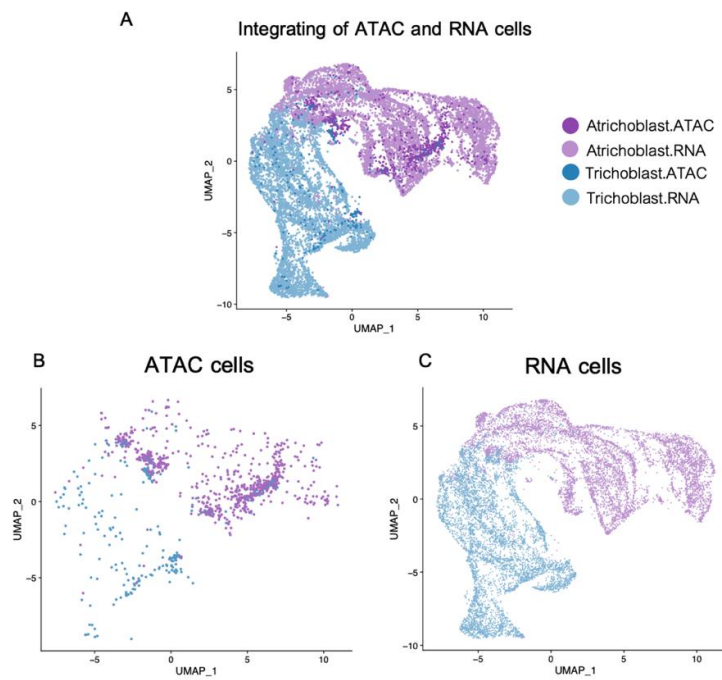
Figure 17 The identification of scATAC-seq epidermal cell types through data transfer from scRNA-seq data.

(A) Combined representation of scATAC-seq and scRNA-seq cells demonstrating co-localization of identical cell types obtained from distinct experiments. UMAP visualization exclusively displaying cells derived from scATAC-seq (B) or scRNA-seq (C) experiments.
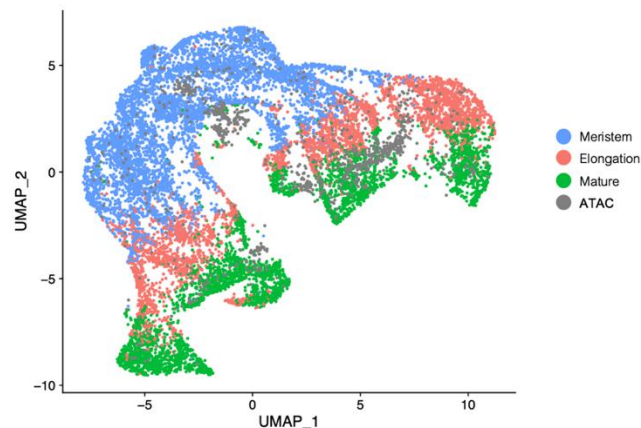
Figure 18 UMAP visualization displaying the integration of scATAC-seq and scRNA-seq epidermal cells.

Cells were color-coded to represent different developmental stages.

4.5.2 Analysis of gene expression patterns and chromatin accessibility associated with epidermal cell development.

We retrieved and annotated accessible region in available scATAC-seq data (Dorrity et al., 2021). Our aim was to investigate patterns of open chromatin and their correlation with gene expression in different epidermal cell types across developmental stages. By using the ChIPpeakAnno package, we annotated accessible peaks and linked them to their nearest genes, providing insights into the regulatory landscape (Zhu et al., 2010). Additionally, we focused on annotating accessible regions associated with BR biosynthesis pathways to gain a better understanding of gene regulation and cell-specific functions related to BR response. In the integrated atlas of scATAC-seq and scRNA-seq data, single cells obtained from scATAC-seq were assigned colors based on peaks associated with their nearest target genes. The objective was to examine the distribution of scATAC-seq cells in relation to the developmental stages of scRNA-seq cells (Figures 19-20).

We conducted peak annotation analysis, which revealed upstream open chromatin regions associated with genes that play crucial roles in the BR biosynthesis pathway, including *DWF4*, *CPD*, *BR6OX1*, and *BR6OX2* (Figure 19). Furthermore, the ChIP-seq peaks corresponding to the BZR1 transcription factor (Oh et al., 2014) were

discovered within the open chromatin regions of these BR-repressed genes, suggesting that these regions may serve as promoter regions. The accessible regions upstream of *DWF4* (Figure 19A) were observed in epidermal cell types at all stages of development. In the elongation stage of the trichoblast, the signal of the upstream region of *DWF4* was smaller than in other stages, but it began to emerge in the maturation stage. This trend aligned with the gene expression pattern across pseudotime (Figure 11A), where the expression of *DWF4* in trichoblasts decreased during the elongation stage and increased during the maturation stage. Surprisingly, despite the decreased expression of *DWF4* during the maturation stage of the atrichoblast, we discovered an accessible region upstream of *DWF4*. This finding suggests that the regulation of *DWF4* expression during this stage may be influenced by the transcriptional control exerted by the BZR1 transcription factor through a negative feedback mechanism (Zhang & Xu, 2018).

Peak signals were detected upstream of the *CPD*, *BR6OX1*, and *BR6OX2* genes in both types of epidermal cells (Figures 19B-D). However, gene expression in these cells was found to be rare (Figures 11C, G, and H). This can be explained by the cell type-specific expression patterns of these genes. Previous studies reported that *CPD-GFP* expression predominantly occurs in procambial and central columella cells, while *BR6OX1-GFP* and *BR6OX2-GFP* expressions are primarily observed in the endodermis and pericycle (Vukašinović et al., 2021).

Accessible regions associated with genes involved in BR catabolism, specifically *BAS1* and *BIA1*, were identified both upstream and overlapping with the gene body (IncludeFeature) (Figure 20). These regions were observed in both types of epidermal cells across all developmental stages. Analysis of gene expression patterns revealed that the process of BR catabolism initiates during the transition from the elongation phase to the maturation stages (Figures 11I-J). Thus, despite the presence of accessible regions of *BIA1* in epidermal cells during the meristematic stage (Figure 20B), the limited levels of BR signaling hinder the activation of gene expression associated with these genes (Youn et al., 2016).

Furthermore, our analysis revealed the presence of peaks associated with *BAS1* in both cell types at the maturation stage (Figure 20A). However, we observed expression of *BAS1* exclusively in the atrichoblast, while its expression decreased in the trichoblast. Trajectory analysis indicated that the atrichoblast exhibited higher endogenous BR levels and increased BR activity during the maturation stage. The elevated BR activity plays a crucial role in regulating the binding of BZR1 to the regulatory elements, thereby recruiting chromatin remodelers to activate the expression of *BAS1*.

As a result, while the chromatin accessibility patterns in these two cell types may exhibit similarities, there are variations in gene expression when responding to BR.
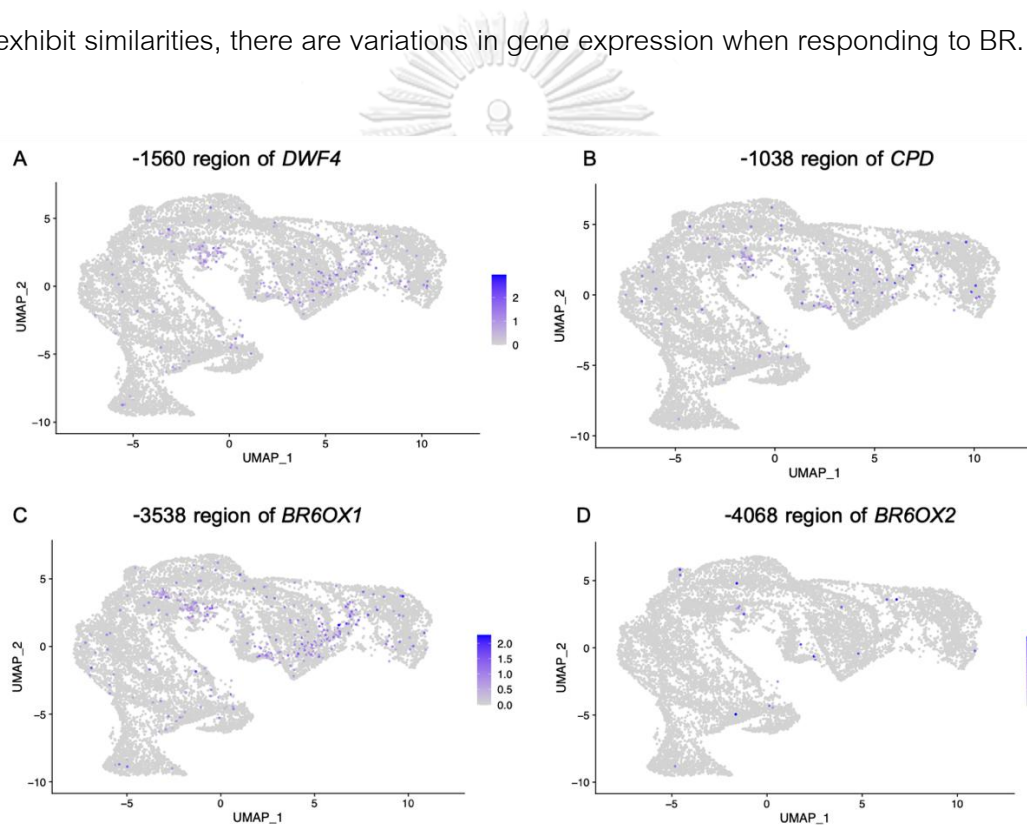
Figure 19 UMAP visualization of all cells, with colors representing the accessibility of genes involved in BR biosynthesis.
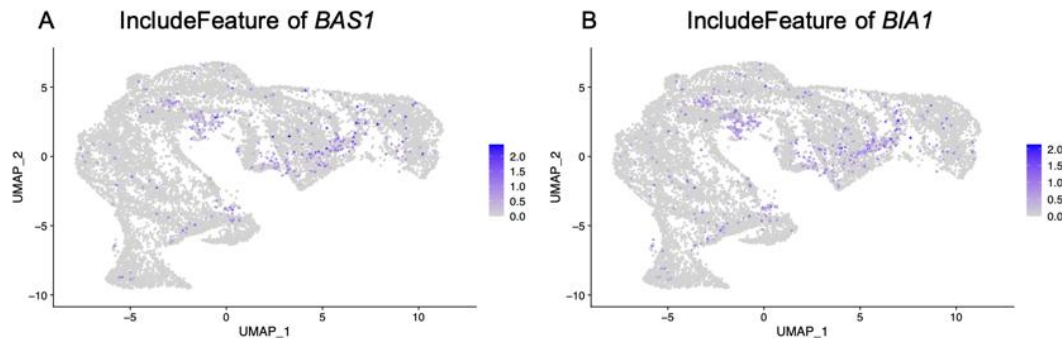
Figure 20 UMAP visualization of all cells, with colors representing the accessibility of genes involved in BR catabolism.

Our study also examined the patterns of open chromatin regions that were annotated in close to genes associated with cell wall organization and extensibility. Regions were identified upstream of and overlapping with the transcription start sites of *EXT6*, *XTH13, XTH17*, and *XTH19*. The region that was located upstream and overlaps the gene body of *XTH4* was also found in this study.

Although *EXT6* exhibited specific expression in trichoblasts during the maturation stage (Figure 13), the presence of peak signals near this gene was infrequently observed (Figure 21A). Similarly, *XTH13* showed specific expression in trichoblasts during maturation stage (Figure 15) while peak signals were detected in both epidermal cell types at all developmental stages (Figure 21B). Signs of peak nearby *XTH4* and *XTH19* were observed in both cell types and all development stages (Figures 21C, E), despite these genes being expressed in the atrichoblast rather than the trichoblast during the maturation stage (Figure 14). However, the accessibility pattern near *XTH17* exhibited consistent with the gene expression level. Accessible regions near *XTH17* were identified throughout all stages of atrichoblast development, with the largest number of regions found during the maturation stage, while being limited in the trichoblasts (Figure 21D).

In this study, we observed that open chromatin regions were annotated in proximity to some genes. However, the presence of open chromatin did not always correlate with gene expression patterns. Interestingly, gene expression can occur

without corresponding open chromatin regions. These findings suggest that chromatin accessibility may not consistently align with gene expression, and the presence of nearby accessible sites might not precisely predict expression levels (Dorrity et al., 2021). The variations in the age of *Arabidopsis* root samples employed in previous studies can result in disparities in chromatin accessibility and gene expression profiles. These characteristics may undergo changes at distinct time points during root growth and development. Thus, this disparity contributes to the observed inconsistency in the integration of chromatin accessibility and gene expression profiles.

In addition to chromatin accessibility, gene expression is regulated by various processes, including those that influence the structure and organization of chromatin, such as histone modification, DNA methylation, and the binding of specific transcription factors to regulatory elements. These interactions facilitate the recruiting of chromatin remodelers and RNA polymerase to enhancers and promoters, contributing to the regulation of gene expression in a cell type-specific manner (Stewart-Morgan et al., 2019).
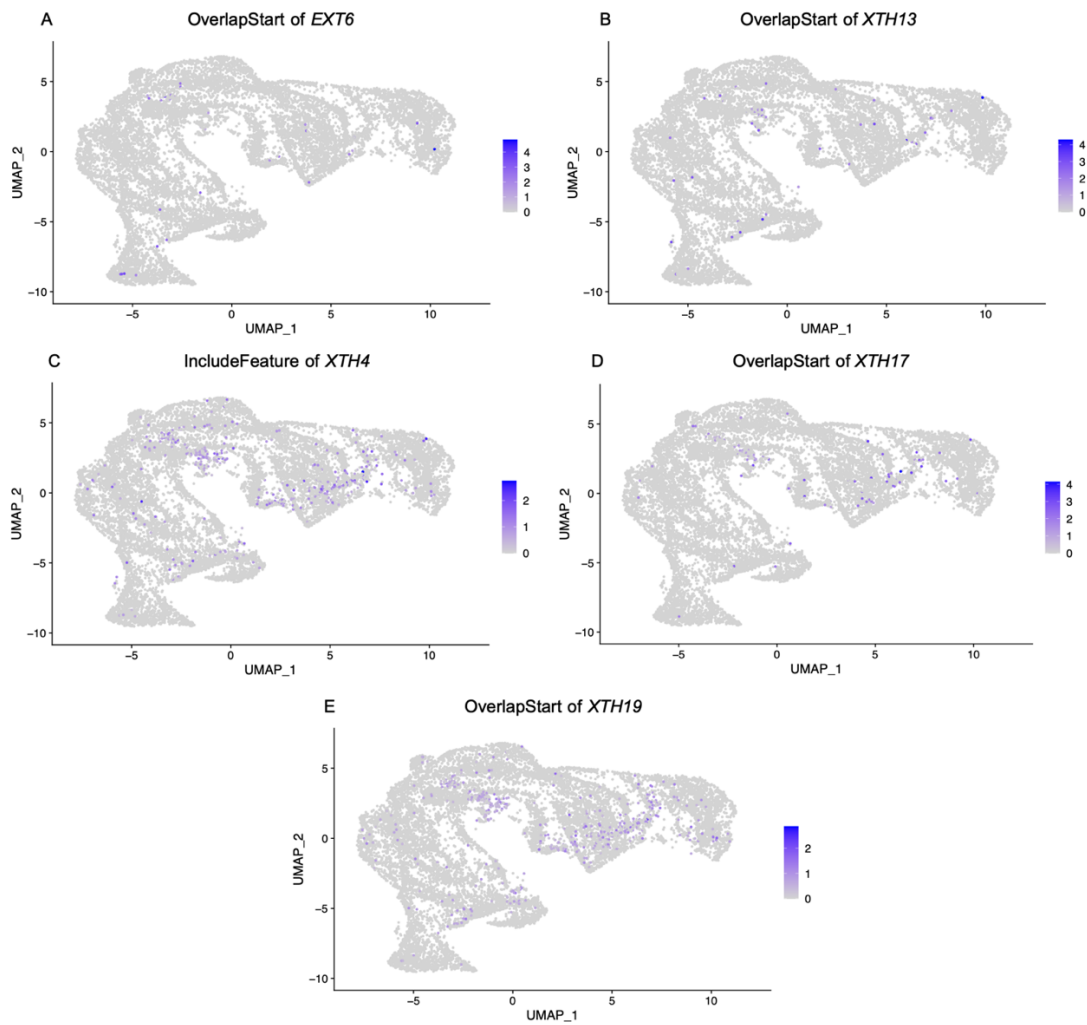
Figure 21 UMAP visualization of all cells, with colors representing the accessibility of genes involved cell wall organization (A) and extensibility (B-F).

4.6 Validation of findings across multiple datasets

The expression patterns of candidate genes identified in the study were examined on additional scRNA-seq datasets (Denyer et al., 2019; Ryu et al., 2019) to ensure the consistency of conclusions across different datasets. Data from Denyer et al. (2019) and Ryu et al. (2019), both of which are part of the Shahan datasets (Shahan et al., 2022), were plotted individually to demonstrate their ability to capture cell types at any developmental stage. Figure 22 demonstrated that the two additional datasets contained epidermal cell types, including trichoblasts and atrichoblasts, across all developmental stages. Comparing the plots revealed that both datasets effectively represented epidermal cell types at different stages. This indicates their sufficiency in validating gene expression patterns compared to the previous findings analyzed from the large dataset of Shahan et al. (2022).
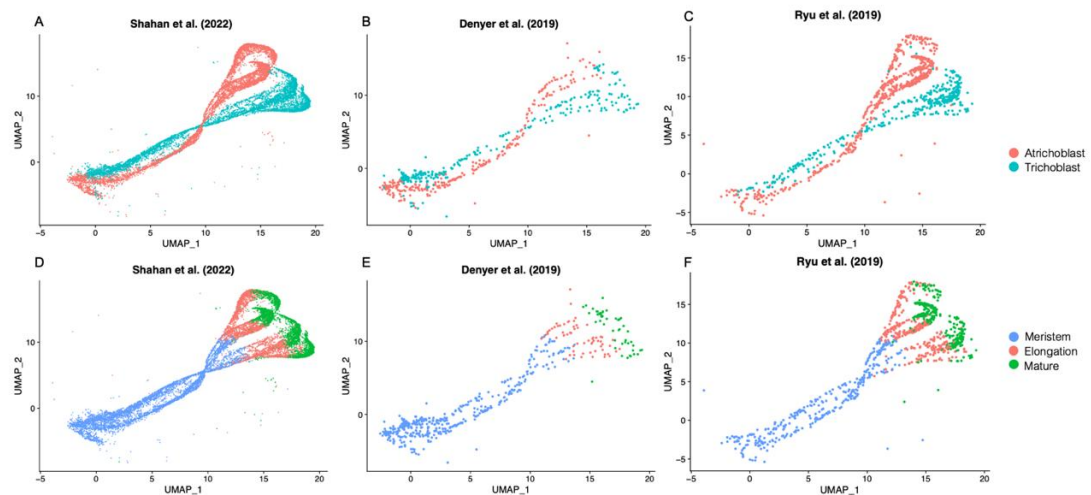
Figure 22 Single-cell RNA-seq datasets of epidermal cells were employed in the study. These included a large dataset from Shahan et al. (2022) (A, D). Additional datasets obtained from Denyer et al. (2019) (B, E) and Ryu et al. (2019) (C, F) were utilized for validation purposes. Cells within these datasets were categorized according to their cell type (A–C) and developmental stage (D–F) annotations.

The expression patterns of genes involved in the BR biosynthesis and signaling pathways were examined in additional datasets (Figures 23-26), and they were found to exhibit considerable consistency with the patterns observed across pseudotime (Figures 11-12). Specifically, the expression of *DWF4* and *DET2* showed an onset at the meristematic stage in both cell types (Figures 23A, C and 25A, C).

The expression of *ROT3* and *BAS1* increased as the developmental stage progressed (Figures 23D, H and 25D, H). However, the other dataset (Ryu et al., 2019) revealed distinct differences between two cell types, with higher levels of gene expression observed in atrichoblasts compared to trichoblasts during the elongation to maturation stages (Figures 25D and H).

BIA1 exhibited increased expression during the maturation stage in both cell types (Figures 23I and 25I), although the expression levels were higher in trichoblasts. *CPD, CYP90D1, BR6OX1,* and *BR6OX2* showed infrequent expressions in both datasets (Figures 23B, E, F, G and 25B, E, F, G). Additionally, the expression levels of genes associated with the BR signaling pathway showed a progressive increase with developmental stages (Figures 24, 26), and the data from the Ryu et al. (2019) dataset demonstrated consistency with the gene expression patterns observed across pseudotime (Figure 26). Especially, *BSK1* exhibited higher expression levels in atrichoblasts compared to trichoblasts (Figures 12C, 26C).

The expression patterns of genes involved in cell wall organization, such as the *EXTs* family, and cell extensibility, including *XTH13* and *XTH14*, showed specificity towards trichoblasts during the maturation stage, as observed in both datasets (Figures 27A-E, J-K, and 28A-E, J-K). These findings were consistent with the visualization results obtained from the dataset of Shahan et al. (2022) (Figure 13).

In contrast, the genes *XTH4, XTH17, XTH18,* and *XTH19* demonstrated a preference for expression in atrichoblasts, as evidenced by their expression patterns in Figures 27F-I and 28F-I. Notably, the dataset of Ryu et al. (2019) displayed higher expression levels of these genes in atrichoblasts during the elongation stage (Figure 28F-

I), which aligned with the expression patterns observed in the dataset of Shahan et al. (2022) (Figure 14).

In summary, the two supplementary datasets demonstrated gene expression visualizations consistent with the patterns observed in the larger datasets (Shahan et al., 2022). The dataset of Ryu et al. (2019) included a greater number of cells compared to the other dataset (Denyer et al., 2019), allowing for visualizations that are closely similar to those of a larger dataset (Shahan et al., 2022). This validation of results helps to ensure consistency and minimize the impact of biases specific to each dataset. Moreover, this analysis validated the distinct responses of different epidermal cell types to BR and highlighted the importance of a specific group of cell wall-related genes in determining the fate of the atrichoblast.
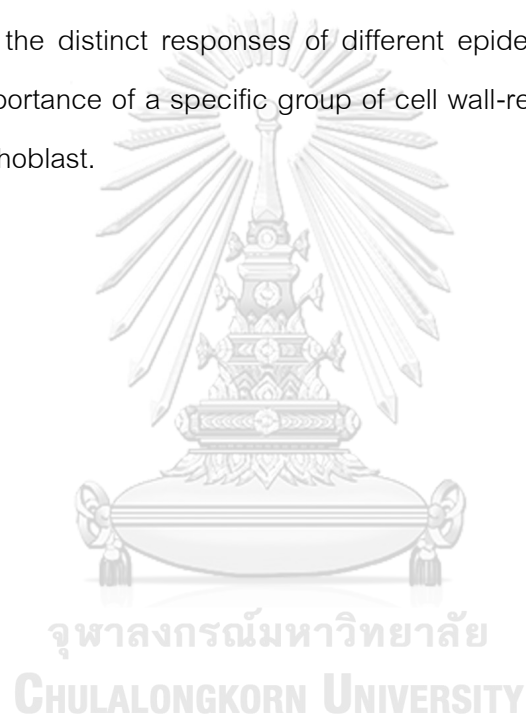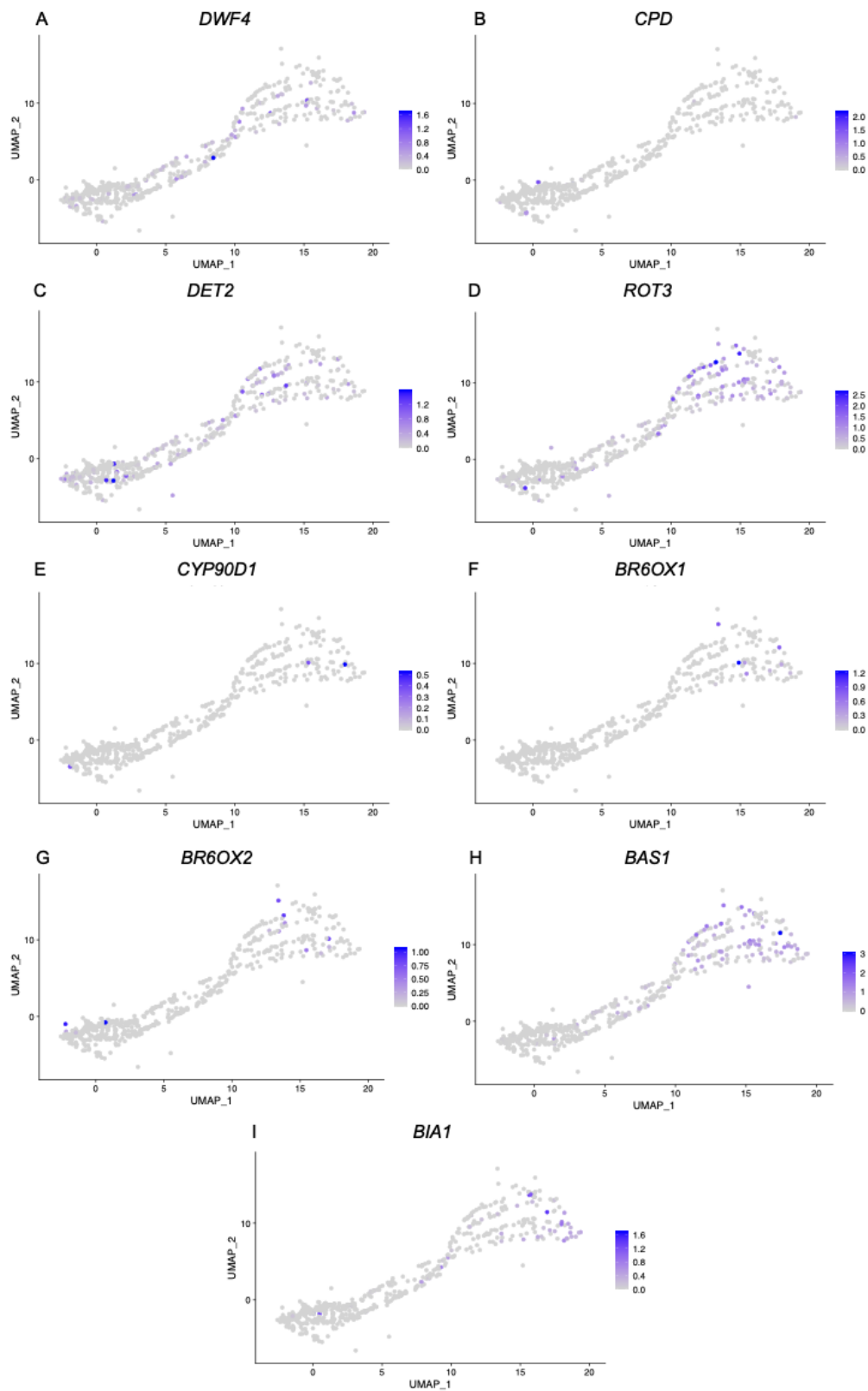
Figure 23 Validation of gene expression patterns in the BR biosynthesis pathway using the dataset from Denyer et al. (2019).
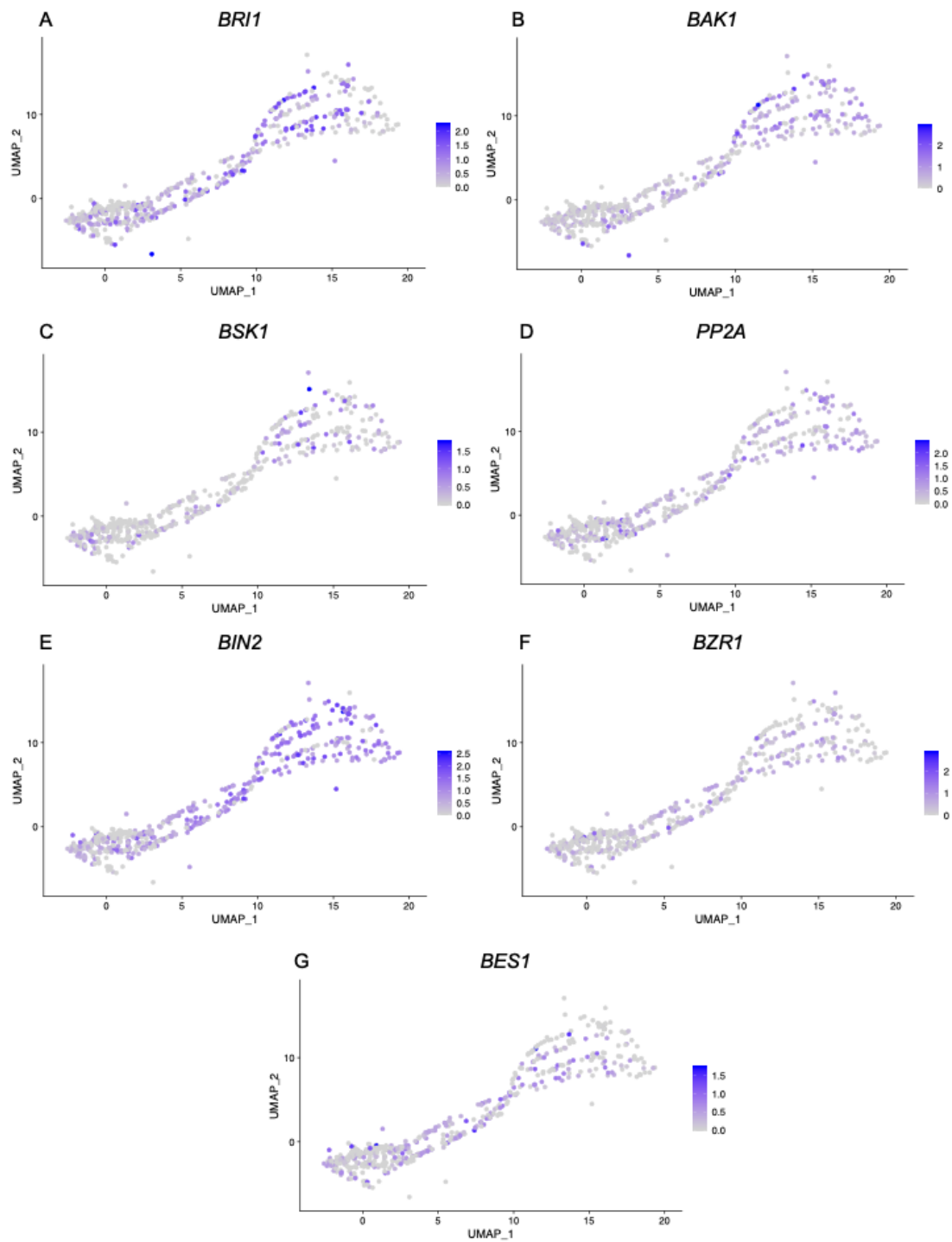
Figure 24 Validation of gene expression patterns in the BR signaling pathway using the dataset from Denyer et al. (2019).
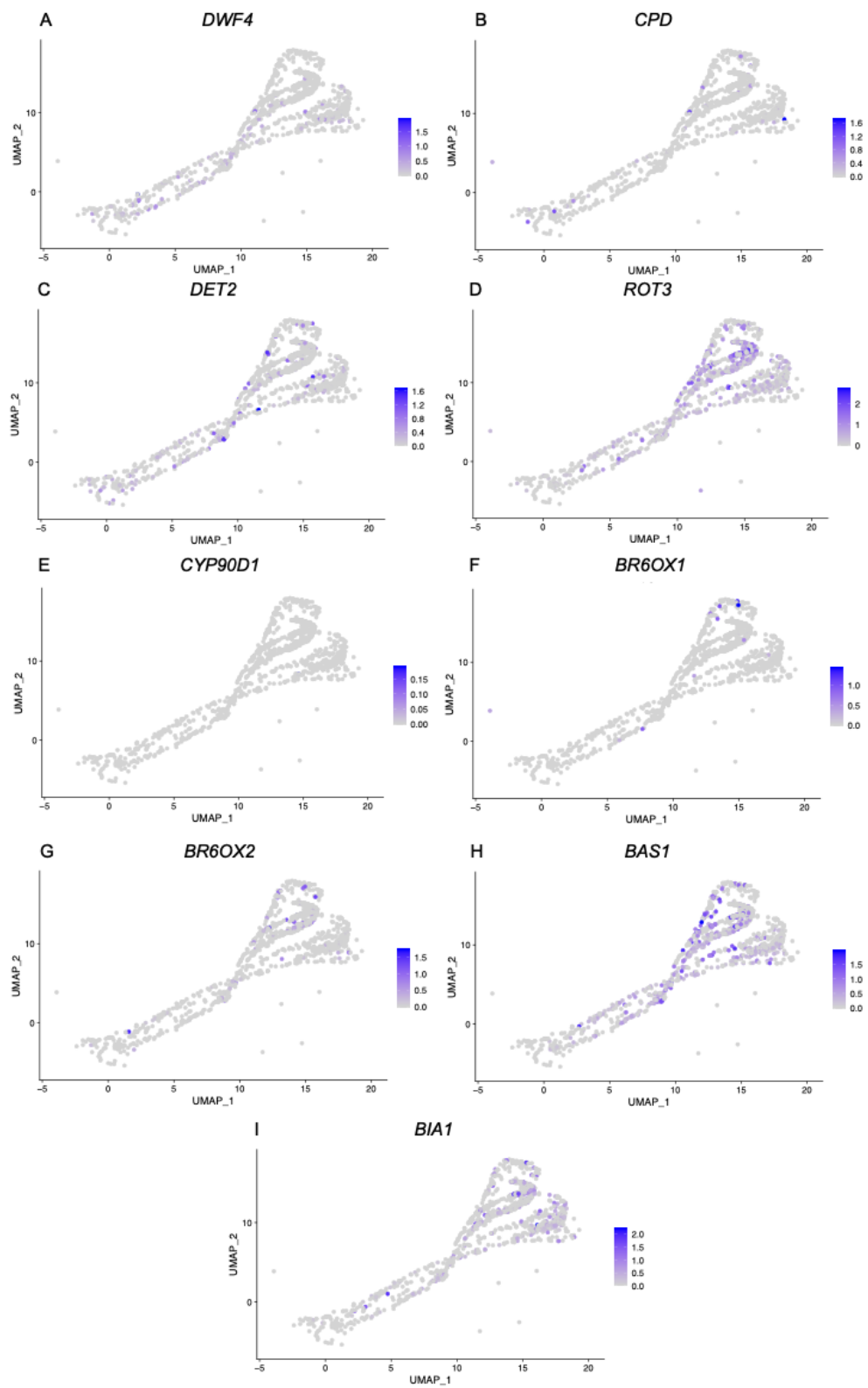
Figure 25 Validation of gene expression patterns in the BR biosynthesis pathway using the dataset from Ryu et al. (2019).
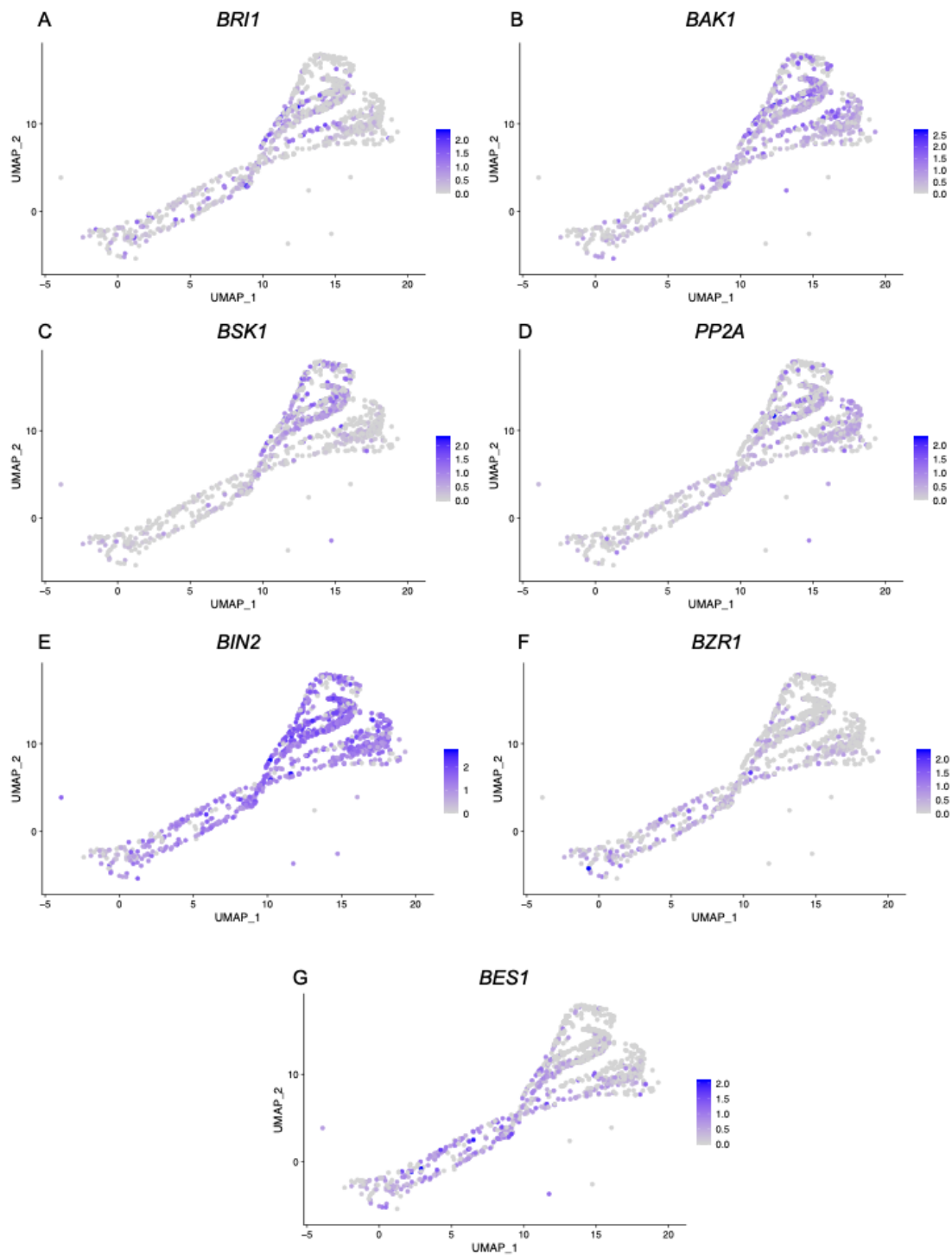
Figure 26 Validation of gene expression patterns in the BR signaling pathway using the dataset from Ryu et al. (2019).

Figure 27 Validation of gene expression patterns involving cell wall organization (A-E) and extensibility (F-K) pathways using the dataset from Denyer et al. (2019).

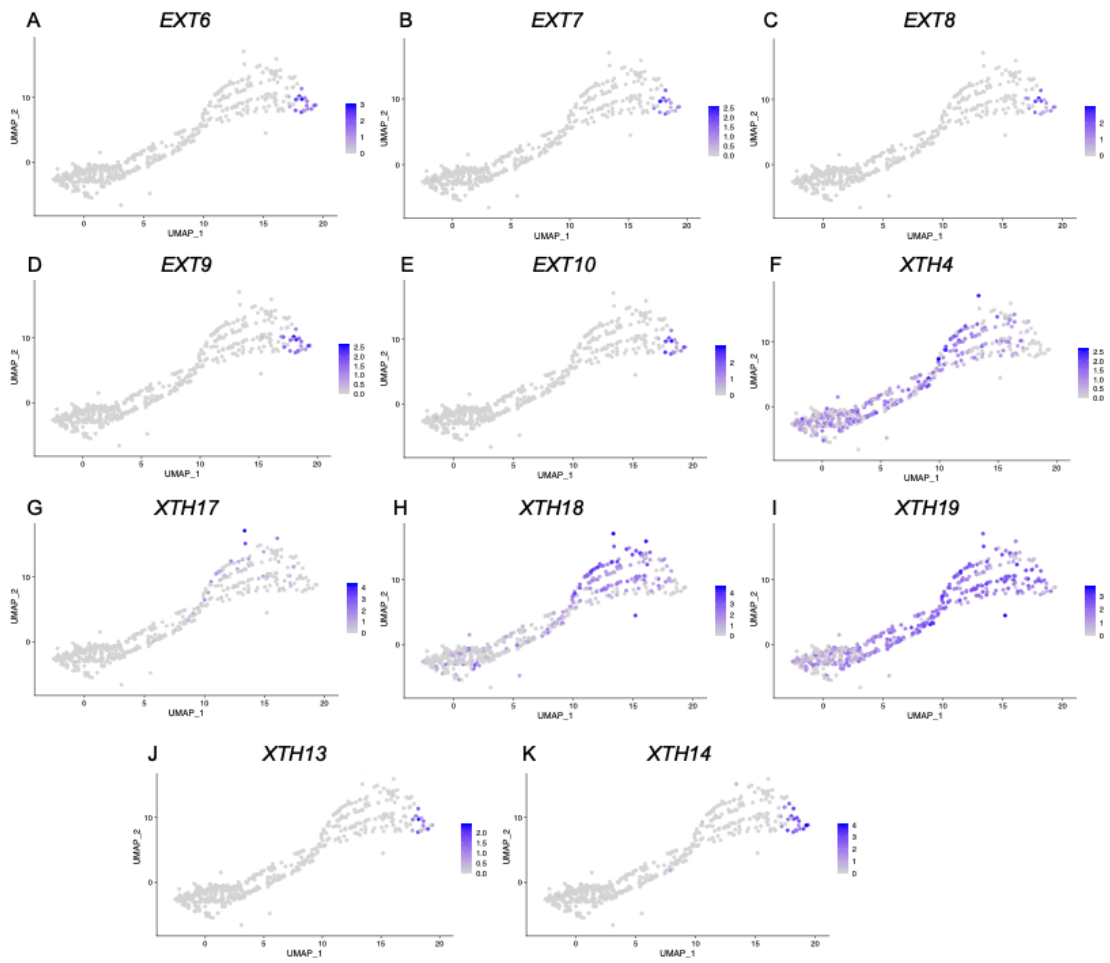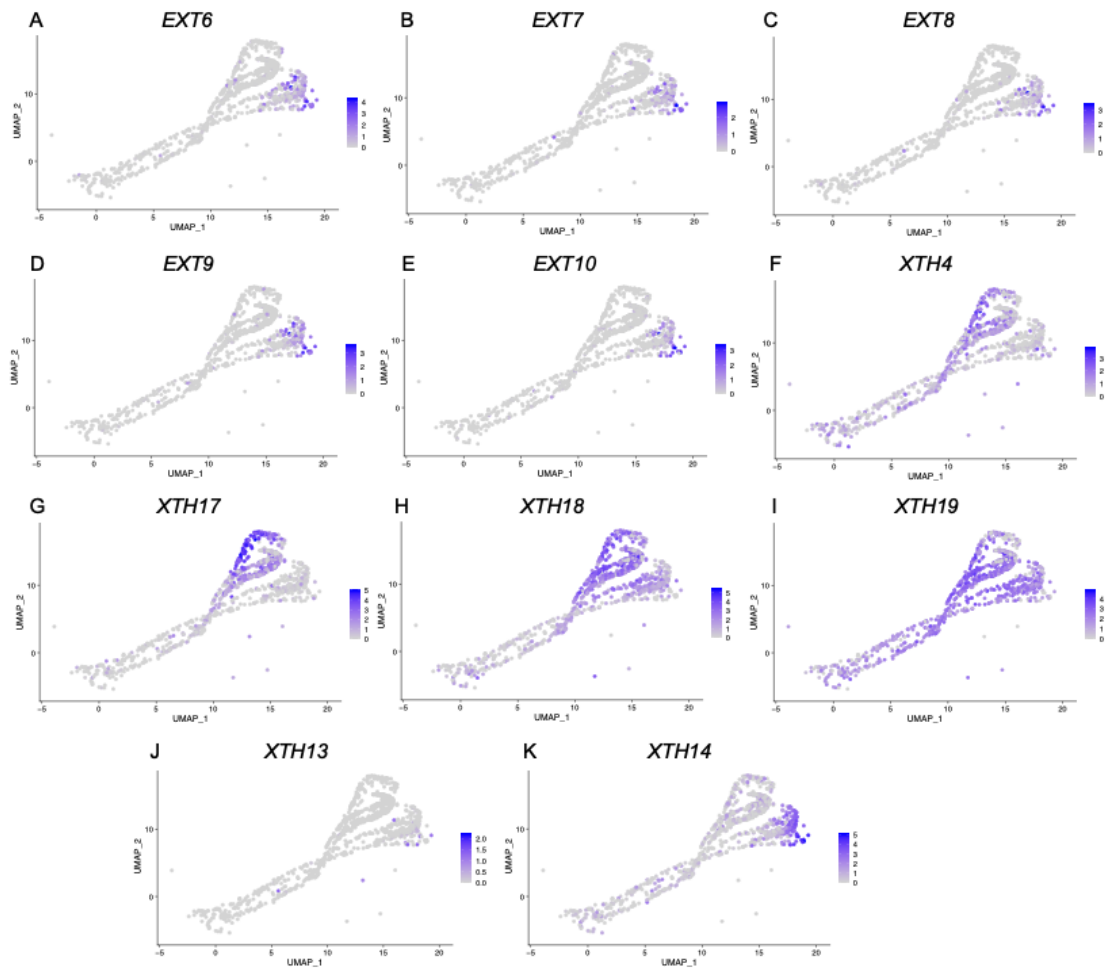**Figure 28** Validation of gene expression patterns involving cell wall organization (A-E) and extensibility (F-K) pathways using the dataset from Ryu et al. (2019).

PART 5

CONCLUSIONS AND RECOMMENDATIONS

In this study, the scRNA-seq data of *Arabidopsis* roots was used to support tissue-specific roles of BR in *Arabidopsis* root development. In the elongation and maturation zones, we showed that the atrichoblast had the highest enrichment of BR-induced genes among all cell types. The utilization of scRNA-seq allowed for the reconstruction of the differentiation trajectory of epidermal cells and provided insights into the expression patterns of signature genes at different developmental stages. We observed variations in the expression of genes involved in BR biosynthesis and signaling pathways between atrichoblasts and trichoblasts throughout the developmental stages. The expression levels of BR biosynthesis genes (*DET2* and *ROT3*) and BR signaling genes (*BRI1* and *BSK1*) were higher in atrichoblast than in trichoblast, particularly during the elongation and maturation stages. This finding may provide evidence that the larger percentages of BR-induced gene expression in the atrichoblast were due to the greater amounts of BR and BR signaling activity in the cell type.

Our analysis revealed distinct patterns of gene expression in response to BR across different clusters of ordering genes that shaped the trajectory of epidermal cells. Gene ontology (GO) term analysis revealed that BR-repressed genes in cluster 3 were associated with plant-type cell wall organization, while BR-induced genes in clusters 4 and 5 were associated with water transport, plant epidermis development, and root morphogenesis. Interestingly, the *EXT* gene family, involved in cell wall reinforcement, exhibited specific expression in trichoblasts at the maturation stage, while the *XTH* gene family, involved in cell wall extensibility, displayed higher expression in atrichoblasts during elongation and maturation stages. This finding suggests that BR signaling plays different roles in regulating the development of atrichoblasts and trichoblasts, with alternative processes possibly involved in the regulation of cell wall processes in trichoblasts.

The integrated scRNA-seq and scATAC-seq approach provided insights into diverse cell types and their specific attributes in the scATAC-seq data, enhancing our understanding of gene regulation and cell-specific functions. However, the correlation between open chromatin regions and gene expression patterns was inconsistent, suggesting that chromatin accessibility alone may not accurately predict expression levels.

Based on the validation of findings using additional data, which are subsets of the comprehensive dataset of Shahan et al. (2022), it was revealed that the combined cell count from these subsets does not represent the majority of the entire dataset. Consequently, the gene expression patterns observed within the two datasets can be compared to the findings obtained from the larger dataset. This validation confirmed the robustness of the findings and minimized dataset-specific bias. The validation results also supported the distinct responses of different epidermal cell types to BR and emphasized the significance of specific cell wall-related genes in the fate determination of atrichoblasts. Nonetheless, to ensure the validity and consistency of the findings, future studies should also incorporate external data for comparative analysis.

The application of scRNA-seq in the investigation of gene expression patterns is a valuable approach that offers significant insights into cellular dynamics and the tissue-specific functions of genes. While the findings have been validated across multiple scRNA-seq datasets, experimental confirmation in the laboratory is necessary to test the distinct response of atrichoblasts and trichoblasts to BR. Further experimental research will strengthen our understanding of the distinct effects of BR on these cell types. This study provides evidence for the differential roles of BR in distinct cell types, particularly in epidermal cells. These variations in BR functions enable precise modulation of BR levels or signaling, presenting opportunities for optimizing the application of BR or BR mimics. This targeted strategy can enhance specific developmental processes and ultimately improve crop production in agriculture.

REFERENCES

Barron, M., & Li, J. (2016). Identifying and removing the cell-cycle effect from single-cell RNA-Sequencing data. **Scientific Reports, 6**(1), 33892.

Bawa, G., Liu, Z., Yu, X., Qin, A., & Sun, X. (2022). Single-cell RNA sequencing for plant research: Insights and possible benefits. **International Journal of Molecular Sciences, 23**(9), 4497.

Berger, F., Haseloff, J., Schiefelbein, J., & Dolan, L. (1998). Positional information in root epidermis is defined during embryogenesis and acts in domains with strict boundaries. **Current Biology, 8**(8), 421-430.

Bruex, A., Kainkaryam, R. M., Wieckowski, Y., Kang, Y. H., Bernhardt, C., Xia, Y., . . .& Schiefelbein, J. (2012). A Gene regulatory network for root epidermis cell differentiation in Arabidopsis. **PLOS Genetics, 8**(1), e1002446.

Butler, A., Hoffman, P., Smibert, P., Papalexi, E., & Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. **Nature Biotechnology, 36**(5), 411-420.

Cajero Sánchez, W., García-Ponce, B., Sánchez, M. P., Álvarez-Buylla, E. R., & Garay-Arroyo, A. (2018). Identifying the transition to the maturation zone in three ecotypes of Arabidopsis thaliana roots. **Communicative & Integrative Biology, 11**(1), e1395993.

Chaiwanon, J., & Wang, Z. Y. (2015). Spatiotemporal brassinosteroid signaling and antagonism with auxin pattern stem cell dynamics in Arabidopsis roots. **Current Biology, 25**(8), 1031-1042.

Cheng, Y., Zhu, W., Chen, Y., Ito, S., Asami, T., & Wang, X. (2014). Brassinosteroids control root epidermal cell fate via direct regulation of a MYB-bHLH-WD40 complex by GSK3-like kinases. **Elife, 3**, e02525.

Clouse, S. D., Langford, M., & McMorris, T. C. (1996). A brassinosteroid-insensitive mutant in Arabidopsis thaliana exhibits multiple defects in growth and development. **Plant Physiology, 111**(3), 671-678.

Cusanovich, D. A., Hill, A. J., Aghamirzaie, D., Daza, R. M., Pliner, H. A., Berletch, J. B., . . .& Shendure, J. (2018). A single-cell atlas of in vivo mammalian chromatin accessibility. **Cell, 174**(5), 1309-1324.e1318.

Deconinck, L., Cannoodt, R., Saelens, W., Deplancke, B., & Saeys, Y. (2021). Recent advances in trajectory inference from single-cell omics data. **Current Opinion in Systems Biology, 27**, 100344.

Denyer, T., Ma, X., Klesen, S., Scacchi, E., Nieselt, K., & Timmermans, M. C. P. (2019). Spatiotemporal developmental trajectories in the Arabidopsis root revealed using high-throughput single-cell RNA sequencing. **Developmental Cell, 48**(6), 840-852.e845.

Dorrity, M. W., Alexandre, C. M., Hamm, M. O., Vigil, A.L., Fields, S., Queitsch, C., & Cuperus, J. T. (2021). The regulatory landscape of Arabidopsis thaliana roots at single-cell resolution. **Nature Communications, 12**(1), 3334.

Duckett, C. M., Grierson, C., Linstead, P., Schneider, K., Lawson, E., Dean, C., . . .& Roberts, K. (1994). Clonal relationships and cell patterning in the root epidermis of Arabidopsis. **Development, 120**(9), 2465-2474.

Farmer, A., Thibivilliers, S., Ryu, K. H., Schiefelbein, J., & Libault, M. (2021). Single-nucleus RNA and ATAC sequencing reveals the impact of chromatin accessibility on gene expression in Arabidopsis roots at the single-cell level. **Molecular Plant, 14**(3), 372-383.

Ferreira-Guerra, M., Marquès-Bueno, M., Mora-García, S., & Caño-Delgado, A. I. (2020). Delving into the evolutionary origin of steroid sensing in plants. **Current Opinion in Plant Biology 57**, 87-95.

Fridman, Y., Elkouby, L., Holland, N., Vragovi**ć**, K., Elbaum, R., & Savaldi-Goldstein, S. (2014). Root growth is modulated by differential hormonal sensitivity in neighboring cells. **Genes & Development, 28**(8), 912-920.

Gampala, S. S., Kim, T. W., He, J. X., Tang, W., Deng, Z., Bai, M. Y., . . .& Wang, Z. Y. (2007). An essential role for 14-3-3 proteins in brassinosteroid signal transduction in Arabidopsis. **Developmental Cell, 13**(2), 177-189.

González-García, M. P., Vilarrasa-Blasi, J., Zhiponova, M., Divol, F., Mora-García, S., Russinova, E., & Caño-Delgado, A. I. (2011). Brassinosteroids control meristem size by promoting cell cycle progression in Arabidopsis roots. **Development, 138**(5), 849-859.

Graeff, M., Rana, S., Marhava, P., Moret, B., & Hardtke, C. S. (2020). Local and systemic effects of brassinosteroid perception in developing phloem. **Current Biology, 30**(9), 1626-1638.e1623.

Grove, M. D., Spencer, G. F., Rohwedder, W. K., Mandava, N., Worley, J. F., Warthen Jr, J. D., . . .& Cook Jr, J. C. (1979). Brassinolide, a plant growth-promoting steroid isolated from Brassica napus pollen. **Nature, 281**(5728), 216-217.

Hacham, Y., Holland, N., Butterfield, C., Ubeda-Tomas, S., Bennett, M. J., Chory, J., & Savaldi-Goldstein, S. (2011). Brassinosteroid perception in the epidermis controls root meristem size. **Development, 138**(5), 839-848.

Hafemeister, C., & Satija, R. (2019). Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. **Genome Biology, 20**(1), 296.

He, J. X., Gendron, J. M., Sun, Y., Gampala, S. S., Gendron, N., Sun, C. Q., & Wang, Z. Y. (2005). BZR1 is a transcriptional repressor with dual roles in brassinosteroid homeostasis and growth responses. **Science, 307**(5715), 1634-1638.

He, J. X., Gendron, J. M., Yang, Y., Li, J., & Wang, Z. Y. (2002). The GSK3-like kinase BIN2 phosphorylates and destabilizes BZR1, a positive regulator of the brassinosteroid signaling pathway in Arabidopsis. **Proceedings of the National Academy of Sciences of the United States of America, 99**(15), 10185-10190.

Hong, R., Koga, Y., Bandyadka, S., Leshchyk, A., Wang, Y., Akavoor, V., . . .& Campbell, J. D. (2022). Comprehensive generation, visualization, and reporting of quality control metrics for single-cell RNA sequencing data. **Nature Communications, 13**(1), 1688.

Huysmans, M., Buono, R. A., Skorzinski, N., Radio, M. C., De Winter, F., Parizot, B., . . .& Nowack, M. K. (2018). NAC transcription factors ANAC087 and ANAC046 control

distinct aspects of programmed cell death in the Arabidopsis columella and lateral root cap. **The Plant Cell, 30**(9), 2197-2213.

Hwang, B., Lee, J. H., & Bang, D. (2018). Single-cell RNA sequencing technologies and bioinformaticspipelines. **Experimental & Molecular Medicine, 50**(8), 1-14.

Jean-Baptiste, K., McFaline-Figueroa, J. L., Alexandre, C. M., Dorrity, M. W., Saunders, L., Bubb, K. L., . . .& Cuperus, J. T. (2019). Dynamics of gene expression in single root cells of Arabidopsis thaliana. **The Plant Cell, 31**(5), 993-1011.

Johansson, I., Karlsson, M., Johanson, U., Larsson, C., & Kjellbom, P. (2000). The role of aquaporins in cellular and whole plant water balance. **Biochimica et Biophysica Acta (BBA) - Biomembranes, 1465**(1), 324-342.

Kaldenhoff, R., Kölling, A., Meyers, J., Karmann, U., Ruppel, G., & Richter, G. (1995). The blue light-responsive AthH2 gene of Arabidopsis thaliana is primarily expressed in expanding as well as in differentiating cells and encodes a putative channel protein of the plasmalemma. **The Plant Journal, 7**(1), 87-95.

Khripach, V. A., Zhabinskii, V., & de Groot, A. E. (1998). *Brassinosteroids: A new class of plant hormones*. Academic Press.

Kim, H. B., Kwon, M., Ryu, H., Fujioka, S., Takatsuto, S., Yoshida, S., . . .& Choe, S. (2006). The regulation of DWARF4 expression is likely a critical mechanism in maintaining the homeostasis of bioactive brassinosteroids in Arabidopsis. **Plant Physiology, 140**(2), 548-557.

Kim, T.W., Guan, S., Burlingame, Alma L., & Wang, Z.Y. (2011). The CDG1 kinase mediates brassinosteroid signal transduction from BRI1 receptor kinase to BSU1 phosphatase and GSK3-like kinase BIN2. **Molecular Cell, 43**(4), 561-571.

Kim, T.W., Guan, S., Sun, Y., Deng, Z., Tang, W., Shang, J.X., . . .& Wang, Z.Y. (2009). Brassinosteroid signal transduction from cell-surface receptor kinases to nuclear transcription factors. **Nature Cell Biology, 11**(10), 1254-1260.

Kim, T.W., & Wang, Z.Y. (2010). Brassinosteroid signal transduction from receptor kinases to transcription factors. **Annual Review of Plant Biology, 61**(1), 681-704.

Kirik, V., Simon, M., Huelskamp, M., & Schiefelbein, J. (2004). The ENHANCER OF TRY AND CPC1 gene acts redundantly with TRIPTYCHON and CAPRICE in trichome and root hair cell patterning in Arabidopsis. **Developmental Biology, 268**(2), 506-513.

Klein, Allon M., Mazutis, L., Akartuna, I., Tallapragada, N., Veres, A., Li, V., . . .& Kirschner, Marc W. (2015). Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. **Cell, 161**(5), 1187-1201.

Kour, J., Kohli, S., Khanna, K., Bakshi, P., Sharma, P., Singh, A., . . .& Sharma, A. (2021). Brassinosteroid signaling, crosstalk and, physiological functions in plants under heavy metal stress. **Frontiers in Plant Science, 12**, 608061.

Kuppusamy, K. T., Chen, A. Y., & Nemhauser, J. L. (2009). Steroids are required for epidermal cell fate establishment in Arabidopsis roots. **Proceedings of the National Academy of Sciences, 106**(19), 8073-8076.

Kwak, S.H., & Schiefelbein, J. (2007). The role of the SCRAMBLED receptor-like kinase in patterning the Arabidopsis root epidermis. **Developmental Biology, 302**(1), 118-131.

Kwak, S.H., Shen, R., & Schiefelbein, J. (2005). Positional signaling mediated by a receptor-like kinase in Arabidopsis. **Science, 307**(5712), 1111-1113.

Lee, M. M., & Schiefelbein, J. (2002). Cell pattern in the Arabidopsis root epidermis determined by lateral inhibition with feedback. **Plant Cell, 14**(3), 611-618.

Li, D.X., Chen, W.Q., Xu, Z.H., & Bai, S.N. (2015). HISTONE DEACETYLASE6-defective mutants show increased expression and acetylation of ENHANCER OF TRIPTYCHON AND CAPRICE1 and GLABRA2 with small but significant effects on root epidermis cellular pattern. **Plant Physiology, 168**(4), 1448-1458.

Li, S., Yamada, M., Han, X., Ohler, U., & Benfey, P. N. (2016). High-resolution expression map of the Arabidopsis root reveals alternative splicing and lincRNA regulation. **Developmental Cell, 39**(4), 508-522.

Li, Z., & He, Y. (2020). Roles of brassinosteroids in plant reproduction. **International Journal of Molecular Sciences, 21**(3), 872.

López-Ruiz, B. A., Zluhan-Martínez, E., Sánchez, M. d. l. P., Álvarez-Buylla, E. R., & Garay-Arroyo, A. (2020). Interplay between hormones and several abiotic stress conditions on Arabidopsis thaliana primary root development. **Cells, 9**(12), 2576.

Mao, J., & Li, J. (2020). Regulation of three key kinases of brassinosteroid signaling pathway. **International Journal of Molecular Sciences, 21**(12), 4340.

Mitchell, J., Mandava, N., Worley, J., Plimmer, J., & Smith, M. (1970). Brassins—a new family of plant hormones from rape pollen. **Nature, 225**, 1065-1066.

Munné-Bosch, S., & Müller, M. (2013). Hormonal cross-talk in plant development and stress responses [Editorial]. **Frontiers in Plant Science, 4**(529).

Müssig, C., Shin, G. H., & Altmann, T. (2003). Brassinosteroids promote root growth in Arabidopsis. **Plant Physiology, 133**(3), 1261-1271.

Nam, K. H., & Li, J. (2002). BRI1/BAK1, a receptor kinase pair mediating brassinosteroid signaling. **Cell, 110**(2), 203-212.

Noguchi, T., Fujioka, S., Takatsuto, S., Sakurai, A., Yoshida, S., Li, J., & Chory, J. (1999). Arabidopsis det2 is defective in the conversion of (24 R)-24-methylcholest-4-en-3-one to (24 R)-24-methyl-5α-cholestan-3-one in brassinosteroid biosynthesis. **Plant Physiology, 120**(3), 833-840.

Nolan, T. M., Vukašinović, N., Liu, D., Russinova, E., & Yin, Y. (2019). Brassinosteroids: Multidimensional regulators of plant growth, development, and stress responses. **The Plant Cell, 32**(2), 295-318.

Oh, E., Zhu, J.Y., Bai, M.Y., Arenhart, R. A., Sun, Y., & Wang, Z.Y. (2014). Cell elongation is regulated through a central circuit of interacting transcription factors in the Arabidopsis hypocotyl. **Elife, 3**, e03031.

Ohnishi, T., Godza, B., Watanabe, B., Fujioka, S., Hategan, L., Ide, K., . . .& Mizutani, M. (2012). CYP90A1/CPD, a brassinosteroid biosynthetic cytochrome P450 of Arabidopsis, catalyzes C-3 oxidation. **Journal of Biological Chemistry, 287**(37), 31551-31560.

Ohnishi, T., Szatmari, A.-M., Watanabe, B., Fujita, S., Bancos, S., Koncz, C., . . .& Sakata, K. (2006). C-23 hydroxylation by Arabidopsis CYP90C1 and CYP90D1 reveals a novel shortcut in brassinosteroid biosynthesis. **The Plant Cell, 18**(11), 3275-3288.

Planas-Riverola, A., Gupta, A., Betegón-Putze, I., Bosch, N., Ibañes, M., & Caño-Delgado, A. I. (2019). Brassinosteroid signaling in plant development and adaptation to stress. **Development, 146**(5), dev151894.

Pliner, H. A., Packer, J. S., McFaline-Figueroa, J. L., Cusanovich, D. A., Daza, R. M., Aghamirzaie, D., . . .& Trapnell, C. (2018). Cicero predicts cis-regulatory DNA interactions from single-cell chromatin accessibility data. **Molecular Cell, 71**(5), 858-871.e858.

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

Ramirez, V. E., & Poppenberger, B. (2020). Modes of brassinosteroid activity in cold stress tolerance. **Frontiers in Plant Science, 11**, 583666.

Rich-Griffin, C., Eichmann, R., Reitz, M. U., Hermann, S., Woolley-Allen, K., Brown, P. E., . . .& Schäfer, P. (2020). Regulation of cell type-specific immunity networks in Arabidopsis roots. **The Plant Cell, 32**(9), 2742-2762.

Richards, S. L., Laohavisit, A., Mortimer, J. C., Shabala, L., Swarbreck, S. M., Shabala, S., & Davies, J. M. (2014). Annexin 1 regulates the H_2O_2-induced calcium signature in A rabidopsis thaliana roots. **The Plant Journal, 77**(1), 136-145.

Roh, H., Jeong, C. W., Fujioka, S., Kim, Y. K., Lee, S., Ahn, J. H., . . .& Lee, J. S. (2012). Genetic evidence for the reduction of brassinosteroid levels by a BAHD acyltransferase-like protein in Arabidopsis. **Plant Physiology, 159**(2), 696-709.

Ryu, K. H., Huang, L., Kang, H. M., & Schiefelbein, J. (2019). Single-cell RNA sequencing resolves molecular relationships among individual plant cells **Plant Physiology, 179**(4), 1444-1456.

Saito, M., Kondo, Y., & Fukuda, H. (2018). BES1 and BZR1 redundantly promote phloem and xylem differentiation. **Plant and Cell Physiology, 59**(3), 590-600.

Schellmann, S., Schnittger, A., Kirik, V., Wada, T., Okada, K., Beermann, A., . . .& Hülskamp, M. (2002). TRIPTYCHON and CAPRICE mediate lateral inhibition during trichome and root hair patterning in Arabidopsis. **The EMBO Journal, 21**(19), 5036-5046.

Schiefelbein, J., Huang, L., & Zheng, X. (2014). Regulation of epidermal cell fate in Arabidopsis roots: the importance of multiple feedback loops. **Frontiers in Plant Science, 5**, 47-47.

Schnabelrauch, L. S., Kieliszewski, M., Upham, B. L., Alizedeh, H., & Lamport, D. T. (1996). Isolation of pl 4.6 extensin peroxidase from tomato cell suspension cultures and identification of Val—Tyr—Lys as putative intermolecular cross-link site. **The Plant Journal, 9**(4), 477-489.

Seyfferth, C., Renema, J., Wendrich, J. R., Eekhout, T., Seurinck, R., Vandamme, N., . . .& Rybel, B. D. (2021). Advances and opportunities in single-cell transcriptomics for plant research. **Annual Review of Plant Biology, 72**(1), 847-866.

Shahan, R., Hsu, C.W., Nolan, T. M., Cole, B. J., Taylor, I. W., Greenstreet, L., . . .& Ohler, U. (2022). A single-cell Arabidopsis root atlas reveals developmental trajectories in wild-type and cell identity mutants. **Developmental Cell, 57**(4), 543-560.e549.

Shulse, C. N., Cole, B. J., Ciobanu, D., Lin, J., Yoshinaga, Y., Gouran, M., . . .& Dickel, D. E. (2019). High-throughput single-cell transcriptome profiling of plant cell types. **Cell Reports, 27**(7), 2241-2247.e2244.

Somssich, M., Khan, G. A., & Persson, S. (2016). Cell wall heterogeneity in root development of Arabidopsis [Review]. **Frontiers in Plant Science, 7**.

Stewart-Morgan, K. R., Reverón-Gómez, N., & Groth, A. (2019). Transcription restart establishes chromatin accessibility after DNA replication. **Molecular Cell, 75**(2), 284-297.e286.

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W. M., . . .& Satija, R. (2019). Comprehensive integration of single-cell data. **Cell, 177**(7), 1888-1902.e1821.

Su, M., Pan, T., Chen, Q.Z., Zhou, W.W., Gong, Y., Xu, G., . . .& Li, Y.S. (2022). Data analysis guidelines for single-cell RNA-seq in biomedical studies and clinical applications. **Military Medical Research, 9**(1), 68.

Sun, Y., Miao, N., & Sun, T. (2019). Detect accessible chromatin using ATAC-sequencing, from principle to applications. **Hereditas, 156**, 29.

Tang, W., Kim, T.W., Oses-Prieto, J. A., Sun, Y., Deng, Z., Zhu, S., . . .& Wang, Z.Y. (2008). BSKs mediate signal transduction from the receptor kinase BRI1 in Arabidopsis. **Science, 321**(5888), 557-560.

Tang, W., Yuan, M., Wang, R., Yang, Y., Wang, C., Oses-Prieto, J. A., . . .& Gampala, S. S. (2011). PP2A activates brassinosteroid-responsive gene expression and plant growth by dephosphorylating BZR1. **Nature Cell Biology, 13**(2), 124-131.

Vert, G., Nemhauser, J. L., Geldner, N., Hong, F., & Chory, J. (2005). Molecular mechanisms of steroid hormone signaling in plants. **Annual Review of Cell and Developmental Biology, 21**, 177-201.

Vissenberg, K., Fry, S. C., & Verbelen, J.P. (2001). Root hair initiation is coupled to a highly localized increase of xyloglucan endotransglycosylase action in Arabidopsis roots. **Plant Physiology**, **127**(3), 1125-1135.

Vissenberg, K., Oyama, M., Osato, Y., Yokoyama, R., Verbelen, J. P., & Nishitani, K. (2005). Differential expression of AtXTH17, AtXTH18, AtXTH19 and AtXTH20 genes in Arabidopsis roots. Physiological roles in specification in cell wall construction. **Plant Cell Physiology, 46**(1), 192-200.

Vragović, K., Sela, A., Friedlander-Shani, L., Fridman, Y., Hacham, Y., Holland, N., . . .& Savaldi-Goldstein, S. (2015). Translatome analyses capture of opposing tissue-specific brassinosteroid signals orchestrating root meristem differentiation. **Proceedings of the National Academy of Sciences, 112**(3), 923-928.

Vukašinović, N., Wang, Y., Vanhoutte, I., Fendrych, M., Guo, B., Kvasnica, M., . . .& Russinova, E. (2021). Local brassinosteroid biosynthesis enables optimal root growth. **Nature Plants, 7**(5), 619-632.

Wang, X., & Chory, J. (2006). Brassinosteroids regulate dissociation of BKI1, a negative regulator of BRI1 signaling, from the plasma membrane. **Science, 313**(5790), 1118-1122.

Wang, Z.Y., Nakano, T., Gendron, J., He, J., Chen, M., Vafeados, D., . . .& Chory, J. (2002). Nuclear-localized BZR1 mediates brassinosteroid-induced growth and feedback suppression of brassinosteroid biosynthesis. **Developmental Cell, 2**(4), 505-513.

Wang, Z.Y., Seto, H., Fujioka, S., Yoshida, S., & Chory, J. (2001). BRI1 is a critical component of a plasma-membrane receptor for plant steroids. **Nature, 410**(6826), 380-383.

Wang, Z. Y., Bai, M. Y., Oh, E., & Zhu, J. Y. (2012). Brassinosteroid signaling network and regulation of photomorphogenesis. **Annual Review of Genetics, 46**, 701-724.

Xuan, W., & Beeckman, T. (2021). Plant signaling: Interplay of brassinosteroids and auxin in root meristems. **Current Biology, 31**(20), R1392-R1395.

Yang, S., & Cui, L. (2009). The action of aquaporins in cell elongation, salt stress and photosynthesis. **Sheng Wu Gong Cheng Xue Bao, 25**(3), 321-327.

Yin, Y., Vafeados, D., Tao, Y., Yoshida, S., Asami, T., & Chory, J. (2005). A new class of transcription factors mediates brassinosteroid-regulated gene expression in Arabidopsis. **Cell, 120**(2), 249-259.

Yin, Y., Wang, Z. Y., Mora-Garcia, S., Li, J., Yoshida, S., Asami, T., & Chory, J. (2002). BES1 accumulates in the nucleus in response to brassinosteroids to regulate gene expression and promote stem elongation. **Cell, 109**(2), 181-191.

Yokota, T., Arima, M., & Takahashi, N. (1982). Castasterone, a new phytosterol with plant-hormone potency, from chestnut insect gall. **Tetrahedron Letters, 23**(12), 1275-1278.

Youn, J.H., Kim, M. K., Kim, E.J., Son, S.H., Lee, J. E., Jang, M.S., . . .& Kim, S.K. (2016). ARF7 increases the endogenous contents of castasterone through suppression of BAS1 expression in Arabidopsis thaliana. **Phytochemistry, 122**, 34-44.

Zhabinskii, V. N., Khripach, N. B., & Khripach, V. A. (2015). Steroid plant hormones: Effects outside plant kingdom. **Steroids, 97**, 87-97.

Zhang, T.-Q., Xu, Z.G., Shang, G.-D., & Wang, J.W. (2019). A single-cell RNA sequencing profiles the developmental landscape of Arabidopsis root. **Molecular Plant, 12**(5), 648-660.

Zhang, W., Tang, Y., Hu, Y., Yang, Y., Cai, J., Liu, H., . . .& Hou, X. (2021). Arabidopsis NF-YCs play dual roles in repressing brassinosteroid biosynthesis and signaling during light-regulated hypocotyl elongation. **The Plant Cell, 33**(7), 2360-2374.

Zhang, Z., & Xu, L. (2018). Arabidopsis BRASSINOSTEROID INACTIVATOR2 is a typical BAHD acyltransferase involved in brassinosteroid homeostasis. **Journal of Experimental Botany, 69**(8), 1925-1941.

Zheng, G. X. Y., Terry, J. M., Belgrader, P., Ryvkin, P., Bent, Z. W., Wilson, R., . . .& Bielas, J. H. (2017). Massively parallel digital transcriptional profiling of single cells. **Nature Communications, 8**(1), 14049.

Zhong, C., & Patra, B. (2021). A transcriptional hub integrating gibberellin-brassinosteroid signals to promote seed germination in Arabidopsis. **Journal of Experimental Botany**.

Zhu, J.Y., Li, Y., Cao, D.M., Yang, H., Oh, E., Bi, Y., . . .& Wang, Z.Y. (2017). The F-box protein KIB1 mediates brassinosteroid-induced inactivation and degradation of GSK3-like kinases in Arabidopsis. **Molecular Cell, 66**(5), 648-657.e644.

Zhu, L. J., Gazin, C., Lawson, N. D., Pagès, H., Lin, S. M., Lapointe, D. S., & Green, M. R. (2010). ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip data. **BMC Bioinformatics, 11**(1), 237.

VITA

NAME   Thanaporn Wongkham

DATE OF BIRTH   20 October 1997

PLACE OF BIRTH   Roi-et, Thailand

INSTITUTIONS ATTENDED   Bachelor of Science (Biology), Khon Kaen university

HOME ADDRESS   31 Village No. 3 Dongdang Sub-district, Chaturaphak Phiman

                         District, Roi Et, 45180

PUBLICATION   -

AWARD RECEIVED   -