Acta Pharmaceutica Sinica B 2026;16(5):2929–2946

Chinese Pharmaceutical Association
Institute of Materia Medica, Chinese Academy of Medical Sciences

Acta Pharmaceutica Sinica B

www.elsevier.com/locate/apsb
www.sciencedirect.com

REVIEW

Exploring specialized metabolic pathways in medicinal plants with single-cell and spatial omics

Guo Li[a,c,†], Miaomiao Cai[a,c,†], Xinhao Liao[a,c], Guang-Lei Ma[a,c,d,*], Xiaohui Fan[a,b,c,d,*]

[a]*College of Pharmaceutical Sciences, Zhejiang University, Hangzhou 310058, China*
[b]*School of Pharmaceutical Sciences, Zhejiang Chinese Medical University, Hangzhou 310053, China*
[c]*State Key Laboratory of Chinese Medicine Modernization, Innovation Center of Yangtze River Delta, Zhejiang University, Jiaxing 314102, China*
[d]*Zhejiang Key Laboratory of Precision Diagnosis and Therapy for Major Gynecological Diseases, Women's Hospital, Zhejiang University School of Medicine, Hangzhou 310006, China*

Received 26 June 2025; received in revised form 16 September 2025; accepted 13 October 2025

KEY WORDS

Medicinal plant;
Single-cell omics;
Specialized metabolite;
Biosynthesis;
Spatial transcriptomics;
Metabolic compartmentation;
Synthetic biology;
Pharmaceutical development

Abstract Medicinal plants synthesize an immense diversity of specialized metabolites that play crucial roles in ecological interactions and serve as valuable pharmaceutical resources. However, the biosynthetic pathways responsible for this chemical diversity remain largely uncharacterized. These pathways are often complex, involving multiple steps that are spatially and temporally orchestrated within highly specialized or rare cell types. Classical bulk omics approaches obscure such cellular heterogeneity by averaging signals across tissues, limiting their utility in resolving cell-specific metabolic processes. Recent advances in single-cell and spatial omics technologies have revolutionized the ability to investigate plant metabolism at high spatiotemporal resolution, as exemplified by monoterpene indole alkaloids and Taxol biosynthesis. In this review, we highlight key technological advances in plant single-cell and spatial omics, examine their applications in pathway discovery and partitioning, and discuss emerging directions for harnessing these tools in plant synthetic biology and metabolic engineering. These developments promise to accelerate the systemic mapping of plant metabolic networks and facilitate their biotechnological exploitation for pharmaceutical development.

*Corresponding authors.
E-mail addresses: guangleima@zju.edu.cn (Guang-Lei Ma), fanxh@zju.edu.cn (Xiaohui Fan).
[†]These authors made equal contributions to this work.
Peer review under the responsibility of Chinese Pharmaceutical Association and Institute of Materia Medica, Chinese Academy of Medical Sciences.

https://doi.org/10.1016/j.apsb.2025.12.023

1. Introduction

Medicinal plants produce a vast array of specialized metabolites (SMs, also referred to as natural products) that underpin ecological resilience and serve as valuable sources of pharmaceuticals[1,2]. Deciphering their biosynthetic pathways is critical not only for understanding plant adaptive strategies but also for enabling the discovery and bioengineering of bioactive natural products[3]. However, despite estimates suggesting plants can produce over 1 million specialized metabolites, only ∼0.1% of the biosynthetic pathways have been fully characterized, exposing a striking knowledge gap in plant metabolism[4-6]. Traditional methodologies, including molecular genetics and analytical biochemistry, have been instrumental in functionally characterizing metabolic genes and pathways[6,7]. Yet, these methods are often labor-intensive and largely confined to genetically tractable species, limiting their scalability. Advances in bulk omics technologies (*i.e.*, genomics, transcriptomics, proteomics, metabolomics) and computational tools have accelerated biosynthetic enzyme identification and pathway elucidation, offering significant insight into plant specialized metabolism[8-10]. A landmark milestone was the recent complete elucidation of the paclitaxel (Taxol) biosynthetic pathway[11-13]. Through integrated multi-omics analyses, researchers identified long-elusive enzymes critical to Taxol biosynthesis, including the C4β-C20 epoxidase, 9α-hydroxylase, 1β-hydroxylase, C2′α hydroxylation and 3′-*N* benzoylation, exemplifying the power of multi-omics approaches in decoding intricate plant biosynthetic pathways.

Despite significant progress, major challenges remain, particularly in unravelling the spatial and cell-type-specific nature of plant specialized metabolism. Many biosynthetic pathways, including those producing monoterpene indole alkaloids (MIAs) and Taxol, function within spatially discrete and highly specialized cell types[14,15]. Conventional bulk omics obscure such compartmentalization by averaging signals across diverse tissues and fail to detect lowly expressed or non-coregulated biosynthetic genes[10,16]. This cell-type heterogeneity and metabolic compartmentalization hence represent a critical bottleneck in fully elucidating specialized metabolic pathways in medicinal plant. The advent of single-cell (sc) and spatial technologies has revolutionized this landscape, offering unprecedented resolution in profiling transcriptomes and metabolomes at the cellular level[17,18]. While early pioneering studies primarily focused on developmental and evolutionary processes in model species like *Arabidopsis thaliana*, recent works have demonstrated the immense potential of these technologies for uncovering medicinal plant specialized metabolism (Supporting Information Table S1). These cutting-edge approaches enable *de novo* pathway discovery, reveal spatial regulatory mechanisms, and profile cell-specific gene-metabolite associations, providing invaluable insights in advancing medicinal plant metabolic engineering and synthetic biology[14,19-21].

In this review, we explore how single-cell and spatial omics technologies are reshaping the discovery, characterization and compartmentation of medicinal plant specialized metabolic pathways. We begin by highlighting recent innovations in these technologies, with a focus on their capacity to resolve cellular heterogeneity and spatial organization. We then discuss applications in correlating cell-type-specific gene expression with metabolite localization to guide biosynthetic pathway/gene discovery, metabolic partitioning, and regulatory network profiling though several remark examples including MIAs and Taxol. We also examine how single-cell insights inform plant bioengineering and synthetic biology for the sustainable production of high-value SMs. Finally, we outline future directions, including extensive multi-omics integration, AI-driven biosynthetic gene mining and pathway design, and single-cell metabolic flux optimization. These advances promise to accelerate the high-throughput discovery and biotechnological exploitation of medicinal plant metabolic pathways (Fig. 1).

2. Emerging toolkits for single-cell and spatial resolution

Over the past decade, advances in single-cell technologies have revolutionized plant science, enabling unprecedented resolution to study cellular heterogeneity and metabolic dynamics (Table 1)[22-34]. Initially developed in animal systems, single-cell transcriptomics (*e.g.*, scRNA-seq, 10x Genomics) rapidly emerged as a fundamental tool, revealing cell-type specific gene expression in model plants, crops, and medicinal plants (Fig. 2A). More recently, spatial transcriptomics (*e.g.*, Stereo-seq, 10x Visium) has extended these insights by mapping gene activity to tissue microarchitecture (Fig. 2C). While single-cell proteomics and metabolomics lag due to technical challenges, recent advances are narrowing this gap (Fig. 2B). The following subsection provides a concise introduction to these single-cell technologies, along with a comparative analysis of their key features and applications (Fig. 2D, Table 1).

2.1. *Single-cell transcriptomics*

Transcriptomics has long played a pivotal role in guiding the discovery of plant biosynthetic pathways. This approach leverages the principle that both gene clusters and dispersed genes involved in specialized metabolism often exhibit coordinated expression patterns across developmental stages and environmental conditions (Fig. 2A). Prior to the advent of droplet-based single-cell RNA sequencing, single-cell studies were constrained by labor-intensive, low-throughput methods such as laser capture microdissection (LCM) and fluorescence-activated cell sorting (FACS). LCM, while precise in isolating individual cells *in situ*, typically captured only 4−40 cells per experiment and required complex manual procedures[35]. FACS enabled the sorting of fluorescently labelled cells at higher throughput, but its reliance on predefined markers limited the discovery of unexpected or subtle cell states[36].

The introduction of droplet-based microfluidic platforms marked a paradigm shift, enabling massive parallel, genome-wide transcriptional profiling of thousands of individual cells in a single experiment[37-39]. Upon this platform, sc sequencing methods have rapidly evolved, with a growing suite of advanced tools now available to accommodate diverse experimental purposes. Both scRNA-seq and single-nucleus RNA-seq (snRNA-seq) have become indispensable tools for dissecting plant tissue complexity,

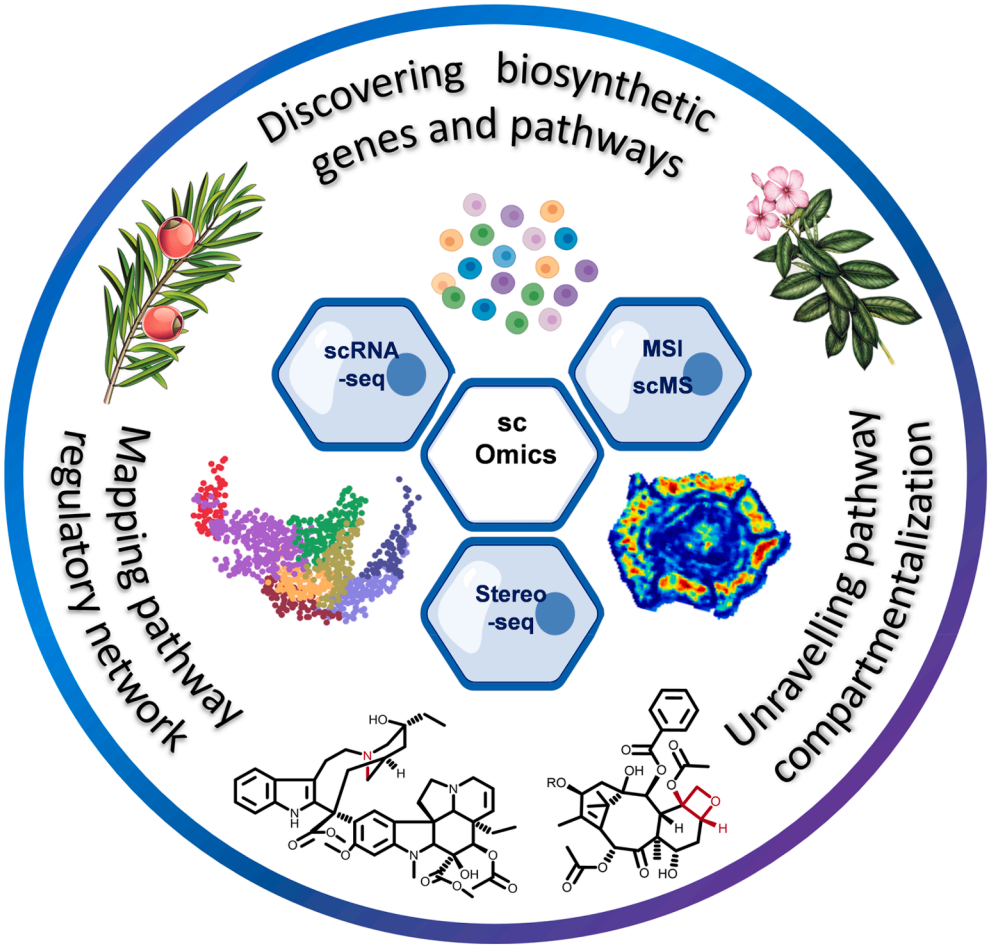

Figure 1 Overview of single-cell and spatial omics applications in plant metabolic pathway studies. The biosynthesis of plant specialized metabolites is a highly complex, spatiotemporally regulated process, and often confined to specific cell types. Single-cell and spatial omics provide powerful means to resolve cellular heterogeneity, enabling high-resolution pathway elucidation, spatial compartmentalization mapping, and regulatory network profiling.

with the latter especially useful for recalcitrant or lignified tissues (Fig. 2A). These techniques offer several key advantages over bulk RNA-seq, including unbiased detection of diverse cell types and states within complex tissues, identification of rare or transient populations, and high-resolution mapping of temporal gene expression dynamics[40,41]. Complementary tools such as single-cell and single-nucleus ATAC-seq further enhance our understanding by revealing chromatin accessibility and the regulatory architecture underpinning gene expression at single-cell resolution (Table 1)[42,43]. Continued innovations in barcoding strategies, combinatorial indexing, and cost-effective workflows have further expanded the scalability and accessibility of these platforms. In the context of plant specialized metabolism, single-cell transcriptomics has been particularly transformative. It enables researchers to identify co-expressed cryptic biosynthetic genes in specific cell types, to map metabolic pathway at cellular resolution, and to discover regulatory nodes and transcriptional activities underlying specialized metabolism.

2.2. *Single-cell proteomics*

While single-cell transcriptomics has revolutionized our understanding of gene expression and cellular diversity, it does not fully capture the functional state of the cell, which is primarily governed by proteins and their post-translational modifications (PTMs). As the direct effectors of cellular function and signalling, proteins provide a more immediate readout of biological processes. However, single-cell proteomics faces substantial technical challenges, particularly in terms of sensitivity, dynamic range, and sample loss, often limiting detection to only high-abundance proteins (Fig. 2B)[44]. Early pioneering efforts in plants demonstrated the feasibility of proteomic analysis at the single-cell level. For example, the identification of glucosinolate biosynthesis enzymes in *Arabidopsis* S-cells was achieved using microcapillary sampling coupled with MALDI-TOF MS and nanoLC−MS/MS, although this approach was constrained to a limited set of proteins[45,46]. Recent advancements in mass spectrometry (MS), particularly improvements in nanodroplet sample preparation, ultra-high-resolution Orbitrap analyser, and data-independent acquisition methods, have greatly expanded the detection capacity. These innovations now enable the profiling of hundreds of proteins from small or even single-cell inputs, facilitating the construction of comprehensive plant proteome atlases and the delineation of cell-type-specific proteomes (Fig. 2B)[47,48].

Despite this progress, plant single-cell proteomics still trails behind animal and microbial studies, largely due to intrinsically technical challenges. The structural and biochemical complexity of plant tissues, including high levels of interfering metabolites,

Table 1 Comparative analysis of single-cell and spatial omics technologies.

Technology	Analytes	Resolution	Throughput	Advantages	Limitations	Key applications	Ref.
scRNA-seq	Whole-cell transcriptome	Single cell	Up to 20,000 cells per lane	Captures full transcriptome	Requires fresh cells; may miss nuclear RNAs	Cell type identification; developmental trajectory analysis	22
snRNA-seq	Nuclear RNA	Single nucleus	Up to 20,000 cells per lane	Works with frozen samples; less dissociation bias	Limited to nuclear transcripts; lower coverage	Transcriptomics of frozen/archived samples; avoiding dissociation bias	23
STRT-seq	5′-End of mRNA transcripts	Single cell	~2000 cells per plate	High-throughput; strand-specific	Limited transcript coverage; 5′ bias	Transcription start site; TSS mapping; promoter activity studies	24
SMART-seq	Full-length transcripts	Single cell	~2000 cells per plate	High sensitivity; isoform analysis	Lower throughput; costly; need high-quality RNA	Isoform analysis; allele-specific expression; rare cell-type analysis	24
MARS-seq	3′-End mRNA	Single cell	~2000 cells per plate	Cost-effective; barcoded; plate-based workflow	Captures only 3′ end; limited dynamic range	Medium-throughput atlas; immune profiling	24, 25
Drop-seq	3′-End mRNA	Single cell	Up to 10,000 cells per run	Low cost; scalable for large populations	Sparse coverage; partial transcripts	High-throughput profiling of heterogeneous populations; cell atlas generation	26
SORT-seq	3′-End mRNA	Single cell	~2000 cells per plate	Allows sorting by surface markers	Moderate throughput; limited to sorted cells	Combined FACS and transcriptomics; rare population studies	24
CROP-seq	mRNA, sgRNA barcode	Single cell	Up to 10,000 cells per run	Links CRISPR perturbations with expression	Requires CRISPR constructs; low transcript depth	CRISPR screens with transcriptomic readout; gene regulatory network dissection	27
SPLIT-seq	mRNA	Single cell	~5000 cells per assay	Labeling without isolation; no microfluidics	Lower sensitivity; 3′ bias	Cost-effective large-scale cell profiling; cell atlas without microfluidics	28
10x Visium	mRNA, spatial barcodes	~55 μm	~10 cells per spot	Preserves spatial context	Not single-cell resolution; costly	Spatial gene expression mapping; tissue architecture studies	24
SLIDE-seq	mRNA	~10 μm (near single-cell)	Varies, depend on sample size	Higher spatial resolution than Visium; cost-effective	Lower transcript capture efficiency	Near-single-cell spatial transcriptomics; fine spatial gradients in tissues architecture analysis	29
Stereo-seq	mRNA (whole transcriptome)	~500 nm (Sub-cellular)	Varies, depend on sample size	Submicron resolution	High complexity; expensive; large data size	Subcellular spatial transcriptomics; 3D tissue mapping	30
MALDI	Metabolites, peptides	~10–50 μm	~1000 samples per day	Minimal sample prep; strong imaging	Matrix interference; moderate resolution	Spatial metabolomics; tissue imaging	31
nanoDESI	Metabolites, lipids	~10–20 μm	Hundred cells per day	*In situ*, label-free, matrix-free	Low sensitivity; complex handling; low throughput	*In situ* tissue/cell metabolomics	32
LDI	Small molecules	~1–100 μm	Depends on instrument and area	No matrix interference; ideal for small molecules	Low sensitivity; poor selectivity	Small molecule or elemental analysis	31
SIMS	Small metabolites	<1 μm	Depends on instrument and area	Highest spatial resolution	Severe fragmentation; complex spectra	Subcellular imaging; elemental localization; spatial metabolomics	31
nanoLC–MS	Peptides, small molecules	Single cell	Depends on run length and method	High sensitivity; good for proteomics and metabolomics	Not *in situ*; requires complex sample prep	Single-cell proteomics and metabolomics	33
CE–MS	Small molecules, short peptides	Single cell	Depends on run length and method	Ultra-high separation efficiency	Low throughput; sensitive to conditions	Amino acid profiling; plant single-cell analysis	34

scRNA-seq, single-cell RNA sequencing; snRNA-seq, single nucleus RNA sequencing; STRT-seq, single cell tagged reverse transcription sequencing; SMART-seq, switching mechanism at 5′ end of RNA template; MARS-seq, massively parallel RNA single-cell sequencing; Drop-seq, droplet-based scRNA-seq; SORT-seq, scRNA-seq after FACS sorting; CROP-seq, CRISPR droplet sequencing; SPLIT-seq, split-pool ligation-based transcriptomics; 10x Visium, 10x Genomics Visium spatial transcriptomics; SLIDE-seq, spatial transcriptomics by bead array; Stereo-seq, SpaTial enhanced resolution omics sequencing; MALDI, matrix-assisted laser desorption/ionization; nanoDESI, nanospray desorption electrospray ionization; LDI, laser desorption ionization; SIMS, secondary ion mass spectrometry; nanoLC–MS, nanoflow liquid chromatography mass spectrometry; CE–MS, capillary electrophoresis mass spectrometry.

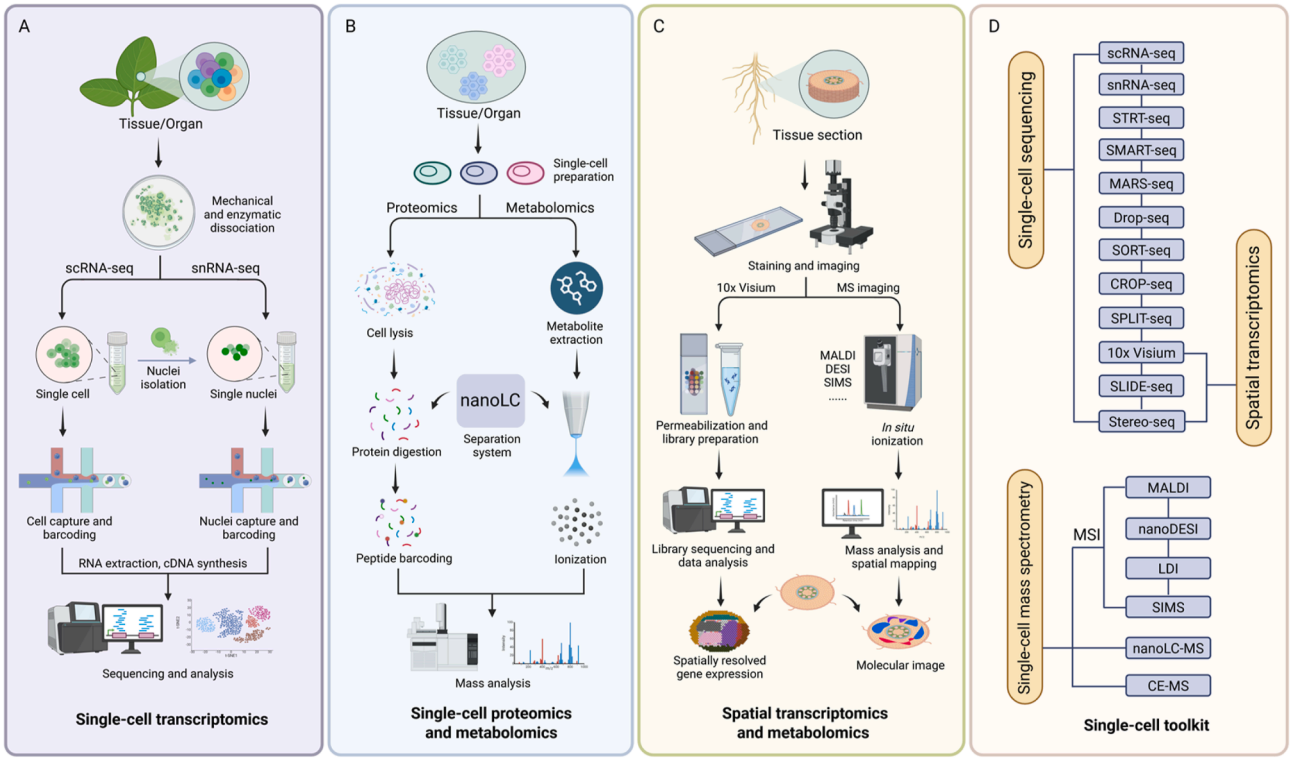

Figure 2 Schematic overview of single-cell and spatial omics technologies. (A) Workflow of single-cell transcriptomics (scRNA-seq, snRNA-seq) analysis in plant system. The process comprises tissue dissection, protoplast or nuclei isolation, RNA extraction and cDNA synthesis, single-cell sequencing, and subsequent data processing for cell clustering and characterization. (B) Diagram of single-cell proteomics and metabolomics, involving tissue dissociation, single-cell isolation, protein digestion or metabolite extraction, and mass spectrometry analysis (UPLC−MS, nanoLC−MS). These methods enable quantification of diverse peptides/metabolites per cell while preserving cellular heterogeneity. Current limitations include lower throughput compared to transcriptomics and challenges in detecting low-abundance analytes. (C) Workflow of spatial transcriptomics and metabolomics, both typically requiring spatial labeling strategies to link expression/molecular profile to precise tissue locations. Spatial transcriptomics is approaching single-cell resolution, while spatial metabolomics generally offers lower spatial resolution and faces challenges in achieving single-cell level analysis. (D) Overview of currently available single-cell and spatial toolkits. Various single-cell sequencing platforms and advanced mass spectrometry technologies have been progressively developed, each offering distinct advantages and collectively enhancing the resolution of plant metabolism at cellular and spatial levels. scRNA-seq, single cell RNA sequencing; snRNA-seq, single nuclei RNA sequencing; nanoLC−MS, nanoflow liquid chromatography mass spectrometry. A detailed comparative analysis of single-cell and spatial omics technologies can be found in Table 1. Created with BioRender.com.

impedes efficient protein extraction and downstream analysis. To overcome these hurdles, undifferentiated plant cell cultures have been employed as a more homogeneous and tractable system. However, findings from such system might not reflect the highly specialized processes and functions of differentiated plant cells. Moving forward, breakthroughs in high-throughput cell isolation, targeted cell lysis, ultra-sensitive MS instrumentation, and advanced computational workflows, will be essential to achieve reliable, quantitative proteomics at true single-cell resolution across diverse plant cell types (Table 1).

2.3. *Single-cell metabolomics*

Single-cell metabolomics remains one of the most technically demanding frontiers in plant science due to the extreme chemical diversity, dynamic range, and small volume of intracellular metabolites. Nevertheless, recent progress in ionization technologies and sample preparation methods has begun to make high-resolution metabolic profiling at single-cell resolution feasible (Fig. 2B). Several mass spectrometry platforms have been adapted for single-cell studies, including matrix-assisted laser desorption/ionization (MALDI) imaging, nanospray desorption electrospray ionization (nanoDESI-MS), laser desorption ionization (LDI-MS), and secondary ion mass spectrometry (SIMS) (Fig. 2B and D, Table 1)[49,50]. Among them, MALDI-MS imaging has become a widely used tool for mapping metabolite distributions across plant tissues with spatial context. NanoDESI-MS has enabled the sensitive detection of specialized metabolites in diverse species. Integration of laser microdissection with LDI-TOF-MS has further improved spatial precision, allowing for targeted metabolic profiling in specific cell types, as demonstrated in *Hypericum* and *Arabidopsis*[51]. More recently, single-cell capillary electrophoresis (CE−MS) platforms are capable in detecting hundreds of metabolites in individual onion cells, making a substantial advancement in sensitivity and resolution[52]. However, most single-cell MS techniques currently operate without chromatographic separation, which limits structural resolution, quantification accuracy, and metabolite identification. To address these gaps, emerging platforms now integrate ultra-high-performance liquid chromatography (UPLC−MS) and ion mobility spectrometry to improve separation and reduce matrix effects. For example, single-cell UPLC−MS profiling of

Catharanthus roseus leaf cells has successfully identified key alkaloids such as vinblastine and catharanthine[15,53].

Indeed, nanoLC systems are improving detection limits and reducing sample volume requirements, which is essential for metabolomics at cellular scale[54]. Innovative live-cell techniques are also expanding capabilities. A notable example is the "Live-MS" platform, which combines video-microscopy with nanoESI to allow real-time extraction and profiling of metabolites from individual plant cells using metal-coated microcapillaries[55]. Such integrations of imaging and MS not only improve spatial accuracy but also open the door to time-resolved studies of dynamic metabolic processes. Despite these advances, significant challenges remain, particularly in preserving cellular integrity during extraction, detecting low-abundance compounds, and achieving comprehensive metabolite coverage due to the vast chemical diversity and lack of universal enrichment strategies.

2.4. *Spatial transcriptomics and metabolomics*

Spatial transcriptomics has emerged as a transformative tool in molecular biology, addressing a key limitation of scRNA-seq, that is the loss of spatial context during cell dissociation (Fig. 2C). Named "Method of the Year 2020" by *Nature Methods*, this technology enables transcriptome-wide profiling while preserving the spatial organization of tissues[56]. Early techniques, such as *in situ* hybridization (ISH) and LCM, offered localized gene expression insights but were time-consuming and limited in scale. For example, ISH depends on custom probes and lacks genome-wide coverage, while LCM requires meticulous laser-guided isolation of individual cells, making it inherently low-throughput and technically demanding[57,58]. Recent innovations, including 10x Genomics Visium, SLIDE-seq, and Stereo-seq integrate next-generation sequencing with spatial barcoding, enabling high-resolution mapping of gene expression within tissue architecture (Fig. 2D, Table 1). Among them, SLIDE-seq achieves near single-cell resolution (~10 μm) using spatially barcoded beads, whereas 10x Visium offers broader tissue coverage with a resolution of ~55 μm, making it more suitable for larger plant organs[29]. Stereo-seq represents a significant breakthrough by combining DNA nanoball arrays with *in situ* RNA capture, achieving submicron resolution (~500 nm) and high transcript capture sensitivity[59]. This method has been successfully applied to *Arabidopsis thaliana*, offering subcellular spatial resolution of gene expression and enabling high-precision mapping of developmental zones and biosynthetic activity[24,30]. As these platforms evolve, their integration into plant research holds great promise for elucidating spatial gene regulation, cell-type-specific interaction, and tissue patterning and development with unprecedented details.

Spatially resolved metabolomics complements transcriptomic data by revealing the localization and distribution of specialized metabolites, which are often tightly restricted to specific cell types or tissue regions. Techniques such as mass spectrometry imaging (MSI) and single-cell MS (scMS) have enabled the mapping of metabolites with spatial context (Table 1)[49,60]. In *Panax ginseng*, integration of MSI with scRNA-seq has shown that ginsenoside biosynthesis is anatomically compartmentalized during leaf development, reflecting the co-evolution of metabolic pathways with tissue specialization[61]. Despite these advancements, single-cell spatial metabolomics in plants faces persistent technical obstacles. The plant cell wall, waxy cuticles, and high intracellular water content complicate sample preparation and metabolite extraction. Moreover, low-abundance metabolites, low throughput detection, and lack of chromatographic separation, remain difficult to detect and resolve using scMS platforms. Recent innovations, such as plate-based scMS, have begun to overcome some of these barriers by enabling high-throughput and highly sensitive profiling of natural products at single-cell resolution[15,53,62]. In *C. roseus*, this chromatography-coupled scMS approach enable identification and quantification of 16 metabolites in individual cells across roots, leaves, and petals at a throughput of approximately 180 cells/day[53]. While scRNA-seq has enabled the identification of biosynthetic gene expression profile, the integration of spatial transcriptomics with spatial metabolomics is crucial to achieving a holistic view of plant SM biosynthesis.

3. Plant specialized metabolism studies enter the single-cell omics era

Medicinal plant SMs, such as MIAs and Taxol, hold immense pharmacological significance and represent promising targets for biotechnological innovation[63]. However, the biosynthetic pathways leading to these compounds are inherently complex, often involving dozens of enzymatic steps distributed across multiple cell types and subcellular compartments. For example, MIAs biosynthesis in *Catharanthus roseus* spans a 38-step metabolic network partitioned among internal phloem-associated parenchyma (IPAP), epidermal cells, laticifers and idioblasts, each contributing specific biosynthetic steps and/or mediating intermediate trafficking[15]. Unlike bacteria and fungi, whose biosynthetic genes are often physically clustered, plant metabolic pathways are typically non-clustered and dispersed throughout the genome. This gene dispersion, along with spatial partitioning of metabolic pathways, pose significant challenges for plant biosynthetic gene mining and pathway discovery. Classical methodologies, including functional genomics, bulk transcriptomics, and targeted metabolomics, have significantly advanced our understanding of plant metabolism by leveraging co-expression analyses and metabolite-gene correlations at the tissue or organ level (Fig. 3A)[64-66]. However, these "guilt-by-association" strategies average gene expression and metabolite profiles across heterogeneous tissues, thereby masking the fine-scale spatial dynamics and obscuring cell-type-specific biosynthetic events.

Recent breakthroughs in single-cell RNA sequencing (scRNA-seq) and spatial metabolomics technologies are overcoming the limitations of bulk omics by providing unprecedented cellular and spatial resolution (Fig. 3B)[53,67,68]. Coupling single-cell transcriptomics with spatial metabolomics will allow researchers to map biosynthetic pathways to specific cell types and uncover how specialized metabolites are synthesized, compartmentalized, and transported across plant tissues (Fig. 3C)[15,53,62,67,69]. This high-resolution insight is particularly valuable for identifying cryptic biosynthetic genes and recognizing *de novo* pathway where bulk analyses fail due to tissue complexity and signal dilution. Furthermore, the cell-specific regulatory networks underlying biosynthetic pathways, along with the cell-cell communication potentially mediating metabolite trafficking, can be systematically profiled using single-cell technologies (Fig. 3C)[70,71].

3.1. *Single-cell and spatial omics in discovering cryptic biosynthetic genes and novel pathways*

Many plant SMs are synthesized in rare or highly specialized cell types, whose gene expression profiles are often imperceptible in

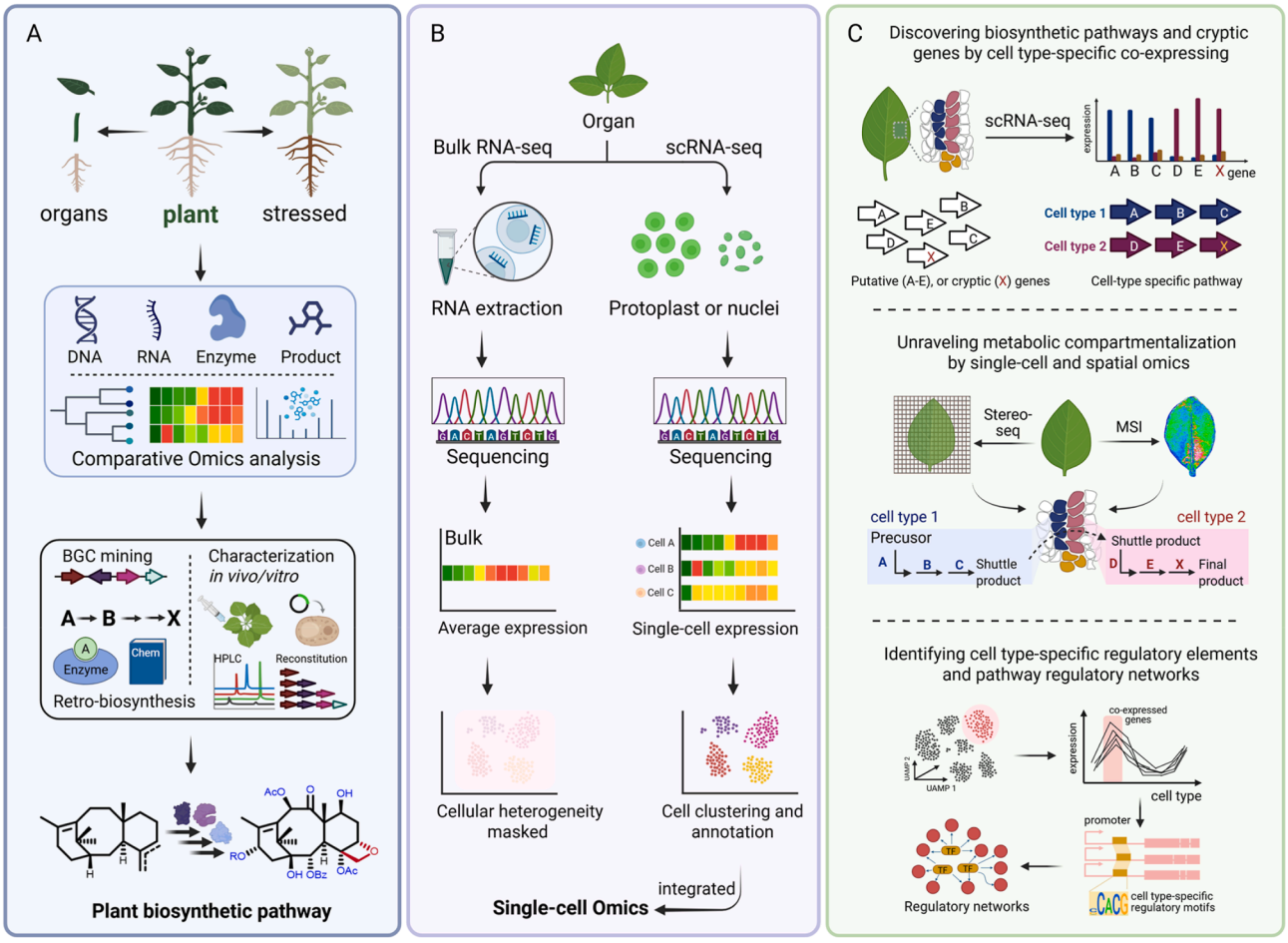

Figure 3 Evolution of pathway elucidation strategies from bulk to single-cell resolution. (A) Bulk omics-guided strategy for deciphering plant specialized metabolic pathways. The workflow integrates comparative multi-omics analysis to identify candidate genes, followed by retro-biosynthetic analysis based on prior knowledge to propose the biosynthetic route. Functional validation of genes is conducted through expression profiling or enzyme assays, and the complete biosynthetic pathway is reconstructed and verified in heterologous hosts. (B) scRNA-seq revealing the cellular heterogeneity of plant biosynthetic processes by profiling gene expression at the individual cell level. This technology enables the detection of rare cell populations and subtle expression differences that are marked by bulk analyses. (C) Integrated single-cell and spatial omics provide high-resolution insights into plant specialized metabolism, which enable precise biosynthetic pathway/gene discovery (top), spatial metabolic compartmentalization (middle), and detailed profiling of regulatory networks across cell types (bottom). As plant multi-omics enters the single-cell era, these technologies are set to redefine our understanding of metabolic networks and accelerate pathway discovery and bioengineering in medicinal plant species. Created with BioRender.com.

bulk RNA-seq analyses. The averaging effect inherent in bulk methods dilutes cell-type-specific signals, obscuring the expression of low-abundance or tightly regulated biosynthetic genes. In contrast, scRNA-seq provides high-resolution insights into the transcriptional landscape of individual cells, allowing for precise correlation of biosynthetic gene expression and metabolite accumulation[72,73]. Moreover, co-expression analyses at the single-cell level significantly enhance the ability to detect missing or uncharacterized genes in biosynthetic pathways by revealing fine-grained gene−gene associations that are otherwise obscured by tissue-level noise (Fig. 3C)[70,74]. The following studies highlight recent successes where scRNA-seq enabled the discovery or completion of complex specialized metabolic pathways.

In *Hypericum perforatum* (St. John's wort), the biosynthesis of hyperforin, a polyprenylated acylphloroglucinol compound with antidepressant activity, had been partially elucidated, with early enzymes such as HpBCAT, HpBCKDHA, HpCCL, and HpPKS characterized for synthesizing phloroisobutyrophenone (PIBP) module (Fig. 4A)[75,76]. However, bulk transcriptomic approaches failed to identify key downstream prenyltransferases (*i.e.*, prenylation module) due to signal masking by surrounding tissues. In a landmark study, Wu et al. used scRNA-seq to generate a leaf cell atlas and identified 12 distinct cell clusters, including a specialized cluster named "hyper cells" that co-expressed all known upstream pathway genes[73] (*e.g.*, PIBP and MEP modules). Gene ontology enrichment of hyper cell cluster revealed upregulated UbiA prenyltransferases, leading to the proposal of four candidates (HpPT1−HpPT4) responsible for the final prenylation steps. Functional validation in *Saccharomyces cerevisiae* and *Nicotiana benthamiana* confirmed the correct roles of these prenyltransferases and enabled full reconstitution of the hyperforin biosynthetic pathway[73]. The findings also enable the establishment of an automated cell-annotation pipeline for non-model medicinal plants by leveraging *Arabidopsis* marker orthologs

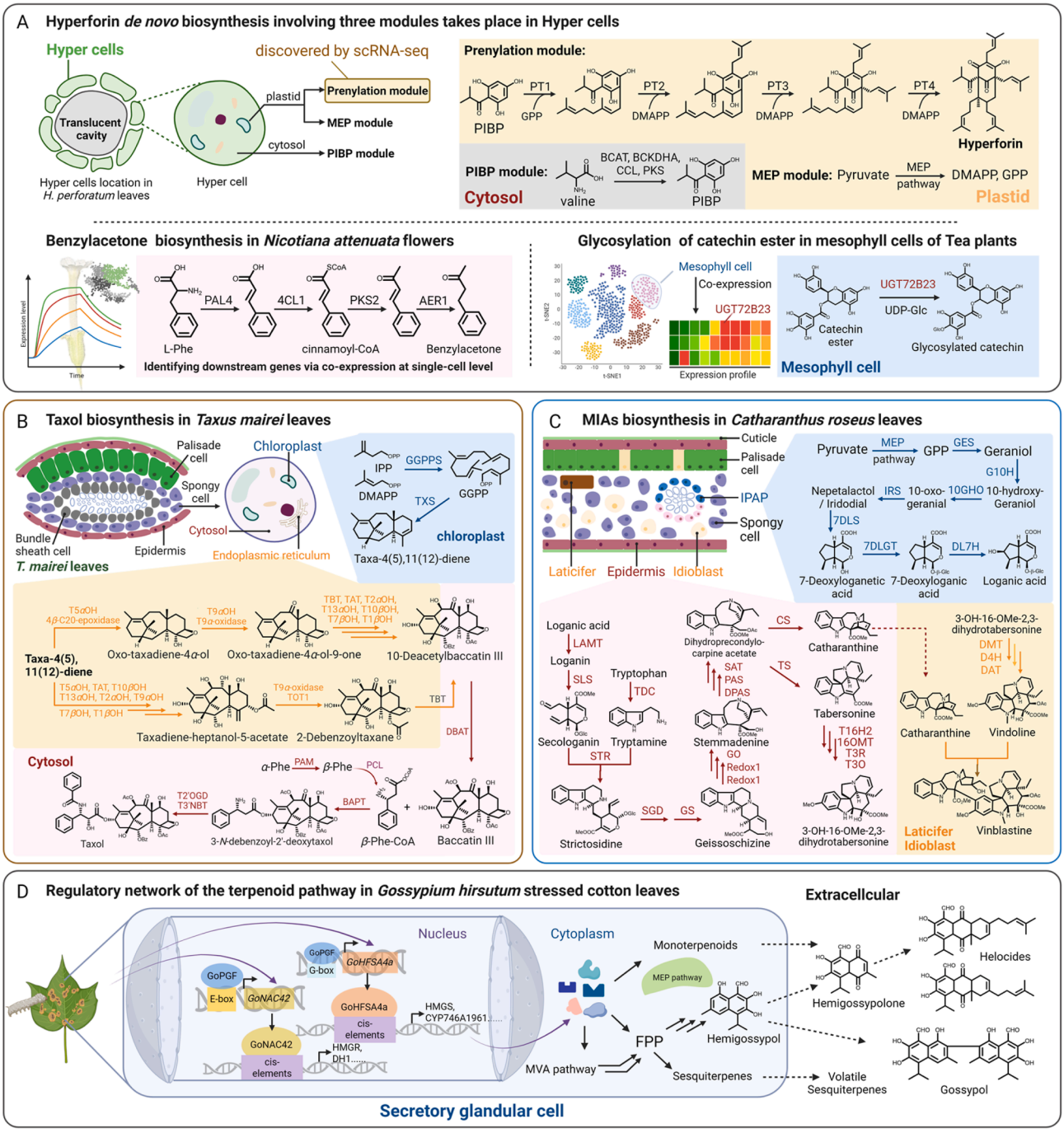

Figure 4 Single-cell technologies in dissecting medicinal plant cellular metabolism. (A) Single-cell omics-assisted discovery of cryptic biosynthetic genes or novel pathways. For hyperforin, scRNA-seq facilitates the discovery and elucidation of *de novo* biosynthesis of the anti-depressant hyperforin. The three modules involved in the pathway all take place in a type of rare cavity-adjacent cell, named "Hyper cells". MEP module, plastid-localized methylerythritol phosphate pathway for providing prenyl unit donors; PIBP module, phloroisobutyrophenone pathway for generating PIBP precursor in cytosol; Prenylation module: the sequential prenylation and tandem cyclization steps for completing the hyperforin biosynthesis; BCKDHA, branched-chain amino acid dehydrogenase; CCL, a coenzyme A (CoA) ligase (CCL); PKS, polyketide synthase; PT1-PT4, prenyltransferases 1−4. For benzylacetone, cell type-specific co-expression discovery of key enzymes for benzylacetone biosynthesis in *Nicotiana attenuata.* Green dots of varying intensity represent different enzymes with high expression across most of these cells. PAL4, phenylalanine ammonia-lyase 4; 4CL1, 4-coumarate-CoA ligase1; PKS2, polyketide synthase 2; AER1, alkenal reductase 1. For catechin ester, mesophyll cells-specific co-expression analysis enables the identification of a novel glycosylation step in the Tea catechin pathway. UGT72B23, UDP-glycosyltransferase. (B) Compartmented biosynthetic pathway of paclitaxel (Taxol) in *T. mairei* leaves. The steps occur in chloroplast, endoplasmic reticulum, and cytosol are coloured in pale blue, yellow and pink arrow, respectively. TXS, taxadiene synthase; T5αOH, taxane 5α-hydroxylase; TAT, taxadiene-5α-ol-*O*-acetyl transferase; T10βOH, taxane 10β-hydroxylase; T13αOH, taxane 13α-hydroxylase; T2αOH, taxane 2α-hydroxylase; T9αOH, taxane 9α-hydroxylase; T7βOH, taxane 7β-hydroxylase; T1βOH, taxane 1β-hydroxylase; TOT1,

identified through genomic synteny analysis. This approach facilitates the rapid identification of cell-type-specific metabolic pathways and offers a generalizable framework for extending such analyses to diverse plant species. Furthermore, the strategy of targeting specialized biosynthetic cells (*e.g.*, hyper cells) effectively resolves the precise synthesis sites of high-value metabolites, while mitigating the transcriptional noise that is often encountered in bulk-tissue profiling.

A similar strategy was used in *Nicotiana attenuata* to elucidate the biosynthesis of benzylacetone (BA), a volatile C6−C4 phenylbutanoid compound involved in nocturnal pollinator attraction[77]. While NaPAL4 and NaCHAL2/3 had previously been implicated in early steps through bulk RNA-seq, the downstream pathway remained elusive[78]. To uncover the remaining biosynthetic route, Kang et al. performed scRNA-seq on corolla protoplasts, revealing distinct cell clusters of epidermal, chlorenchyma, and parenchyma. By correlating gene expression with NaPAL4 at the single-cell level, they narrowed down Na4CL1, NaPKS2, and NaAER1 as key enzymes and envisioned the entire BA biosynthetic pathway (Fig. 4A)[72]. The roles of these genes were further validated through *in vitro* biochemical assays and gene silencing experiments. The study highlights the superiority of scRNA-seq to bulk RNA-seq in prioritizing functionally relevant genes. For example, while bulk RNA-seq suggested alternative enzymes, single-cell analysis correctly identified NaPKS2 as the dominant synthase and revealed the previously overlooked NaAER1 in BA pathway[72]. In another study, Wang et al. applied scRNA-seq to *Camellia sinensis*, constructing a high-resolution leaf cell atlas and identifying a previously uncharacterized catechin ester biosynthetic pathway (Fig. 4A)[74]. Through single-cell gene−metabolite correlation networks, they identified eight UDP-glycosyltransferases (UGTs) co-expressed with catechin biosynthetic genes. Among them, UGT72B23 stood out for its higher expression in mature mesophyll cells. Subsequent biochemical validation confirmed its role in catechin modification, demonstrating how scRNA-seq can resolve subtle spatial expression differences and reveal missing biosynthetic genes[74].

In a very recent study, McClune et al. employed single-nucleus RNA sequencing of *Taxus* tissues under multiplexed perturbations (mpXsn), revealing three distinct transcriptional modules that govern Taxol biosynthesis[13]. Systematic analysis of these modules enabled the discovery of eight previously unknown biosynthetic genes and a novel nuclear transport factor 2-like protein, FoTO1, which had not been previously linked to plant metabolism. Remarkably, FoTO1 was shown to resolve a long-standing bottleneck in pathway reconstitution by suppressing side products formation during the critical first taxane oxidation step, boosting yields by 10−17-fold. Combined with 11 known genes, these discoveries allowed the team to achieve *de novo* biosynthesis of baccatin III in *Nicotiana benthamiana* at levels comparable to native *Taxus* production. This work highlights how perturbation-driven single-cell transcriptomics can uncover cryptic metabolic genes by resolving precise gene co-expression patterns across diverse cellular states, even within genomically complex systems[13].

3.2. *Single-cell and spatial omics in unraveling multiscale compartmentalization of specialized metabolism*

Compartmentalization is a fundamental feature of plant specialized metabolism, enabling the spatial and temporal separation of biosynthetic steps to enhance metabolic efficiency, regulation, and adaptability[7,14]. By segregating competing or sequential pathways into specific cells, organelles, or tissues, plants can avoid deleterious metabolic cross-talk, minimize the buildup of toxic intermediates, and fine-tune enzyme activity through specialized microenvironments[79]. For instance, terpenoid precursor biosynthesis can be split into the cytosolic mevalonate (MVA) and plastidic methylerythritol phosphate (MEP) pathways, while toxic intermediates (*e.g.*, alkaloids) are often sequestered in specific organelle (*e.g.*, vacuoles) to prevent cellular toxicity[80,81]. Additionally, spatial organization supports efficient resource allocation, stabilizes labile intermediates, and facilitates final modifications such as glycosylation or methylation[79]. Although subcellular compartmentation has been relatively well characterized, the multicellular coordination of biosynthetic pathways across different cell types remains limitedly understood. Recent advances in single-cell and spatial multi-omics technologies are bridging this gap by providing previously unrecognized layers of cellular-level organization. The recent highly explored studies of Taxol and monoterpene indole alkaloids (MIAs) biosynthesis in *T. mairei* and *C. roseus*, are selected herein to exemplify how these integrative approaches are redefining our understanding of metabolic compartmentalization[15,43,67,71,82].

In *Taxus* species, the biosynthesis of the well-known anticancer drug Taxol involves more than 18 enzymatic steps that begin with the cyclization of geranylgeranyl diphosphate (GGPP) to taxadiene by taxadiene synthase (TXS), followed by a series of hydroxylations, acetylations, and side-chain attachments

taxane oxetanase 1; TBT, taxane-2α-O-benzoyltransferase; DBAT, 10-deacetyl baccatin III-10-*O*-acetyltransferase; BAPT, baccatin III-3-amino,13-phenylpropanoyltransferase; T2′OGD, taxane 2′α-hydroxylase; T3′NBT, *N*-debenzoyl-taxol-3′-*N*-benzoyltransferase; PAM, phenylalanine aminomutase; PCL, β-phenylalanine coenzyme A ligase. (C) Cellular organization of MIAs biosynthesis in *C. roseus* leaves. Recent studies have provided the first comprehensive genomics, sc transcriptomics, and sc metabolomics overview of MIAs biosynthesis, which offer valuable insights into the architecture and regulation of MIAs biosynthetic routes. MEP and early secoiridoid pathways occur in IPAP cells, while vindoline intermediates are formed in epidermal cells, with final steps restricted to idioblasts and laticifers. The colour of the arrows represents the catalysed steps are distributed in corresponding cell types. GES, geraniol synthase; IRS, iridoid synthase; LAMT, loganic acid *O*-methyltransferase; SLS, secologanin synthase; TDC, tryptophan decarboxylase; STR, strictosidine synthase; SGD, strictosidine β-glucosidase; GS, geissoschizine synthase; GO, geissoschizine oxidase; SAT, stemmadenine *O*-acetyltransferase; PAS, precondylocarpine acetate synthase; DPAS, dehydroprecondylocarpine acetate synthase; CS, catharanthine synthase; TS, tabersonine synthase; T16H2, tabersonine 16-hydroxylase 2; 16OMT, 16-hydroxytabersonine *O*-methyltransferase; T3R, tabersonine 3-reductase; T3O, tabersonine 3-oxygenase; DMT, 2,3-dihydrotabersonine *N*-methyltransferase; D4H, deacetoxyvindoline 4-hydroxylase; DAT, deacetylvindoline *O*-acetyltransferase. (D) Regulatory network of the terpenoid pathway in *Gossypium hirsutum* secretory glandular cell. *GoHSFA4a* and *GoNAC42* are constitutively expressed *via* GoPGF binding, while their expression can also be directly induced by pathogen and insect herbivore attack. These transcription factors subsequently activate genes in the MVA pathway, gossypol pathway and terpene synthases, thereby initiating synthesis of hemigossypol and volatile terpenes. Terpene volatilization is represented by curved arrowheads. Created with BioRender.com.

(Fig. 4B)[11,83]. While conventional bulk omics studies have identified the gene set involved in Taxol biosynthesis, the spatial dynamics and regulatory mechanisms governing their expression remained unresolved until recent advances in single-cell technologies. Using scRNA-seq combined with MALDI-MSI, Shen et al. mapped the transcriptional and metabolic landscape of *Taxus* leaves and stems at single-cell resolution, respectively[43,71,82]. These studies revealed that most Taxol biosynthesis-related genes predominantly express in leaf mesophyll cells, while stems exhibit a distinct pattern, with gene activity distributed across epidermal, endodermal, and xylem parenchyma cells, suggesting a different pattern of biosynthetic organization between leaves and stems[43,71,82]. In the leaves of *T. mairei*, Taxol biosynthesis pathway initiates in chloroplast with the synthesis of the terpene precursor GGPP *via* MEP pathway. GGPP is then converted to core diterpene skeleton of Taxol, taxa-4(5),11(12)-diene, by taxadiene synthase (TXS). From this point, the pathway diverges into two proposed routes for key intermediate baccatin III synthesis. Taxa-4(5),11(12)-diene in the first route is transported to endoplasmic reticulum (ER), where it undergoes multistep hydroxylation at C1, C2, C5, C7, C9, C10, C13 and further oxidation at C9, along with oxetane ring formation and acetylation, ultimately yielding baccatin III. In another route, taxadiene is likely shuttled to the cytoplasm *via* plastid-ER contact sites and undergoes synergistic catalysis by sequential ER-anchored oxidases (T5αOH, T10βOH, T13αOH, T2αOH, T9αOH, T7βOH, TOT1) and three cytoplasmic acyltransferases (TAT, TBT, DBAT) to form baccatin III. The cytoplasmic side chain β-Phe-CoA synthesized by PAM and PCL from α-Phe are then assembled with baccatin III *via* BAPT, followed by further modifications by T2′OGD and T3′NBT to produce Taxol (Fig. 4B)[84]. Although the exact pathway portioning in *T. mairei* stem remains to be fully characterized, proteomic analysis using LCM coupled with MS/MS localized several Taxol biosynthetic enzymes to the endodermis of young stems, and identified MYB1L and its regulatory microRNA miR858b as key players in endodermis-specific pathway regulation[85]. Furthermore, regulatory insights from scATAC revealed differential chromatin accessibility across cell types in leaves, with transcription factor (TF) motifs (*e.g.*, TCP4, KAN4, CCA1, TCP19) differently enriched in mesophyll and epidermal cells, likely suggesting that the cell-type specific TFs may act distinct roles in control of spatial pathway activation and accumulation of SMs[43].

A parallel example is seen in *Catharanthus roseus*, a medicinal plant that produces a structurally diverse group of MIAs, including the dimeric vinblastine and vincristine, which are important anticancer agents[86]. Over the past three decades, a combination of biochemical assays and bulk omics approaches led to the identification of 38 MIA biosynthetic genes and several associated transcription factors[87]. MIA biosynthesis in *C. roseus* is organ-specific, with vinblastine and vincristine produced in the leaves and hörhammercine synthesized in the roots[88]. Recent work by Caputi, O'Connor, Buell, and colleagues has significantly advanced the field through the integration of telomere-to-telomere genome assembly, single-cell transcriptomics, and spatial metabolomics[15]. The complete genome revealed over 200 paralogs of known MIA genes, emphasizing gene duplication as a central mechanism driving metabolic diversification. While linear clustering of MIA genes was limited, 3D chromatin conformation analyses uncovered topologically associated domains that organize biosynthetic genes into spatial neighbourhoods across different chromosomes, suggesting a chromatin-level mechanism for their coordinated regulation. This architectural insight enabled the prediction and discovery of novel genes, including key secologanin transporters[15]. Importantly, single-cell transcriptomic profiling of *C. roseus* leaves revealed that MIA biosynthesis is sequentially compartmentalized across multiple specialized cell types[15,53]. Specifically, the entire MEP pathway and the initial seven steps of the secoiridoid pathway are localized in internal phloem-associated parenchyma (IPAP) cells. Transcripts for loganic acid *O*-methyltransferase (LAMT) to tabersonine synthase (TS) and catharanthine synthase (CS), as well as tryptophan decarboxylase (TDC), the final enzyme in the indole biosynthesis pathway, are highly concentrated in epidermal cells. The upstream steps in the vindoline pathway, from tabersonine to 3-OH-16-OMe-2,3-dihydrotabersonine, predominantly occur in epidermal cells, while the last three steps (*i.e.*, DMT, D4H, DAT) are restricted to idioblasts and laticifers (Fig. 4C)[36,67]. This spatial expression pattern was also validated by single-cell metabolomics, which pinpointed metabolite accumulation at the corresponding cellular sites. Surprisingly, although genes directly involved in the catharanthine biosynthesis are predominantly expressed in epidermal cells, cathranthine itself was found to accumulate primarily in the idioblasts, reaching a concentration as high as 100 mmol/L. This unexpected distribution suggests the existence of a rapid and directed transport mechanism, likely mediated by specific transporters, which facilitates the movement of catharanthine from the epidermis to idioblasts rather than to the cuticle as previously described (Fig. 4C)[53,89]. The findings challenge earlier models of metabolite localization and highlight the dynamic nature of intercellular trafficking in specialized metabolism. Moreover, the integrated multi-omics approach enabled the identification of previously uncharacterized biosynthetic enzymes, including an anhydrovinblastine-forming reductase, thereby resolving a long-standing gap in the vinblastine biosynthetic pathway[15,53].

Collectively, these studies underscore that metabolic pathway compartmentalization in plants is a highly organized, multi-scale phenomenon. It operates not only at the subcellular level, but also across distinct cell types, tissues, and even through higher-order chromatin architecture. These insights offer valuable guidance for investigating metabolic pathways in other medicinal plants, emphasizing the necessity of integrating multi-omics approaches—including single-cell transcriptomics, spatial metabolomics, and 3D genome organization. Importantly, they also call for careful interpretation of discrepancies between gene expression and metabolite accumulation, as such inconsistencies may reveal complex regulatory mechanisms or transport processes that are essential to fully understanding specialized metabolism.

3.3. *Single-cell and spatial omics in mapping pathway regulatory networks and intercellular communication*

The biosynthesis of specialized metabolites in medicinal plants is orchestrated by intricate gene regulatory networks (GRNs) that integrate developmental cues, environmental signals, and metabolic demands. With the advent of single-cell transcriptomics and chromatin accessibility profiling, researchers can now resolve these networks at single-cell resolution, uncovering the TFs, *cis*-regulatory elements, and epigenetic landscapes that dictate cell-type-specific expression of biosynthetic genes. Spatial transcriptomics further enables the mapping of these regulatory modules across distinct sub-tissues and developmental zones, offering insights into the spatial dynamics of regulation that are obscured in bulk tissue analyses. These tools have been

extensively applied in the context of abiotic stress response in both model plants and crops, and a suite of computational pipelines has been developed to infer GRNs from plant single-cell data, such as SCENIC[90], and LINGER[91].

In *Gossypium hirsutum*, scRNA-seq analysis revealed that secretory glandular cells (SGCs) function as dedicated metabolic hubs, exclusively expressing the complete biosynthetic machinery for both volatile terpenes (*e.g.*, *β*-caryophyllene) and non-volatile gossypol-type terpenoids (Fig. 4D)[70]. Critically, the sc transcriptional profiling uncovered a hierarchical regulatory network in SGCs, with newly identified transcription factors GoHSFA4a and GoNAC42 acting downstream of gland formation genes to directly coordinate terpenoid pathway genes (*e.g.*, HMGS, HMGR, CYP736A196, DH1)[70]. This cell-type-restricted regulatory architecture enables cotton to spatially segregate defence strategies, producing volatile terpenes for indirect defence while accumulating non-volatile gossypol for direct pest deterrence (Fig. 4D). A similar cell-type-specific regulatory mechanism was uncovered in *Cinnamomum camphora* leaves, where scRNA-seq analysis resolved eight transcriptionally distinct cell populations and localized terpenoid biosynthetic genes (*e.g.*, *GPPS1*, *4CLL9*) primarily to mesophyll and epidermal cells[92]. Through pseudotime trajectory analysis, the authors reconstructed mesophyll developmental progression, linking cell fate transitions with dynamic shifts in terpenoid and lipid metabolism. Single-cell co-expression networks further identified transcriptional regulators, including MYB and bHLH family TFs, as well as ethylene-responsive factors involved in hormonal regulation of the pathway[92]. Most recently, Li and colleagues integrated single-cell gene expression, chromatin accessibility profiles, and TF-binding site analysis to construct the first cell-type-resolved gene regulatory network for MIA biosynthesis[93]. Their analysis identified *Idioblast MYB1* (*CrIDM1*) as a novel idioblast-specific TF that activates late-stage MIA biosynthetic genes, representing the first known case of a cell-type-specific TF coordinating a defined metabolic regulon in *C. roseus*[93]. This highlights the power of single-cell approaches to uncover previously inaccessible regulatory logic underlying plant metabolic diversity.

Beyond transcriptional control, cell−cell communication (CCC) plays a pivotal role in coordinating metabolism and development across spatially distinct cell types. At the cellular level, plasmodesmata facilitate the exchange of signalling molecules, proteins, and small RNAs between adjacent cells. Well-known examples include auxin and cytokinin transport, which mediate lateral root initiation and legume-rhizobia symbiosis through autoregulation of nodulation pathways[94,95]. These communication channels are essential not only for developmental patterning but also for enabling metabolic division of labor, a hallmark of many plant biosynthetic pathways. Emerging computational approaches now leverage ligand−receptor gene pair expression in single-cell datasets to infer intercellular communication networks[96]. For instance, using the PlantPhoneDB resource, Xu et al. analysed scRNA-seq data of *Arabidopsis* root under heat stress and identified the AT1G28290−AT2G14890 ligand−receptor pair as a potential mediator signalling between atrichoblasts and cortex cells, offering insights into how stress responses are spatially regulated at the single-cell level[97]. Importantly, CCC also contributes to specialized metabolism by coordinating intercellular metabolite exchange in pathways with cell-type segregated steps. This is exemplified by MIAs pathway in *C. roseus*, where vindoline and catharanthine are synthesized in separated idioblasts and epidermal cells, hence requiring the coordinate transport into the same cell types for coupling reaction to yield the final vinblastine and vincristine products[15,67]. Such transport-mediated coordination is likely underpinned by spatially defined ligand−receptor signalling, though the molecular mediators remain unknown.

Together, these studies underscore that plant specialized metabolism is shaped not only by cell-intrinsic transcriptional programs, but also by extrinsic intercellular communication systems. The integration of single-cell transcriptomics, chromatin accessibility mapping, and cell−cell signalling inference is transforming our understanding of metabolic regulation, facilitating precise engineering of biosynthetic pathways through targeted modulation of regulatory hierarchies and communication circuits.

4. Application of single-cell and spatial omics in plant bioengineering and synthetic biology

Plants produce a diverse repertoire of specialized metabolites with pharmaceutical value, yet their low abundance, tissue-specific distribution, and complex cellular localization hinder efficient and sustainable production. Although synthetic biology and metabolic engineering offer scalable solutions, these approaches depend on a precise understanding of biosynthetic genes, transporters, and regulatory networks. Despite advances enabled by bulk omics, elucidating plant biosynthetic pathways remains challenging due to gene dispersion and cellular heterogeneity.

Recent advances in single-cell and spatial omics have transformed genome-to-pathway discovery by resolving transcriptional and metabolic landscapes at cellular resolution. Single-cell transcriptomics enables the deconvolution of complex tissues, allowing for the assignment of candidate biosynthetic genes to specific cell types where the metabolites accumulate. Integration with spatial metabolomics and chromatin accessibility profiling further allows the identification of co-expressed gene modules, cell-type-specific transporters, transcription factors, and *cis*-regulatory elements that govern pathway activation. These high-resolution insights accelerate the discovery of biosynthetic pathway/genes and regulatory circuits, facilitating the rational design of heterologous production platforms for medicinal plant-derived natural products[21,60].

Medicinal plants are composed of a complex mosaic of specialized cell types, each performing distinct metabolic and physiological roles within spatially organized tissues. This cellular heterogeneity presents both a challenge and an opportunity for plant metabolic engineering and synthetic biology[21,98,99]. Traditional gene manipulation in plant often rely on constitutive promoters, such as the CaMV *35S* promoter, which may introduce pleiotropic effects that impair growth or stress responses[100]. In contrast, cell-type-specific promoters provide precise control over gene expression, enabling targeted metabolic rewiring while minimizing off-target effects. For example, to alleviate the growth penalties associated with overexpression of a lignin biosynthesis repressor in *Populus*, Gui et al.[101] employed a xylem-fiber-specific promoter to restrict gene expression to fiber cells. This approach improved wood biomass quality and saccharification efficiency, demonstrating that the cell-type-specific gene expression can enhance desired traits without compromising overall plant health. Such promoters have become indispensable tools in plant metabolic engineering, particularly for tailoring specialized metabolism in specific tissues or organs (Fig. 5A and D)[98,99]. Zhu and

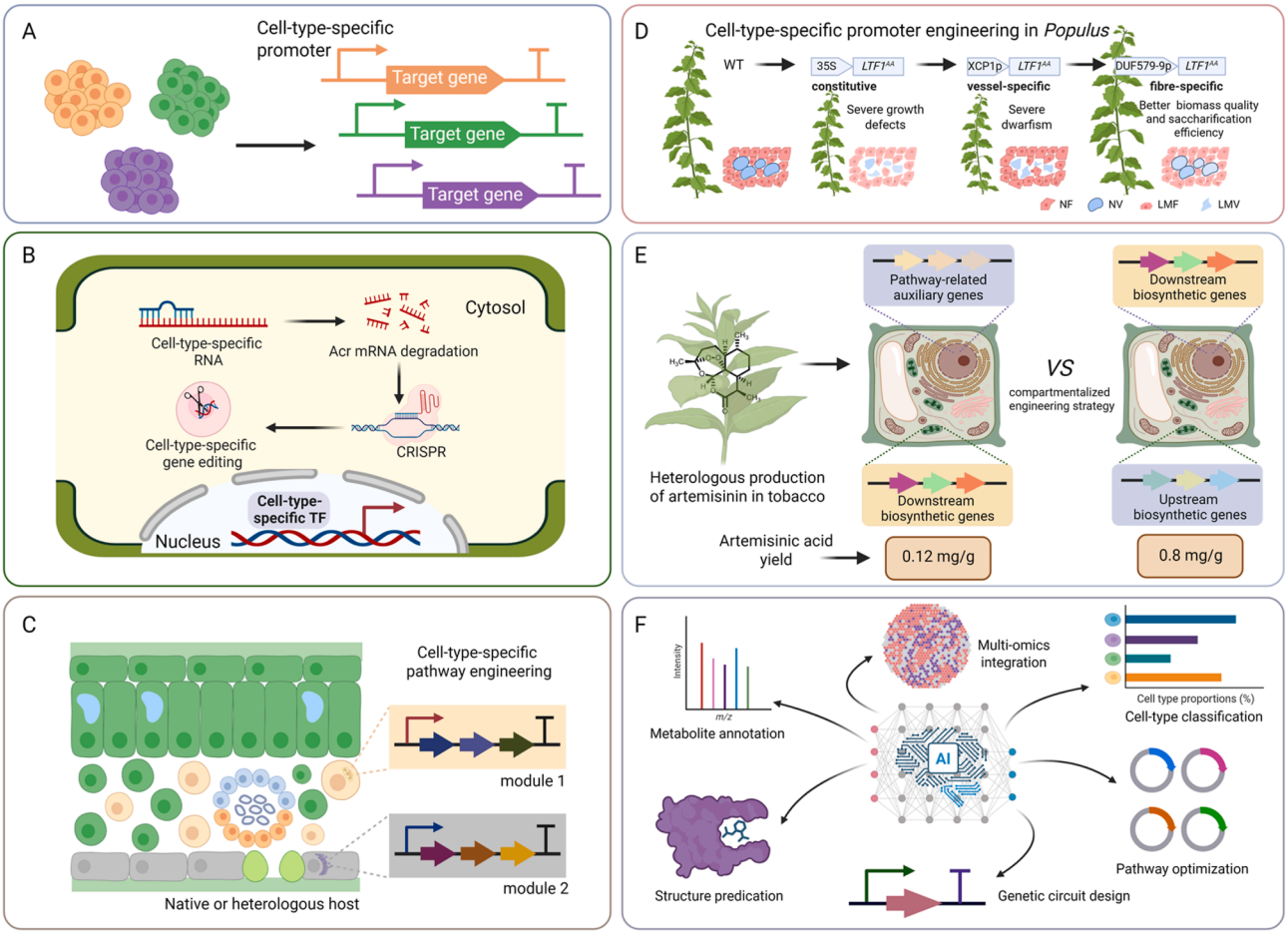

Figure 5 Harnessing single-cell and spatial omics to advance precision plant bioengineering and synthetic biology. (A) Cell-type-specific promoter discovery and engineering for spatially controlled gene expression. (B) Discovery of cell-type-specific transcription factors for cell-targeted CRISPR mediated gene editing. (C) Cell-type-specific promoter engineering and compartmentation engineering in plants. (D) Precise lignin pathway engineering in *Populus* resulted in improved biomass quality and saccharification efficiency. Cell-specific promoter engineering was applied to both vessel and fibre cells. The results demonstrated that selective suppression of lignin biosynthesis in fibre cells alone achieved the desired growth performance. *35S*, constitutive expression promoter; $LTF1^{AA}$, a repressor of lignin biosynthesis; XCP1P, vessel-specific promoter; DUF579-9P, fibre-specific promoter; LMF, lignin-modified fibres; LMV, lignin-modified vessels; NF, normal fibers; NV, normal vessels. (E) Heterologous biosynthesis of artemisinin in tobacco. Overexpression of the AMA pathway in plastid yielded 0.12 mg/g dry weight of artemisinin, whereas integrating MVA genes into chloroplast and AMA genes into the nuclear genome increased production to 0.8 mg/g. AMA, artemisinic acid. (F) AI accelerates the application of single-cell omics in plant bioengineering and synthetic biology by enhancing cell-type classification, multi-omics integration, metabolite annotation, rational design of synthetic circuit and pathway optimization. Created with BioRender.com.

co-workers utilized eight different endosperm-specific promoters to drive anthocyanin biosynthesis, resulting in the development of purple endosperm rice with high anthocyanin content and antioxidant activity specifically in the endosperm. This was achieved by the resulting activation of 13 endogenous anthocyanin biosynthetic genes, which are normally silenced or expressed at low levels in wild-type rice endosperm[102]. In addition to metabolic engineering, CRISPR/Cas9-based genome editing tools are increasingly important in plant biotechnology. Employing tissue- or cell-specific factors (*e.g.*, TF, miRNA) to drive *Cas9* expression has proven effective in enhancing editing efficiency while minimizing off-target effects (Fig. 5B)[99,103]. In *Arabidopsis*, promoters such as the meristematic tissue-specific *YAO* and *CLAVATA3*, root cap-specific *SOMBRERO*, stomatal lineage-specific *TMM* and *FAMA*, and lateral root primordia-specific *GATA23* have been used to drive *Cas9* expression, resulting in improved mutation efficiency[103,104].

Beyond promoter engineering, single-cell and spatial omics technologies are emerging as foundational tools in guiding plant synthetic biology, particularly in reconstructing specialized metabolite pathways in heterologous hosts (Fig. 5C). Plants and microbes are increasingly recognised as valuable chassis for sustainable production of bioactive specialized metabolites. In this context, insights into the pathway cellular/subcellular compartmentalization gained through single cell (*e.g.*, Stereo-seq) and imaging (*e.g.*, MERFISH) technologies are instrumental. A landmark achievement in this field was the industrial-scale biosynthesis of artemisinic acid (AMA), a precursor to the antimalarial drug artemisinin, in engineered yeast[105]. Upon this success, an earlier attempt in *N. benthamiana* employs a compartmentalized

engineering strategy by introducing the downstream AMA pathway into the plastid genome and the flux enhancement genes into the nuclear genome, which achieved an AMA yield of 0.12 mg/g dry weight[106]. Later, Malhotra et al.[107] implemented a swapped approach by targeting upstream MVA genes to chloroplasts and downstream AMA genes to nuclear genome. This optimized spatial separation boosted metabolic flux, achieving clinically relevant yields (∼0.8 mg/g dry weight) and enabling oral delivery of artemisinin through bioencapsulated plant cells (Fig. 5E). Together, these advances underscore the importance of precise spatial compartmentalization in metabolic pathway engineering and optimization. Additionally, transporter engineering provides an alternative strategy to optimize specialized metabolite biosynthesis, particularly for complex plant pathways that span multiple organelles or tissues. Srinivasan et al.[108] demonstrated that incorporating vacuolar-specific transporters (*i.e.*, *AbPUP1*, *AbLP1*) identified from *Atropa belladonna* into yeast system substantially increased tropane alkaloids production over 100-fold, highlighting the critical role of cellular transport processes in recapitulating complex plant biosynthesis. More recently, single-cell omics studies have unravelled the cellular complexity and compartmentation of MIAs and Taxol biosynthetic pathways, and several cell-type-specific transporters were identified and proposed as key components for future metabolic engineering efforts in these valuable pathways[67,82].

5. Technical challenges and current solutions in plant single-cell omics

Despite the transformative potential of single-cell and spatial omics in plant science, several technical and biological challenges continue to hinder their broad application, particularly in non-model medicinal species and complex tissues. A major obstacle lies in the difficulty of isolating intact plant cells, largely due to rigid, lignified cell walls that are often recalcitrant to enzymatic digestion (Table 2)[20,69,109,110]. Cells with secondary cell wall thickening, such as xylem elements and sclerenchyma, are especially resistant to digestion, leading to their underrepresentation in dissociation-based single-cell transcriptomes[26]. Protoplast isolation, while widely used, introduces additional complications. The enzymatic and osmotic conditions required for protoplasting can trigger stress-related transcriptional responses and even cause cell lysis, compromising both biological accuracy and data quality[111]. To circumvent these issues, snRNA-seq has emerged as a promising alternative, particularly for frozen or fixed samples[23]. Bypassing the need for cell wall digestion, snRNA-seq enables access to otherwise intractable tissues. However, it captures fewer transcripts than cytoplasmic RNA-seq and includes a higher proportion of immature pre-mRNAs, complicating the interpretation of transcriptional dynamics[23]. While snRNA-seq reflects nascent transcriptional activity, scRNA-seq captures cumulative gene expression, underscoring the complementary nature of the two approaches. Spatial transcriptomics provides an orthogonal strategy by preserving tissue architecture and enabling *in situ* localization of gene expression, offering particular value for dissecting plant specialized metabolism. However, current spatial platforms often suffer from limited resolution, which prevents single-cell resolution in densely packed or morphologically complex plant tissues. Additionally, high operational costs limit throughput and accessibility, especially for large-scale or comparative studies across diverse species and conditions[112].

Another critical bottleneck is the limited availability of well-characterized cell-type-specific marker genes in most plant species (Table 2). Unlike *A. thaliana*, many medicinal and non-model plants lack validated markers for defining distinct cell populations, complicating cluster annotation and hindering biological interpretation of single-cell data[113]. The issue is further compounded by the lack of reference genome information for many species, or the poor quality of existing genome annotations (*e.g.*, fragmentation, high repetitive sequences, high heterozygosity, or polyploid complexity), which impair accurate gene annotation and cell type identification[114,115]. Fortunately, recent advancements in long-read sequencing technologies (*e.g.*, PacBio HiFi, Nanopore), chromatin conformation capture (Hi-C), and genome assembly algorithms are driving the generation of high-quality plant genomes, thereby supporting deeper investigation into plant cellular heterogeneity, development, and specialized metabolism[114]. In addition to experimental challenges, computational challenges also limit progress (Table 2). scRNA-seq generates inherently sparse and noisy datasets, prone to batch effects, transcript dropout, and amplification biases[116]. These artifacts may result in misleading interpretations, such as inflated cell-to-cell variability or spurious gene co-expression networks. Moreover, most existing bioinformatic pipelines were developed for animal systems, and often require significant adaptation to accommodate plant-specific

Table 2 Technical challenges and current solutions in plant single-cell omics.

Category	Specific challenge	Current solution	Ref.
Sample preparation	Rigid cell walls resist enzymatic digestion; Stress responses during protoplasting; Secondary metabolites interference; Low protoplast yields	Optimized enzyme cocktails; Minimizing protoplasting duration, using transcription inhibitors; SnRNA-seq, chemical quenching; Protocol optimization, snRNA-seq	19, 20
Reference resources	Lack of cell-type markers in non-models; Incomplete genome assemblies; Limited annotation in many plants	Cross-species marker transfer; PacBio, Nanopore, Hi-C; Ortholog-based function inference	20, 109
Computational analysis	Data sparsity and dropout noise; Data heterogeneity and scale discrepancies; Incompatibility of animal-centric pipelines Low detection of proteins/metabolites	Deep sequencing, replicate integration; Data standardization Plant-optimized pipelines; NanoLC−MS platform	20, 110

snRNA-seq: single nucleus RNA sequencing; Hi-C: high-throughput chromosome conformation capture; nanoLC−MS: nano-liquid chromatography mass spectrometry.

complexities, including variable ploidy, cell wall-related gene expression, and unique transcriptional programs[117]. Finally, the development of single-cell proteomics and metabolomics in medicinal plants remains in its infancy. Technical hurdles in isolating sufficient quantities of intact cells, combined with the low abundance and instability of proteins and metabolites at the single-cell level, have made it exceedingly difficult to construct integrated, multi-omic, cell-type-resolved profiles. This limits our ability to fully elucidate specialized metabolite biosynthesis in complex plant systems.

In summary, while single-cell and spatial omics technologies are reshaping plant science, their application in plant metabolic biosynthesis remains constrained by biological barriers, technical hurdles, and computational limitations. Addressing these challenges will require continued innovation in cell isolation methods, sequencing platforms, genome assembly and annotation, and the development of plant-specific analytical tools, as well as expanded efforts to generate integrated multi-omics datasets across diverse plant lineages.

6. Concluding remarks and future perspectives

Recent advances in single-cell and spatial omics technologies have revolutionized our ability to dissect medicinal plant specialized metabolism at high spatiotemporal resolution. These tools enable the identification of cell-type-specific biosynthetic genes and reveal the spatial organization of metabolic processes that are often obscured by traditional bulk approaches. By profiling gene expression patterns across diverse cell types, researchers can map cell-type-specific regulatory networks, characterize cell-cell communication and signalling, and pinpoint critical promoter and transporter elements for unravelling the complexity of specialized metabolism and encouraging plant biotechnological application. The integration of sc transcriptomic with other omics technologies (*e.g.*, proteomics and metabolomics) have facilitated a more comprehensive understanding of cellular metabolism. This multi-omics approach provides unprecedented insights into biosynthetic gene function, regulatory logic, and metabolite flux, thereby advancing plant metabolic engineering and synthetic biology. These tools are now facilitating the rational design of plants or heterologous hosts to enhance specialized metabolite production. Examples such as the dissection of the monoterpene indole alkaloid (MIA) and Taxol pathways illustrate the growing impact of single-cell and spatial omics in uncovering complex metabolic networks[15,52,71,118]. Looking forward, expanding single-cell multi-omics beyond model species and well-known medicinal plants will be crucial to uncovering lineage-specific metabolic innovations and conserved regulatory circuits across plant taxa. As for the specific structural classes, terpenoids and alkaloids are particularly promising targets for single-cell omics studies. Complex terpenoids such as highly oxidized diterpenoids (*e.g.*, triptolide) and glycosylated triterpenoid saponins (*e.g.*, astragalosides) are usually biosynthesized *via* extended, modular pathways that are likely compartmentalized across distinct cell types. Single-cell and spatial omics can thus reveal their biosynthetic logic and enable cell-type-specific engineering. Likewise, pharmacologically important alkaloids like huperzine A and homoharringtonine remain largely uncharacterized at both bulk-tissue and single-cell levels, presenting valuable opportunities for future exploration.

As single-cell datasets grow in size and complexity, robust artificial intelligence (AI)-driven data integration and analytical tools are becoming essential. Machine learning (ML), both unsupervised and supervised, is emerging as a powerful approach for mining high-dimensional datasets, enabling the discovery of cryptic biosynthetic genes and novel metabolic pathways. Recent advances in this area include the development of AI tools such as scVI and Scanpy for dimensionality reduction and cell-type classification, Cell2location and Tangram for integrating scRNA-seq with spatial transcriptomics, and DeepCell and PlantSeg for image-based phenotypic analysis. In metabolomics, deep learning models applied to MS/MS spectra (*e.g.*, SIRIUS, Spec2Vec) are improving metabolite annotation and structure prediction[119-122]. These computational frameworks are accelerating the elucidation of biosynthetic networks, pathway prioritization, and the rational design of synthetic circuits in plant biosystems (Fig. 5F). For instance, ML has successfully identified biosynthetic enzymes and transporters in *C. roseus* and *Rauvolfia serpentina*[123,124], while AlphaFold3-driven structural predictions now enable enzyme−substrate interaction modeling and functional annotation of orphan genes[125,126]. Nevertheless, robust AI applications depend on high-quality, well-annotated multi-omics datasets. Public databases such as the Phytozome, PlantscRNAdb, scPlantDB, and SingPro, as well as curated datasets (*e.g.*, SpatialDB) from spatial and temporal multi-omics studies, could serve as valuable resources for training and validation[127-129].

The convergence of single-cell omics, synthetic biology, and AI-driven design is setting to transform plant metabolic engineering by enabling more precise, efficient, and predictable manipulation of complex biosynthetic pathways. High-resolution data from scRNA-seq will play a central role in this transformation, providing critical insights into cellular processes and gene regulation that support the Design-Build-Test-Learn cycle. Further efforts will increasingly emphasize compartmentalized metabolic engineering, wherein biosynthetic pathways are targeted to specific subcellular organelles to enhance metabolic flux, catalytic efficiency, and product yields. AI will be integral to this process, facilitating large-scale prediction of gene function, entire pathway optimization, and *de novo* design of synthetic routes through the strategic combination of enzymes and localization signals. As costs decrease and accessibility improves, scRNA-seq is expected to become a routine tool for analyzing gene expression in engineered plants. Together, these advances will accelerate the development of programmable, high-performance plant platforms for sustainable production of pharmaceuticals and other valuable SMs.

Acknowledgments

This research was supported by the National Key Research and Development Program of China (No. 2024YFC3506504), the National Natural Science Foundation of China (Nos. U23A20513, Xiaohui Fan; 82404455, Guang-Lei Ma), “Pioneer” and “Leading Goose” R&D Program of Zhejiang (No. 2024C03106, Xiaohui Fan, China), the Natural Science Foundation of Zhejiang (Nos. LZ24H300001 and LD25H280001, China), the Starlit South Lake Leading Elite Program (2023A303005, Xiaohui Fan, China), Joint TCM science & Technology Projects of National Demonstration Zones for Comprehensive TCM reform (GZY-KJS-ZJ-2025-078, China), China Postdoctoral Science Foundation (2024M762891, Guo Li), and the Fundamental Research Funds for the Central Universities (226-2024-00001, Xiaohui Fan; 226-2024-00149, Guang-Lei Ma, China).

Author contributions

Xiaohui Fan and Guang-Lei Ma conceptualized the review topic and supervised the project. Guo Li, Miaomiao Cai, and Xinhao Liao conducted literature research and synthesized key insights. Xiaohui Fan, Guang-Lei Ma, Guo Li, and Miaomiao Cai drafted the manuscript. All authors contributed to revising and refining the intellectual content.

Conflicts of interest

The authors declare no conflict of interest.

Appendix A. Supporting information

Supporting information to this article can be found online at https://doi.org/10.1016/j.apsb.2025.12.023.

References

1. Newman DJ, Cragg GM. Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019. *J Nat Prod* 2020;**83**:770−803.
2. Erb M, Kliebenstein DJ. Plant secondary metabolites as defenses, regulators, and primary metabolites: the blurred functional trichotomy. *Plant Physiol* 2020;**184**:39−52.
3. Cravens A, Payne J, Smolke CD. Synthetic biology strategies for microbial biosynthesis of plant natural products. *Nat Commun* 2019;**10**:2142.
4. Afendi FM, Okada T, Yamazaki M, Hirai-Morita A, Nakamura Y, Nakamura K, et al. KNApSAcK family databases: integrated metabolite−plant species databases for multifaceted plant research. *Plant Cell Physiol* 2012;**53**:e1.
5. Caspi R, Billington R, Keseler IM, Kothari A, Krummenacker M, Midford PE, et al. The MetaCyc database of metabolic pathways and enzymes-a 2019 update. *Nucleic Acids Res* 2020;**48**:D445−53.
6. Schläpfer P, Zhang P, Wang C, Kim T, Banf M, Chae L, et al. Genome-wide prediction of metabolic enzymes, pathways, and gene clusters in plants. *Plant Physiol* 2017;**173**:2041−59.
7. Jacobowitz JR, Weng JK. Exploring uncharted territories of plant specialized metabolism in the postgenomic era. *Annu Rev Plant Biol* 2020;**71**:631−58.
8. Li Q, Dai Y, Huang XC, Sun L, Wang K, Guo X, et al. The chromosome-scale assembly of the *Notopterygium incisum* genome provides insight into the structural diversity of coumarins. *Acta Pharm Sin B* 2024;**14**:3760−73.
9. Xu Z, Chen S, Wang Y, Tian Y, Wang X, Xin T, et al. Crocus genome reveals the evolutionary origin of crocin biosynthesis. *Acta Pharm Sin B* 2024;**14**:1878−91.
10. Singh KS, van der Hooft JJJ, van Wees SCM, Medema MH. Integrative omics approaches for biosynthetic pathway discovery in plants. *Nat Prod Rep* 2022;**39**:1876−96.
11. Jiang B, Gao L, Wang HJ, Sun YP, Zhang XL, Ke H, et al. Characterization and heterologous reconstitution of biosynthetic enzymes leading to baccatin III. *Science* 2024;**383**:622−9.
12. Liang F, Xie Y, Zhang C, Zhao Y, Motawia MS, Kampranis SC. Elucidation of the final steps in Taxol biosynthesis and its biotechnological production. *Nat Synth* 2025. Available from: https://doi.org/10.1038/s44160-025-00800-z.
13. McClune CJ, Liu JCT, Wick C, De La Peña R, Lange BM, Fordyce PM, et al. Discovery of FoTO1 and Taxol genes enables biosynthesis of baccatin III. *Nature* 2025;**643**:582−92.
14. Shen X, Guan Z, Zhang C, Yan Z, Sun C. The multicellular compartmentation of plant specialized metabolism. *Curr Opin Plant Biol* 2024;**81**:102616.
15. Li C, Wood JC, Vu AH, Hamilton JP, Rodriguez Lopez CE, Payne RME, et al. Single-cell multi-omics in the medicinal plant *Catharanthus roseus*. *Nat Chem Biol* 2023;**19**:1031−41.
16. Wolters FC, Del Pup E, Singh KS, Bouwmeester K, Schranz ME, van der Hooft JJJ, et al. Pairing omics to decode the diversity of plant specialized metabolism. *Curr Opin Plant Biol* 2024;**82**:102657.
17. Shulse CN, Cole BJ, Ciobanu D, Lin J, Yoshinaga Y, Gouran M, et al. High-throughput single-cell transcriptome profiling of plant cell types. *Cell Rep* 2019;**27**:2241−7.
18. Tenorio Berrío R, Verstaen K, Vandamme N, Pevernagie J, Achon I, Van Duyse J, et al. Single-cell transcriptomics sheds light on the identity and metabolism of developing leaf cells. *Plant Physiol* 2022;**188**:898−918.
19. Singh S, Praveen A, Dudha N, Sharma VK, Bhadrecha P. Single-cell transcriptomics: a new frontier in plant biotechnology research. *Plant Cell Rep* 2024;**43**:294.
20. Rhaman MS, Ali M, Ye W, Li B, Wang X. Opportunities and challenges in advancing plant research with single-cell omics. *Genom Proteom Bioinform* 2024;**22**:qzae026.
21. Islam MT, Liu Y, Hassan MM, Abraham PE, Merlet J, Townsend A, et al. Advances in the application of single-cell transcriptomics in plant systems and synthetic biology. *BioDesign Res* 2024;**6**:29.
22. Tang F, Barbacioru C, Wang Y, Nordman E, Lee C, Xu N, et al. mRNA-Seq whole-transcriptome analysis of a single cell. *Nat Methods* 2009;**6**:377−82.
23. Cervantes-Pérez SA, Thibivillliers S, Tennant S, Libault M. Review: challenges and perspectives in applying single nuclei RNA-seq technology in plant biology. *Plant Sci* 2022;**325**:111486.
24. Bawa G, Liu Z, Yu X, Tran LSP, Sun X. Introducing single cell stereo-sequencing technology to transform the plant transcriptome landscape. *Trends Plant Sci* 2024;**29**:249−65.
25. Jaitin DA, Kenigsberg E, Keren-Shaul H, Elefant N, Paul F, Zaretsky I, et al. Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell Types. *Science* 2014;**343**:776−9.
26. Macosko Evan Z, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M, et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* 2015;**161**:1202−14.
27. Datlinger P, Rendeiro AF, Schmidl C, Krausgruber T, Traxler P, Klughammer J, et al. Pooled CRISPR screening with single-cell transcriptome readout. *Nat Methods* 2017;**14**:297−301.
28. Rosenberg AB, Roco CM, Muscat RA, Kuchina A, Sample P, Yao Z, et al. Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. *Science* 2018;**360**:176−82.
29. Rodriques SG, Stickels RR, Goeva A, Martin CA, Murray E, Vanderburg CR, et al. Slide-seq: a scalable technology for measuring genome-wide expression at high spatial resolution. *Science* 2019;**363**:1463−7.
30. Xia K, Sun H-X, Li J, Li J, Zhao Y, Chen L, et al. The single-cell stereo-seq reveals region-specific cell subtypes and transcriptome profiling in Arabidopsis leaves. *Dev Cell* 2022;**57**:1299−310.
31. Zou Y, Tang W, Li B. Exploring natural product biosynthesis in plants with mass spectrometry imaging. *Trends Plant Sci* 2025;**30**:69−84.
32. Iqfath M, Wali SN, Amer S, Hernly E, Laskin J. Nanospray desorption electrospray ionization mass spectrometry imaging (nano-DESI MSI): a tutorial review. *ACS Meas Sci Au* 2024;**4**:475−87.
33. Kawashima Y, Ohara O. Development of a nanoLC−MS/MS system using a nonporous reverse phase column for ultrasensitive proteome analysis. *Anal Chem* 2018;**90**:12334−8.
34. Bagwe K, Gould N, Johnson KR, Ivanov AR. Single-cell omic molecular profiling using capillary electrophoresis-mass spectrometry. *Trends Anal Chem* 2023;**165**:117117.
35. Emmert-Buck MR, Bonner RF, Smith PD, Chuaqui RF, Zhuang Z, Goldstein SR, et al. Laser capture microdissection. *Science* 1996;**274**:998−1001.
36. Carter AD, Bonyadi R, Gifford ML. The use of fluorescence-activated cell sorting in studying plant development and environmental responses. *Int J Dev Biol* 2013;**57**:545−52.

37. Moragues T, Arguijo D, Beneyton T, Modavi C, Simutis K, Abate AR, et al. Droplet-based microfluidics. *Nat Rev Methods Primers* 2023;**3**:32.
38. Whitesides GM. The origins and the future of microfluidics. *Nature* 2006;**442**:368−73.
39. Zilionis R, Nainys J, Veres A, Savova V, Zemmour D, Klein AM, et al. Single-cell barcoding and sequencing using droplet microfluidics. *Nat Protoc* 2016;**12**:44−73.
40. Ziegenhain C, Vieth B, Parekh S, Reinius B, Guillaumet-Adkins A, Smets M, et al. Comparative analysis of single-cell RNA sequencing methods. *Mol Cell* 2017;**65**:631−43.
41. Grindberg RV, Yee-Greenbaum JL, McConnell MJ, Novotny M, O'Shaughnessy AL, Lambert GM, et al. RNA-sequencing from single nuclei. *Proc Natl Acad Sci U S A* 2013;**110**:19802−7.
42. Baek S, Lee I. Single-cell ATAC sequencing analysis: from data preprocessing to hypothesis generation. *Comput Struct Biotechnol J* 2020;**18**:1429−39.
43. Zhan X, Zhang H, Liang X, Kailin H, Lin W, Ma R, et al. Single-cell ATAC sequencing illuminates the *cis*-regulatory differentiation of Taxol biosynthesis between leaf mesophyll and leaf epidermal cells in *Taxus mairei*. *Ind Crops Prod* 2023;**205**:117411.
44. Montes C, Zhang J, Nolan TM, Walley JW. Single-cell proteomics differentiates *Arabidopsis* root cell types. *New Phytol* 2024;**244**:1750−9.
45. Koroleva OA, Cramer R. Single-cell proteomic analysis of glucosinolate-rich S-cells in *Arabidopsis thaliana*. *Methods* 2011;**54**:413−23.
46. Schad M, Lipton MS, Giavalisco P, Smith RD, Kehr J. Evaluation of two-dimensional electrophoresis and liquid chromatography-tandem mass spectrometry for tissue-specific protein profiling of laser-microdissected plant samples. *Electrophoresis* 2005;**26**:2729−38.
47. Bennett HM, Stephenson W, Rose CM, Darmanis S. Single-cell proteomics enabled by next-generation sequencing or mass spectrometry. *Nat Methods* 2023;**20**:363−74.
48. van Wijk KJ, Leppert T, Sun Q, Boguraev SS, Sun Z, Mendoza L, et al. The Arabidopsis PeptideAtlas: harnessing worldwide proteomics data to create a comprehensive community proteomics resource. *Plant Cell* 2021;**33**:3421−53.
49. Pang H, Hu Z. Metabolomics in drug research and development: the recent advances in technologies and applications. *Acta Pharm Sin B* 2022;**12**:4011−39.
50. Duncan KD, Fyrestam J, Lanekoff I. Advances in mass spectrometry based single-cell metabolomics. *Analyst* 2019;**144**:782−93.
51. Hölscher D, Shroff R, Knop K, Gottschaldt M, Crecelius A, Schneider B, et al. Matrix-free UV-laser desorption/ionization (LDI) mass spectrometric imaging at the single-cell level: distribution of secondary metabolites of *Arabidopsis thaliana* and *Hypericum* species. *Plant J* 2009;**60**:907−18.
52. Huang L, Fang M, Cupp-Sutton KA, Wang Z, Smith K, Wu S. Spray-capillary-based capillary electrophoresis mass spectrometry for metabolite analysis in single cells. *Anal Chem* 2021;**93**:4479−87.
53. Vu AH, Kang M, Wurlitzer J, Heinicke S, Li C, Wood JC, et al. Quantitative single-cell mass spectrometry provides a highly resolved analysis of natural product biosynthesis partitioning in plants. *J Am Chem Soc* 2024;**146**:23891−900.
54. Sanders KL, Edwards JL. Nano-liquid chromatography-mass spectrometry and recent applications in omics investigations. *Anal Methods* 2020;**12**:4404−17.
55. Rivera ES, Djambazova KV, Neumann EK, Caprioli RM, Spraggins JM. Integrating ion mobility and imaging mass spectrometry for comprehensive analysis of biological tissues: a brief review and perspective. *J Mass Spectrom* 2020;**55**:e4614.
56. Marx V. Method of the Year: spatially resolved transcriptomics. *Nat Methods* 2021;**18**:9−14.
57. Jiang J, Gill BS. Nonisotopic *in situ* hybridization and plant genome mapping: the first 10 years. *Genome* 1994;**37**:717−25.
58. Espina V, Wulfkuhle JD, Calvert VS, VanMeter A, Zhou W, Coukos G, et al. Laser-capture microdissection. *Nat Protoc* 2006;**1**:586−603.
59. Chen A, Liao S, Cheng M, Ma K, Wu L, Lai Y, et al. Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays. *Cell* 2022;**185**:1777−92.
60. Chen C, Zhang X, Yue M. Spatial multi-omics in medicinal plants: from biosynthesis pathways to industrial applications. *Trends Plant Sci* 2024;**29**:510−3.
61. Wang R, Zhang L, Li X, Sun W, Wang Y, Huang L, et al. Single-cell sequencing and mass spectrometry imaging reveal the multicellular compartmentalization map of plant triterpenes: a ginsenoside *in Panax ginseng* example. *Plant Biotechnol J* 2025. Available from: https://doi.org/10.1111/pbi.70198.
62. Burlat V, Papon N, Courdavault V. Medicinal plants enter the single-cell multi-omics era. *Trends Plant Sci* 2023;**28**:1205−7.
63. Atanasov AG, Zotchev SB, Dirsch VM, Orhan IE, Banach M, Rollinger JM, et al. Natural products in drug discovery: advances and opportunities. *Nat Rev Drug Discov* 2021;**20**:200−16.
64. Xu Z, Tian Y, Wang J, Ma Y, Li Q, Zhou Y, et al. Convergent evolution of berberine biosynthesis. *Sci Adv* 2024;**10**:eads3596.
65. Xu B, Huang JP, Peng G, Cao W, Liu Z, Chen Y, et al. Total biosynthesis of the medicinal triterpenoid saponin astragalosides. *Nat Plants* 2024;**10**:1826−37.
66. Zhao Y, He Y, Han L, Zhang L, Xia Y, Yin F, et al. Two types of coumarins-specific enzymes complete the last missing steps in pyran- and furanocoumarins biosynthesis. *Acta Pharm Sin B* 2024;**14**:869−80.
67. Sun S, Shen X, Li Y, Li Y, Wang S, Li R, et al. Single-cell RNA sequencing provides a high-resolution roadmap for understanding the multicellular compartmentation of specialized metabolism. *Nat Plants* 2022;**9**:179−90.
68. Yu X, Liu Z, Sun X. Single-cell and spatial multi-omics in the plant sciences: technical advances, applications, and perspectives. *Plant Commun* 2023;**4**:100508.
69. Chen C, Ge Y, Lu L. Opportunities and challenges in the application of single-cell and spatial transcriptomics in plants. *Front Plant Sci* 2023;**14**:1185377.
70. Lin JL, Chen L, Wu WK, Guo XX, Yu CH, Xu M, et al. Single-cell RNA sequencing reveals a hierarchical transcriptional regulatory network of terpenoid biosynthesis in cotton secretory glandular cells. *Mol Plant* 2023;**16**:1990−2003.
71. Zhan X, Qiu T, Zhang H, Hou K, Liang X, Chen C, et al. Mass spectrometry imaging and single-cell transcriptional profiling reveal the tissue-specific regulation of bioactive ingredient biosynthesis in *Taxus* leaves. *Plant Commun* 2023;**4**:100630.
72. Kang M, Choi Y, Kim H, Kim SG. Single-cell RNA-sequencing of *Nicotiana attenuata* corolla cells reveals the biosynthetic pathway of a floral scent. *New Phytol* 2022;**234**:527−44.
73. Wu S, Morotti ALM, Yang J, Wang E, Tatsis EC. Single-cell RNA sequencing facilitates the elucidation of the complete biosynthesis of the antidepressant hyperforin in St. John's wort. *Mol Plant* 2024;**17**:1439−57.
74. Wang Q, Wu Y, Peng A, Cui J, Zhao M, Pan Y, et al. Single-cell transcriptome atlas reveals developmental trajectories and a novel metabolic pathway of catechin esters in tea leaves. *Plant Biotechnol J* 2022;**20**:2089−106.
75. Wu S, Malaco Morotti AL, Wang S, Wang Y, Xu X, Chen J, et al. Convergent gene clusters underpin hyperforin biosynthesis in St John's wort. *New Phytol* 2022;**235**:646−61.
76. Adam P, Arigoni D, Bacher A, Eisenreich W. Biosynthesis of hyperforin in *Hypericum perforatum*. *J Med Chem* 2002;**45**:4786−93.
77. Xu S, Baldwin IT, Wang Y, Bing J, Li R, Köllner TG, et al. Evolution of a novel and adaptive floral scent in wild tobacco. *Mol Biol Evol* 2020;**37**:1090−9.
78. Widhalm Joshua R, Dudareva N. A familiar ring to it: biosynthesis of plant benzoic acids. *Mol Plant* 2015;**8**:83−97.
79. Bar-Peled L, Kory N. Principles and functions of metabolic compartmentalization. *Nat Metab* 2022;**4**:1232−44.
80. Shitan N, Yazaki K. Dynamism of vacuoles toward survival strategy in plants. *Biochim Biophys Acta* 2020;**1862**:183127.

81. Vranová E, Coman D, Gruissem W. Network analysis of the MVA and MEP pathways for isoprenoid synthesis. *Annu Rev Plant Biol* 2013;**64**:665−700.
82. Yu C, Hou K, Zhang H, Liang X, Chen C, Wang Z, et al. Integrated mass spectrometry imaging and single-cell transcriptome atlas strategies provide novel insights into taxoid biosynthesis and transport in *Taxus mairei* stems. *Plant J* 2023;**115**:1243−60.
83. Zhao Y, Liang F, Xie Y, Duan YT, Andeadelli A, Pateraki I, et al. Oxetane ring formation in Taxol biosynthesis is catalyzed by a bifunctional cytochrome P450 enzyme. *J Am Chem Soc* 2023;**146**: 801−10.
84. Yin JY, Lai M, Yu XY, Su DD, Xiong XY, Li YL. Comprehensive strategies for paclitaxel production: insights from plant cell culture, endophytic microorganisms, and synthetic biology. *Hortic Res* 2025; **12**:uhae346.
85. Yu C, Zhang D, Zhang L, Fang Z, Zhang Y, Lin W, et al. Role of an endodermis-specific miR858b-MYB1L module in the regulation of Taxol biosynthesis in *Taxus mairei*. *Plant J* 2025;**122**:e70135.
86. O'Connor SE, Maresh JJ. Chemistry and biology of monoterpene indole alkaloid biosynthesis. *Nat Prod Rep* 2006;**23**:532−47.
87. Caputi L, Franke J, Farrow SC, Chung K, Payne RME, Nguyen TD, et al. Missing enzymes in the biosynthesis of the anticancer drug vinblastine in *Madagascar periwinkle*. *Science* 2018;**360**:1235−8.
88. Kulagina N, Méteignier LV, Papon N, O'Connor SE, Courdavault V. More than a Catharanthus plant: a multicellular and pluri-organelle alkaloid-producing factory. *Curr Opin Plant Biol* 2022;**67**:102200.
89. Roepke J, Salim V, Wu M, Thamm AMK, Murata J, Ploss K, et al. Vinca drug components accumulate exclusively in leaf exudates of Madagascar periwinkle. *Proc Natl Acad Sci* 2010;**107**:15287−92.
90. Bravo González-Blas C, De Winter S, Hulselmans G, Hecker N, Matetovici I, Christiaens V, et al. SCENIC+: single-cell multiomic inference of enhancers and gene regulatory networks. *Nat Methods* 2023;**20**:1355−67.
91. Yuan Q, Duren Z. Inferring gene regulatory networks from single-cell multiome data using atlas-scale external data. *Nat Biotechnol* 2024;**43**:247−57.
92. Qin Z, Chen C, Zhang T, Wu Y, Zheng Y. Single-cell RNA sequencing reveals transcriptional regulation and metabolic pathways of terpenoid biosynthesis in developing *Cinnamomum camphora* leaf cells. *Curr Plant Biol* 2025;**42**:100467.
93. Li C, Colinas M, Wood JC, Vaillancourt B, Hamilton JP, Jones SL, et al. Cell-type-aware regulatory landscapes governing monoterpene indole alkaloid biosynthesis in the medicinal plant *Catharanthus roseus*. *New Phytol* 2024;**245**:347−62.
94. Thibivilliers S, Libault M. Enhancing our understanding of plant cell-to-cell interactions using single-cell omics. *Front Plant Sci* 2021;**12**: 696811.
95. Yue K, Beeckman T. Cell-to-cell communication during lateral root development. *Mol Plant* 2014;**7**:758−60.
96. Armingol E, Officer A, Harismendy O, Lewis NE. Deciphering cell−cell interactions and communication from gene expression. *Nat Rev Genet* 2020;**22**:71−88.
97. Xu C, Ma D, Ding Q, Zhou Y, Zheng HL. PlantPhoneDB: a manually curated pan-plant database of ligand-receptor pairs infers cell−cell communication. *Plant Biotechnol J* 2022;**20**:2123−34.
98. Gondalia N, Quiroz LF, Lai L, Singh AK, Khan M, Brychkova G, et al. Harnessing promoter elements to enhance gene editing in plants: perspectives and advances. *Plant Biotechnol J* 2025;**23**:1375−95.
99. Brooks EG, Elorriaga E, Liu Y, Duduit JR, Yuan G, Tsai CJ, et al. Plant promoters and terminators for high-precision bioengineering. *BioDes Res* 2023;**5**:13.
100. Wang Y, Yang K, Jing F, Li M, Deng T, Huang R, et al. Cloning and characterization of trichome-specific promoter of cpr71av1 gene involved in artemisinin biosynthesis in *Artemisia annua* L. *Mol Biol* 2011;**45**:751−8.
101. Gui J, Lam PY, Tobimatsu Y, Sun J, Huang C, Cao S, et al. Fibre-specific regulation of lignin biosynthesis improves biomass quality in Populus. *New Phytol* 2020;**226**:1074−87.
102. Zhu Q, Yu S, Zeng D, Liu H, Wang H, Yang Z, et al. Development of "purple endosperm rice" by engineering anthocyanin biosynthesis in the endosperm with a high-efficiency transgene stacking system. *Mol Plant* 2017;**10**:918−29.
103. Singha DL, Das D, Sarki YN, Chowdhury N, Sharma M, Maharana J, et al. Harnessing tissue-specific genome editing in plants through CRISPR/Cas system: current state and future prospects. *Planta* 2021; **255**:28.
104. Wolter F, Klemm J, Puchta H. Efficient in planta gene targeting in Arabidopsis using egg cell-specific expression of the Cas9 nuclease of *Staphylococcus aureus*. *Plant J* 2018;**94**:735−46.
105. Ro D-K, Paradise EM, Ouellet M, Fisher KJ, Newman KL, Ndungu JM, et al. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* 2006;**440**:940−3.
106. Fuentes P, Zhou F, Erban A, Karcher D, Kopka J, Bock R. A new synthetic biology approach allows transfer of an entire metabolic pathway from a medicinal plant to a biomass crop. *eLife* 2016;**5**:e13664.
107. Malhotra K, Subramaniyan M, Rawat K, Kalamuddin M, Qureshi MI, Malhotra P, et al. Compartmentalized metabolic engineering for artemisinin biosynthesis and effective malaria treatment by oral delivery of plant cells. *Mol Plant* 2016;**9**:1464−77.
108. Srinivasan P, Smolke CD. Engineering cellular metabolite transport for biosynthesis of computationally predicted tropane alkaloid derivatives in yeast. *Proc Natl Acad Sci* 2021;**118**:e2104460118.
109. Adema K, Schon MA, Nodine MD, Kohlen W. Lost in space: what single-cell RNA sequencing cannot tell you. *Trends Plant Sci* 2024; **29**:1018−28.
110. Conte MI, Fuentes-Trillo A, Domínguez Conde C. Opportunities and tradeoffs in single-cell transcriptomic technologies. *Trends Genet* 2024;**40**:83−93.
111. Shaw R, Tian X, Xu J. Single-cell transcriptome analysis in plants: advances and challenges. *Mol Plant* 2021;**14**:115−26.
112. Yin R, Xia K, Xu X. Spatial transcriptomics drives a new era in plant research. *Plant J* 2023;**116**:1571−81.
113. Chau TN, Timilsena PR, Bathala SP, Kundu S, Bargmann BOR, Li S. Orthologous marker groups reveal broad cell identity conservation across plant single-cell transcriptomes. *Nat Commun* 2025;**16**:201.
114. Marks RA, Hotaling S, Frandsen PB, VanBuren R. Representation and participation across 20 years of plant genome sequencing. *Nat Plants* 2021;**7**:1571−8.
115. Cheng LT, Wang ZL, Zhu QH, Ye M, Ye CY. A long road ahead to reliable and complete medicinal plant genomes. *Nat Commun* 2025; **16**:2150.
116. Andrews TS, Kiselev VY, McCarthy D, Hemberg M. Tutorial: guidelines for the computational analysis of single-cell RNA sequencing data. *Nat Protoc* 2020;**16**:1−9.
117. Zappia L, Theis FJ. Over 1000 tools reveal trends in the single-cell RNA-seq analysis landscape. *Genome Biol* 2021;**22**:301.
118. Wang H, Zhang BY, Gong T, Chen TJ, Chen JJ, Yang JL, et al. Construction of acetyl-CoA and DBAT hybrid metabolic pathway for acetylation of 10-deacetylbaccatin III to baccatin III. *Acta Pharm Sin B* 2021;**11**:3322−34.
119. Gao F, Huang K, Xing Y. Artificial intelligence in omics. *Genom Proteom Bioinform* 2022;**20**:811−3.
120. Kleshchevnikov V, Shmatko A, Dann E, Aivazidis A, King HW, Li T, et al. Cell2location maps fine-grained cell types in spatial transcriptomics. *Nat Biotechnol* 2022;**40**:661−71.
121. Dührkop K, Fleischauer M, Ludwig M, Aksenov AA, Melnik AV, Meusel M, et al. SIRIUS 4: a rapid tool for turning tandem mass spectra into metabolite structure information. *Nat Methods* 2019;**16**:299−302.
122. Lopez R, Regier J, Cole MB, Jordan MI, Yosef N. Deep generative modeling for single-cell transcriptomics. *Nat Methods* 2018;**15**:1053−8.
123. Payne RME, Xu D, Foureau E, Teto Carqueijeiro MIS, Oudin A, Bernonville TDd, et al. An NPF transporter exports a central monoterpene indole alkaloid intermediate from the vacuole. *Nat Plants* 2017;**3**:16208.
124. Nguyen TAM, Nguyen TD, Leung YY, McConnachie M, Sannikov O, Xia Z, et al. Discovering and harnessing oxidative

enzymes for chemoenzymatic synthesis and diversification of anti-cancer camptothecin analogues. *Commun Chem* 2021;**4**:177.
125. Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* 2024;**630**:493−500.
126. Yao Y, Chen F, Wu C, Chang X, Cheng W, Wang Q, et al. Structure-based virtual screening aids the identification of glycosyltransferases in the biosynthesis of salidroside. *Plant Biotechnol J* 2025;**23**: 1725−35.
127. He Z, Luo Y, Zhou X, Zhu T, Lan Y, Chen D. scPlantDB: a comprehensive database for exploring cell types and markers of plant cell atlases. *Nucleic Acids Res* 2023;**52**:D1629−38.
128. Chen H, Yin X, Guo L, Yao J, Ding Y, Xu X, et al. PlantscRNAdb: a database for plant single-cell RNA analysis. *Mol Plant* 2021;**14**: 855−7.
129. Lian X, Zhang Y, Zhou Y, Sun X, Huang S, Dai H, et al. SingPro: a knowledge base providing single-cell proteomic data. *Nucleic Acids Res* 2023;**52**:D552−61.