

Single-nucleus transcriptomics reveal the morphogenesis and artemisinin biosynthesis in *Artemisia annua* glandular trichomes

Received: 3 April 2025

Accepted: 28 August 2025

Published online: 30 September 2025

 Check for updatesMinghui Zhang^{1,3}, Mingyu Li^{1,3}, Yanyan An^{1,3}, Chang Liu^{1,3}, Qiaojuan Zhao¹, Danfeng Zhang¹, Baiwei Pan¹ & Hexin Tan^{1,2} 

Artemisinin, the key antimalarial drug, is synthesized in *Artemisia annua* glandular secretory trichomes (GSTs), yet their development and artemisinin's precise cellular origins are unclear. Utilizing single-nucleus RNA sequencing and spatial transcriptomics, we construct a high-resolution cellular atlas mapping metabolic dynamics across GST development. We define three developmental states: the initiation phase, transcriptional activation of core metabolic pathways establishes fundamental cellular machinery; the intermediate phase, marked lipid metabolism activation with coordinated fatty acid and wax biosynthesis, accompanied by active photosynthetic activity; the terminal differentiation phase, metabolite specialized through spatial compartmentalization of terpenoid and lipid biosynthetic pathways. Notably, we discover that six specific secretory cells within the 10-cell GSTs constitute the primary site for artemisinin production. We identify hundreds of hub genes potentially contributing to trichome development or artemisinin biosynthesis. Overall, this study systematically elucidates GST development and artemisinin biosynthesis, revealing its spatial production mechanism and providing essential cellular and genetic foundations for metabolic engineering and fundamental trichome biology.

Artemisinin is a sesquiterpene lactone characterized by its endoperoxide bridge. As the cornerstone of artemisinin-based combination therapies (ACTs), it's also the WHO-recommended first-line treatment for *Plasmodium falciparum* malaria¹. Beyond its antimalarial efficacy, emerging evidence demonstrates artemisinin and its derivatives' therapeutic potential against diverse pathologies^{2–4}, substantially expanding its clinical relevance. This multifaceted bioactive compound is derived from *A. annua* (Asteraceae), within which specialized glandular secretory trichomes (GSTs) act as the biosynthetic factories for the production of artemisinin⁵. These epidermal structures are multifunctional sites for artemisinin synthesis, secretion, and storage^{6,7}, therefore GST density, developmental maturity, and metabolic efficiency critically govern yield.

The artemisinin biosynthetic pathway originates from the precursor molecules isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP), synthesized via the spatially compartmentalized methylerythritol phosphate (MEP) pathway in plastids and the mevalonate (MVA) pathway in the cytosol, respectively^{8,9}. Although the enzymatic cascade from farnesyl pyrophosphate (FPP) to dihydroartemisinic acid (DHAA) has been partially elucidated - involving amorpha-4,11-diene synthase (ADS)¹⁰, cytochrome P450 monooxygenase CYP71AV1 (CYP71AV1)¹¹, alcohol dehydrogenase 1 (ADH1)¹², double bond reductase 2 (DBR2)¹³, and aldehyde dehydrogenase 1 (ALDH1)¹⁴, recently, a dihydroartemisinic acid dehydrogenase (AaDHAADH) has been identified, which catalyzes the bidirectional conversion between artemisinic acid (AA) and DHAA¹⁵ - the terminal

¹Department Chinese Medicine Authentication, College of Pharmacy, Naval Medical University, Shanghai 200433, China. ²Shanghai Key Laboratory for Pharmaceutical Metabolite Research, Shanghai 200433, China. ³These authors contributed equally: Minghui Zhang, Mingyu Li, Yanyan An, Chang Liu.

✉ e-mail: hexintan@163.com

conversion from DHAA to artemisinin remains mechanistically ambiguous. Particularly contentious is whether the final conversion involves exclusively non-enzymatic photooxidation or requires undiscovered enzymatic catalysis^{16–18}, a key knowledge gap hindering microbial bioengineering approaches^{19–21}.

Beyond artemisinin biosynthesis, the GSTs of *A. annua* function as evolutionarily optimized metabolic center, producing commercially pivotal specialized metabolites including monoterpenes, sesquiterpenes, flavonoids, aliphatic compounds, coumarins, and essential oils^{22,23}. Notably, these secretory micro-organs represent a phylogenetically conserved adaptation, serving as dedicated biosynthetic factories for bioactive compounds essential to human societies. This functional convergence is exemplified by cannabinoids in *Cannabis sativa*., nicotine in *Nicotiana tabacum*., menthol in *Mentha canadensis*., and α -tomatine in tomato (*Solanum lycopersicum*). Despite two decades of research dissecting artemisinin biosynthesis within *A. annua* GST, fundamental understanding of these multicellular ‘chemical factories’ remains fragmented, particularly regarding their developmental programming and metabolic regulatory networks²⁴. While non-glandular trichome development is well characterized in model systems like *Arabidopsis thaliana* (which lacks GST), the mechanism of GST morphogenesis and metabolic specialization remains obscure. Consequently, a systematic understanding of GST development that integrates morphogenesis with spatiotemporal control of specialized metabolism is imperative. Such knowledge is essential both to maximize artemisinin production and to unlock the biotechnological potential of uncharacterized GST-derived metabolites across plant species.

The structural complexity and diversity of GSTs are evident across plants. For instance, tomato possesses seven distinct trichome types, four of which (types I, IV, VI, and VII) are GSTs. Tobacco contains two types of GSTs (long glandular and short glandular trichomes). In *A. annua*, a single GST type exists, yet it comprises ten cells organized into specialized substructures. Specifically, these GSTs consist of two stalk cells (SCs), two basal cells (BCs), and six secretory cells partitioned into two apical secretory cells (ACs) and four subapical secretory cells (SAs) surrounding a subcuticular cavity (Fig. 1a, Supplementary Fig. 1a)^{6,25}. This intricate cellular architecture suggests functional specialization that conventional bulk tissue analyses fail to resolve. Earlier laser microdissection studies produced conflicting conclusions: initial work detected no artemisinin pathway genes (*ADS*, *CYP71AV1*, *DBR2*) in SAs²⁶, whereas subsequently refined protocols identified biosynthetic activity across both ACs and SAs^{27,28}. These discrepancies underscore the limitations of low-resolution cellular sampling and emphasize the necessity for single-cell interrogation.

Single-cell RNA sequencing (scRNA-seq) has emerged as a powerful high-throughput technology enabling the comprehensive analysis of gene expression profiles at the single cell resolution, effectively addressing the inherent cellular heterogeneity in complex tissues and organs^{29,30}. This innovative methodology has been effectively implemented in plant systems research, driving significant progress in model organisms and agronomically important species^{31–35}. Recent advancements in single-nucleus RNA sequencing (snRNA-seq) circumvent the technical artifacts induced by protoplast isolation, preserving native transcriptional states while addressing cellular heterogeneity^{36–39}. When integrated with spatial transcriptomics and metabolomics, this approach enables unprecedented resolution in mapping biosynthetic activities to specific cellular niches⁴⁰. While single-cell omics are now widely applied in plant biology^{41–43}, their implementation in medicinal plants remains challenging due to specific technical hurdles, including secondary metabolite interference, cell wall complexity, or limited reference genomes.

Our study pioneers integrating snRNA-seq with spatial transcriptome to deconvolute the developmental-metabolic continuum in *A. annua* GST, establishing a transformative framework for plant

secretory structure research. We verified four GST-specific marker genes and identified hundreds of promising hub genes potentially regulating artemisinin biosynthesis or trichome development, providing mechanistic insights into the functional diversification and compartmentalization of plant secretory architectures and the metabolomic orchestration underlying specialized secondary metabolite biogenesis.

Results

Developmental dynamics of GST and artemisinin biosynthesis in *A. annua* leaves

Bright-field and epifluorescence microscopy (488 nm excitation) of *A. annua* leaves across four developmental stages (30-day-old plants) revealed stage-specific morphological and metabolic transitions in GSTs (Fig. 1b). Stage 1 (S1) leaves exhibited curled morphology and predominantly immature GSTs with red autofluorescence. Transition to Stage 2 (S2), leaf showed partial expansion accompanied by a 2.6-fold increase in GST density, and the autofluorescence of GSTs mainly changed to yellow. Mature Stage 3 (S3) leaves exhibited fully expanded morphology with GSTs demonstrating predominant yellow autofluorescence, while residual red-fluorescent GSTs represented only a minor fraction of the total GST population. Concurrently, the GSTs density underwent a 45% reduction compared to S2 peak levels. Senescing Stage 4 (S4) leaves exhibited dark-green pigmentation and completely loss of red-fluorescent GSTs, retaining only yellow-fluorescent GSTs. Our preliminary studies, encompassing systematic observations of numerous intact leaves, consistently revealed this developmental shift in GST autofluorescence from red (immature) to yellow (mature). Figure 1b presents a representative image from these observations. Quantitative LC-MS/MS analysis revealed dynamic artemisinin metabolite profiles across these developmental stages (Fig. 1c). Artemisinin content increased 7.2-fold from S1 to S3 (from 0.5 to 3.6 mg/g), directly correlating with GST maturation. Subsequently, the senescence of GSTs in S4 led to a 22% reduction in artemisinin (2.8 mg/g). Notably, S2 exhibited a peak in GST density compared to adjacent developmental stages, defining the S1-to-S2 transition as the pivotal morphogenetic window for GST proliferation. This developmental transition aligns spatiotemporally with leaf expansion dynamics and provides an ideal model to investigate the molecular mechanisms underlying GST initiation, particularly the temporal coordination of specialized metabolite biosynthesis.

Generation of the GST atlas for *A. annua*

Obtaining high-quality single-nucleus transcriptomic data from GSTs encountered particular technical challenges compared to whole-leaf analyses. In *A. annua* leaves, GSTs constitute rare epidermal protrusions (<1% of total leaf cells), with each GSTs comprising a 10-cell structure. Mature GSTs measure ~50 μ m in diameter (Supplementary Fig. 1a) and are encased in a cuticular sheath that resists conventional enzymatic protoplast isolation protocols. Whole leaf snRNA-seq approaches are at risk of GSTs exclusion during quality filtering due to their scarcity and fragility. A mild mechanical extraction method was designed to circumvent these constraints, enabling efficient isolation of intact GSTs from over 400 leaves across S1 and S2 developmental stages (Fig. 2a, Supplementary Fig. 1b–j). Nuclear extraction was optimized for both GST-enriched and whole-leaf samples, followed by droplet-based snRNA-seq. Flow cytometry eliminated cellular debris before chromium controller loading (Supplementary Fig. 2). Sequencing on the Illumina platform yielded 688 million reads. Post-quality control, 8,334 leaf nuclei (936 mean genes per nucleus) and 7,995 GST-enriched nuclei (565 mean genes per nucleus) were retained (Supplementary Data 1).

Integrated analysis revealed 15 transcriptionally distinct clusters via UMAP visualization (Fig. 2b). GST-enriched clusters (7,8,9,11,12,14,15) exhibited spatial segregation from leaf-dominant

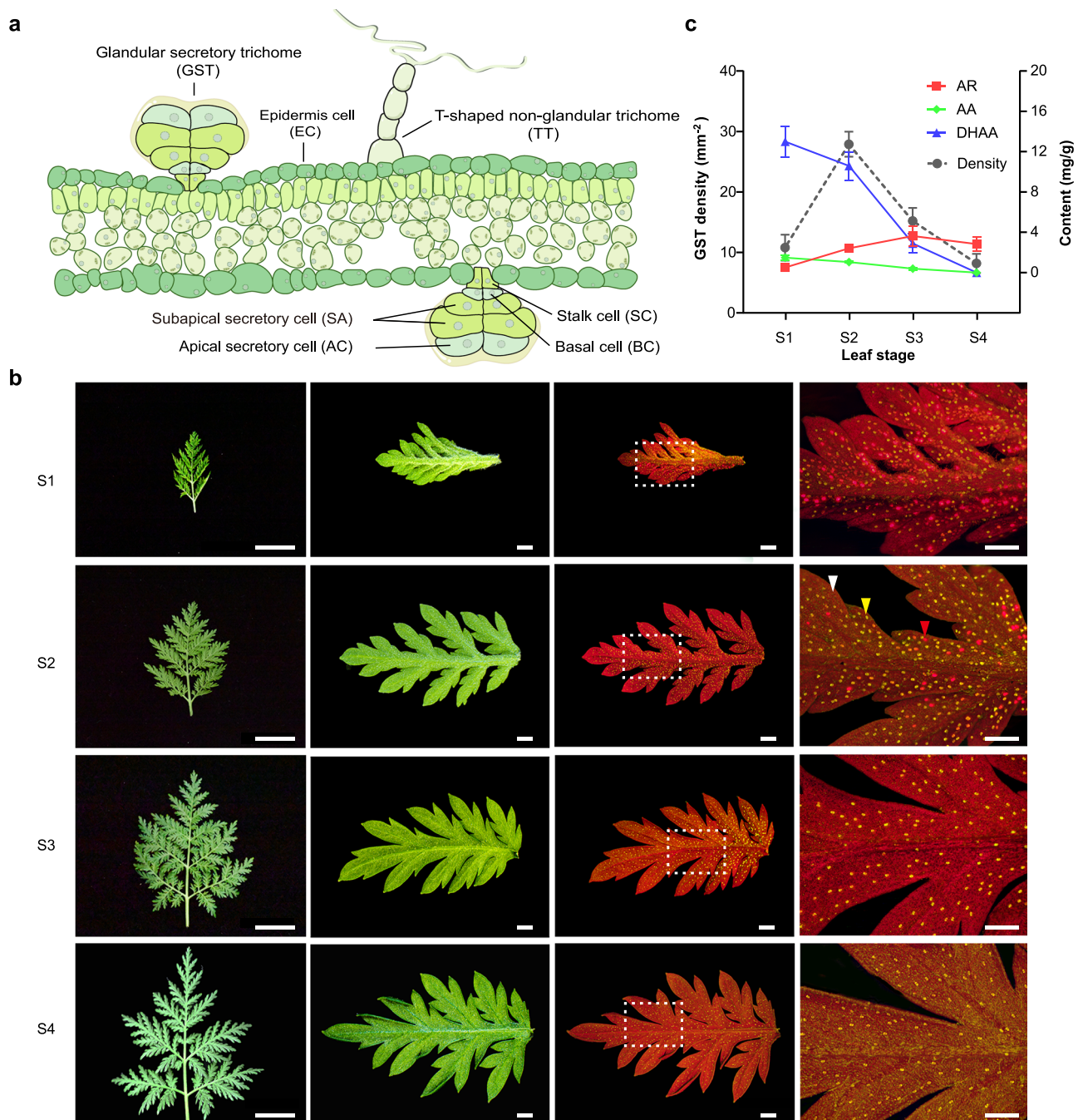


Fig. 1 | Artemisinin dynamics in *A. annua* leaves. a Schematic of leaf cross-section showing GST structure and TT. **b** Leaf development stages (S1-S4) in one-month-old plants. Left: Leaf morphology under bright field (scale bars: 2 cm, 1 mm). Right: Leaf morphology under 488 nm excitation, and the autofluorescence signals indicate chloroplasts (white triangles), mature GSTs (yellow triangles), and immature GSTs

(red triangles) (scale bars: 1 mm, 0.5 mm). **c** GST density (mean $\pm$ SD, $n = 5$ biologically independent samples) and LC-MS/MS quantification of artemisinin (AR) and its precursors artemisinic acid (AA), dihydroartemisinic acid (DHAA) (mean $\pm$ SD, $n = 3$ biologically independent samples).

clusters (1,2,4,6,10), with differential abundance patterns (Supplementary Fig. 3a, b). Differential expression analysis revealed 1,341 genes with significant transcriptional divergence, comprising 685 upregulated and 656 downregulated genes in GST-enriched sample. Gene Ontology (GO) functional annotation and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis of the 685 upregulated DEGs in GST-enriched sample revealed that these genes are primarily involved in the biosynthesis of fatty acids, cutin, lignin, terpenoids, wax, sphingolipids, and flavonoids, as well as biological processes such as fatty acid metabolism, cuticle development, and lipid catabolism (Supplementary Fig. 3c). These findings highlight the

functional relevance of these genes to specialized metabolic pathways and developmental processes in GST.

The 15 clusters were precisely annotated through a multi-evidence integration strategy combining: (1) β -Glucuronidase (GUS) staining validated genes exclusive to GST in *A. annua*, (2) evolutionarily conserved marker genes from orthologous systems, and (3) Pearson correlation-based cluster kinship analysis (Fig. 2c, Supplementary Data 2, Supplementary Fig. 4). Four clusters (7,11,12,14) were unequivocally identified as GST through two hallmark features: Spatial enrichment in trichome-associated zones and elevated expression of GST-specific genes (*AaGSW1*⁴⁴, *AaGSW2*⁴⁵, *ADS*, *DBR2*, *CYP71AV1*,

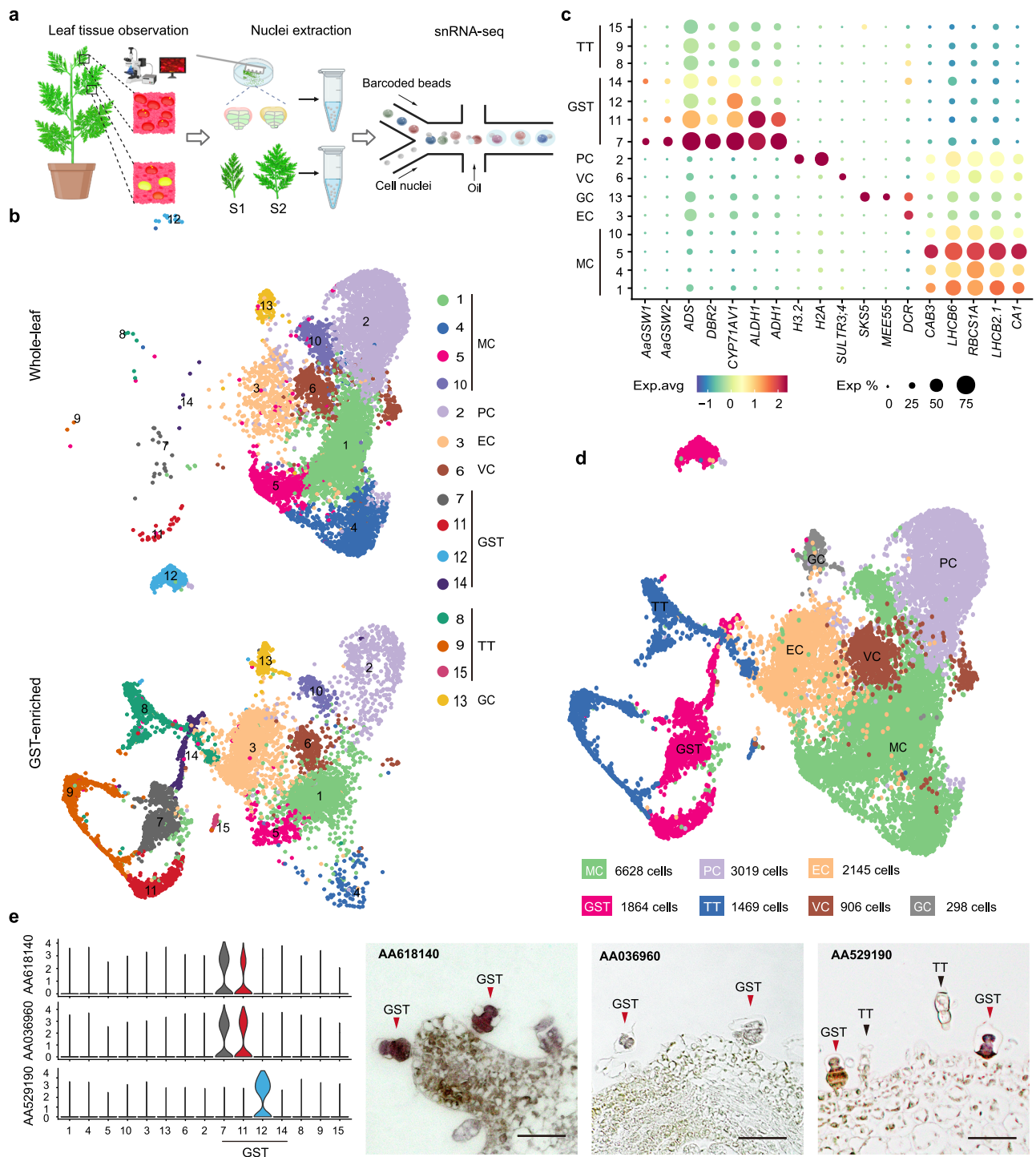


Fig. 2 | Single-nucleus atlas resolves cellular heterogeneity and validates GST clusters in *A. annua* leaves. **a** Workflow for snRNA-seq: nuclei from GST-enriched sample (up) and the whole leaves (down) were isolated, sorted, and processed for library construction. **b** Integrated UMAP of 15 transcriptionally distinct clusters, color-coded by cell identities: GST, TT, EC, mesophyll cell (MC), guard cell (GC), vascular cell (VC), and proliferating cell (PC). **c** Dot plot of cell type-specific marker genes. Dot size represents expression proportion, and color intensity shows mean

normalized expression (log (TPM + 1)). **d** Consensus annotation of 15 clusters into seven broad cell populations: GST, TT, MC, EC, PC, VC, and GC. **e** GST identity validation: Left, Violin plots of GST-specific genes (AA618140, AA036960, AA529190); Right, RNA in situ hybridization experiments were performed with three independent biological replicates, all yielding consistent results. Representative images are shown. Red arrows indicate GST-specific signals; black arrows mark TT regions lacking expression (scale bar: 50 μ m).

ALDH1, *ADH1*). Clusters 8, 9, 15 displayed developmental proximity to GST but lacked characteristic GST markers (*AaGSW1*, *AaGSW2*), defining them as T-shaped non-glandular trichome (TT). Based on the reported marker gene expression patterns in *Arabidopsis thaliana*, clusters 3, 13, 6, and 2 were annotated as epidermal cell (EC), guard cell (GC), vascular cell (VC), and proliferating cell (PC), respectively.

Clusters 1, 4, 5, and 10 were identified as mesophyll cell (MC). Consequently, seven distinct cell populations were ultimately delineated (Fig. 2d).

Spatially resolved RNA in situ hybridization of three GST-specific markers (AA618140, AA036960, and AA529190) confirmed the annotation accuracy (Fig. 2e, Supplementary Fig. 5). Furthermore, multiple

cell types marker genes were identified through the analysis of differentially expressed genes across distinct cell populations and validated by qRT-PCR (Supplementary Fig. 6a, b; Supplementary Data 3). These findings provide valuable molecular resources for future investigations into the developmental processes of *A. annua* leaf and GST, as well as the specialized metabolic pathways associated with these cell types.

Spatial transcriptomics reveals the distribution and molecular signatures of GSTs in *A. annua* leaves

Spatial transcriptomics mapping of *A. annua* leaves resolved GST/TT spatial architectures and validated seven snRNA-seq clusters (Fig. 3a). Given the unique anatomical features of *A. annua* leaves, we developed specific protocols for cryosections and tissue permeabilization, establishing an optimized workflow for spatial transcriptomics in this medicinal species (Supplementary Fig. 7a). From the leaf samples, we obtained 849 million sequencing reads, with 96.1% of barcodes identified as valid, ensuring high data quality. Four leaves exhibiting dense GSTs with intact morphology were selected for high-resolution spatial transcriptomic profiling at 14 μ m resolution, yielding 1,178 spatially resolved transcriptomic spots and 25,162 detected genes (Supplementary Data 4, Supplementary Fig. 7b).

Spatial clustering analysis revealed five distinct tissue domains (Fig. 3b): EC/TT (Cluster 0), MC (Cluster 1), GST (Cluster 2), and VC/PC (Clusters 3 and 4). UMAP visualization revealed spatially segregated GST domains proximal to TT/EC but distinct from other cell types, with distribution patterns concordant with snRNA-seq projections (Supplementary Fig. 7c). Differential gene expression analysis identified cluster-specific markers, with GST-enriched Cluster 2 showing striking high expression of terpenoids biosynthetic pathway components. Four key enzymes (*ADS*, *ALDH1*, *DBR2*, *ADH1*) involved in artemisinin (sesquiterpene) biosynthesis and two camphene synthases⁴⁶ (AA068290, AA251590) responsible for monoterpene biosynthesis ranked among the top 10 GST markers (Fig. 3c). Spatial validation through RNA in situ hybridization revealed specific expression of a marker gene AA050580 of GST, and conclusively confirmed the GST identity of Cluster 2 (Fig. 3d, Supplementary Fig. 5).

Through anchor-based integration of snRNA-seq and spatial transcriptomics, we mapped fine-grained subpopulations to tissue coordinates (Supplementary Fig. 8). Notably, GST spatial zonation patterns were consistent across modalities (Fig. 3e), and artemisinin biosynthetic genes (*FPS*, *ADS*, *CYP71AV1*, *ADH1*, *ALDH1*, *DBR2*, *CPR*) within GSTs demonstrated concordant expression (Fig. 3f), validating biologically relevant niche capture by our cross-modal alignment. Besides, four analytical tools were used to a computational framework to link spatial and single-nucleus transcriptomic datasets. SingleR and SciBet analyses revealed that the GST cluster in snRNA-seq data exhibited the strongest correlation with Cluster 2 in the spatial transcriptomics data (Fig. 3g). However, cell type classification diverged between the two methods: SingleR grouped GST with PC, whereas SciBet associated GST with EC. RCTD and CARD analyses yielded consistent results, with GST spatial localization aligning with both SingleR and SciBet findings (Fig. 3g). These integrative analyses consuming confirmed the accuracy of the GST cluster annotation in the snRNA-seq data, providing a reliable foundation for subsequent investigations.

Pseudotime analysis unravels the differentiation trajectory and functional specialization of GST

The differentiation mechanisms governing both GST and TT, which are epidermal derivatives, remain incompletely understood, especially regarding metabolic specialization during GST maturation⁴⁷. To address this gap, pseudotime analysis was used to reconstruct cellular differentiation hierarchies and identify regulatory gene networks⁴⁸. Applying this approach to *A. annua* EC, GST, and TT revealed a

bifurcating developmental trajectory. Epidermal progenitor cells diverged into GST and TT lineages along the pseudo-temporal axis (Fig. 4a). Gene expression dynamics across this developmental trajectory segregated into four functional modules (Fig. 4b, Supplementary Fig. 9). Module 4 genes exhibited high expression in the mid-pseudotime axis, primarily involved in cell wall organization, cellulose biosynthesis process, and lipid catabolic process. Notably, Module 1, functionally tethered to GST morphogenesis, exhibited hierarchical activation of terpenoid biosynthetic process and fatty acid metabolic networks, with de novo fatty acid biosynthesis emerging as the most enriched pathway. This metabolic signature highlights fatty acid metabolism as a central orchestrator of GST development, likely coordinating precursor provisioning for GST terpenoid synthesis and dynamic membrane biogenesis during secretory cell differentiation.

Our refined reclustering of GST-associated clusters identified seven distinct subclusters (G1-G7) (Fig. 4c). Notably, G6 and G7 (21 and 22 cells, respectively) displayed markedly lower cellular abundance relative to other clusters, suggesting their role as short-lived intermediates in pivotal developmental transitions (Supplementary Fig. 10a). Analysis of Pearson correlation coefficients based on average gene expression profiles across cell clusters (Fig. 4d) revealed that G1, G3, G5, and G7 exhibited strong co-expression patterns, as evidenced by their high correlation coefficients, indicative of shared developmental origins and conserved functional characteristics. G2, G4, and G6 formed another cluster.

Seven GST subclusters (G1-G7) were annotated through the integration of photosynthetic gene expression profiles (Supplementary Fig. 10b) and intercellular correlation networks, building upon Duke et al.'s developmental framework⁶. G5 was annotated as AC due to minimal photosynthetic gene expression (1.9%). G1, identified as SA, was the most abundant (554 cells) and expressed more photosynthetic genes (10.6%). G7, strongly correlated with G1, was also annotated as SA. G3 was identified as BC, lacking detectable photosynthetic gene expression and clustering with secretory cells, potentially serving as support cells for GST structure maintenance and substance transport. G2 was annotated as SC, showing a low correlation with secretory cells but expressing some photosynthetic genes (1.7%). G6 specifically expressed *H2A*, a classic proliferation marker (Supplementary Fig. 10c). Meanwhile, GO enrichment analysis revealed that G4 primarily functions in translation, translational elongation, cutin biosynthesis, fatty acid metabolism, and plant epidermis morphogenesis. The strong Pearson correlation coefficient between G4 and G6 further supported their biological similarity (Fig. 4d). Therefore, G4 and G6 were annotated as immature GST cells (IGCs).

Pseudotime trajectory analysis elucidated the developmental progression of GST subpopulations, revealing at least three distinct functional states during GST development (Fig. 4e). The initiation phase (G4/G6, IGC) was predominantly associated with fundamental cellular processes, including translation (GO: 0006412), nucleosome assembly (GO: 0006334), and cytoplasmic translation (GO: 0002181), establishing the essential cellular machinery for subsequent development. The intermediate phase (G2/G3/G1, SC, BC, and SA) exhibited a metabolic shift toward lipid metabolism, marked by the activation of the fatty acid biosynthetic process (GO: 0006633), wax biosynthetic process (GO: 0010025), and very long-chain fatty acid biosynthetic process (GO: 0042761). Notably, this phase co-expressed active photosynthetic pathways (GO: 0018298, GO: 0009768, GO: 0015979) alongside these biosynthetic processes, demonstrating a concurrent metabolic capacity during this transitional stage. Most significantly, the late developmental stage (G5/G7 clusters, SA and AC) exhibited upregulation of specialized metabolic pathways, particularly terpenoid biosynthesis (GO: 0016114) and fatty acid metabolism (GO: 0006633). This molecular profile aligns with the evolutionary role of GSTs as specialized biosynthetic compartments for secondary metabolite production (Fig. 4f, Supplementary Fig. 11).

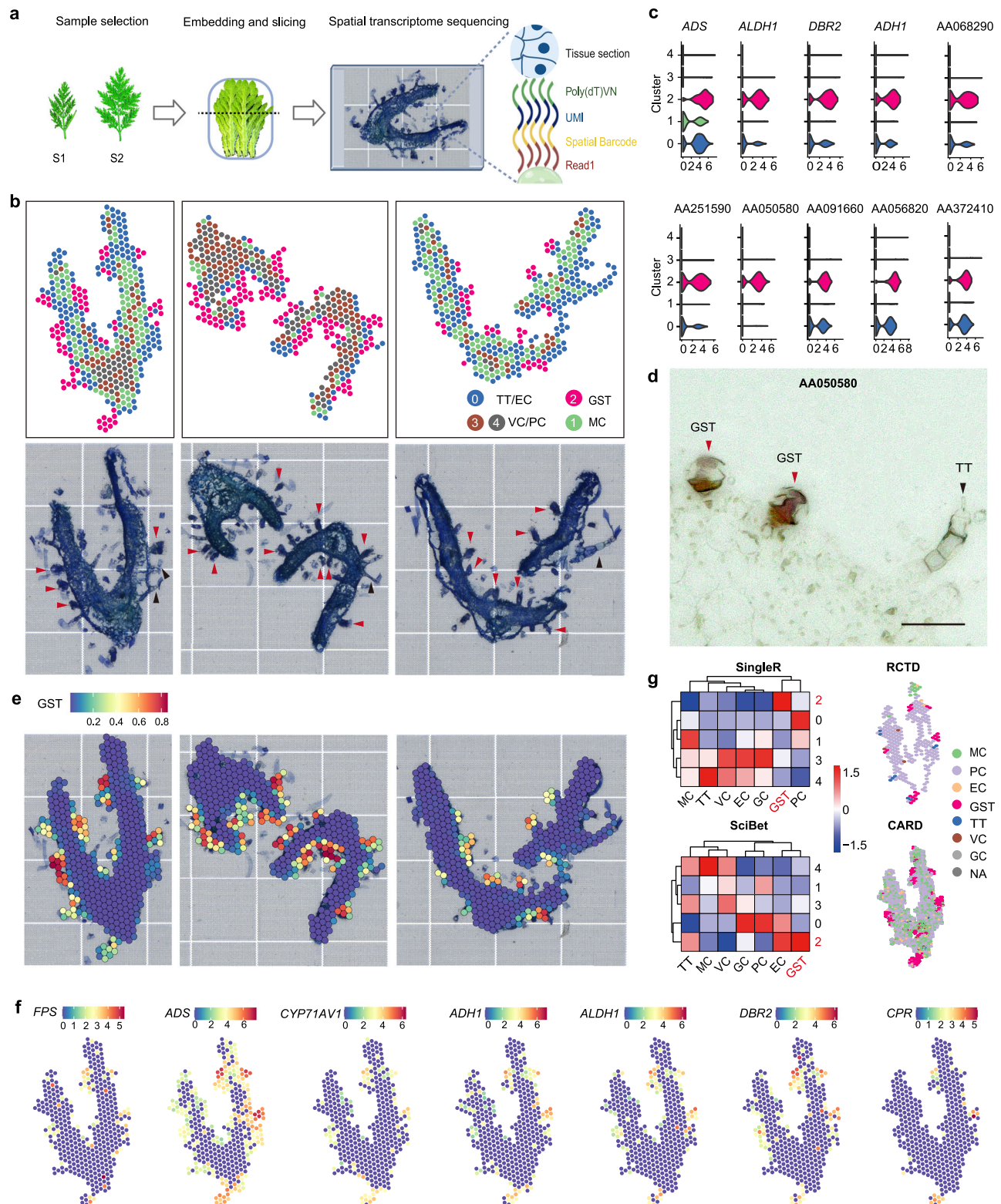


Fig. 3 | Spatial mapping of cell types and artemisinin biosynthesis in *A. annua*.

a Workflow: Fresh leaf tissues were embedded, vertically sectioned, stained with TBO, and processed for spatial transcriptome sequencing. **b** Spatial distribution of cell clusters in the leaves of *A. annua*. Representative images from four independent leaves are shown. Top, overlay of spatial clusters (14 μ m resolution). Bottom, toluidine blue-stained leaf section. Red arrows indicate GSTs, black arrows mark TTs. **c** Spatial expression of GST top ten marker genes (violin plots: log (CPM + 1)). **d** RNA in situ hybridization experiments of the AA050580 gene were performed

with three independent biological replicates, all yielding consistent results. Representative images are shown. Red arrows indicate GST-specific signals; black arrows mark TT regions lacking expression (scale bar: 50 μ m). **e** Integrating snRNA-seq with spatial transcriptomics via Seurat. **f** Co-localization of artemisinin biosynthetic genes in GST domains (color gradient: blue $\rightarrow$ red, low $\rightarrow$ high). **g** Integrative validation of snRNA-seq and spatial transcriptomic concordance using SciBet, SingleR, RCTD, and CARD algorithms.

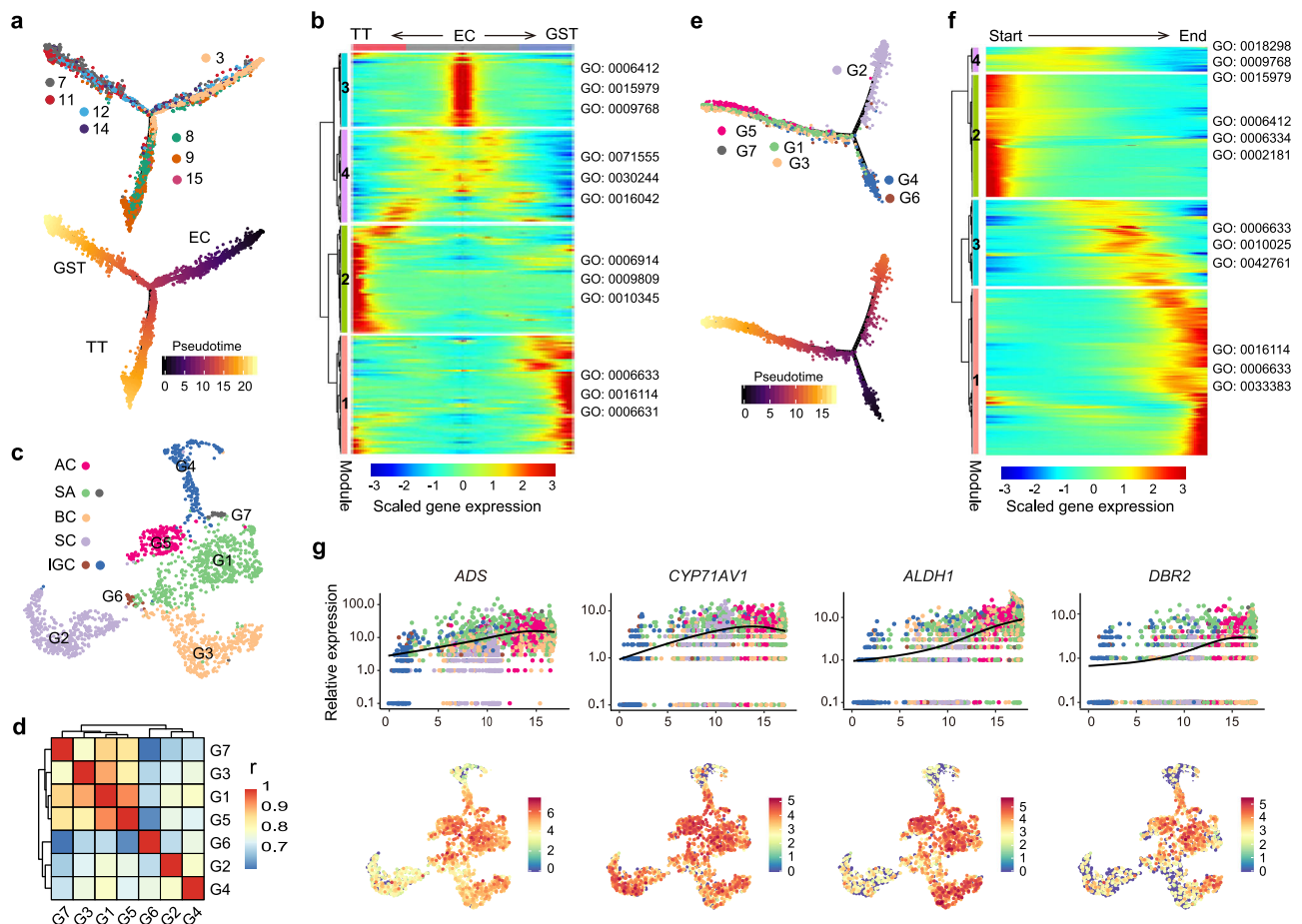


Fig. 4 | Developmental trajectory reconstruction in spatiotemporal coupling of artemisinin biosynthesis and GST differentiation. **a** Pseudotemporal trajectory of EC differentiation into GST and TT. Top, UMAP colored by cell identity. Bottom, Pseudotime axis (dark → light: early → late). **b** Functional annotation of gene modules along differentiation trajectories. Heatmaps show expression dynamics of EC → TT and EC → GST branches (color gradients: blue → red, gene expression levels Z-score-normalized). Representative GO terms are labeled. **c** UMAP of reclustered GST subpopulations ($k = 7$): G5 (apical cells, AC), G1/G7 (subapical cells, SA), G3 (basal cells, BC), G2 (stalk cells, SC), G4/G6 (immature GST cells, IGC). **d** Transcriptional correlation heatmap of GST subclusters. The color gradient

represents correlation coefficients (r), with increasing red intensity indicating higher degrees of inter-cluster similarity. **e** Pseudotime reconstruction of GST development. Root anchored at IGC (G4/G6), with a bifurcation point G1/G5/G7 and G3 from SC (G2). **f** Heatmap of pseudotime-dependent gene expression across GST subclusters (color gradients: blue → red, gene expression levels Z-score-normalized). Representative GO terms are labeled. **g** Spatiotemporal validation of representative genes. Top, Pseudotime-dependent upregulation of *ADS* (AA415620), *CYP71AV1*, *ALDH1* and *DBR2*. Bottom, UMAP showing spatial enrichment in secretory subclusters (G1/G5/G7; color intensity (blue → red) indicates the relative transcript level for the indicated gene in each cell).

Reconstruction of the developmental trajectory via pseudotime analysis enabled determination of the dynamic expression of key genes in artemisinin biosynthesis during GST development (Fig. 4g). The expression of four downstream key genes in artemisinin synthesis gradually increased with trichome development. By resolving the pseudotemporal architecture of GST maturation, our work provides unprecedented insights into the dynamic functional transitions during GST development, offering a refined framework for understanding the coordination mechanisms underlying their differentiation and maturation.

Artemisinin biosynthesis exhibits strict spatial confinement to six secretory cells within GST of *A. annua*

While GSTs are recognized as the primary biosynthetic sites for artemisinin in *A. annua*, the precise cellular specialization within their 10-cell architecture has remained unresolved. Single-nucleus transcriptomic profiling of all known key genes in the artemisinin biosynthetic pathway was performed to determine cell-type-specific biosynthesis (Supplementary Data 5). GST clusters (C7, C11, C12, C14) displayed coordinated high expression of upstream terpenoid precursor pathways (MVA and MEP) and downstream biosynthetic genes (*FPS*⁹, *ADS*, *CYP71AV1*, *CPR*⁹,

ALDH1, *DBR2*, *ADH1*). In contrast, TT clusters exhibited only basal expression of *DXS*⁹, *HDS*⁹, *FPS*, and *CPR*, while epidermal and mesophyll cells showed negligible pathway activity (Fig. 5a). These results conclusively demonstrate evolutionary compartmentalization of artemisinin biosynthesis within GST cells of *A. annua*.

Integrated analysis of the artemisinin biosynthetic pathway, which spans terpenoid precursor pathways (MVA/MEP) and downstream genes, across GST subclusters revealed cell-type-specific metabolic partitioning. ACs and SAs exhibited coordinated upregulation of both precursor pathways and all known downstream enzymatic components (*ADS*, *CYP71AV1*, *ALDH1*, *DBR2*) (Fig. 5a), indicating their autonomous capacity for the production of enzymatic precursors in artemisinin biosynthesis. Conversely, basal cells (BCs) showed attenuated expression of late-stage genes (*FPS*, *ADS*, *CYP71AV1*, *DBR2*) but maintained active transcription of early MVA pathway genes (*HMGS*, *HMGR*)⁹, suggesting specialized roles in precursor provisioning for neighboring secretory cells. Strikingly, SCs and IGCs displayed minimal transcriptional activity across the entire pathway, confirming their metabolic exclusion from artemisinin production. High-resolution single-nucleus RNA-seq spatial mapping definitively restricted artemisinin production to six specialized secretory cells within *A. annua*

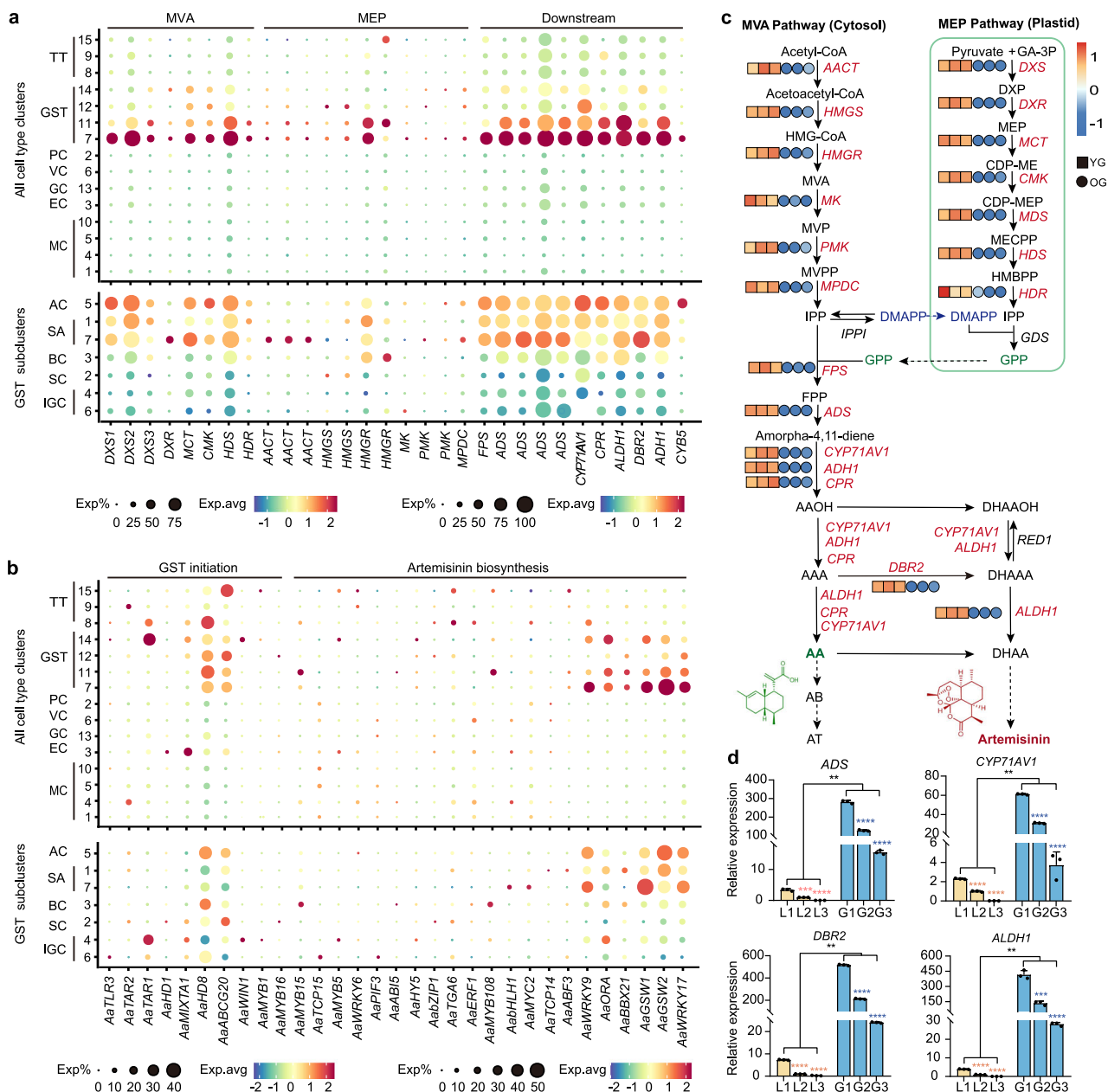


Fig. 5 | Cell type-specific expression of artemisinin biosynthetic genes across GST subclusters. a Single-nucleus resolution of artemisinin pathway gene expression. Dot size: proportion of expressing cells; color gradient: mean normalized expression (log (TPM + 1)). Corresponding legend panels (left for All cell type clusters, right for GST subclusters) specify visualization parameters. **b** Single-nucleus resolution of transcription factors expression. **c** Stage-specific transcriptional regulation of the artemisinin pathway. Bulk RNA-seq (log₂(FPKM + 1), Z-score

normalized) shows coordinated upregulation of MVA/MEP and downstream genes in young GSTs (YG, S1/S2; n = 3) versus old GSTs (OG, S3/S4; n = 3). **d** qRT-PCR validation of core biosynthetic genes in whole leaf (L) versus GST-enriched (G) samples across stages (S1-S3). Data: mean ± SD (n = 3 biologically independent samples; *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001, two-tailed t-test). Actin (EU531837) is used for normalization.

GSTs, establishing these compartments as the principally biosynthetic site for this antimalarial compound.

In addition, we analyzed the expression patterns of 48 previously reported transcription factors (TFs) associated with GST initiation or artemisinin biosynthesis at single-nucleus resolution (Supplementary Data 6). Figure 5b illustrates representative TF expression profiles across all *A. annua* leaf cell types and within GST subclusters. Key regulators of GST initiation, such as *AaTLR3*⁴⁹, *AaTARI*⁵⁰, *AaMIXTA*⁵¹, and *AaWINI*⁵², showed high expression in IGCs and/or SCs. Conversely, TFs modulating artemisinin biosynthesis, including *AaHLH1*⁵³, *AaMYC2*⁵⁴, *AaWRKY9*⁵⁵, *AaBBX21*⁵⁶, *AaGSW1*, *AaWRKY17*⁵⁷, exhibited high expression in ACs and SAs. Notably, BCs co-expressed TFs

implicated in both GST development (*AaHDS*⁵⁸) and artemisinin synthesis (*AaABIS*⁵⁹, *AaERF1*⁶⁰, *AaMYB108*⁶¹).

Our snRNA-seq data corroborated with published GUS localization studies. TFs validated by GUS assays for GST-specific expression, such as *AaGSW1* and *AaGSW2* also showed cluster-specific enrichment in GST within our scRNA-seq data. TFs documented to express in TT (*AaMYB1*⁶², *AaTARI*, *AaABF3*⁶³) showed high expression in corresponding snRNA-seq clusters. Similarly, *AaTCP15*⁶⁴ expression, previously observed in mesophyll cells via GUS staining, was confirmed in the MC cluster. While *AaTCP14*⁶⁵ was reported to exhibit non-specific GUS staining but high trichome expression, our analysis precisely localized its high expression to SA and PC (Fig. 5b).

Spatiotemporal patterns were resolved at single-nucleus resolution. The artemisinin regulator *AaHYS*⁶⁶ and *AabZIP1*⁶⁶ both showed high expression in IGCs. The salicylic acid-associated bZIP TF *TGA6*⁶⁷ was highly expressed in both IGCs and SAs. Intriguingly, while *AaGSW2* (a reported GST initiation factor) showed high expression within the overall GST cluster, subcluster analysis revealed its specific enrichment in ACs and SAs rather than stalk or IGCs. This suggests *AaGSW2* likely functions as a multifunctional TF, potentially regulating secretory cells metabolism (such as artemisinin biosynthesis) beyond its role in GST initiation. A notable observation concerns *AaMYB15*⁶⁸, a negative regulator of artemisinin biosynthesis previously localized to both GSTs and TTs by GUS staining. Our data confirmed its high expression in both trichome types, with particularly strong enrichment in BCs. This expression pattern further supports the hypothesis that BCs may supply precursors for artemisinin biosynthesis in secretory cells.

Bulk RNA-seq profiling of *A. annua* GSTs across developmental stages revealed pronounced enrichment of artemisinin biosynthetic genes during the young GSTs compared to old GSTs (Fig. 5c). This transcriptional signature was substantiated by developmental stage resolved metabolomic profiling (Fig. 1c), revealing precisely coordinated artemisinin accumulation dynamics that indicate biosynthesis is most active during the young secretory phase of GSTs, with a progressive decline in synthesis rate as the GST age. To establish the robustness of our single-nucleus and bulk transcriptomic datasets, we conducted independent validation using qRT-PCR to quantify expression dynamics of core artemisinin biosynthetic genes across *A. annua* leaf tissues and GST-enriched samples (Fig. 5d). Quantitative analysis demonstrated coordinated expression of four core artemisinin biosynthetic genes (*ADS*, *CYP71AV1*, *ALDH1*, *DBR2*), with pronounced enrichment in GST relative to leaf tissues and progressive attenuation from young to old GSTs. It confirms the technical reproducibility of our sequencing datasets while offering robust evidence for cell-type-specific metabolic specialization in artemisinin biosynthesis.

hdWGCNA reveals hub genes associated with GST development and artemisinin biosynthesis in *A. annua*

Integration of snRNA-seq with high-definition weighted gene co-expression network analysis (hdWGCNA) uncovered co-expressed gene modules linked to GST development and artemisinin biosynthesis, revealing functional coordination between these processes through distinct co-expression networks. Scale-free network analysis determined an optimal soft threshold of 2 (network fit index = 0.9; Supplementary Fig. 12a), yielding four functional modules (Fig. 6a, Supplementary Data 7). Core genes within each module were identified via gene connectivity (kME), with hub module eigengenes (hMEs) visualized (Fig. 6b). The orange module showed predominant expression in EC, while the brown module was enriched in GSTs, indicating spatial specialization.

Module-trait relationships confirmed a strong association of the brown module with GST clusters (Fig. 5c). Functional enrichment revealed the module's core enrichment in specialized metabolic pathways critical to sesquiterpene biosynthesis: terpenoid biosynthesis, fatty acid metabolism, isopentenyl diphosphate metabolism, and flavonoid biosynthetic process (Supplementary Fig. 12b), suggesting its close association with the artemisinin biosynthetic pathway. Differently, the orange module associated broadly with EC, TT, and GST clusters (Fig. 5c) and was enriched for cuticle biosynthesis, very long-chain fatty acid (VLCFA) metabolism, lipid catabolism, and plant epidermal morphogenesis (Supplementary Fig. 12b). This spatial partitioning, where cuticle-related processes preceding GST specialization, suggests metabolic preparation may facilitate subsequent developmental transitions.

Using our developed scoring system that integrates intramodular connectivity (kME) and association with known biological pathways,

we established interaction networks for the top 30% hub genes in the brown and orange modules to identify key genes involved in artemisinin biosynthesis and GST development (Fig. 6d). Strikingly, the brown module network integrated both upstream (MEP) pathway genes (*DXS*, *HDS*) and six downstream enzymatic genes (*ADS*, *CYP71AV1*, *ADH1*, *ALDH1*, *DBR2*, *CPR*) that collectively span the complete artemisinin biosynthetic cascade. Furthermore, the network encompassed an ABC transporter gene (*PDR1*, AA090620) functionally validated in *Nicotiana tabacum* for artemisinic acid transport implicated⁶⁹. This comprehensive coverage of key genes in the artemisinin biosynthetic pathway significantly enhances the reliability and biological relevance of the constructed network.

Six genes from brown and orange modules were prioritized as key candidates influencing artemisinin biosynthesis or GST initiation/development, particularly given the unique structural and functional specialization of plastids within GST cells. In the brown module, AA013090 (annotated as a photosystem antenna protein-like based on sequence homology) is co-expressed with core artemisinin biosynthetic genes. Although its exact role in GST plastids remains unverified, this gene may contribute to light-dependent processes that indirectly support artemisinin production. This role is supported by tobacco mutant studies, where loss of light-harvesting chlorophyll a/b-binding proteins (LHCB) fully abolished glandular secretory capacity⁷⁰. Although AA013090 is not a direct ortholog of tobacco *LHCB*, the convergent GST dysfunction in both systems underscores the critical dependence of GST activity on photosystem components. Moreover, AA244300 (annotated as a dimethylaniline monooxygenase) may be involved in oxidative modifications of artemisinin intermediates. Two aldehyde dehydrogenase family genes (AA036960 and AA405270) may influence artemisinin synthesis by modulating oxidoreductase activity (acting on the aldehyde or oxo group of donors, with NAD or NADP as acceptor). The orange module candidates comprised AA393400 (annotated as Lipase, GDSL) and AA372410 (non-specific lipid-transfer protein 3). Two genes potentially functioning in processes critical for cuticle and wax metabolism^{71,72}, thereby influencing GST initiation and physical integrity. Spatial expression profiling (Fig. 6e) confirmed functional relevance; brown module candidates showed GST enrichment, while orange module candidates exhibited broad expression across EC/TT/GST cell types. qRT-PCR validation further demonstrated that AA036960, AA013090, and AA244300 exhibited expression patterns consistent with four core artemisinin genes (Fig. 6f). Notably, AA036960 was validated as a marker gene for GST via RNA in situ hybridization (Fig. 2e, Supplementary Fig. 5). More comprehensive functional validation studies using CRISPR-Cas9 knockout are currently underway.

Discussion

Advances in single-cell transcriptomics have revolutionized molecular genetics, offering unprecedented opportunities to understand the cellular compartmentalization of specialized metabolism^{73–75}. Here, we established a single-nucleus atlas of GST in *A. annua*, resolving spatial organization of the artemisinin biosynthetic pathway. Integration of snRNA-seq with spatial transcriptomics yielded a high-resolution spatiotemporal dataset of *A. annua* GST at the single-nucleus level. This dataset provides a foundation for uncovering unknown pathways in artemisinin biosynthesis and elucidating the developmental mechanisms of GSTs, while also serving as a model for understanding the morphogenesis and regulatory mechanisms of GSTs in other species. It also offers insights into cell-specific synthesis of natural products.

scRNA-seq of *A. annua* GSTs presented a major challenge. GSTs constitute less than 1% of leaf cells. Mature GSTs are small, encapsulated in a cuticular shell, and show reduced transcriptional activity typical of metabolite-storing cells^{47,76}. Conventional enzymatic methods failed to isolate intact GST protoplasts due to structural resistance and low RNA access, leading us to develop an optimized mechanical-

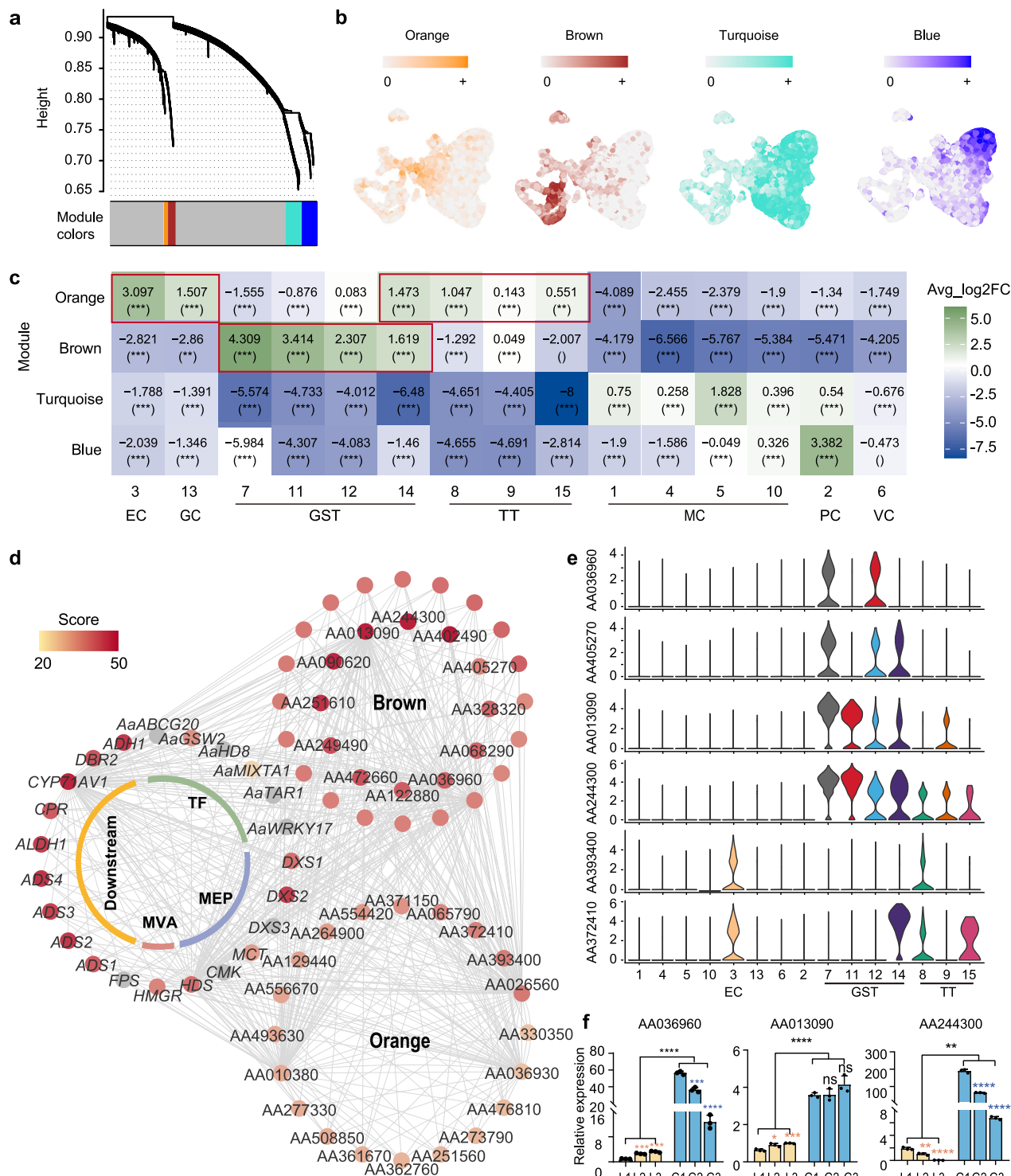


Fig. 6 | hdWGCNA identified hub genes of trichome development and artemisinin biosynthesis. a Hierarchical clustering dendrogram of co-expression modules (hdWGCNA; soft threshold power = 2, $R^2 = 0.9$). 3,727 genes partitioned into four modules (colored branches); grey: unassigned genes. **b** Spatiotemporal distribution of module eigengenes. UMAP colored by module eigengene scores (kME values). **c** Cell type-specific expression dynamics of co-expression modules. Heatmap: average log₂(fold change) in GST vs. other cell types. Significance: *FDR < 0.05, **FDR < 0.01, ***FDR < 0.001.

0.01, ***FDR < 0.001. **d** Co-expression networks between selected unreported hub genes and known artemisinin biosynthetic genes or transcription factors. **e** Violin Plots of expression patterns for six hub genes in artemisinin biosynthesis and trichome development. **f** qRT-PCR validation of three artemisinin biosynthesis genes in whole leaf (L) versus GST-enriched (G) samples across stages (S1-S3). Data: mean $\pm$ SD (n = 3 biologically independent samples; * P < 0.05, ** P < 0.01, *** P < 0.001, **** P < 0.0001, two-tailed t-test). Actin (EU531837) used for normalization.

enzymatic workflow. Focusing on the S1-S2 developmental stage where GST formation and peak biosynthesis coincide, we processed over 400 leaves. This yielded a high-resolution single-nucleus atlas with 1,864 GST nuclei and 1,469 TT nuclei, resolving seven distinct cell

populations representing known subtypes. While spatial transcriptomics captured only about 200 trichome cells from limited sections, snRNA-seq profiled 1,864 cells across stages. This discrepancy arises from inherent biological constraints. GSTs are

stochastically distributed, smaller than standard section thickness, and hindered by dense metabolites and cuticles, limiting capture efficiency despite optimization efforts. RNA in situ hybridization confirmed dataset accuracy and identified four GST marker genes. This integrated approach overcomes technical barriers, establishing *A. annua* as a cellular resolution model for specialized metabolism.

Our atlas reveals unprecedented functional specialization within GSTs. Artemisinin biosynthesis is evolutionarily compartmentalized to secretory cells (SAs and ACs), specifically to just six specialized secretory cells possessing the all currently known enzymatic machinery for de novo production. BCs express early MVA genes but show reduced late-stage gene expression, suggesting a potential role in supplying precursors to adjacent secretory cells. This indicates an evolutionary strategy optimizing resource allocation via pathway compartmentalization. Notably, SCs and IGCs showed minimal artemisinin pathway gene expression, confirming metabolic activity is strictly confined to mature secretory cells. Using single-nucleus resolution, we clarified the distinct spatial specificity of transcription factor expression and localization. For example, *AaGSW2*, previously thought broadly expressed, is specifically enriched in apical/subapical cells, indicating potential dual roles in regulating secretory metabolism. High expression of the repressor *AaMYB15* in basal cells further supports their involvement in precursor supply. These findings resolve the cellular basis of artemisinin synthesis and provide precise targets for metabolic engineering.

Pseudotemporal trajectory analysis defined three GST developmental stages. The initiation phase establishes basic cellular machinery by activating fundamental metabolic pathways. The intermediate stage activates lipid metabolism, coordinating fatty acid and wax biosynthesis alongside photosynthesis. The terminal differentiation phase achieves metabolic specialization by spatially partitioning terpenoid and lipid biosynthetic pathways. This cascade aligns with the GST role as a specialized metabolite factory⁷⁷.

Applying hdWGCNA identified two key modules linked to GST development and artemisinin biosynthesis. The brown module, co-expressed with GST clusters, contained all known core artemisinin pathway enzymes. Significantly, this module encompasses three of the five candidate hub genes previously identified through WGCNA analysis of bulk RNA-seq data from *A. annua* trichomes⁷⁸, reinforcing its central role in participating in the pathway and suggesting strong concordance between bulk and single-nucleus approaches. This convergence strongly suggests the module orchestrates artemisinin biosynthesis.

The orange module is associated with EC and trichomes. Its enrichment in cuticle biosynthesis, VLCFA metabolism, lipid catabolism, and epidermal morphogenesis aligns with the reconstructed GST trajectory and known GST regulation. Cuticle layer composition and activation of its biosynthesis by *AaMIXTA1*⁵¹, which enhances GST density, confirm this module's role. *TARI*⁵⁰ regulates artemisinin genes while maintaining cuticular wax, and *TLR3*⁴⁹ constrains metabolic overflow by negatively regulating cuticle organization and branching. Together, those validated the orange module's functions in epidermal structural integrity and lipid control. We speculate that flux through these cuticle/lipid biosynthetic pathways may generate signals (such as lipid-derived molecules, mechanical stress from cuticle deposition) or establish a permissive epidermal environment, thereby facilitating or influencing the downstream transcriptional regulatory cascade controlling GST fate specification. To functionally investigate these modules and test this hypothesis, six high-scoring genes within them are undergoing functional characterization via CRISPR-Cas9 knockout, and their detailed functions will be elucidated in subsequent publications.

This study used an integrated multi-omics approach to reconstruct the spatiotemporal development of *A. annua* GSTs at single-cell resolution and identify hub genes regulating artemisinin biosynthesis and GST development. To overcome limitations of prior trichome studies of other species, which used whole-leaf samples with few

GST cells^{34,79}. We created a GST atlas using enriched GST samples and leaf controls. This strategy addresses the current lack of similarly enriched GST atlases for other key species. Such atlases are essential for cross-species integration to evaluate the universality and evolutionary conservation of GST mechanisms identified here. Future work should apply this enrichment strategy to create high-resolution atlases for diverse GST metabolite-producing species. Integrating these datasets will define conserved developmental programs and species-specific adaptations. Functional characterization of the hub genes using gene editing is crucial to elucidate their roles in GST morphogenesis and artemisinin biosynthesis and test their conservation. This comparative and functional genomics work forms our core follow-up, with full validation to be detailed later.

Methods

Plant materials and growth conditions

Seeds of *A. annua* (cultivar 'Huhao 1') were sourced from the School of Agriculture and Biology, Shanghai Jiao Tong University. Plants were grown in a 2:1 (v/v) nutrient soil-perlite mixture under controlled greenhouse conditions (25 °C, 16/8-hour light/dark photoperiod) for one month.

Nuclei isolation and 10× single-nucleus RNA-seq library construction

GSTs were enriched from 100 one-month-old *A. annua* plants by gently brushing leaf surfaces in 0.4% D-mannitol, filtering through a 70 µm strainer, and centrifuging (1000 rpm, 5 min). GST-enriched and whole leaf samples were lysed in pre-chilled buffer (0.44 M sucrose, 1.25% Ficoll, 2.5% Dextran T40, 20 mM HEPES, 10 mM MgCl₂, 0.5% Triton X-100, 1 mM DTT, 1× protease inhibitor, 1000 U/mL RNase inhibitor), homogenized 20 min, and filtered through 40 µm and 20 µm sieves. Nuclei were collected by centrifugation (500 g, 5 min, 4 °C) and resuspended in 500 µl buffer (1× DPBS, 0.04% BSA, 1000 U/mL RNase inhibitor).

For sorting, nuclei were stained with propidium iodide (PI) and isolated by flow cytometry based on PI signal and size. After pelleting (1000 g, 5 min, 4 °C), nuclei were resuspended in PBS, assessed by fluorescence microscopy, and 30,000 nuclei were loaded onto a 10X Genomics chip. Libraries were constructed using the 10X Chromium Single Cell 3' Solution v3.1 kit for Illumina sequencing.

snRNA-seq data processing

Raw sequencing data were processed with Cell Ranger (v8.0.1) for quality control and alignment to the reference genome (PRJNA416223). Further quality filtering was performed using Seurat⁸⁰ (v4.0.0), excluding cells with gene or UMI counts beyond mean ± 2 standard deviations. Doublets were removed using DoubletFinder⁸¹ (v2.0.3). Data normalization and identification of the top 2000 highly variable genes (HVGs) were conducted using Seurat's NormalizeData and FindVariableGenes functions, respectively. Dimensionality reduction was achieved via principal component analysis (PCA) and uniform manifold approximation and projection (UMAP), followed by clustering using the shared nearest neighbor (SNN) algorithm (resolution = 0.4). Marker genes were identified with FindAllMarkers (logfc.threshold > 0, min.pct > 0.25), and differentially expressed genes (DEGs) were detected using FindMarkers (p-value < 0.05, fold change > 1.5). Functional enrichment analysis was performed using GO⁸² and KEGG⁸³ databases.

Spatial transcriptomic experiments and library construction

Spatial transcriptomics was conducted using the BMKMANU S3000 platform. Leaves from the same developmental stage as snRNA-seq samples were co-embedded in OCT compound within single embedding molds (multiple leaves per block). During embedding, all leaves were uniformly oriented with abaxial epidermis facing upward and

adaxial epidermis facing downward, with positional markings applied to the molds. Samples were flash-frozen in liquid nitrogen and stored at -80°C . Sectioning was performed by obtaining transverse sections (10 μm thickness) through the mid-region of each leaf using 90° vertical cuts. Sections were mounted on BMKMANU S3000 Tissue Optimization and Gene Expression Slides. Tissue optimization determined the optimal permeabilization time, followed by Toluidine Blue O (TBO) staining, imaging, permeabilization, and cDNA synthesis. Libraries were prepared following the manufacturer's protocol and sequenced on an Illumina platform.

stRNA-seq data processing

Clean data were processed with BSTMatrix software for quality control. Cell boundary segmentation and fluorescent staining were performed using Cellpose⁸⁴ (v2.0). Spots overlapping tissue sections were selected for analysis. Data were aligned to the reference genome (PRJNA416223), and a spot-gene expression matrix was generated. Subsequent analysis in Seurat⁸⁰ included normalization via the SCTransform (SCT) algorithm, identification of the top 3,000 highly variable genes (HVGs), principal component analysis (PCA), graph-based clustering, and UMAP visualization.

Integration analysis of stRNA-seq and snRNA-seq

We integrated snRNA-seq and spatial transcriptomics data using anchor-based integration (Seurat) to transfer cell-type annotations from single-cell clusters to spatially resolved spots. This enabled deconvolution of spot-level compositions and revealed subpopulation spatial distributions at near-cellular resolution. To align clusters from single-nucleus and spatial transcriptomic data, we employed multiple computational tools: SingleR⁸⁵, SciBet⁸⁶, RCTD⁸⁷, and CARD⁸⁸. SingleR assigns cell type labels by calculating gene expression similarity between spatial transcriptomic spots and single-nucleus transcriptomic cell types, selecting the highest similarity score. SciBet uses supervised learning and Kullback-Leibler (KL) divergence to predict cell-type probabilities based on reference distributions. RCTD, implemented via the spacexr R package (doublet_mode = 'doublet'), estimates cell type proportions in spatial data and visualizes them on tissue sections. CARD applies a non-negative matrix factorization model to integrate snRNA-seq reference data with spatial transcriptomic data, inferring cell type composition at each spatial location. The code for SingleR (<https://github.com/dviraran/SingleR>) and SciBet (<https://github.com/PaulingLiu/scibet>) is publicly available.

Pseudotime analysis

Pseudotime analysis was performed using Monocle⁸⁹. Highly variable genes across cells were selected for dimensionality reduction. A minimum spanning tree (MST) was constructed, with the longest path representing the differentiation trajectory of transcriptionally similar cells. The trajectory root was determined based on cell cluster or sub-cluster identity. Genes significantly varying along pseudotime ($q\text{-value} < 0.1$) were identified using the different Gene Test function and clustered via plot_pseudotime_heatmap. For GST clusters, dynamically expressed genes were visualized using plot_genes_in_pseudotime.

High-dimensional weighted gene co-expression network analysis (hdWGCNA)

hdWGCNA was applied to high-dimensional single-cell RNA-seq data using the hdWGCNA R package⁹⁰ (v0.2.19). Genes expressed in $\geq 5\%$ of cells were selected via the SetupForWGCNA function. The k-nearest neighbors (KNN) algorithm aggregated similar cells to generate a meta-cell gene expression matrix⁹¹. Parameter scanning and soft threshold determination were performed using the TestSoftPowers and ConstructNetwork functions, respectively. Module eigengenes (MEs) summarized co-expression module profiles, and gene connectivity (kME) was quantified using the ModuleConnectivity function.

Genes with high kME values were identified as hub genes. Downstream visualizations and analyses were conducted to explore module functions. To identify the top candidate genes within these functionally relevant modules, we employed a dual strategy focusing on both intramodular connectivity and association with known key players: We identified the Top 30% genes within each module (Brown and Orange) based on their kME values. For each of these Top 30% genes, we calculated their TOM similarity to a curated set of known key genes involved in artemisinin biosynthesis and GST initiation. We developed a scoring system to integrate these two important aspects:

$$\text{Score} = (50\% \times \text{kME} + 50\% \times \text{TOM_to_KnownGenes}) \times 100$$

kME = Module membership value for the candidate gene (range 0 to 1). TOM_to_KnownGenes = Mean Topological Overlap Measure between the candidate gene and the set of known artemisinin/trichome genes (range 0 to 1).

RNA in situ hybridization

Fresh *A. annua* leaves (stages 1 and 2) were fixed in FAA solution at 4°C for 24 h, dehydrated through an ethanol gradient, embedded in paraffin, and sectioned into 10 μm slices using a Leica microtome. Sections were baked at 62°C for 2 h, dewaxed in xylene and ethanol, rehydrated in DEPC water, and digested with proteinase K (Beyotime, Cat# ST537). After washing with PBS, sections were hybridized overnight with DIG-labeled RNA probes, incubated with mouse anti-DIG-alkaline phosphatase (Jackson, Cat# 200-052-156), and developed using BCIP/NBT substrate (Boster, Cat# AR1023). Images were acquired with a Nikon microscope. Probes were 5'-DIG-labeled oligonucleotides, with details provided in the supplementary Data 8.

GST morphological analysis

GST morphology and developmental dynamics were analyzed in *A. annua* leaves using a stereomicroscope (Motic SM7). High-resolution imaging was performed under standardized conditions using the ImageView software to ensure consistency across biological replicates. Leaves were systematically sampled from identical nodal positions of three individual plants per developmental stage (S1-S4) to minimize developmental-stage variability. GST density was quantified by manual enumeration of all GSTs across the abaxial leaf surfaces and normalized to GSTs per unit area. Whole-leaf phenotyping was performed with a Nikon (Z MC 105 mm) camera under controlled LED lighting.

Metabolite profiling by LC-MS/MS

Artemisinin and its precursors (artemisinic acid [AA], dihydroartemisinic acid [DHAA]) were quantified using an optimized LC-MS/MS protocol⁵⁰. Three independent biological replicates were analyzed for each sample ($n = 3$). Leaf samples were dried at 60°C , and 20 mg of powdered tissue was extracted in 1 mL methanol via ultrasonication (40 min) and centrifugation (14,000 rpm, 10 min). The supernatant was filtered (0.25 μm) and diluted 500-fold for analysis. LC-MS/MS analysis was performed on an Agilent 1200 Infinity LC system coupled to a 6410 Triple Quadrupole mass spectrometer. Separation was achieved using a ZORBAX SB-C18 column (2.1 $\times$ 100 mm, 3.5 μm) with isocratic elution (acetonitrile/0.1% formic acid, 65:35 v/v) at 0.3 mL/min. The mass spectrometer operated in positive electrospray ionization mode with multiple reaction monitoring (MRM), with parameters set to: nebulizer gas pressure (40 psi), source temperature (350°C), gas flow rate (10 L/min), and capillary voltage (4000 V). Data were acquired and processed using MassHunter Workstation software (vB.08.00).

Bulk RNA sequencing

Total RNA was extracted from leaves (stages 1-3, one-month-old plants, 3 biological replicates per stage) and GST-enriched samples (stages 1-2:

young GSTs; stages 3-4: old GSTs; 3 replicates of 100 plants each). Libraries were prepared using the HiEff NGS Ultima Dual-mode mRNA Library Prep Kit (Yeasen Biotechnology) and sequenced on an Illumina NovaSeq 6000 platform (150 bp paired-end reads). Raw data were processed and quality-controlled using the BMKCloud platform (www.biocloud.net).

RNA isolation and quantitative real-time polymerase chain reaction (qRT-PCR)

Total RNA was extracted using the TransZol Up Plus RNA Kit (Tiangen Biotech) according to the manufacturer's protocol. First-strand cDNA was synthesized from 1 µg RNA using the TransScript First-strand cDNA Synthesis SuperMix (TransGen Biotech). qRT-PCR was performed with SYBR-Green PCR Mastermix (Takara) on a Thermal Cycler Dice system. Gene expression levels were normalized to actin (EU531837) as an internal control. All reactions were conducted in triplicate, with primer sequences detailed in Supplementary Data 8.

Statistical analysis

Statistical analyses were conducted using GraphPad Prism (v9.8.0.200). Data normality and homogeneity of variance were assessed before analysis. Two-tailed Student's t-tests were used for comparisons between the two groups. Significance levels were denoted as: * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, and **** $P < 0.0001$.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The snRNA-seq, stRNA-seq, and bulk RNA data from this study can be found in CNGBdb (<https://db.cngb.org/>) and under project accession code [CNP0006977](https://db.cngb.org/). Source data are provided with this paper.

References

- Venkatesan, P. WHO World Malaria Report 2024. *Lancet Microbe*, 101073 (2025).
- Liu, Y. et al. Artemisinins ameliorate polycystic ovarian syndrome by mediating LONP1-CYP11A1 interaction. *Science* **384**, eadk5382 (2024).
- Zhang, H. et al. Multiscale drug screening for cardiac fibrosis identifies MD2 as a therapeutic target. *Cell* **187**, 7143–7163 (2024).
- Wu, Y. et al. Therapeutic effects of the artemisinin analog SM934 on lupus-prone MRL/lpr mice via inhibition of TLR-triggered B-cell activation and plasma cell formation. *Cell. Mol. Immunol.* **13**, 379–390 (2016).
- Tu, Y. The discovery of artemisinin (qinghaosu) and gifts from Chinese medicine. *Nat. Med.* **17**, 1217–1220 (2011).
- Duke, S. O. & Paul, R. N. Development and fine structure of the glandular trichomes of *Artemisia annua* L. *Int. J. Plant Sci.* **154**, 107–118 (1993).
- Duke, M. V., Paul, R. N., Elsohly, H. N., Sturtz, G. & Duke, S. O. Localization of Artemisinin and Artemisitene in Foliar tissues of glanded and glandless biotypes of *Artemisia annua* L. *Int. J. Plant Sci.* **155**, 365–372 (1994).
- Schramek, N. et al. Artemisinin biosynthesis in growing plants of *Artemisia annua*. A ^{13}C study. *Phytochemistry* **71**, 179–187 (2010).
- Vranová, E., Coman, D. & Grisse, W. Network analysis of the MVA and MEP pathways for isoprenoid synthesis. *Annu. Rev. Plant Biol.* **64**, 665–700 (2013).
- Mercke, P., Bengtsson, M., Bouwmeester, H. J., Posthumus, M. A. & Brodelius, P. E. Molecular cloning, expression, and characterization of amorpho-4,11-diene synthase, a key enzyme of artemisinin biosynthesis in *Artemisia annua* L. *Arch. Biochem. Biophys.* **381**, 173–180 (2000).
- Teoh, K. H., Polichuk, D. R., Reed, D. W., Nowak, G. & Covello, P. S. *Artemisia annua* L. (Asteraceae) trichome-specific cDNAs reveal CYP71AV1, a cytochrome P450 with a key role in the biosynthesis of the antimalarial sesquiterpene lactone artemisinin. *FEBS Lett.* **580**, 1411–1416 (2006).
- He, Q. et al. Glandular trichome-specific expression of alcohol dehydrogenase 1 (*ADH1*) using a promoter-GUS fusion in *Artemisia annua* L. *Plant Cell Tiss. Org.* **130**, 61–72 (2017).
- Zhang, Y. et al. The molecular cloning of artemisinic aldehyde Delta11(13) reductase and its role in glandular trichome-dependent biosynthesis of artemisinin in *Artemisia annua*. *J. Biol. Chem.* **283**, 21501–21508 (2008).
- Teoh, K. H., Polichuk, D. R., Reed, D. W. & Covello, P. S. Molecular cloning of an aldehyde dehydrogenase implicated in artemisinin biosynthesis in *Artemisia annua*. *Botany* **87**, 635–642 (2009).
- Guo, Z. et al. Dihydroartemisinic acid dehydrogenase-mediated alternative route for artemisinin biosynthesis. *Nat. Commun.* **16**, 3888 (2025).
- Brown, G. D. & Sy, L.-K. In vivo transformations of dihydroartemisinic acid in *Artemisia annua* plants. *Tetrahedron* **60**, 1139–1159 (2004).
- Sy, L.-K. & Brown, G. D. The role of the 12-carboxylic acid group in the spontaneous autoxidation of dihydroartemisinic acid. *Tetrahedron* **58**, 909–923 (2002).
- Czechowski, T. et al. *Artemisia annua* mutant impaired in artemisinin synthesis demonstrates importance of nonenzymatic conversion in terpenoid metabolism. *Proc. Natl Acad. Sci. USA.* **113**, 15150–15155 (2016).
- Paddon, C. J. & Keasling, J. D. Semi-synthetic artemisinin: a model for the use of synthetic biology in pharmaceutical development. *Nat. Rev. Microbiol.* **12**, 355–367 (2014).
- Paddon, C. J. et al. High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature* **496**, 528–532 (2013).
- Ro, D. K. et al. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **440**, 940–943 (2006).
- Schillmiller, A. L., Last, R. L. & Pichersky, E. Harnessing plant trichome biochemistry for the production of useful compounds. *Plant J.* **54**, 702–711 (2008).
- Soni, R., Shankar, G., Mukhopadhyay, P. & Gupta, V. A concise review on *Artemisia annua* L.: A major source of diverse medicinal compounds. *Ind. Crop. Prod.* **184**, 115072 (2022).
- Schuurink, R. & Tissier, A. Glandular trichomes: micro-organs with model status?. *N. Phytol.* **225**, 2251–2266 (2020).
- Ferreira, J. F. S. & Janick, J. Floral Morphology of *Artemisia annua* with Special Reference to Trichomes. *Int. J. Plant Sci.* **156**, 807–815 (1995).
- Olsson, M. E. et al. Localization of enzymes of artemisinin biosynthesis to the apical cells of glandular secretory trichomes of *Artemisia annua* L. *Phytochemistry* **70**, 1123–1128 (2009).
- Soetaert, S. S. et al. Differential transcriptome analysis of glandular and filamentous trichomes in *Artemisia annua*. *BMC Plant Biol.* **13**, 220 (2013).
- Olofsson, L., Lundgren, A. & Brodelius, P. E. Trichome isolation with and without fixation using laser microdissection and pressure catapulting followed by RNA amplification: expression of genes of terpene metabolism in apical and sub-apical trichome cells of *Artemisia annua* L. *Plant Sci.* **183**, 9–13 (2012).
- Macosko, E. Z. et al. Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell* **161**, 1202–1214 (2015).
- Tang, F. et al. mRNA-Seq whole-transcriptome analysis of a single cell. *Nat. Methods* **6**, 377–382 (2009).
- Liew, L. C. et al. Establishment of single-cell transcriptional states during seed germination. *Nat. Plants* **10**, 1418–1434 (2024).

32. Zong, J. et al. A rice single cell transcriptomic atlas defines the developmental trajectories of rice floret and inflorescence meristems. *N. Phytol.* **234**, 494–512 (2022).
33. Xu, X. et al. Single-cell RNA sequencing of developing maize ears facilitates functional analysis and trait candidate gene discovery. *Dev. Cell* **56**, 557–568 (2021).
34. Lin, J. L. et al. Single-cell RNA sequencing reveals a hierarchical transcriptional regulatory network of terpenoid biosynthesis in cotton secretory glandular cells. *Mol. Plant* **16**, 1990–2003 (2023).
35. Birnbaum, K. et al. A gene expression map of the *Arabidopsis* root. *Science* **302**, 1956–1960 (2003).
36. Habib, N. et al. Massively parallel single-nucleus RNA-seq with DroNc-seq. *Nat. Methods* **14**, 955–958 (2017).
37. Liu, Z. et al. Single-nucleus transcriptomes reveal spatiotemporal symbiotic perception and early response in *Medicago*. *Nat. Plants* **9**, 1734–1748 (2023).
38. Wang, Q. et al. Histological and single-nucleus transcriptome analyses reveal the specialized functions of ligular sclerenchyma cells and key regulators of leaf angle in maize. *Mol. Plant* **17**, 920–934 (2024).
39. Sun, G. et al. The maize single-nucleus transcriptome comprehensively describes signaling networks governing movement and development of grass stomata. *Plant Cell* **34**, 1890–1911 (2022).
40. Longo, S. K., Guo, M. G., Ji, A. L. & Khavari, P. A. Integrating single-cell and spatial transcriptomics to elucidate intercellular tissue dynamics. *Nat. Rev. Genet.* **22**, 627–644 (2021).
41. Serrano, K. et al. Spatial co-transcriptomics reveals discrete stages of the arbuscular mycorrhizal symbiosis. *Nat. Plants* **10**, 673–688 (2024).
42. Li, X. et al. Single-cell and spatial RNA sequencing reveal the spatiotemporal trajectories of fruit senescence. *Nat. Commun.* **15**, 3108 (2024).
43. Liu, Z. et al. Integrated single-nucleus and spatial transcriptomics captures transitional states in soybean nodule maturation. *Nat. Plants* **9**, 515–524 (2023).
44. Chen, M. et al. GLANDULAR TRICHOME-SPECIFIC WRKY 1 promotes artemisinin biosynthesis in *Artemisia annua*. *N. Phytol.* **214**, 304–316 (2017).
45. Xie, L. et al. The WRKY transcription factor AaGSW2 promotes glandular trichome initiation in *Artemisia annua*. *J. Exp. Bot.* **72**, 1691–1701 (2021).
46. Lu, S. et al. Cloning and functional characterization of a beta-pinene synthase from *Artemisia annua* that shows a circadian pattern of expression. *Plant Physiol.* **130**, 477–486 (2002).
47. Tissier, A. Glandular trichomes: what comes after expressed sequence tags?. *Plant J.* **70**, 51–68 (2012).
48. Haghverdi, L., Büttner, M., Wolf, F. A., Büttner, F. & Theis, F. J. Diffusion pseudotime robustly reconstructs lineage branching. *Nat. Methods* **13**, 845–848 (2016).
49. Dong, B. et al. TrichomeLess Regulator 3 is required for trichome initial and cuticle biosynthesis in *Artemisia annua*. *Mol. Hortic.* **4**, 10 (2024).
50. Tan, H. et al. Trichome and Artemisinin Regulator 1 is required for trichome development and artemisinin biosynthesis in *Artemisia annua*. *Mol. Plant* **8**, 1396–1411 (2015).
51. Shi, P. et al. The roles of AaMIXTA1 in regulating the initiation of glandular trichomes and cuticle biosynthesis in *Artemisia annua*. *N. Phytol.* **217**, 261–276 (2018).
52. Wang, C. et al. AaWIN1, an AP2/ERF protein, positively regulates glandular secretory trichome initiation in *Artemisia annua*. *Plant Sci.* **329**, 111602 (2023).
53. Ji, Y. et al. Cloning and characterization of AabHLH1, a bHLH transcription factor that positively regulates artemisinin biosynthesis in *Artemisia annua*. *Plant Cell Physiol.* **55**, 1592–1604 (2014).
54. Shen, Q. et al. The jasmonate-responsive AaMYC2 transcription factor positively regulates artemisinin biosynthesis in *Artemisia annua*. *N. Phytol.* **210**, 1269–1281 (2016).
55. Fu, X. et al. AaWRKY9 contributes to light- and jasmonate-mediated to regulate the biosynthesis of artemisinin in *Artemisia annua*. *N. Phytol.* **231**, 1858–1874 (2021).
56. He, W. et al. The AaBBX21-AaHY5 module mediates light-regulated artemisinin biosynthesis in *Artemisia annua* L. *J. Integr. Plant Biol.* **66**, 1735–1751 (2024).
57. Chen, T. et al. AaWRKY17, a positive regulator of artemisinin biosynthesis, is involved in resistance to *Pseudomonas syringae* in *Artemisia annua*. *Hortic. Res.* **8**, 217 (2021).
58. Yan, T. et al. A novel HD-ZIP IV/MIXTA complex promotes glandular trichome initiation and cuticle development in *Artemisia annua*. *N. Phytol.* **218**, 567–578 (2018).
59. Li, Y. et al. AaABI5 transcription factor mediates light and abscisic acid signaling to promote anti-malarial drug artemisinin biosynthesis in *Artemisia annua*. *Int. J. Biol. Macromol.* **253**, 127345 (2023).
60. Yu, Z. X. et al. The jasmonate-responsive AP2/ERF transcription factors AaERF1 and AaERF2 positively regulate artemisinin biosynthesis in *Artemisia annua* L. *Mol. Plant* **5**, 353–365 (2012).
61. Liu, H. et al. AaMYB108 is the core factor integrating light and jasmonic acid signaling to regulate artemisinin biosynthesis in *Artemisia annua*. *N. Phytol.* **237**, 2224–2237 (2023).
62. Matías-Hernández, L. et al. AaMYB1 and its orthologue AtMYB61 affect terpene metabolism and trichome development in *Artemisia annua* and *Arabidopsis thaliana*. *Plant J.* **90**, 520–534 (2017).
63. Zhong, Y. et al. AaABF3, an abscisic acid-responsive transcription factor, positively regulates Artemisinin Biosynthesis in *Artemisia annua*. *Front. Plant Sci.* **9**, 1777 (2018).
64. Ma, Y. N. et al. Jasmonate and abscisic acid-activated AaGSW1-AaTCP15/AaORA transcriptional cascade promotes artemisinin biosynthesis in *Artemisia annua*. *Plant Biotechnol. J.* **19**, 1412–1428 (2021).
65. Ma, Y. N. et al. Jasmonate promotes artemisinin biosynthesis by activating the TCP14-ORA complex in *Artemisia annua*. *Sci. Adv.* **4**, eaas9357 (2018).
66. Zhang, F. et al. A basic leucine zipper transcription factor, AabZIP1, connects abscisic acid signaling with artemisinin biosynthesis in *Artemisia annua*. *Mol. Plant* **8**, 163–175 (2015).
67. Lv, Z. et al. Interaction of bZIP transcription factor TGA6 with salicylic acid signaling modulates artemisinin biosynthesis in *Artemisia annua*. *J. Exp. Bot.* **70**, 3969–3979 (2019).
68. Wu, Z. et al. AaMYB15, an R2R3-MYB TF in *Artemisia annua*, acts as a negative regulator of artemisinin biosynthesis. *Plant Sci.* **308**, 110920 (2021).
69. Wang, B. et al. Transient production of artemisinin in *Nicotiana benthamiana* is boosted by a specific lipid transfer protein from *A. annua*. *Metab. Eng.* **38**, 159–169 (2016).
70. Dlugé, K. L. et al. Characterization of *Nicotiana tabacum* genotypes possessing deletion mutations that affect potyvirus resistance and the production of trichome exudates. *BMC Genomics* **19**, 484 (2018).
71. Shen, G. et al. Plant GDSL Esterases/Lipases: evolutionary, physiological and molecular functions in plant development. *Plants* **11**, 468 (2022).
72. Liu, F. et al. Non-specific lipid transfer proteins in plants: presenting new advances and an integrated functional analysis. *J. Exp. Bot.* **66**, 5663–5681 (2015).
73. Sun, S. et al. Single-cell RNA sequencing provides a high-resolution roadmap for understanding the multicellular compartmentation of specialized metabolism. *Nat. Plants* **9**, 179–190 (2023).
74. Wu, S., Morotti, A. L. M., Yang, J., Wang, E. & Tatsis, E. C. Single-cell RNA sequencing facilitates the elucidation of the complete

- biosynthesis of the antidepressant hyperforin in *St. John's wort*. *Mol. Plant* **17**, 1439–1457 (2024).
75. Seyfferth, C. et al. Advances and Opportunities in Single-Cell Transcriptomics for Plant Research. *Annu. Rev. Plant Biol.* **72**, 847–866 (2021).
 76. Huchelmann, A., Boutry, M. & Hachez, C. Plant Glandular Trichomes: Natural Cell Factories of High Biotechnological Interest. *Plant Physiol.* **175**, 6–22 (2017).
 77. Zhao, Q., Li, M., Zhang, M. & Tan, H. Glandular trichomes: the factory of artemisinin biosynthesis. *Med. Plant Biol.* **3**, e019 (2024).
 78. Huang, D. et al. Trichome-specific analysis and Weighted Gene Co-Expression Correlation Network Analysis (WGCNA) reveal potential regulation mechanism of Artemisinin biosynthesis in *Artemisia annua*. *Int. J. Mol. Sci.* **24**, 8473 (2023).
 79. Zhou, P. et al. Single-cell transcriptome of *Nepeta tenuifolia* leaves reveal differentiation trajectories in glandular trichomes. *Front. Plant Sci.* **13**, 988594 (2022).
 80. Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587.e3529 (2021).
 81. McGinnis, C. S., Murrow, L. M. & Gartner, Z. J. DoubletFinder: Doublet detection in single-cell RNA Sequencing Data Using Artificial Nearest Neighbors. *Cell Syst.* **8**, 329–337 (2019).
 82. The Gene Ontology Consortium The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Res* **47**, 330–338 (2019).
 83. Kanehisa, M. et al. KEGG for linking genomes to life and the environment. *Nucleic Acids Res* **36**, 480–484 (2008).
 84. Stringer, C., Wang, T., Michaelos, M. & Pachitariu, M. Cellpose: a generalist algorithm for cellular segmentation. *Nat. Methods* **18**, 100–106 (2021).
 85. Aran, D. et al. Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nat. Immunol.* **20**, 163–172 (2019).
 86. Li, C. et al. SciBet as a portable and fast single cell type identifier. *Nat. Commun.* **11**, 1818 (2020).
 87. Cable, D. M. et al. Robust decomposition of cell type mixtures in spatial transcriptomics. *Nat. Biotechnol.* **40**, 517–526 (2022).
 88. Ma, Y. & Zhou, X. Spatially informed cell-type deconvolution for spatial transcriptomics. *Nat. Biotechnol.* **40**, 1349–1359 (2022).
 89. Trapnell, C. et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* **32**, 381–386 (2014).
 90. Morabito, S., Reese, F., Rahimzadeh, N., Miyoshi, E. & Swarup, V. hdWGCNA identifies co-expression networks in high-dimensional transcriptomics data. *Cell Rep. Methods* **3**, 100498 (2023).
 91. Beygelzimer, A. et al. FNN: Fast Nearest Neighbor Search Algorithms and Applications. R package version 1.1.4.1, <https://CRAN.R-project.org/package=FNN>. (2024).
- 82474033 (T.), the Natural Science Foundation of Shanghai no. 24ZR1493300 (T.), and the Traditional Chinese Medicine Research Project of Shanghai no.2024PT010 (T.). The drafts of Figs. 1a, 2a, and 3a were created using BioRender (<https://biorender.com>) and further edited with Adobe Illustrator.

Author contributions

H.T. conceived and designed the project. M.Z., M.L., Y.A., C.L., Q.Z. and B.P. performed the experiments. M.Z., M.L., Y.A. and D.Z. analyzed the data. M.Z., Y.A. and Q.Z. wrote the manuscript draft. H.T., M.L. and C.L. revised the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-025-63770-y>.

Correspondence and requests for materials should be addressed to Hexin Tan.

Peer review information *Nature Communications* thanks Xiao-Ya Chen, who co-reviewed with Jin-Quan Huang, Young Jin Lee, Bosheng Li and the other anonymous reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025

Acknowledgements

This work was made possible through the support of grants from the National Natural Science Foundation of China no. 82122072 (T.) and no.