Single-cell RNA sequencing facilitates the elucidation of the complete biosynthesis of the antidepressant hyperforin in St. John's wort

Song Wu[1,2], Ana Luisa Malaco Morotti[1], Jun Yang[1], Ertao Wang[1] and Evangelos C. Tatsis[1,3,*]

[1]National Key Laboratory of Plant Molecular Genetics, CAS Center for Excellence in Molecular Plant Sciences, Shanghai Institute of Plant Physiology and Ecology, Chinese Academy of Sciences, Shanghai 200032, China

[2]University of Chinese Academy of Sciences, Shanghai 200032, China

[3]CEPAMS - CAS–JIC Centre of Excellence for Plant and Microbial Science, Shanghai 200032, China

*Correspondence: Evangelos C. Tatsis (etatsis@cemps.ac.cn)

https://doi.org/10.1016/j.molp.2024.08.003

ABSTRACT

Hyperforin is the compound responsible for the effectiveness of St. John's wort (*Hypericum perforatum*) as an antidepressant, but its complete biosynthetic pathway remains unknown. Gene discovery based on co-expression analysis of bulk RNA-sequencing data or genome mining failed to discover the missing steps in hyperforin biosynthesis. In this study, we sequenced the 1.54-Gb tetraploid *H. perforatum* genome assembled into 32 chromosomes with the scaffold N50 value of 42.44 Mb. By single-cell RNA sequencing, we identified a type of cell, "Hyper cells", wherein hyperforin biosynthesis *de novo* takes place in both the leaves and flowers. Through pathway reconstitution in yeast and tobacco, we identified and characterized four transmembrane prenyltransferases (HpPT1–4) that are localized at the plastid envelope and complete the hyperforin biosynthetic pathway. The hyperforin polycyclic scaffold is created by a reaction cascade involving an irregular isoprenoid coupling and a tandem cyclization. Our findings reveal how and where hyperforin is biosynthesized, enabling synthetic-biology reconstitution of the complete pathway. Thus, this study not only deepens our comprehension of specialized metabolism at the cellular level but also provides strategic guidance for elucidation of the biosynthetic pathways of other specializied metabolites in plants.

Key words: hyperforin, *Hypericum perforatum*, *de novo* genome sequencing, single-cell RNA sequencing, biosynthetic pathway, meroterpenoids

INTRODUCTION

St. John's wort (SJW, *Hypericum perforatum*, Hypericaceae) is a high-value medicinal plant with annual global sales exceeding US$6 billion (Rizzo et al., 2020). In classical antiquity, extracts from aerial SJW organs were used to treat skin ailments due to antiseptic and anti-inflammatory properties (Istikoglou et al., 2010), while nowadays they are prescribed to treat mild to moderate depression and are associated with fewer adverse reactions than common antidepressants (Linde et al., 1996; Di Carlo et al., 2001). Its primary potent bioactive ingredient, hyperforin, acts similarly to traditional antidepressants by inhibiting neurotransmitter reuptake like serotonin, though via a different mechanism (Chatterjee et al., 1998; Schempp et al., 1999; Müller et al., 2001; Leuner et al., 2007; Ng et al., 2017). Unlike conventional antidepressants, hyperforin induces Ca^{2+} and Na^+ by uniquely activating TRPC6 receptors, disrupting

cation gradients and reducing neurotransmitter uptake. This unique process makes hyperforin a promising agent for treating depression (El Hamdaoui et al., 2022).

H. perforatum is endemic in Eurasia and North Africa and was introduced into the Americas and Oceania by human migration. It originates from an ancient hybridization event in Siberia between *Hypericum maculatum* and *Hypericum attenuatum* (Robson, 2003). Typically tetraploid, with a chromosome number of $2n = 4x = 32$ (Robson, 2003), the plant also reproduces asexually through facultative apomixis, leading to populations with diploid, triploid, and hexaploid cytotypes (Barcaccia et al., 2006), with tetraploids by far the most

---

common and likely the ancestral type (Alan et al., 2015). Diploid and hexaploid cytotypes are produced by tetraploid plants via haploid parthenogenesis and fertilization of unreduced egg cells, respectively (Matzk et al., 2001). Besides hyperforin, SJW also biosynthesizes diverse specialized metabolites including essential oils rich in complex terpenoids (Nahrstedt and Butterweck, 1997), flavonoids (Tatsis et al., 2007), biflavonoids (Tatsis et al., 2007), phenolic acids (Tatsis et al., 2007), and naphtodianthrones (Skalkos et al., 2002; Tatsis et al., 2007, 2008).

Hyperforin is the first polycyclic polyprenylated acylphloroglucinol (PPAP) to be isolated (Gurevich et al., 1971) (Figure 1A). PPAPs are a diverse class of natural products with various biological activities that are taxonomic markers for the family Hypericaceae (Ciochina and Grossman, 2006; Richard, 2014; Yang et al., 2018). Hyperforin's unique structure features an isoprenyl group at C8 (Figure 1A), thus posing a challenging task for enantioselective synthesis, inspiring chemists to pursue novel organic-synthesis strategies (Ciochina and Grossman, 2006; Richard, 2014; Yang et al., 2018). Such synthesis routes involve multiple steps and can result in total yield of 1.64% (Ji et al., 2022). Despite hyperforin's significance, its biosynthesis *de novo* has up to now been unresolved. With reconstitution of the pathway in microbial or plant synthetic-biology systems, obtaining hyperforin from such sources is a sustainable and environmentally superior strategy.

To date, details surrounding the biosynthesis of specialized metabolites in SJW, and for the majority of the hyperforin biosynthesis steps, are sparse and insubstantial. We have previously characterized three enzymes, namely a branched-chain amino acid dehydrogenase (BCKDHA), a coenzyme A (CoA) ligase (CCL), and an aromatic polyketide synthase (PKS) from two biosynthetic gene clusters (*HpBGC1* and *HpBGC2*) that encode the first three steps dedicated to biosynthesis of the polyketide hyperforin precursor, phloroisobutyrophenone (PIBP) (Wu et al., 2022) (Figure 1A). Provision of ^{13}C-labeled substrates to *H. perforatum* seedlings revealed that the rest of the hyperforin carbon atoms comprise five isoprenoid units derived from the plastid-localized methyl-D-erythritol-4-phosphate (MEP) pathway (Adam et al., 2002) (Figure 1A). Chemical logic dictates that the isoprenoid units derived from the plastid-localized MEP pathway (Adam et al., 2002; Dewick, 2011; Walsh and Tang, 2017) are joined to the polyketide PIBP by electrophilic substitution, involving an additional ring closure to biosynthesize the bicyclo[3.3.1]nonane scaffold (Supplemental Figure 1) of hyperforin. Transfer of isoprenoid moieties from pyrophosphate donor substrates to aromatic acceptor substrates in plants is typically catalyzed by membrane-bound aromatic prenyltransferases (Tello et al., 2008; Vickery et al., 2016; Christianson, 2017; de Bruijn et al., 2020); however, there are no reports of enzymes catalyzing cyclization reactions in prenylated aromatic compounds involving C–C bond formation.

High-quality genome assemblies are of paramount importance for complete pathway elucidation of metabolites with complicated structures (Zhao et al., 2019; Reed et al., 2023). Two fragmented and incomplete annotations of a diploid *H. perforatum* genome are available (Galla et al., 2019; Zhou et al., 2021), but these are insufficient for hyperforin biosynthetic pathway elucidation (Supplemental Figure 2).

In this study, to fully elucidate the hyperforin biosynthetic pathway, we sequenced the tetraploid genome of *H. perforatum* and generated single-cell atlases from leaves and flowers. Based on mapping of previously characterized upstream genes, we identified a new type of cell—"Hyper cells"—wherein *de novo* biosynthesis of hyperforin occurs. By using two synthetic-biology platforms for heterologous expression (yeast and tobacco), we discovered and characterized four new UbiA-family prenyltransferases that complete the hyperforin biosynthetic pathway. The last step for hyperforin biosynthesis involves an unusual prenyltransferase (HpPT4) catalyzing a non-canonical branching prenylation and a rare tandem cyclization.

RESULTS

Sequencing and assembly of the tetraploid *H. perforatum* genome

To enable the discovery of the missing steps in hyperforin biosynthesis, we generated transcriptome and genome-sequence resources *de novo* for *H. perforatum* using a suite of Illumina, PacBio HiFi, and Hi-C sequencing (Supplemental Table 1). We used Illumina paired-end short-read sequencing to survey the genome and, based on k-mer analysis, the genome size was estimated as 1.71 Gb (Supplemental Table 2) with a heterozygosity rate of 0.12%. PacBio HiFi long-read sequencing was mainly used for genome assembly in conjunction with Illumina sequencing for contig error correction, resulting in a 1.54-Gb (Supplemental Table 3) draft assembly comprising 3710 scaffolds and an N50 value of 42.44 Mb. A Hi-C (*in vivo* fixation of chromosomes) library refined the first version of the reference genome and the scaffolds aligned into 32 groups hereafter referred to as chromosomes (Figure 1B, Supplemental Table 4, and Supplemental Figures 3 and 4), which represented a tetraploid accession with a haploid chromosome number of 8. Coverage of the final assembly was evaluated by Illumina short-read mapping, CEGMA, and BUSCO (Supplemental Tables 5–7) that indicate a high-quality assembly with 97.2% and 97.6% completeness, respectively. A total of 98 005 genes (Supplemental Table 8) were annotated by a pipeline combining *de novo* and homology-based predictions, with bulk RNA-sequencing (RNA-seq) data (Li et al., 2023b). Transposable elements account for 56.95% of the genome (Supplemental Table 9).

Functional annotation of genes was performed by searching against publicly available databases including SwissProt, NCBI-NR, Pfam, KEGG, and InterPro. The best match from each database was assigned to the corresponding protein-coding gene. Following successful anchoring of 32 chromosomes and phasing of four subgenomes by Hi-C (Supplemental Figure 5A), we mapped the location of upstream genes within the known biosynthetic gene cluster *HpBGC1* (Wu et al., 2022). This cluster is located on Chr4A within a topologically associating domain (Nützmann et al., 2020) (Supplemental Figure 5B), highlighting the co-regulation of the cluster's genes for proper biosynthesis of the polyketide precursor PIBP. Intraspecies syntenic analysis revealed the presence of five allele counterparts of *HpBGC1* on Chr4A–D and on the non-homologous chromosome Chr5A (Supplemental Figure 6A). These copies of *HpBGC1a–e* are not identical, but bear variation in metabolic-gene
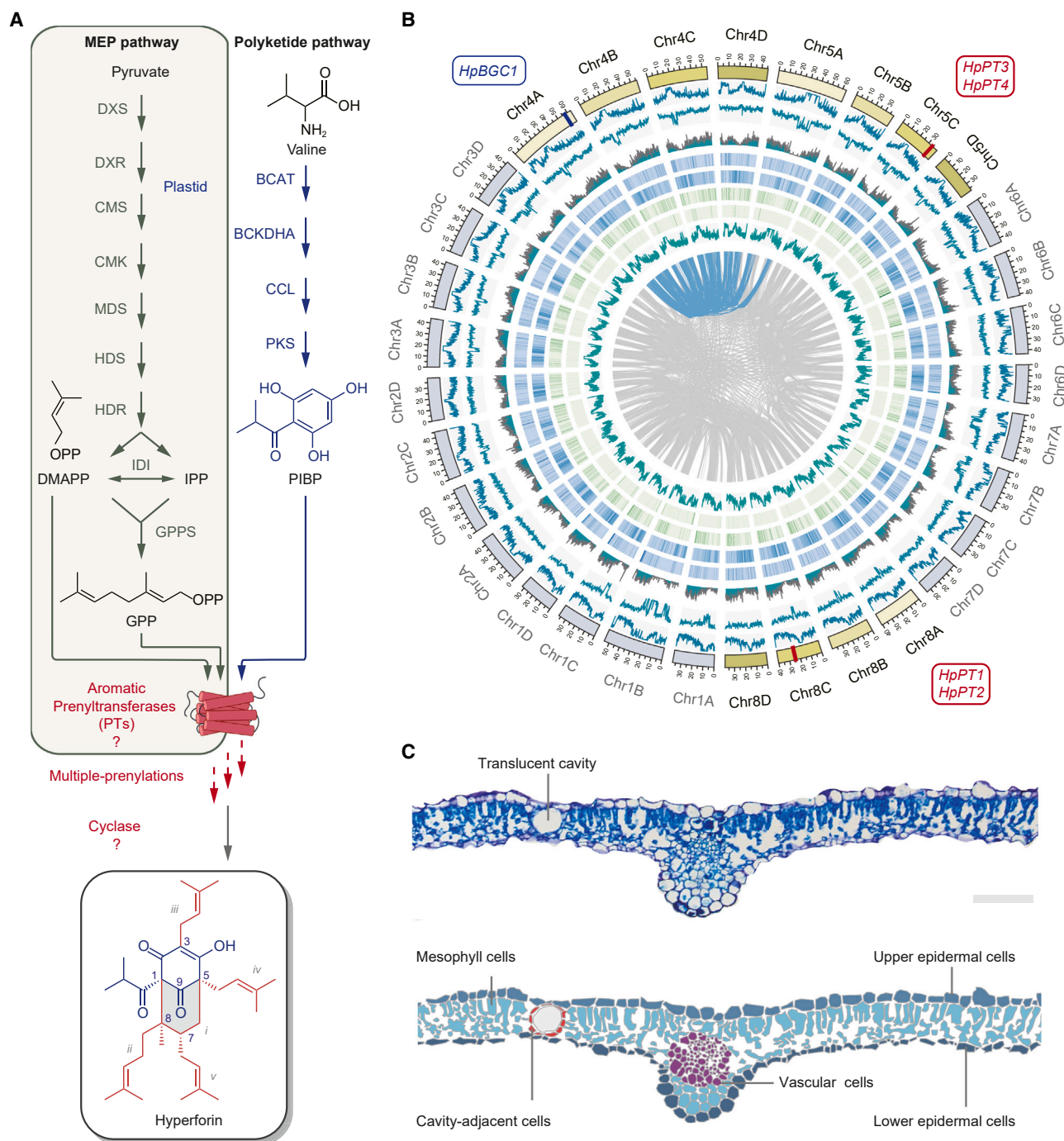
Figure 1. Hyperforin biosynthesis, translucent cavities, and genomic architecture of *Hypericum perforatum*.

(A) Overview of the proposed model of hyperforin biosynthesis *de novo* in St. John's wort. The pathway involves three metabolic modules. The first module involves the biosynthesis of polyketide phloroisobutyrophenone (PIBP) from valine (Wu et al., 2022). The plastid-localized methylerythritol 4-phosphate (MEP) pathway biosynthesizes prenyl units such as dimethylallyl diphosphate (DMAPP) and geranyl diphosphate (GPP) (Adam et al., 2002). The last module is hypothesized to involve prenylation of PIBP by a plastid-localized transmembrane aromatic prenyltransferase(s) and an enigmatic cyclization involving the formation of a C1–C8 bond (Adam et al., 2002; Dewick, 2011; Walsh and Tang, 2017).

(B) Overview of *H. perforatum* genome assembly. The CIRCOS plot represents the features of the *H. perforatum* genome (size 1.39 Gb). From inner to outer layers: the gray links represent intraspecies syntenic relationships, densities of simple sequence repeats, rRNA, tRNA, snRNA, miRNA, genes, GC content, and transposable elements in 100-kb sliding windows, respectively. In the outer layer, bars represent individual chromosomes. Blocks marked in blue on Chr4A (*HpBGC1*), in red on Chr5C (*HpPT3–4*) and Chr8C (*HpPT1–2*) indicate metabolic gene locations involved in hyperforin biosynthesis. Blue links highlight the syntenic relationships of *HpBGC1* copies among different chromosomes.

(C) Cross-section of *H. perforatum* leaf, highlighting the translucent pockets. Top: a tissue section stained with tolonium chloride, with a schematic depiction (bottom) depicting the spatial profile of the indicated cell types. Scale bar, 100 μm.

co-linearity, gene-copy numbers, and tandem duplications (Supplemental Figure 6B). By contrast, *HpBGC1d* on Chr4D has undergone substantial gene losses and retained only one predicted *PKS*. These *HpBGC1* polymorphisms are possibly the result of segmental rearrangements within the tetraploid genome.

Genome mining and co-expression analysis to identify aromatic prenyltransferases

Several instances are reported of three or more genes comprising the same biosynthetic pathway as being physically co-localized; thus, genome mining can aid gene discovery (Nützmann et al., 2020; Polturak and Osbourn, 2021; Polturak et al., 2022; Méteignier et al., 2023). Because of the polyketide-module biosynthesis genes clustered within *HpBGC1* (Wu et al., 2022), clustering may well also be the case for the downstream genes (Weng, 2014; Leong and Last, 2017; Li et al., 2023b). Based on functional annotation, 405 genes were predicted to encode UbiA-family aromatic prenyltransferases (Supplemental Figure 7) from the four sets of chromosomes. By mining the *H. perforatum* genome, we identified a cluster comprising 11 annotated aromatic prenyltransferase genes bearing high homology to known enzymes acting in biosynthesis of prenylated xanthones (Sayed et al., 2023). A blueprint of prenylated-xanthone metabolism in *Hypericum* species (Fiesel et al., 2015; Nagia et al., 2019; Sayed et al., 2023) indicates analogies with the proposed model of hyperforin biosynthesis involving the recruitment of dedicated CoA ligases, type-III polyketide synthases and aromatic prenyltransferases that differ in the starting polyketide unit, and the role of oxidative cyclizations in the xanthone pathway (El-Awaad et al., 2016; Wang et al., 2022a) (Supplemental Figure 8). Toward determining their functionality and potential contribution to hyperforin biosynthesis in the absence of available functional-characterization approaches in *H. perforatum*, the 11 predicted prenyltransferase genes were individually cloned and co-expressed in yeast together with the known CoA ligase (*HpCCL2*) and polyketide synthase (*HpPKS2*) from *HpBGC1* (Wu et al., 2022). No prenylated products were detected by liquid chromatography–mass spectrometry (LC–MS) analysis, indicating that these prenyltransferases do not participate in hyperforin biosynthesis.

Genes in plant specialized-metabolite biosynthesis are usually co-regulated as part of the same biological pathway; thus, discovery of unknown metabolic genes in such pathways is usually based on co-expression analysis of bulk RNA-seq data across plant samples (e.g., various tissues or different treatments) (Nett et al., 2020, 2023; De La Peña et al., 2023; Reed et al., 2023). Co-expression analysis with upstream genes from *HpBGC1* as bait did not yield likely prenyltransferase candidates with high correlation efficiency (Nett et al., 2020; De La Peña et al., 2023) (Pearson correlation coefficient >0.90) among the 405 annotated prenyltransferases (Supplemental Figure 9).

Single-cell RNA-seq enables the discovery of missing hyperforin biosynthetic genes

The eponym *perforatum* for St John's wort is derived from the visible translucent spots that resemble perforations found on both leaves and flowers that also are present in certain other *Hypericum* species (Maggi et al., 2004; Kucharíková et al., 2016)

(Figure 1C). These translucent cavities are spherical cellular structures spanning the upper to lower epidermis, delimited as described (Ciccarelli et al., 2001; Soelberg et al., 2007) by cell layers consisting of flattened secretory-like cells and turgid parenchymatous cells (Ciccarelli et al., 2001; Soelberg et al., 2007). Hyperforin seems to accumulate within these structures in *H. perforatum* leaves (Soelberg et al., 2007). In contrast, hyperforin is undetectable in non-secretory tissues, proposing that hyperforin's presence in the secretory structures originates either from the translucent cavities or from the immediately adjacent cells (Soelberg et al., 2007). Since there are no supporting biochemical or molecular data, it remains unclear whether the complete hyperforin biosynthetic pathway, or step(s) along it, is active around the translucent cavities or whether they simply serve as storage vessels.

High-throughput single-cell RNA-seq (scRNA-seq) can distinguish cell types or cell populations based on cell-to-cell heterogeneity as regards gene expression (Li et al., 2023a; Liao and Wang, 2023; Xu and Jackson, 2023; Grones et al., 2024), but its mainstream exploitation in plant science is hindered by substantial technical barriers such as the preparation of protoplasts and the heterogeneity of plant tissues (Liao and Wang, 2023; Xu and Jackson, 2023; Grones et al., 2024). Despite this, scRNA-seq has been applied to cellular differentiation and communication (Liao and Wang, 2023; Xu and Jackson, 2023), while use in plant specialized metabolism is sparse (Kang et al., 2022; Li et al., 2023a).

We hypothesize that since hyperforin biosynthesis correlates spatially with translucent cavities and their associated cell types (Figure 1C), the expression signal of hyperforin biosynthesis genes might be concealed by noise typical of bulk RNA-seq, thus necessitating scRNA-seq to carefully examine the cellular architecture of *H. perforatum* leaves. To generate a single-cell transcriptome atlas of *H. perforatum* leaves, we successfully isolated viable protoplasts from young leaves to capture the cellular heterogeneity within the tissue. Droplet-based 10× Genomics was used for RNA-seq (Mazutis et al., 2013) to capture gene-expression data at the single-cell level. Reads were mapped to the reference genome, quantifying gene-expression levels in individual cells with 92.2% of the reads mapped to the reference genome, thus validating the quality and accuracy of the transcriptome sequence. We identified a total of 15 074 cells expressing a median of 1405 genes across different cell types. To interpret this dataset and categorize the cell types, low-quality and doublet cells were removed, followed by normalization of the expression data to eliminate technical noise and standardization for downstream analysis. Principal component analysis (PCA) was used to reduce the dimensionality of scRNA-seq data and distill the complex gene-expression profiles, capturing the most significant variation among the assayed cells. To categorize the cells into clusters based on shared transcriptome profiles, we applied the unbiased clustering algorithm uniform manifold approximation and projection (UMAP) (Hao et al., 2021) to visualize and identify distinct cell populations (Figure 2A and Supplemental Figure 10).

Annotation of cell types in a plant single-cell atlas is based on single or a set of marker genes, although such markers are absent for *H. perforatum*. A common approach for cell annotation relies
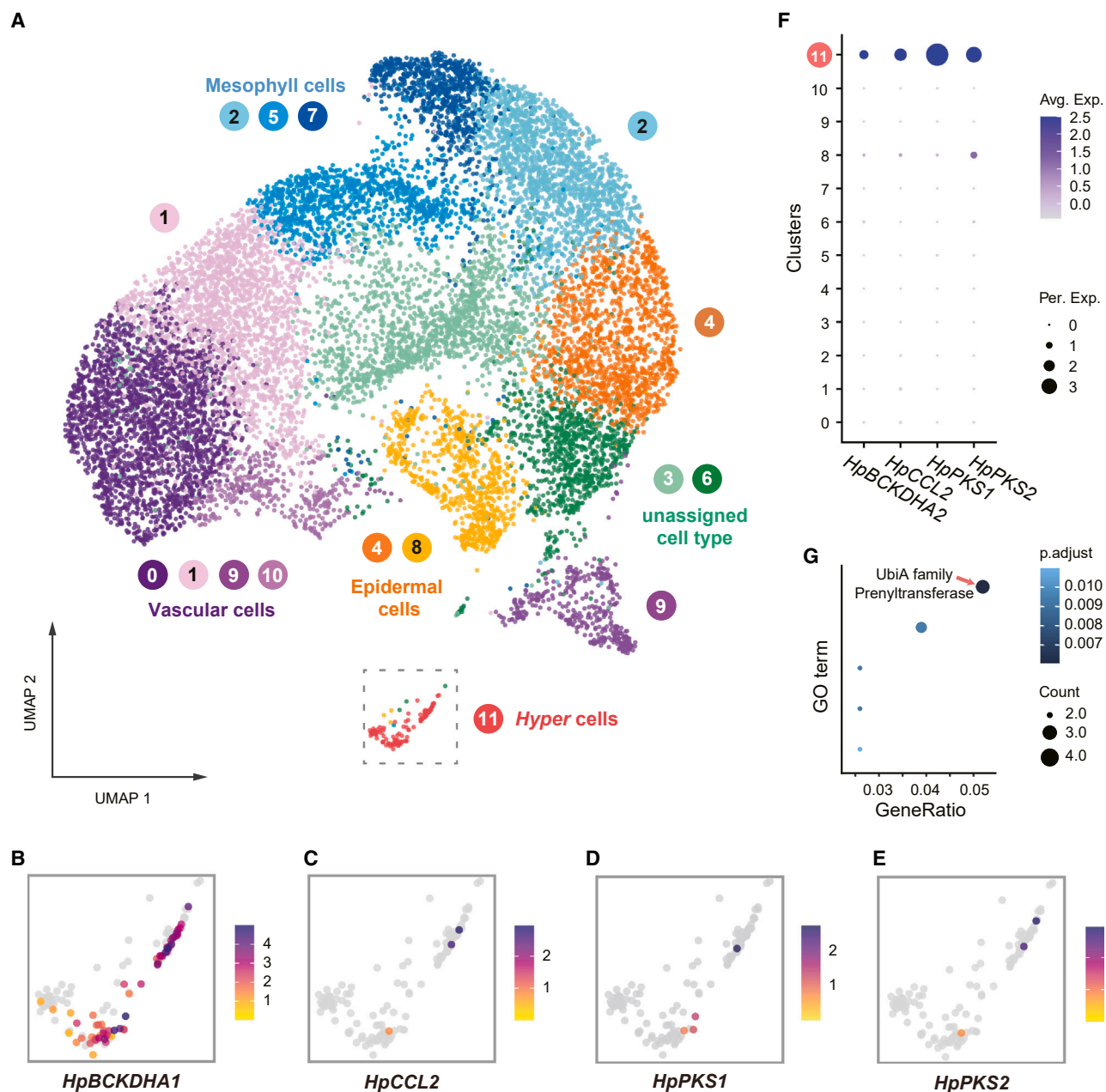
Figure 2. A single-cell atlas of an *H. perforatum* leaf and identification of Hyper cells.

(A) Dimensionally reduced UMAP visualization of single leaf cells with each color representing a distinct cluster. Twelve different cell clusters were identified. The four clusters of vascular cells (0, 1, 9, and 10) are visualized in shades of purple/pink; the three clusters of mesophyll cells (2, 5, and 7) are visualized in shades of blue/cyan; the two clusters of epidermal cells (4 and 8) in shades of blue/gray; and two clusters of unassigned cells (3 and 6) in shades of green. Cells of cluster 11, visualized in red and enclosed within a dashed box, are assigned as Hyper cells. Each dot represents one cell.

(B–E) Close-up view of the expression profiles in Hyper cells of genes previously identified in hyperforin biosynthesis, *HpBCKDHA1* **(B)**, *HpCCL2* **(C)**, *HpPKS1* **(D),** and *HpPKS2* **(E),** from *HpBGC1*. The heatmap color scale depicts the relative expression of each gene in each cell.

(F) Dotplot illustrating the predominant expression of metabolic genes in the polyketide module of hyperforin biosynthesis (*HpBCKDHA2*, *HpCCL2*, *HpPKS1*, and *HpPKS2* from *HpBGC1*) across all clusters in this atlas. The color scale indicates the average scaled expression of each gene in each cell type. Dot sizes depict the fraction of cells wherein a given gene is expressed in a given cell type.

(G) GO enrichment analysis of cluster 11, revealing "UbiA-family prenyltransferase" as the top-enriched protein family.

upon homologous marker genes (Li et al., 2023a; Sun et al., 2023; Zhan et al., 2023); however, such an approach proved unsuitable for our datasets (Supplemental Figure 11). Extensive databases for genes from model plants can instead be exploited as cell-type markers (Chen et al., 2021; Jin et al., 2022; He

et al., 2024). Identification of functional orthologs between *Arabidopsis thaliana* and *H. perforatum* by sequence alignment and homology may be prone to false positives. Synteny is one of the criteria for establishing the functional relationships between orthologous genes or genomic regions across different
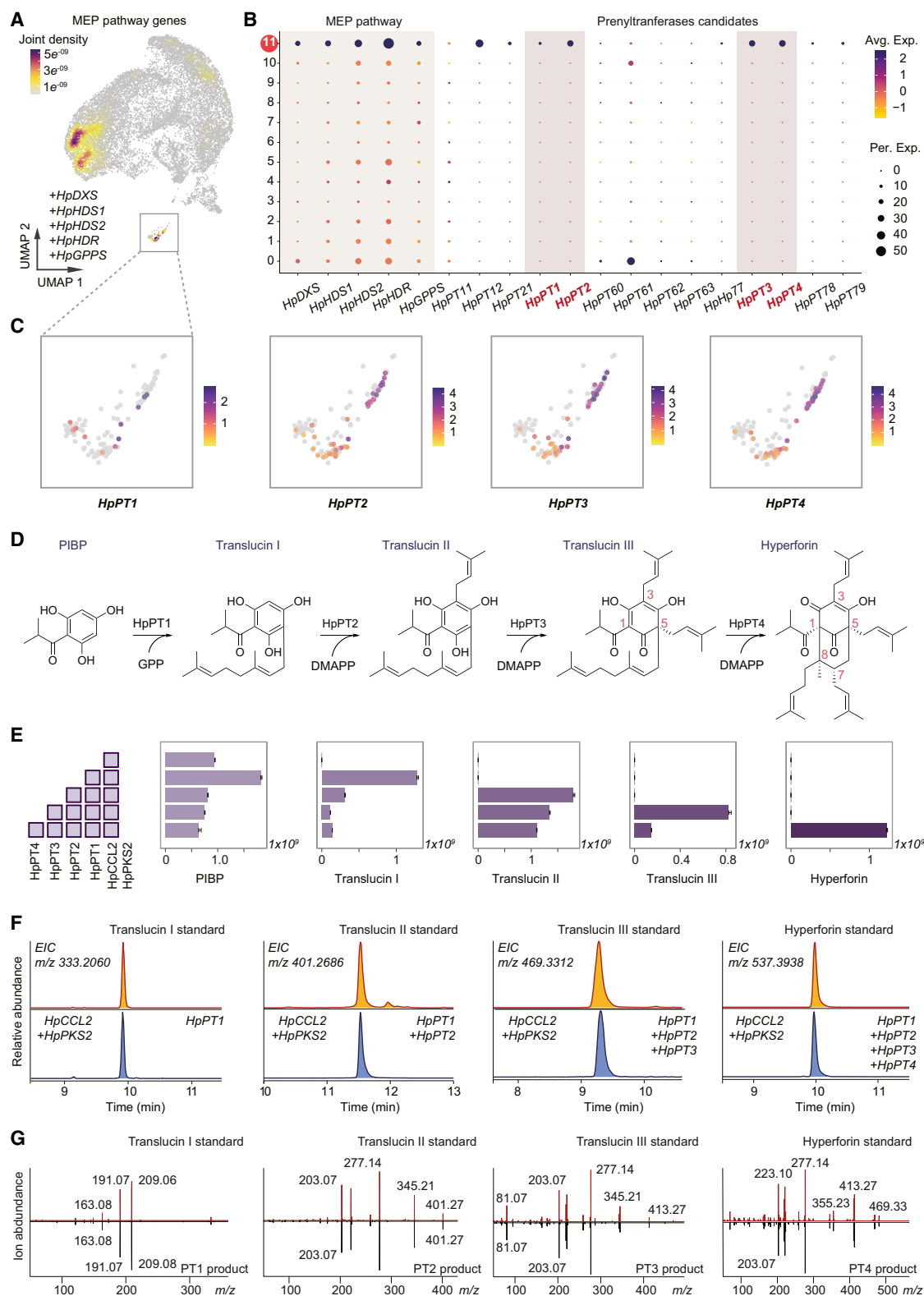
Figure 3. Elucidation of the catalytic steps to complete the formation of hyperforin.

(A) Overlaid gene expression of MEP pathway genes (*HpDXS*, *HpHDS1*, *HpHDS2*, *HpHDR*, and *HpGPPS*) across the *H. perforatum* leaf single-cell atlas demonstrates high expression in cell clusters 0 (vascular) and 11 (Hyper cells). The heatmap color scale indicates the sum of expression from the five MEP genes in each cell.

(legend continued on next page)

species (Tang et al., 2008; Zeng et al., 2017; Cervantes-Pérez et al., 2022; Li et al., 2023b). To annotate the cell clusters, we performed a fully automated annotation pipeline comprising the following. (1) We compiled an extensive marker-gene dataset for *A. thaliana* leaves (Chen et al., 2021. (2) Through syntenic analysis between *A. thaliana* and *H. perforatum*, we correlated the orthologous marker genes for different cell types in leaf tissue between the two species and generated a marker-gene dataset for *H. perforatum*. (3) Using the ScType (Ianevski et al., 2022) platform, an automated cell-type annotation was performed based on the new set of markers (Supplemental Figure 12). The 12 different cell clusters (Figure 2A, Supplemental Figure 10, and Supplemental Table 10) were annotated, with the cells from clusters 0, 1, 9, and 10 annotated as vascular cells, the cells from clusters 2, 5, and 7 as mesophyll cells, and cells from clusters 4 and 8 as epidermal cells. There were no adequate markers to support annotation for cells from clusters 3 and 6, while the minority and topologically distant cell cluster 11 was annotated with low confidence (i.e., low ScType score) as myrosin cells (Shirakawa et al., 2016; Chen et al., 2021).

Visualization and quantification of expression of known early hyperforin biosynthetic-pathway genes (Wu et al., 2022) revealed that *HpBCKDHA1*, *HpBCKDHA2*, *HpCCL2*, *HpPKS1*, and *HpPKS2* are distinctively and highly expressed in cluster 11 (Figures 2B–2F). This indicates that biosynthesis of the precursor PIBP takes place predominantly in cell cluster 11 and, in parallel, raises the question of whether genes encoding downstream steps are co-expressed in these cells. Toward addressing this, upon gene ontology (GO) enrichment analysis (Paysan-Lafosse et al., 2023), we found that genes belonging to the UbiA prenyltransferase family are over-represented in this cluster (Figure 2G and Supplemental Figure 13). The isoprenoid carbon atoms in hyperforin are derived from the MEP plastidial pathway (Adam et al., 2002), and UbiA-type transmembrane aromatic prenyltransferases are expected to catalyze the prenylation reactions by transferring isoprenoid units to PIBP. MEP pathway genes *HpDXS*, *HpHDS1*, *HpHDS2*, *HpHDR*, and *HpGPPS* are all highly expressed in vascular cells (cluster 0) and in cluster 11 (Figures 3A and 3B). We therefore focused on aromatic prenyltransferases expressed in cell cluster 11. Based on this profile, four genes annotated as aromatic prenyltransferases (*Hper_g146976*, *Hper_g147016*, *Hper_g96819*, and *Hper_g96821*) were selected as candidates based on their high expression in cluster 11 (Figures 3B and 3C; Supplemental Figure 14), and successfully cloned. All the selected candidate prenyltransferases have plastid transit peptides at the N termini, which were removed for expression in yeast based on ChloroP predictions (Emanuelsson et al., 1999).

To screen the candidate prenyltransferases involved in hyperforin biosynthesis (Figures 3D and 3E), we assembled the PIBP precursor biosynthetic module in *Saccharomyces cerevisiae*. The CoA ligase HpCCL2 and the type-III PKS HpPKS2 (Wu et al., 2022) were cloned into pESC-Leu, and transcripts lacking the corresponding sequence encoding plastidial transit peptides of the four candidate prenyltransferases were cloned into pESC-His and transformed into the isoprenoid-rich yeast strain DD104 (Li et al., 2015). After galactose induction, the culture was extracted with ethyl acetate and analyzed by LC–MS. Only the prenyltransferase encoded by *Hper_g146976* demonstrated activity toward PIBP with a new peak eluting at 9.9 min (Figure 3F and Supplemental Figure 15) under the extracted ion chromatogram (EIC) mode at m/z 333.2060, suggesting that the enzyme uses a geranyl diphosphate molecule as prenyl donor and attaches the geranyl chain to PIBP. The geranyl-PIBP synthesized standard elutes at 9.9 min with m/z = 330.2060 in positive mode, which aligned with the peak of the Hper_g146976 product (Figure 3F) and is in agreement with the MS/MS (MS^2) spectrum (Figure 3G and Supplemental Figure 15). We conclude that the prenyltransferase encoded by *Hper_g146976* acts on PIBP and catalyzes the first prenylation step in hyperforin biosynthesis by way of a geranyl transferase. It is therefore assigned as HpPT1 and its product as translucin I. Further analysis of enzyme-assay LC–MS data showed that HpPT1 uses only geranyl diphosphate (GPP) as a prenyl donor and not dimethylallyl diphosphate (DMAPP) (Supplemental Figure 16).

To elucidate the outstanding biosynthesis steps involving PIBP prenylations, we cloned the three remaining candidate prenyltransferase genes into the second cloning site of pESC-His and assayed their activity in yeast. Ethyl acetate (EtOAc) extracts from galactose-induced strains were analyzed by LC–MS, and a new peak at EIC m/z 401.2686 was detected at 11.6 min (Figures 3E and 3F; Supplemental Figure 17) in a strain expressing *Hper_g147016*. By comparison with the translucin II synthetic standard and alignment of chromatograph (based on retention time, Figure 3F) and MS data (MS and MS^2, Figure 3G), we concluded that *Hper_g147016* encoding HpPT2 catalyzes the second prenylation in hyperforin biosynthesis.

(B) Dotplot highlighting elevated expression levels of MEP pathway genes in cluster 11, along with potential prenyltransferase (*PT*) candidate genes. Genes with confirmed activity are marked in red. The color scale indicates the average scaled expression of each gene in each cell type. Dot sizes reflect the fraction of cells wherein a given gene is expressed in a given cell type.

(C) Cell-level expression profiles of prenyltransferases (*HpPT1:Hper_g146976*, *HpPT2:Hper_147016*, *HpPT3:Hper_96819*, and *HpPT4:Hper_96821*) in Hyper cells (cluster 11). The heatmap color scale shows the expression of each gene in each cell.

(D) The sequential prenylation and tandem cyclization steps leading to hyperforin formation from phloroisobutyrophenone (PIBP), based on the verified activity of prenyltransferases HpPT1, HpPT2, HpPT3, and HpPT4. The intermediates translucins I–III were chemically synthesized and detected in *H. perforatum* extracts (Supplemental Figures 14 and 16–18).

(E) LC–MS peak area of the exact ion mass $[M + H]^+$ for each intermediate produced in yeast after co-expression of the indicated genes. Boxes to the left of the plots denote combinations of biosynthetic genes in each co-expression experiment. Data were obtained from three independent biological replicates, and the bars represent the mean ± SE.

(F) Comparison of extracted ion chromatogram (EIC) of each PIBP prenylated product from *in vivo* co-expression of biosynthetic genes (in red/gold) with EICs from corresponding synthetic standards (translucins I–III) and hyperforin (in black/blue).

(G) Comparative MS/MS fragmentation patterns (MS^2 spectra) of prenylated products and standards from **(F)**.
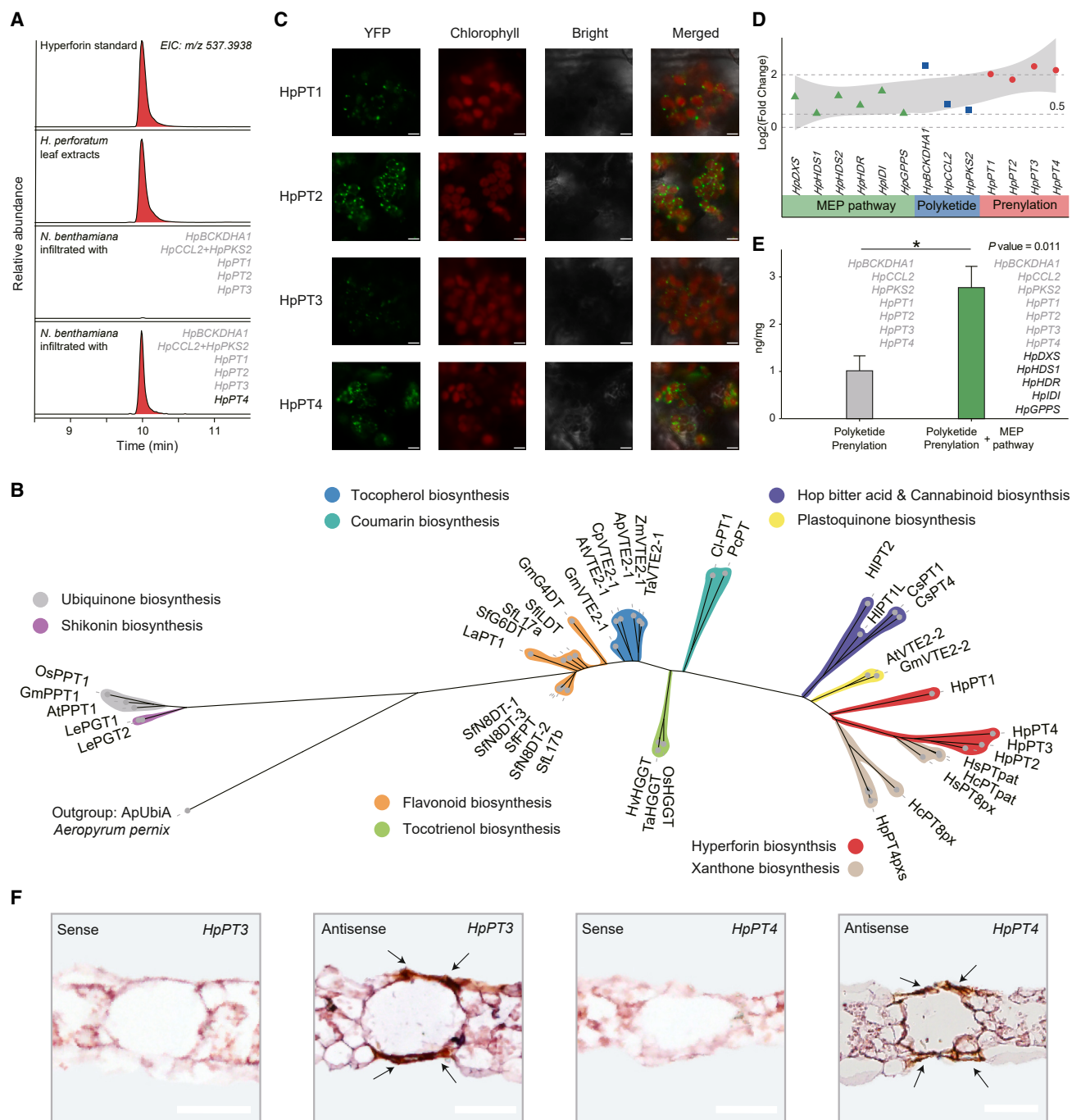
Figure 4. Plastids are prenyl-unit donors for hyperforin biosynthesis in *H. perforatum*.

(A) Analysis of transient co-expression of *HpBCKDHA1*, *HpCCL2*, *HpPKS2*, *HpPT1*, *HpPT2*, *HpPT3*, and *HpPT4* in tobacco for biosynthesis of hyperforin by LC–MS. Extracted ion chromatograms (EIC, *m/z* 537.3938) from LC–MS analysis of a hyperforin standard and *H. perforatum* extract are shown for comparison.

(B) Phylogenetic tree of characterized aromatic prenyltransferases from UbiA family from plants. The phylogeny was built from published, confirmed-active UbiA enzymes in plant metabolism and the prenyltransferases characterized in this study (highlighted in red shading) using 1000 bootstrap values. UbiA from the fungus *Aeropyrum pernix* was used as the outgroup. The different color shades indicate the different pathways in which the respective prenyltransferases are acting. Phylogenetic analysis illustrates the close relation of prenyltransferases in hyperforin biosynthesis with the ones from plastid-localized plastoquinone biosynthesis (shaded in yellow).

(C) Subcellular localization of HpPT1, HpPT2, HpPT3, and HpPT4 fused at the C terminus to YFP in transiently transgenic *N. benthamiana* leaves imaged by confocal laser-scanning microscopy. Scale bars, 10 μm.

(legend continued on next page)

To discover further the remaining prenylation steps, we cloned the remaining two candidates into pESC-Ura and assayed their activity using the same workflow. When we co-expressed *Hper_g96819* with HpCCL2, HpPKS2, HpPT1, and HpPT2, a new peak appeared in the LC–MS chromatogram of EIC *m/z* 469.3312 at 9.4 min (Figures 3D–3F and Supplemental Figure 18). This new product was identified as translucin III, based on comparison of retention time and MS and MS^2 spectra with a synthetic standard (Figures 3D–3F and Supplemental Figure 18), indicating the germinal prenylation on C5 (Figure 3D) as HpPT1 and the de-aromatization of the polyketide ring. Therefore, Hper_g96819 was assigned HpPT3, which catalyzes the third prenylation step on the scaffold of PIBP polyketide.

Based on chemical intuition (Figure 3D), a prenylation step and a cyclization step are expected for the transformation of translucin III to hyperforin. To determine whether the only remaining candidate gene encodes such functionalities, *Hper_g96821* was cloned into the second site of pESC-Ura and co-expressed with HpCCL2, HpPKS2, and HpPT1–3. Hper_g96821 is enzymatically active, transforming translucin III to hyperforin, with the product eluting at 10.0 min (the same time as the hyperforin standard), and the MS and MS^2 spectra were identical to those of hyperforin (Figures 3D–3G and Supplemental Figure 19). Hper_g96821 catalyzes the fourth sequential prenylation step and was thus named HpPT4.

Although HpPT4 is a UbiA prenyltransferase, it does not apparently catalyze a typical Friedel–Crafts alkylation (Tello et al., 2008; de Bruijn et al., 2020) as aromatic prenyltransferases are expected to, but instead acts as an irregular prenyltransferase catalyzing the 1′–2 coupling (branching) (Thulasiram et al., 2007; Christianson, 2017) between two isoprenyl units (Supplemental Figure 1A). Moreover, it appears that HpPT4 does not only catalyze the transfer of a prenyl unit to translucin III but also catalyzes bond formation between C(1) and C(8), resulting in cyclization and biosynthesis of the iconic bicyclo[3.3.1]nonane (Supplemental Figure 1A) scaffold of PPAPs (Ciochina and Grossman, 2006; Yang et al., 2018). Such cyclizations generating complex polycyclic structures are hallmarks of natural-product biosynthesis (Supplemental Figure 1B) (Dang et al., 2018).

To ensure host enzyme activities do not bias the functional characterization of candidates (Czechowski et al., 2022), we assayed the activity of full-length HpPT1–4 by transient expression in *Nicotiana benthamiana* leaves with co-expression with *HpBCKDHA1*, HpCCL2, and *HpPKS2* followed by LC–MS analysis of leaf methanol extracts. Hyperforin was detected *in planta* only when all the enzymes in the pathway were expressed (Figure 4A). This supports the functional characterization in

yeast and safeguards the proposed role of HpPT4 as a hyperforin synthase to transfer a prenyl group and catalyze the cyclization step to complete hyperforin biosynthesis.

To further validate the proposed catalytic capacity of HpPT4 acting as branching 1′–2 prenyltransferase and cyclase, we tested the catalytic activity of HpPT4 against the translucin III analog colupulone (Li et al., 2015) (Supplemental Figure 20). We co-expressed HpPT4 with HpCCL2, HpPKS2, HlPT1L, and HlPT2 (Li et al., 2015) in yeast. LC–MS analysis of extracts showed that HpPT4 could transform colupulone to secohyperforin (Charchoglyan et al., 2007; 8.0 min; *m/z* [M + H]$^+$ 469.3312) in a similar manner as it transforms translucin III to hyperforin (Supplemental Figure 17). To screen for additional prenyltransferases that have the same activity with HpPT1–4, we cloned, expressed, and assayed HpPT12, HpPT21, and HpPT77 (Supplemental Figures 15 and 17–19). No activity was observed.

Hyperforin biosynthesis *de novo* involves all three modules and takes place in Hyper cells

HpPT1 and *HpPT2* are located close by on chromosome 8 (Supplemental Figures 21A and 22A), and *HpPT3* and *HpPT4* are also located near one another on chromosome 5 Supplemental Figures 21B and 22B). Genomic proximity can imply that the two pairs are results of tandem duplications; however, the low sequence identity between these two pairs suggests that this is not the case (Supplemental Table 11). Phylogenetic analysis of annotated prenyltransferases based on amino acid sequences reveals that HpPT1–4 are members of the homogentisate clade of transmembrane plant aromatic prenyltransferases with plastid localization, similar to enzymes acting in the plastoquinone pathway (Figure 4B). HpPT1–4 fused at the C terminus to yellow fluorescent protein (YFP) when transiently expressed in *N. benthamiana* leaves and *A. thaliana* protoplasts localize to plastids (Figure 4C and Supplemental Figure 23), respectively, supporting conclusions from stable-isotope supplementation experiments indicating that isoprenyl units are derived from the plastidial MEP pathway (Adam et al., 2002).

Hyperforin biosynthesis can be divided into three metabolic modules involving the biosynthesis of the polyketide PIBP localized to the cytosol, the MEP pathway in plastids supplying prenyl donors (DMAPP and GPP), and plastid-localized transmembrane prenylation (Supplemental Figure 24). We have previously shown that genes encoding PIBP biosynthetic enzymes are upregulated 24 h after methyl jasmonate treatment (Wu et al., 2022). Genes in the MEP pathway and *HpPT1–4* are also responsive to methyl jasmonate treatment, indicating co-regulation of the whole hyperforin biosynthetic

---

(D) Dotplot depicting the logarithmic differential enhanced expression of genes in leaves from the hyperforin biosynthetic modules MEP (green), polyketide/PIBP (blue), and prenylations (red) after 24 h of treatment with 400 μM methyl jasmonate. All the genes in hyperforin biosynthesis *de novo* show enhancement of their expression by methyl jasmonate.

(E) LC–MS quantification of hyperforin production by transient co-expression in tobacco (hyperforin [ng] per leaf sample dry weight [mg]) of the genes from polyketide and prenylation modules (in gray) and co-expression of the genes from MEP, polyketide, and prenylation modules (in green). Data were obtained from five independent biological replicates. Bars represent the mean ± SE.

(F) RNA *in situ* hybridization reveals the localization of *HpPT3* and *HpPT4* expression in proximity to the translucent cavities in *H. perforatum* leaf cross-sections. Scale bars, 50 μm.

pathway (Figure 4D and Supplemental Figure 24). To confirm the role of *H. perforatum* genes from the MEP pathway in supplying precursors for hyperforin biosynthesis, we transiently co-expressed the *H. perforatum* MEP pathway genes (*HpDXS*, *HpDXR*, *HpHDS1*, *HpHDR*, *HpIDI*, and *HpGPPS*) together with genes from the PIBP module and *HpPT1–4* in *N. benthamiana*, resulting in over 2-fold hyperforin production compared to co-expression of genes from polyketide and prenylation modules (Figure 4E).

Because all the newly identified genes involved in hyperforin biosynthesis are exclusively and highly expressed in cell cluster 11, we therefore nickname these cells "Hyper cells." Using Seurat (Hao et al., 2021), we identified 15 high-confidence genes that can be used as markers for this type of cell (Supplemental Table 12 and Supplemental Figure 26). Of these, *HpBCKDHA1*, *HpPT2*, *HpPT3*, and *HpPT4* feature with expression at least 4-fold higher than other cell types. To visualize the position of Hyper cells in leaf cellular architecture, we performed RNA *in situ* hybridization using probes against *HpPT3* and *HpPT4*. Detection of *HpPT3* and *HpPT4* transcripts imply that Hyper cells are adjacent to translucent cavities (Figure 4F; Supplemental Figures 27 and 28).

Biosynthesis of hyperforin in flowers also takes place in specific cells

Because hyperforin mainly accumulates in both leaves and open flowers (Tekel'ová et al., 2000; Wu et al., 2022), the description of Hyper cells and elucidation of hyperforin biosynthesis *de novo* in leaves raises the question about activity of the hyperforin biosynthetic pathway in flowers. To determine whether hyperforin biosynthesis *de novo* in floral organs is organized at the cellular level in a similar way as leaves, protoplasts were isolated from flowers for scRNA-seq. The floral single-cell atlas reveals 11 different cell clusters (Figure 5A and Supplemental Table 13). We applied the same strategy we employed for the leaf single-cell atlas and used flower-specific cell-type markers for *A. thaliana* to generate a customized database for cell types in *H. perforatum* flower. *HpPT1–4* are specifically and highly expressed in cell cluster 0 (Figures 5B–5F and Supplemental Figure 29) Moreover, expression of genes from all three modules of hyperforin biosynthesis (MEP, PIBP/polyketide, and prenylation) predominantly occur in cell cluster 0 in flowers (Figures 5B–5F and Supplemental Figure 24).

To explore the correlation and similarity of Hyper cells in leaves and cell cluster 0 in flowers, we compared the respective scRNA-seq datasets using MetaNeighbor (Crow et al., 2018; Fischer et al., 2021), which through gene filtering, neighbor voting, and visualization quantifies cell-type replicability, determines gene sets that contribute to cell-type identity, and enables rapid identification of clusters with high similarity. MetaNeighbor provides a performance score, which is the mean area under the receiver-operator characteristic curve (AUROC). AUROC values vary between 0 and 1, with 1 a perfect classification and values over 0.9 for 1-vs.-all and over 0.7 for 1-vs.-best considered to be of high confidence (Fischer et al., 2021). A high level of similarity between cluster 11 from leaves with cluster 0 from flowers (1-vs.-all AUROC score 0.91 and 1-vs.-best AUROC score 0.77) lead to us to conclude that similar types of specialized

cells are present in both these tissues of SJW (Figure 5G) that biosynthesize hyperforin.

DISCUSSION

By sequencing the genome of *H. perforatum* at chromosome level and generating single-cell atlases from leaves and flowers, we demonstrate how the combination of these approaches enabled the elucidation of hyperforin biosynthesis *de novo* and reconstitution of the complete pathway in yeast and *N. benthamiana*. We anticipate these findings will serve as a blueprint to elucidate the biosynthesis of plant specialized metabolites, particularly other prominent bioactive PPAPs such as the anticancer molecule garcinol (Oike et al., 2012) and neuroactive garsubellin A (Fukuyama et al., 1997) (Supplemental Figure 1B). Our results illustrate the recruitment of a UbiA-family aromatic prenyltransferase to catalyze the 1′–2 "head-to-middle" irregular prenylation reaction (Christianson, 2017) and also act as a terpene cyclase for the formation of the bicyclo[3.3.1]nonane scaffold found in hyperforin and certain other PPAPs in medicinal plants (Ciochina and Grossman, 2006; Yang et al., 2018). This complex bicyclic scaffold is a common structural feature among natural products with different biosynthetic origins (such as alkaloids, terpenoids, polyketides, and meroterpenoids; Supplemental Figure 1B) (Matsuda and Abe, 2016; Shen et al., 2020; Nett et al., 2023) across plants and microorganisms. Hyperforin synthase (HpPT4) is one example of how nature's enzymatic versatility permits different enzymes to biosynthesize this abundant yet complex bicyclic scaffold. Heterologous biosynthesis of hyperforin highlights synthetic-biology applications for production of this valuable molecule and other bioactive PPAPs.

ScRNA-seq serves as a powerful, alternative approach for gene discovery and rapid pathway elucidation in plant specialized metabolism, offering high resolution in gene expression at the cellular level (Kang et al., 2022; Li et al., 2023a). To address the challenge (Grones et al., 2024) of performing annotation of the different cell types in non-model plant species, as the majority of medicinal plants are, we present an automated annotation pipeline using ScType (Ianevski et al., 2022) with a customized marker-gene database based on identification of functional ortholog genes through syntenic analysis. Another challenge imposed by the high number of genes with high degree of similarity, especially in a tetraploid genome, is the accurate quantification of gene expression at the single-cell level. In mapping of monoterpene indole alkaloid metabolism in *Catharansus roseus* at single-cell level (Sun et al., 2023), the expression of the key gene *CrGO* (Tatsis et al., 2017; Wang et al., 2022b) was depicted differently with the use of Alevin (Srivastava et al., 2019) over Cellranger (Zheng et al., 2017). Either the use of the mapping algorithm Alevin or the t-SNE (t-distributed stochastic neighbor embedding) clustering algorithm as alternatives to handle the complexities of gene-expression quantification (e.g., *HpCCL2*, *HpPKS2*) show similar gene-expression profiles for the genes of polyketide and prenylation modules (Supplemental Figures 30–35).

These conclusions should be considered in light of certain limitations. The observed ratio of vascular-bundle cells to mesophyll cells in our dataset deviates from common patterns seen in
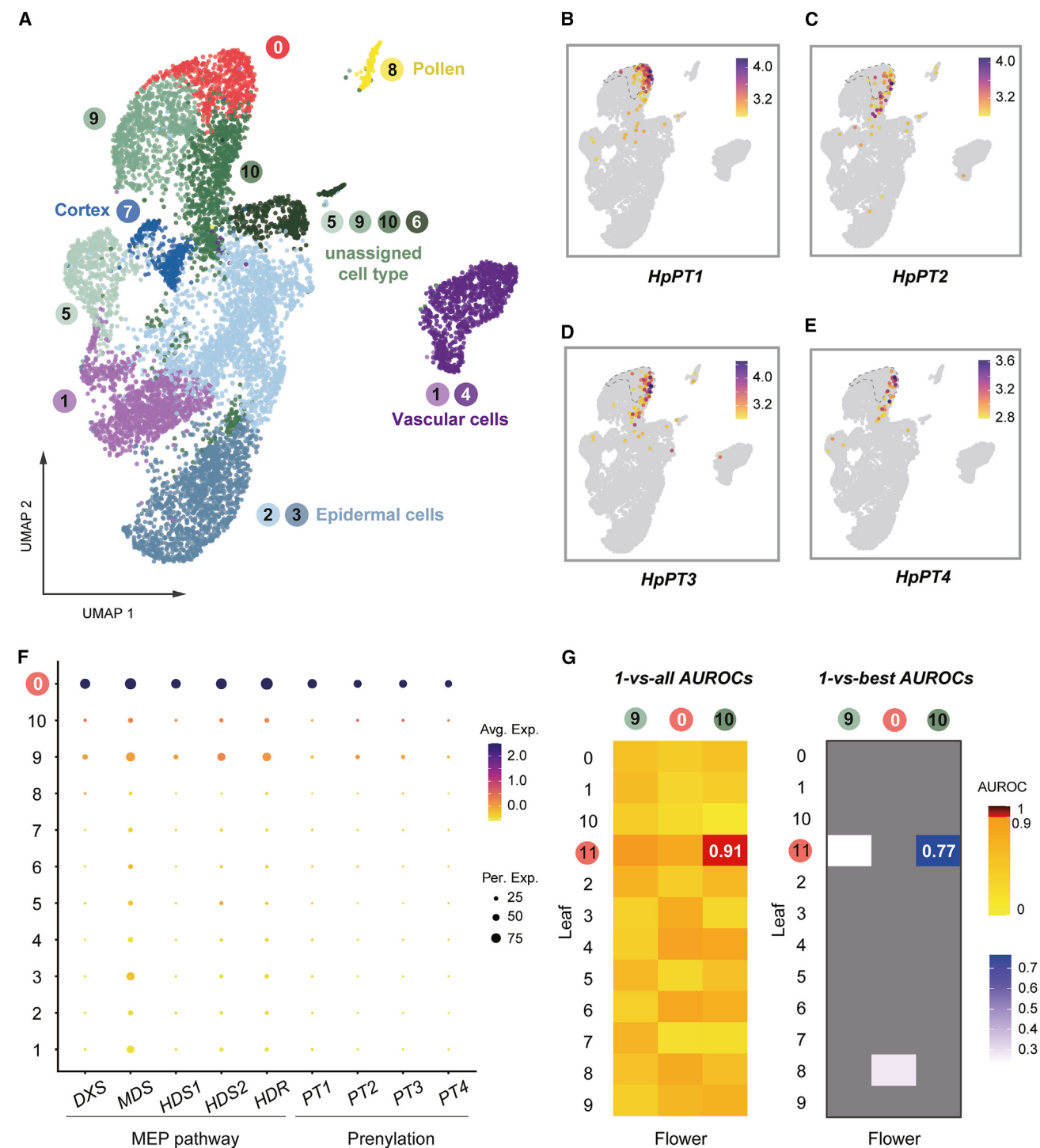
Figure 5. A single-cell atlas reveals hyperforin biosynthesis in *H. perforatum* flowers.
(A) Dimensionally reduced UMAP visualization depicting various single-cell clusters within the flower, with each color denoting a distinct cluster.
(B–E) Expression profiles of the indicated prenyltransferases (PT) at the single-cell level within the flower.
(F) Dotplot showcasing the predominant expression patterns of metabolic genes involved in hyperforin biosynthesis *de novo* across identified clusters.
(G) Heatmap based on MetaNeighbor AUROCs applied to cluster 0 and neighbor clusters of cells 9 and 10 in flower atlas, and clusters in leaf atlas. The AUROC values that passed the threshold for both 1-vs.-all (>0.9) and 1-vs.-best (>0.7) analysis are shown.

more tractable model plants, potentially reflecting species-specific cellular organization, anatomical differences, or challenges in protoplast preparation. The use of sucrose centrifugation, while ensuring high-quality protoplasts, may inadvertently

enrich certain cell types, contributing to atypical cell ratios. The presence of numerous highly similar genes in the tetraploid genome poses challenges to accurate quantification of expression levels of poly-copy metabolic genes. Precise annotation of

cell types can be affected by cell-clustering resolution, while non-model plants such as *H. perforatum* present fundamental challenges especially due to the lack of established reference data. To correlate gene markers between model plants and less-studied species, genomic synteny is a suitable bioinformatics tool to weigh the gene orthology with a consideration to their taxonomic divergence, gene expansion due to polyploidy, whole-genome duplications, and tandem gene duplications. However, these challenges are counterbalanced by the discovery of missing steps and high-confidence elucidation of the complete biosynthetic pathway for hyperforin.

Our findings reveal a major role of specialized Hyper cells to function as cellular factories for the biosynthesis *de novo* of hyperforin that combines three biosynthetic modules (MEP, PIBP/polyketide, and prenylation) in one cell type. The finding that Hyper cells are adjacent to translucent cavities where hyperforin is stored at high levels in leaves, together with the identification of biosynthetic genes with our next steps toward identifying regulatory and transport mechanisms, will deepen the understanding of specialized metabolism in plants at the single-cell level (Weng et al., 2021). Our results not only provide a synthetic-biology approach to produce the hallmark polycyclic scaffold of hyperforin but also underscore a powerful strategy for accelerating discovery and manipulation of natural-product biosynthesis in medicinal plants that are otherwise not readily amenable to functional characterization in conventional model organisms.

METHODS

Plant materials

H. perforatum subsp. *chinense* were collected in Enshi Tujia and Miao Autonomous Prefectures in Hubei, China, delivered, and grown under a 16-h:8-h light/dark cycle at 20°C in the CEMPS Phytotron. *A. thaliana* and *N. benthamiana* plants for transient expression assay were grown under a 16-h:8-h light/dark cycle at 22°C.

Bulk RNA sequencing

A total of 17 plant tissues, namely untreated leaves, MeJA-treated leaves (Wu et al., 2022), young flowers, old flowers, young buds, middle-aged buds, old buds, leaves near buds, young leaves, middle-aged leaves, old leaves, stressed leaves, young stems, old stems, old roots, primary roots, and secondary roots were frozen on liquid nitrogen and homogenized using a tissue lyser (Shanghai Jingxin 40 JXFSTPRP-48). Total RNA was then extracted using the RNAprep Pure Plant Plus Kit (Tiangen, cat. #DP441), following the manufacturer's instructions.

The purity of the extracted RNA was assessed based on the $OD_{260\ nm}$/$OD_{280\ nm}$ ratio determined by a NanoDrop instrument (Thermo Fisher Scientific, USA), and agarose gel electrophoresis was performed to test for RNA degradation and potential contamination. RNA integrity was evaluated using a 2100 Bioanalyzer (Agilent Technologies, USA).

RNA samples that met quality-control criteria were used as input for preparation of Illumina short-read libraries. The NEBNext Ultra RNA Library Prep Kit (NEB, USA, cat. #E7530L) was used according to the manufacturer's recommendations, and index codes were incorporated into the libraries to attribute sequences to their respective samples. Qualified libraries were pooled together and subjected to sequencing on the Illumina HiSeq X Ten platform to produce 150-bp pair-end reads.

Mixed RNA from 11 different plant tissues (except untreated and MeJA-treated leaves) was prepared to construct continuous long-read libraries using the SMRTbell Express Template Prep Kit 2.0 and sequenced on PacBio Sequel IIe to generate a full-length transcriptome.

Genome sequencing, assembly, and annotation

For the long-read sequencing of the SJW genome, genomic DNA was extracted using a phenol–chloroform method (Gautam, 2022). We constructed single-molecule, real-time (SMRT) sequencing libraries following Pacific Biosciences standard protocols (www.pacb.com) by shearing DNA to a target size of 15–18 kb, followed by damage repair, end repair, blunt-end ligation, and size selection to create a large-insert SMRTbell library. This library was then sequenced on the PacBio Sequel platform. The assembly of the long reads was carried out using Hifiasm (Cheng et al., 2021). Additionally, a crosslinking step was employed using 4% (v/v) formaldehyde for the genomic DNA followed by vacuum treatment. After quenching the crosslinking reaction, samples were digested overnight with MboI. The fragment ends were marked with biotin, and the chromatin DNA was religated. The DNA underwent purification via the phenol–chloroform method, and biotin was removed from non-ligated fragment ends. A-tails were then added to the fragment ends and ligated to Illumina paired-end sequencing adapters. The resulting Hi-C sequencing libraries were amplified by PCR and sequenced on the Illumina NovaSeq platform. The data from Hi-C sequencing were used to anchor the long-read assembly into pseudochromosomes.

Repetitive genomic sequences were identified using TRF (Benson, 1999), RepeatMasker (Saha et al., 2008), and LTR_FINDER(Xu and Wang, 2007). Structural annotation was performed using TblastN (Altschul et al., 1990) and GeneWise (Birney et al., 2004) for homolog prediction and tools Augustus (Stanke et al., 2008), Geneid (Alioto et al., 2018), and Genescan (Burge and Karlin, 1997) for *ab initio* gene prediction. Integration of RNA-seq data facilitated accurate gene-model prediction. Functional annotation involved aligning proteins to the Swiss-Prot database and domains to InterProScan. The prediction of non-coding RNAs, including tRNAs, rRNAs, microRNAs, and small nuclear RNAs, was achieved using tRNAscan-SE (Chan et al., 2021) and alignment with the Rfam database. Genomic features were presented by different tracks in a CIRCOS plot (Krzywinski et al., 2009).

Reference-genome-based read mapping

To acquire high-quality, clean RNA-seq reads, the raw data were trimmed using Trimmomatic v0.36 (Bolger et al., 2014) with the following parameters: HEADCROP: 13; LEADING: 3; TRAILING: 3; SLIDINGWINDOW: 4:15; and MINLEN: 36. Clean data were then aligned to the genome assembly using hisat2 (Kim et al., 2019) with default settings. Fragments per kilobase of transcript per million mapped reads (FPKM) values were then generated through the cuffdiff command within the Cufflinks suite (Trapnell et al., 2010).

For co-expression analysis of bulk transcriptomic data by Pearson correlation coefficient, the raw gene-expression data from 15 different samples (young flowers, old flowers, young buds, middle-aged buds, old buds, leaves near buds, young leaves, middle-aged leaves, old leaves, stressed leaves, young stems, old stems, old roots, primary roots, and secondary roots from 1-year-old plants kept in the CEMPS Phytotron) were obtained in FPKM format, representing normalized expression levels across various samples. The data were initially loaded into R and transposed using the transpose function. A pairwise correlation matrix was generated, which computed Pearson correlation coefficients between all pairs of genes across all samples. From these, correlations specific to known genes were extracted. This subset included only the correlation values between the gene of interest and all other genes including prenyltransferase candidates.

Preparation of St. John's wort protoplasts for single-cell RNA-seq

Fresh SJW leaves and buds from three individual 1-year-old plants grown in the CEMPS Phytotron were collected and thoroughly washed with distilled water to remove any surface contaminants. Anthers were removed from bud samples because of their negative influence on protoplast purity, leaving only the petals and ovary. Tissues were then finely chopped and placed in a 10-cm culture dish. Ten milliliters of digestion solution (pH 5.7) containing cellulase (1.25% w/v), pectinase (0.3% w/v), D-mannitol (0.4 M), 2-(*N-morpholino*)ethanesulfonic acid *(*MES) (20 mM), and KCl (20 mM) was added to the culture dish containing the macerated tissues and incubated at room temperature (25°C) with shaking at 80 rpm for 2 h. The digested solution was centrifuged at room temperature (500 *g*) for 10 min to separate the protoplast-containing pellet from the tissue debris. The resulting protoplast pellet was carefully resuspended in 500 μl of washing buffer consisting of 0.4 mM mannitol, 20 mM MES (pH 5.7), 20 mM KCl, 10 mM CaCl$_2$, and 0.1% (w/v) bovine serum albumin. The protoplast suspension was filtered through a 40-μm cell sieve to remove any remaining tissue debris and obtain a clarified protoplast solution, followed by centrifugation at room temperature (200 *g*) for 6 min to further remove debris and obtain a purified protoplast pellet. The pellet was gently resuspended in 40 μl of fresh washing buffer to obtain a concentrated protoplast suspension suitable for subsequent experiments. The viability of the isolated protoplasts was evaluated using trypan blue staining.

Single-cell RNA-seq

The library construction process used the Chromium Single Cell 3' Library and Single Cell 3' v3 Gel Beads kit. To generate single-cell GEMs (gel beads in emulsion), a suspension of protoplasts was loaded onto the automated Chromium Controller instrument. Ensuring the integrity of each GEM, reverse transcription occurred within them, resulting in cDNA that carried a unique cell barcode. Following cDNA synthesis, GEMs were broken to release the cDNA from individual cells. Liberated cDNA was then pooled and amplified using PCR. After amplification, the cDNA was fragmented and sequencing adapters were attached to construct libraries. Using the Illumina HiSeq X Ten platform, paired-end short reads of 150 bp were generated.

scRNA-seq data analysis

Raw sequencing data were preprocessed to transform it into a structured format by Cellranger (Zheng et al., 2017). Initial steps in Seurat (Hao et al., 2021) involved filtering the dataset to exclude cells with aberrantly high gene counts (>3000), potentially indicative of artifacts or multiplets. This was followed by normalization to adjust for variations in sequencing depth across individual cells using 10 000 as scale factor. To focus the analysis on genes most likely to reveal biologically significant variation, a feature-selection process based on the top 2000 variable genes was used. PCA dimensionality reduction was then applied to the data to distill the high-dimensional gene-expression profiles. The selection of dimensions for this reduction was determined through jackstraw and elbow plot (Hao et al., 2021). The presence of potential doublets was assessed and removed by DoubletFinder (McGinnis et al., 2019). In total, 15 074 and 12 382 protoplast cells were isolated from leaves and flowers, respectively and, after the DoubletFinder removal, 13 957 and 10 176 cells from leaves and flowers, respectively, were used for downstream analysis. Clustering algorithms were then used to identify distinct cell populations within the dataset. The resolution of clustering was fine-tuned by clustree (Zappia and Oshlack, 2018) ranging from 0.1 to 1. For visual representation, the cell populations were projected by UMAP reduction. Marker genes were identified as genes predominantly expressed in each cell cluster. UCell (Andreatta and Carmona, 2021) was used to score gene signatures based on the gene data matrix, and the cell-type replicability between the two datasets from the two tissues (leaves and flowers) was asserted by MetaNeighbor (Crow et al., 2018; Fischer et al., 2021) enabling rapid identification of clusters with high similarity.

MetaNeighbor (Crow et al., 2018; Fischer et al., 2021) has the advantage of quantifying cell-type similarity between different tissues or datasets by a three-step procedure comprising gene filtering, neighbor voting, and visualization.

We adapted a customized cell-type annotation approach using the ScType pipeline (Ianevski et al., 2022), due to the limitation of the standard ScType marker-gene database being specific to human and mouse. We compiled an extensive marker-gene dataset for *A. thaliana* from PlantscRNAdb (Chen et al., 2021) and mapped these genes to their syntenic counterparts in *H. perforatum*. This tailored database enabled us to perform automated and unbiased cell-type classification. A schematic of this annotation is detailed in Supplemental Figure 12.

Phylogenetic analysis of prenyltransferases

The amino acid sequences of characterized prenyltransferases were aligned with sequences of genes obtained from synteny analysis using MUSCLE (Edgar, 2004). JTT+I+G was selected as the best-fit model of amino acid replacement using prottest (Darriba et al., 2011) based on alignment results. The maximum-likelihood method was then used to infer the phylogenetic tree by RaxML-ng (Kozlov et al., 2019) with 1000 bootstrap values.

Gene isolation and plasmid construction

Total RNA was extracted from the aerial parts of *H. perforatum*, and 5 μg was transcribed into cDNA using the SuperScript IV First-Strand Synthesis system (ThermoFisher, Invitrogen). Candidate gene open reading frames were PCR-amplified (Supplemental Table 14) and cloned into various pESC vectors (pESC-LEU for *HpCCL2* and *HpPKS2*, pESC-HIS for *HpPT1* and *HpPT2*, pESC-URA for *HpPT3* and *HpPT4*), pEAQ-HT (for *N. benthamiana* expression), and PA7-YFP (YFP fused to the C terminus of prenyltransferase signal peptides) using the ClonExpress II One Step Cloning Kit (Vazyme). Restriction enzymes and overhangs are detailed in Supplemental Table 15. Following transformation into *Escherichia coli* DH5α (DE3), colonies were selected on antibiotic-supplemented Luria–Bertani (LB) agar. Positive colonies were confirmed through colony PCR using primers listed in Supplemental Table 14, and plasmids were isolated from cultures grown overnight at 37°C for verification by sequencing.

Expression and enzyme activity of prenyltransferases in yeast

The *S. cerevisiae* DD104 strain (engineered for isoprenoid production) (Fischer et al., 2011) was transformed using lithium acetate to introduce combinations of *HpCCL2*, *HpPKS2* and *PT* candidate genes inserted into pESC vectors (Agilent). Three days after transformation, colonies were grown in 5 ml of dextrose dropout medium for 2 days. Following this, cells were harvested and resuspended in 15 ml of dropout medium containing 2% (w/v) D-galactose for 3 days. After induction, cultures were extracted using an equal volume of EtOAc. The organic layer obtained was separated and dried under reduced pressure. The dried extract was subsequently resuspended in 400 μl of MeOH followed by filtration using a 0.22-μm pore-size syringe filter.

Transient expression in *N. benthamiana*

For *Agrobacterium*-mediated transient expression in *N. benthamiana*, pEAQ-HT plasmids (Plant Bioscience, Norwich, UK) (Sainsbury et al., 2009) harboring genes of interest were transformed into *Agrobacterium tumefaciens* GV3101 with selection. PCR-positive colonies were cultured overnight in 15 ml of LB medium with relevant selection. Cells were centrifuged and resuspended in MMA buffer (10 mM MgCl$_2$, 10 mM MES, 150 μM acetosyringone), adjusted to OD$_{600\,nm}$ = 0.2, and kept in darkness for 2 h. Leaves of 4- to 6-week-old *N. benthamiana* were infiltrated with the indicated combinations of each suspension. After 3 days, plants were supplemented with 2 mM valine solution, and 2 days later leaves were collected, flash-frozen, and stored at −80°C. Frozen leaves were then lyophilized and ground into a fine powder before extraction in 10 ml of

methanol. The extraction was concentrated under vacuum, resuspended in 500 µl of methanol, filtered through a 0.22-µm organic syringe filter, and stored at −20°C.

LC–MS method for metabolite analysis

Metabolites from co-expression experiments in yeast and tobacco samples were subjected to analysis using a Q-Exactive system, integrated with a Dionex UltiMate 3000 UHPLC system. Three microliters of each processed sample was injected into a UPLC system (Kinetex 2.6 µm C18 100 Å, measuring 100 × 2.1 mm, Phenomenex). This was carried out at a flow rate of 0.35 ml per minute. The gradient with solvent A (2 mM NH_4Ac) and solvent B (MeCN) was from 0 to 0.5 min with maintained 10% B, followed by a linear increase from 10% to 98% B between 0.5 and 11 min. This was sustained at 98% B from 11 to 13 min. Mass spectra acquisition was performed in the positive mode under the following parameters: spray voltage of 3500 V, capillary temperature of 320°C, sheath gas of 40, and aux gas of 10.

RNA *in situ* hybridization

RNA *in situ* hybridization was conducted using an optimized protocol (Dong et al., 2021). Young *H. perforatum* leaves were fixed in FAA solution (3.7% v/v paraformaldehyde, 5% v/v acetic acid, and 50% v/v ethanol), dehydrated by Leica HistoCore PEARL, and embedded in paraffin by Leica HistoCore Arcadia H. Thin sections (8 µm) were prepared using the HistoCore MULTICUT. The sections underwent toluidine blue staining for morphological examination. Additionally, a schematic representation of the leaf was generated by ggplantmap (Jo and Kajala, 2023) for detailed visualization. Sense and antisense probes were synthesized through T_7-promoter-based *in vitro* transcription, incorporating digoxigenin-labeled uridine triphosphate for subsequent detection. Hybridization was performed overnight at room temperature, allowing for optimal probe–target annealing. Post-hybridization washes were conducted to remove excess probe using a series of saline sodium citrate buffers of decreasing concentration. The hybridized probes were then detected using an anti-digoxigenin antibody conjugated to alkaline phosphatase, with NBT/BCIP (nitro blue tetrazolium/5-bromo-4-chloro-3-indolyl phosphate) as the chromogenic substrate. The specificity of the hybridization signal was confirmed by comparing the results with sense probe controls. The hybridized sections were visualized using an Olympus BX53 microscope.

Subcellular localization of prenyltransferases

The full-length prenyltransferases were fused with YFP at C termini (Guirimand et al., 2020). *A. tumefaciens* transformation and infiltration of *N. benthamiana* leaves with the bacterial suspension were carried out as described above. The lower epidermis of leaf disks of *N. benthamiana* was visualized 3 days after infiltration. The images were captured with a Zeiss LSM880 confocal microscope.

A. thaliana protoplasts were prepared from Columbia-0 ecotype and transformed with PA7-YFP plasmids as described previously (Wu et al., 2022). YFP was fused to the C terminus of predicted prenyltransferase signal peptides alone, giving rise to HpPT1-YFP (comprising amino acids 1–76), HpPT2-YFP (1–75), HpPT3-YFP (1–71), and HpPT4-YFP (1–79). The images were captured with a Zeiss LSM880 confocal microscope.

Chemical synthesis of prenylated intermediates

For translucin I synthesis, PIBP (2.55 mmol) and geraniol (2.55 mmol) were dissolved in anhydrous CH_2Cl_2 (5 ml), followed by the addition of $BF_3 \cdot Et_2O$ (0.25 mmol) at 35°C. The reaction mixture was stirred for 1 h at the same temperature. Post reaction, it was quenched with saturated $NaHCO_3$ solution (10 ml) and extracted with CH_2Cl_2 (3 × 20 ml). The dried organic phases (over Na_2SO_4) were then purified using silica gel column chromatography with EtOAc–hexane to yield translucin I (Dethe and Nirpal, 2021).

For the synthesis of translucin II and III, translucin I (0.2 mmol) was dissolved in H_2O (2 ml) and tetrahydrofuran (1 ml) under nitrogen stream at 0°C. KOH (0.6 mmol) was added, followed by prenyl bromide (0.8 mmol). After stirring for 1 h at 0°C, the reaction was quenched with saturated $NaHCO_3$ solution (10 ml) and extracted with CH_2Cl_2 (3 × 10 ml). The combined organic phases were dried over Na_2SO_4 and purified using silica gel column chromatography with EtOAc–hexane, resulting in translucin II and III (George et al., 2010).

Compound characterization

High-resolution (HR) electrospray ionization MS spectra were measured with a Thermo Q-Exactive Orbitrap mass spectrometer. NMR spectra were acquired using an Agilent DD2 NMR instrument operating at 500 MHz for ^{1}H and a Bruker Advance NMR instrument operating at 600 MHz for ^{1}H. The number of scans depended on the concentration of the sample.

Translucin I

HR-MS (Orbitrap): found for $[M + H]^+$: $[C_{20}H_{29}O_4]^+$ = 333.2057: calcd 333.206585; ^{1}H-NMR (500 MHz, $CDCl_3$): δ 11.70 (br, s, 1H), 9.03 (br, s, 1H), 6.44 (s, 1H), 5.87 (s, 1H), 5.25 (t, J = 7.4 Hz, 1H), 5.05 (t, J = 7.4 Hz, 1H), 3.90 (p, J = 6.8 Hz, 1H), 3.37 (d, J = 7.1 Hz, 2H), 2.13–2.05 (m, 4H), 1.81 (s, 3H), 1.66 (s, 3H), 1.59 (s, 3H) 1.17 (d, J = 6.8 Hz, 6 H).

Translucin II

HR-MS (Orbitrap): found for $[M + H]^+$: $[C_{25}H_{37}O_4]^+$ = 401.2691: calcd 401.269185; ^{1}H-NMR (500 MHz, $CDCl_3$): δ 10.14 (s, 2H), 6.30 (s, 1H), 5.26–5.20 (m, 2H), 5.08–5.02 (m, 1H), 3.91 (p, J = 6.8 Hz, 1H), 3.40 (d, J = 6.8 Hz, 2H), 3.38 (d, J = 6.8 Hz, 2H), 2.30–2.05 (m, 4H), 1.84 (s, 3H), 1.82 (s, 3H), 1.78 (s, 3H), 1.68 (s, 3H), 1.60 (s, 3H), 1.16 (d, J = 6.7 Hz, 6H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 211.1, 159.4, 159.3, 140.4, 136.6, 132.4, 123.7, 123.6, 122.5, 121.9, 121.8, 105.0, 104.5, 39.8, 39.4, 32.1, 26.3, 26.3, 26.0, 25.9, 25.8, 23.6, 22.0, 21.9, 21.7, 19.5, 18.0, 17.9, 16.3.

Translucin III

HR-MS (Orbitrap): found for $[M + H]^+$: $[C_{30}H_{45}O_4]^+$ = 469.3315: calcd 469.331784; ^{1}H-NMR (600 MHz, $CDCl_3$): δ 7.02 (s, 1H), 5.16 (t, J = 7.7 Hz, 1H), 4.98 (br m, 1H), 4.86 (t, J = 7.7 Hz, 1H), 4.81 (t, J = 7.5 Hz, 1H), 4.02 (p, J = 6.8 Hz, 1H), 3.19 (br m, 2H), 2.75–2.64 (m, 4H), 2.66–2.56 (m, 2H), 2.50 (ddd, J = 13.9, 8.1, 8.1 Hz, 2H), 1.79 (s, 3H), 1.77 (s, 3H), 1.63 (s, 3H), 1.58 (s, 3H), 1.56 (s, 3H), 1.55 (s, 3H), 1.54 (s, 3H), 1.23 (dd, J = 6.8, 2.0 Hz, 6H). ^{13}C NMR (151 MHz, $CDCl_3$) δ 207.2, 195.5, 189.9, 172.5, 139.0, 135.3, 135.1, 131.8, 124.2, 121.1, 118.4, 118.2, 109.9, 107.4, 57.4, 40.0, 37.8, 37.6, 35.6, 26.9, 26.1, 26.0, 25.9, 21.2, 19.2, 19.1, 19.0, 18.1, 16.4.

DATA AND CODE AVAILABILITY

Reference genome data and transcriptome sequence reads are deposited in the GenBank under project number PRJNA1060879. The raw sequencing data are deposited in the Sequence Read Archive with the accession numbers listed in Supplemental Table 16. The NCBI accession numbers of the genes reported in this study can be found in Supplemental Table 17.

The whole-genome sequence data reported in this paper have been deposited in the China National Genomics Data Center under accession number GWHEUWF00000000.1 (BioProject: PRJCA028830) that is publicly accessible at https://ngdc.cncb.ac.cn/gwh. The genome annotation file has also been deposited at https://github.com/TatsisLab/ScRNA-seq-pipeline.

All custom code used to analyze the data and generate figures can be found at https://github.com/TatsisLab/ScRNA-seq-pipeline.

ACKNOWLEDGMENTS

AUTHOR CONTRIBUTIONS

S.W. and E.C.T. conducted the study design. S.W., A.L.M.M., and J.Y. performed the research. E.W. provided resources. S.W. and E.C.T. analyzed data and wrote the manuscript with input from the other authors.

SUPPLEMENTAL INFORMATION

Supplemental information is available at *Molecular Plant Online*.

REFERENCES

Adam, P., Arigoni, D., Bacher, A., and Eisenreich, W. (2002). Biosynthesis of hyperforin in Hypericum perforatum. J. Med. Chem. **45**:4786–4793. https://doi.org/10.1021/jm0209782.

Alan, A.R., Murch, S.J., and Saxena, P.K. (2015). Evaluation of ploidy variations in Hypericum perforatum L. (St. John's wort) germplasm from seeds, in vitro germplasm collection, and regenerants from floral cultures. In Vitro Cell Dev. Biol. Plant **51**:452–462. https://doi.org/10.1007/s11627-015-9708-7.

Alioto, T., Blanco, E., Parra, G., and Guigó, R. (2018). Using geneid to Identify Genes. Curr. Protoc. Bioinformatics **64**:e56. https://doi.org/10.1002/cpbi.56.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J. Mol. Biol. **215**:403–410. https://doi.org/10.1016/s0022-2836(05)80360-2.

Andreatta, M., and Carmona, S.J. (2021). UCell: Robust and scalable single-cell gene signature scoring. Comput. Struct. Biotechnol. J. **19**:3796–3798. https://doi.org/10.1016/j.csbj.2021.06.043.

Barcaccia, G., Arzenton, F., Sharbel, T.F., Varotto, S., Parrini, P., and Lucchin, M. (2006). Genetic diversity and reproductive biology in ecotypes of the facultative apomict Hypericum perforatum L. Heredity **96**:322–334. https://doi.org/10.1038/sj.hdy.6800808.

Benson, G. (1999). Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res. **27**:573–580. https://doi.org/10.1093/nar/27.2.573.

Birney, E., Clamp, M., and Durbin, R. (2004). GeneWise and Genomewise. Genome Res. **14**:988–995. https://doi.org/10.1101/gr.1865504.

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics **30**:2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

Burge, C., and Karlin, S. (1997). Prediction of complete gene structures in human genomic DNA. J. Mol. Biol. **268**:78–94. https://doi.org/10.1006/jmbi.1997.0951.

Cervantes-Pérez, S.A., Thibivilliers, S., Laffont, C., Farmer, A.D., Frugier, F., and Libault, M. (2022). Cell-specific pathways recruited for symbiotic nodulation in the Medicago truncatula legume. Mol. Plant **15**:1868–1888. https://doi.org/10.1016/j.molp.2022.10.021.

Chan, P.P., Lin, B.Y., Mak, A.J., and Lowe, T.M. (2021). tRNAscan-SE 2.0: improved detection and functional classification of transfer RNA genes. Nucleic Acids Res. **49**:9077–9096. https://doi.org/10.1093/nar/gkab688.

Charchoglyan, A., Abrahamyan, A., Fujii, I., Boubakir, Z., Gulder, T.A.M., Kutchan, T.M., Vardapetyan, H., Bringmann, G., Ebizuka, Y., and Beerhues, L. (2007). Differential accumulation of hyperforin and secohyperforin in Hypericum perforatum tissue cultures. Phytochemistry **68**:2670–2677. https://doi.org/10.1016/j.phytochem.2007.06.004.

Chatterjee, S.S., Bhattacharya, S.K., Wonnemann, M., Singer, A., and Müller, W.E. (1998). Hyperforin as a possible antidepressant component of hypericum extracts. Life Sci. **63**:499–510. https://doi.org/10.1016/s0024-3205(98)00299-9.

Chen, H., Yin, X., Guo, L., Yao, J., Ding, Y., Xu, X., Liu, L., Zhu, Q.H., Chu, Q., and Fan, L. (2021). PlantscRNAdb: A database for plant single-cell RNA analysis. Mol. Plant **14**:855–857. https://doi.org/10.1016/j.molp.2021.05.002.

Cheng, H., Concepcion, G.T., Feng, X., Zhang, H., and Li, H. (2021). Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. Nat. Methods **18**:170–175. https://doi.org/10.1038/s41592-020-01056-5.

Christianson, D.W. (2017). Structural and Chemical Biology of Terpenoid Cyclases. Chem. Rev. **117**:11570–11648. https://doi.org/10.1021/acs.chemrev.7b00287.

Ciccarelli, D., Andreucci, A.C., and Pagni, A.M. (2001). Translucent Glands and Secretory Canals in Hypericum perforatum L. (Hypericaceae): Morphological, Anatomical and Histochemical Studies During the Course of Ontogenesis. Ann. Bot. **88**:637–644. https://doi.org/10.1006/anbo.2001.1514.

Ciochina, R., and Grossman, R.B. (2006). Polycyclic polyprenylated acylphloroglucinols. Chem. Rev. **106**:3963–3986. https://doi.org/10.1021/cr0500582.

Crow, M., Paul, A., Ballouz, S., Huang, Z.J., and Gillis, J. (2018). Characterizing the replicability of cell types defined by single cell RNA-sequencing data using MetaNeighbor. Nat. Commun. **9**:884. https://doi.org/10.1038/s41467-018-03282-0.

Czechowski, T., Forestier, E., Swamidatta, S.H., Gilday, A.D., Cording, A., Larson, T.R., Harvey, D., Li, Y., He, Z., King, A.J., et al. (2022). Gene discovery and virus-induced gene silencing reveal branched pathways to major classes of bioactive diterpenoids in <i>Euphorbia peplus</i&gt. Proc. Natl. Acad. Sci. USA **119**:e2203890119. https://doi.org/10.1073/pnas.2203890119.

Dang, T.T.T., Franke, J., Carqueijeiro, I.S.T., Langley, C., Courdavault, V., and O'Connor, S.E. (2018). Sarpagan bridge enzyme has substrate-controlled cyclization and aromatization modes. Nat. Chem. Biol. **14**:760–763. https://doi.org/10.1038/s41589-018-0078-4.

Darriba, D., Taboada, G.L., Doallo, R., and Posada, D. (2011). ProtTest 3: fast selection of best-fit models of protein evolution. Bioinformatics **27**:1164–1165. https://doi.org/10.1093/bioinformatics/btr088.

de Bruijn, W.J.C., Levisson, M., Beekwilder, J., van Berkel, W.J.H., and Vincken, J.-P. (2020). Plant aromatic prenyltransferases: Tools for microbial cell factories. Trends Biotechnol. **38**:917–934.

De La Peña, R., Hodgson, H., Liu, J.C.-T., Stephenson, M.J., Martin, A.C., Owen, C., Harkess, A., Leebens-Mack, J., Jimenez, L.E., Osbourn, A., et al. (2023). Complex scaffold remodeling in plant triterpene biosynthesis. Science **379**:361–368. https://doi.org/10.1126/science.adf1017.

Dethe, D.H., and Nirpal, A.K. (2021). Enantiospecific Total Synthesis of (-)-Japonicol C. Org. Lett. **23**:2648–2653. https://doi.org/10.1021/acs.orglett.1c00560.

Dewick, P.M. (2011). Medicinal Natural Products: A Biosynthetic Approach (John Wiley & Sons).

Di Carlo, G., Borrelli, F., Ernst, E., and Izzo, A.A. (2001). St John's wort: Prozac from the plant kingdom. Trends Pharmacol. Sci. **22**:292–297.

Dong, W., Zhu, Y., Chang, H., Wang, C., Yang, J., Shi, J., Gao, J., Yang, W., Lan, L., Wang, Y., et al. (2021). An SHR–SCR module specifies legume cortical cell fate to enable nodulation. Nature **589**:586–590. https://doi.org/10.1038/s41586-020-3016-z.

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. **32**:1792–1797. https://doi.org/10.1093/nar/gkh340.

El Hamdaoui, Y., Zheng, F., Fritz, N., Ye, L., Tran, M.A., Schwickert, K., Schirmeister, T., Braeuning, A., Lichtenstein, D., Hellmich, U.A., et al. (2022). Analysis of hyperforin (St. John's wort) action at TRPC6 channel leads to the development of a new class of antidepressant drugs. Mol. Psychiatry **27**:5070–5085. https://doi.org/10.1038/s41380-022-01804-3.

El-Awaad, I., Bocola, M., Beuerle, T., Liu, B., and Beerhues, L. (2016). Bifunctional CYP81AA proteins catalyse identical hydroxylations but alternative regioselective phenol couplings in plant xanthone biosynthesis. Nat. Commun. **7**:11472. https://doi.org/10.1038/ncomms11472.

Emanuelsson, O., Nielsen, H., and von Heijne, G. (1999). ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. Protein Sci. **8**:978–984. https://doi.org/10.1110/ps.8.5.978.

Fiesel, T., Gaid, M., Müller, A., Bartels, J., El-Awaad, I., Beuerle, T., Ernst, L., Behrends, S., and Beerhues, L. (2015). Molecular Cloning and Characterization of a Xanthone Prenyltransferase from Hypericum calycinum Cell Cultures. Molecules **20**:15616–15630. https://doi.org/10.3390/molecules200915616.

Fischer, M.J.C., Meyer, S., Claudel, P., Bergdoll, M., and Karst, F. (2011). Metabolic engineering of monoterpene synthesis in yeast. Biotechnol. Bioeng. **108**:1883–1892. https://doi.org/10.1002/bit.23129.

Fischer, S., Crow, M., Harris, B.D., and Gillis, J. (2021). Scaling up reproducible research for single-cell transcriptomics using MetaNeighbor. Nat. Protoc. **16**:4031–4067. https://doi.org/10.1038/s41596-021-00575-5.

Fukuyama, Y., Kuwayama, A., and Minami, H. (1997). Garsubellin A, a novel polyprenylated phloroglucin derivative, increasing choline acetyltransferase (ChAT) activity in postnatal rat septal neuron cultures. Chem. Pharm. Bull. **45**:947–949.

Galla, G., Basso, A., Grisan, S., Bellucci, M., Pupilli, F., and Barcaccia, G. (2019). Ovule Gene Expression Analysis in Sexual and Aposporous Apomictic Hypericum perforatum L. (Hypericaceae) Accessions. Front. Plant Sci. **10**:654. https://doi.org/10.3389/fpls.2019.00654.

Gautam, A. (2022). Phenol-Chloroform DNA Isolation Method. In DNA and RNA Isolation Techniques for Non-experts (Springer International Publishing), pp. 33–39. https://doi.org/10.1007/978-3-030-94230-4_3.

George, J.H., Hesse, M.D., Baldwin, J.E., and Adlington, R.M. (2010). Biomimetic synthesis of polycyclic polyprenylated acylphloroglucinol natural products isolated from Hypericum papuanum. Org. Lett. **12**:3532–3535. https://doi.org/10.1021/ol101380a.

Grones, C., Eekhout, T., Shi, D., Neumann, M., Berg, L.S., Ke, Y., Shahan, R., Cox, K.L., Jr., Gomez-Cano, F., Nelissen, H., et al. (2024). Best practices for the execution, analysis, and data storage of plant single-cell/nucleus transcriptomics. Plant Cell **36**:812–828. https://doi.org/10.1093/plcell/koae003.

Guirimand, G., Guihur, A., Perello, C., Phillips, M., Mahroug, S., Oudin, A., Dugé de Bernonville, T., Besseau, S., Lanoue, A., Giglioli-Guivarc'h, N., et al. (2020). Cellular and Subcellular Compartmentation of the 2C-Methyl-D-Erythritol 4-Phosphate Pathway in the Madagascar Periwinkle. Plants **9**:462.

Gurevich, A.I., Dobrynin, V.N., Kolosov, M.N., Popravko, S.A., and Riabova, I.D. (1971). Antibiotic hyperforin from Hypericum perforatum L. Antibiotiki **16**:510–513.

Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W.M., 3rd, Zheng, S., Butler, A., Lee, M.J., Wilk, A.J., Darby, C., Zager, M., et al. (2021). Integrated analysis of multimodal single-cell data. Cell **184**:3573–3587.e29. https://doi.org/10.1016/j.cell.2021.04.048.

He, Z., Luo, Y., Zhou, X., Zhu, T., Lan, Y., and Chen, D. (2024). scPlantDB: a comprehensive database for exploring cell types and markers of plant cell atlases. Nucleic Acids Res. **52**:D1629–D1638. https://doi.org/10.1093/nar/gkad706.

Ianevski, A., Giri, A.K., and Aittokallio, T. (2022). Fully-automated and ultra-fast cell-type identification using specific marker combinations from single-cell transcriptomic data. Nat. Commun. **13**:1246. https://doi.org/10.1038/s41467-022-28803-w.

Istikoglou, C.I., Mavreas, V., and Geroulanos, G. (2010). History and therapeutic properties of Hypericum perforatum from antiquity until today. Psychiatriki **21**:332–338.

Ji, Y., Hong, B., Franzoni, I., Wang, M., Guan, W., Jia, H., and Li, H. (2022). Enantioselective Total Synthesis of Hyperforin and Pyrohyperforin. Angew. Chem. Int. Ed. Engl. **61**:e202116136. https://doi.org/10.1002/anie.202116136.

Jin, J., Lu, P., Xu, Y., Tao, J., Li, Z., Wang, S., Yu, S., Wang, C., Xie, X., Gao, J., et al. (2022). PCMDB: a curated and comprehensive resource of plant cell markers. Nucleic Acids Res. **50**:D1448–D1455. https://doi.org/10.1093/nar/gkab949.

Jo, L., and Kajala, K. (2023). ggPlantmap: an R package for creation of informative and quantitative ggplot maps derived from plant images. Preprint at bioRxiv. https://doi.org/10.1101/2023.11.30.569429.

Kang, M., Choi, Y., Kim, H., and Kim, S.G. (2022). Single-cell RNA-sequencing of Nicotiana attenuata corolla cells reveals the biosynthetic pathway of a floral scent. New Phytol. **234**:527–544. https://doi.org/10.1111/nph.17992.

Kim, D., Paggi, J.M., Park, C., Bennett, C., and Salzberg, S.L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat. Biotechnol. **37**:907–915. https://doi.org/10.1038/s41587-019-0201-4.

Kozlov, A.M., Darriba, D., Flouri, T., Morel, B., and Stamatakis, A. (2019). RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. Bioinformatics **35**:4453–4455. https://doi.org/10.1093/bioinformatics/btz305.

Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., Jones, S.J., and Marra, M.A. (2009). Circos: an information aesthetic for comparative genomics. Genome Res. **19**:1639–1645. https://doi.org/10.1101/gr.092759.109.

Kucharíková, A., Kusari, S., Sezgin, S., Spiteller, M., and Čellárová, E. (2016). Occurrence and Distribution of Phytochemicals in the Leaves of

17 In vitro Cultured Hypericum spp. Adapted to Outdoor Conditions. Front. Plant Sci. **7**:1616. https://doi.org/10.3389/fpls.2016.01616.

Leong, B.J., and Last, R.L. (2017). Promiscuity, impersonation and accommodation: evolution of plant specialized metabolism. Curr. Opin. Struct. Biol. **47**:105–112.

Leuner, K., Kazanski, V., Müller, M., Essin, K., Henke, B., Gollasch, M., Harteneck, C., and Müller, W.E. (2007). Hyperforin–a key constituent of St. John's wort specifically activates TRPC6 channels. Faseb j **21**:4101–4111. https://doi.org/10.1096/fj.07-8110com.

Li, C., Wood, J.C., Vu, A.H., Hamilton, J.P., Rodriguez Lopez, C.E., Payne, R.M.E., Serna Guerrero, D.A., Gase, K., Yamamoto, K., Vaillancourt, B., et al. (2023a). Single-cell multi-omics in the medicinal plant Catharanthus roseus. Nat. Chem. Biol. **19**:1031–1041. https://doi.org/10.1038/s41589-023-01327-0.

Li, H., Ban, Z., Qin, H., Ma, L., King, A.J., and Wang, G. (2015). A heteromeric membrane-bound prenyltransferase complex from hop catalyzes three sequential aromatic prenylations in the bitter acid pathway. Plant Physiol. **167**:650–659. https://doi.org/10.1104/pp.114.253682.

Li, H., Wu, S., Lin, R., Xiao, Y., Malaco Morotti, A.L., Wang, Y., Galilee, M., Qin, H., Huang, T., Zhao, Y., et al. (2023b). The genomes of medicinal skullcaps reveal the polyphyletic origins of clerodane diterpene biosynthesis in the family Lamiaceae. Mol. Plant **16**:549–570. https://doi.org/10.1016/j.molp.2023.01.006.

Liao, R.-Y., and Wang, J.-W. (2023). Analysis of meristems and plant regeneration at single-cell resolution. Curr. Opin. Plant Biol. **74**:102378. https://doi.org/10.1016/j.pbi.2023.102378.

Linde, K., Ramirez, G., Mulrow, C.D., Pauls, A., Weidenhammer, W., and Melchart, D. (1996). St John's wort for depression—an overview and meta-analysis of randomised clinical trials. Bmj **313**:253–258.

Maggi, F., Ferretti, G., Pocceschi, N., Menghini, L., and Ricciutelli, M. (2004). Morphological, histochemical and phytochemical investigation of the genus Hypericum of the Central Italy. Fitoterapia **75**:702–711. https://doi.org/10.1016/j.fitote.2004.09.009.

Matsuda, Y., and Abe, I. (2016). Biosynthesis of fungal meroterpenoids. Nat. Prod. Rep. **33**:26–53. https://doi.org/10.1039/c5np00090d.

Matzk, F., Meister, A., Brutovská, R., and Schubert, I. (2001). Reconstruction of reproductive diversity in Hypericum perforatum L. opens novel strategies to manage apomixis. Plant J. **26**:275–282. https://doi.org/10.1046/j.1365-313x.2001.01026.x.

Mazutis, L., Gilbert, J., Ung, W.L., Weitz, D.A., Griffiths, A.D., and Heyman, J.A. (2013). Single-cell analysis and sorting using droplet-based microfluidics. Nat. Protoc. **8**:870–891.

McGinnis, C.S., Murrow, L.M., and Gartner, Z.J. (2019). DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors. Cell Syst. **8**:329–337.e4. https://doi.org/10.1016/j.cels.2019.03.003.

Méteignier, L.V., Nützmann, H.W., Papon, N., Osbourn, A., and Courdavault, V. (2023). Emerging mechanistic insights into the regulation of specialized metabolism in plants. Nat. Plants **9**:22–30. https://doi.org/10.1038/s41477-022-01288-7.

Müller, W.E., Singer, A., and Wonnemann, M. (2001). Hyperforin–antidepressant activity by a novel mechanism of action. Pharmacopsychiatry **34**:S98–S102. https://doi.org/10.1055/s-2001-15512.

Nagia, M., Gaid, M., Biedermann, E., Fiesel, T., El-Awaad, I., Hänsch, R., Wittstock, U., and Beerhues, L. (2019). Sequential regiospecific gem-diprenylation of tetrahydroxyxanthone by prenyltransferases from Hypericum sp. New Phytol. **222**:318–334. https://doi.org/10.1111/nph.15611.

Nahrstedt, A., and Butterweck, V. (1997). Biologically active and other chemical constituents of the herb of Hypericum perforatum L. Pharmacopsychiatry **30**:129–134.

Nett, R.S., Lau, W., and Sattely, E.S. (2020). Discovery and engineering of colchicine alkaloid biosynthesis. Nature **584**:148–153. https://doi.org/10.1038/s41586-020-2546-8.

Nett, R.S., Dho, Y., Tsai, C., Passow, D., Martinez Grundman, J., Low, Y.-Y., and Sattely, E.S. (2023). Plant carbonic anhydrase-like enzymes in neuroactive alkaloid biosynthesis. Nature **624**:182–191. https://doi.org/10.1038/s41586-023-06716-y.

Ng, Q.X., Venkatanarayanan, N., and Ho, C.Y.X. (2017). Clinical use of Hypericum perforatum (St John's wort) in depression: A meta-analysis. J. Affect. Disord. **210**:211–221. https://doi.org/10.1016/j.jad.2016.12.048.

Nützmann, H.W., Doerr, D., Ramírez-Colmenero, A., Sotelo-Fonseca, J.E., Wegel, E., Di Stefano, M., Wingett, S.W., Fraser, P., Hurst, L., Fernandez-Valverde, S.L., et al. (2020). Active and repressed biosynthetic gene clusters have spatially distinct chromosome states. Proc. Natl. Acad. Sci. USA **117**:13800–13809. https://doi.org/10.1073/pnas.1920474117.

Oike, T., Ogiwara, H., Torikai, K., Nakano, T., Yokota, J., and Kohno, T. (2012). Garcinol, a histone acetyltransferase inhibitor, radiosensitizes cancer cells by inhibiting non-homologous end joining. Int. J. Radiat. Oncol. Biol. Phys. **84**:815–821. https://doi.org/10.1016/j.ijrobp.2012.01.017.

Paysan-Lafosse, T., Blum, M., Chuguransky, S., Grego, T., Pinto, B.L., Salazar, G.A., Bileschi, M.L., Bork, P., Bridge, A., Colwell, L., et al. (2023). InterPro in 2022. Nucleic Acids Res. **51**, D418-d427. https://doi.org/10.1093/nar/gkac993.

Polturak, G., and Osbourn, A. (2021). The emerging role of biosynthetic gene clusters in plant defense and plant interactions. PLoS Pathog. **17**:e1009698. https://doi.org/10.1371/journal.ppat.1009698.

Polturak, G., Dippe, M., Stephenson, M.J., Chandra Misra, R., Owen, C., Ramirez-Gonzalez, R.H., Haidoulis, J.F., Schoonbeek, H.-J., Chartrain, L., Borrill, P., et al. (2022). Pathogen-induced biosynthetic pathways encode defense-related molecules in bread wheat. Proc. Natl. Acad. Sci. USA **119**:e2123299119. https://doi.org/10.1073/pnas.2123299119.

Reed, J., Orme, A., El-Demerdash, A., Owen, C., Martin, L.B.B., Misra, R.C., Kikuchi, S., Rejzek, M., Martin, A.C., Harkess, A., et al. (2023). Elucidation of the pathway for biosynthesis of saponin adjuvants from the soapbark tree. Science **379**:1252–1264. https://doi.org/10.1126/science.adf3727.

Richard, J.A. (2014). Chemistry and biology of the polycyclic polyprenylated acylphloroglucinol hyperforin. European J. Org. Chem. **2014**:273–299.

Rizzo, P., Altschmied, L., Ravindran, B.M., Rutten, T., and D'Auria, J.C. (2020). The Biochemical and Genetic Basis for the Biosynthesis of Bioactive Compounds in Hypericum Perforatum L., One of the Largest Medicinal Crops in Europe. Genes **11**:1210. https://doi.org/10.3390/genes11101210.

Robson, N.K.B. (2003). Hypericum Botany. In Hypericum: The Genus Hypericum, E. Ernst, ed. (CRC Press), pp. 1–22.

Saha, S., Bridges, S., Magbanua, Z.V., and Peterson, D.G. (2008). Empirical comparison of ab initio repeat finding programs. Nucleic Acids Res. **36**:2284–2294. https://doi.org/10.1093/nar/gkn064.

Sainsbury, F., Thuenemann, E.C., and Lomonossoff, G.P. (2009). pEAQ: versatile expression vectors for easy and quick transient expression of heterologous proteins in plants. Plant Biotechnol. J. **7**:682–693. https://doi.org/10.1111/j.1467-7652.2009.00434.x.

Sayed, H.M.B., Nassar, S., Kaufholdt, D., Beerhues, L., Liu, B., and El-Awaad, I. (2023). Biosynthesis of polyprenylated xanthones in Hypericum perforatum roots involves 4-prenyltransferase. Plant Physiol. **192**:2971–2988. https://doi.org/10.1093/plphys/kiad219.

Schempp, C.M., Pelz, K., Wittmer, A., Schöpf, E., and Simon, J.C. (1999). Antibacterial activity of hyperforin from St John's wort, against multiresistant Staphylococcus aureus and gram-positive bacteria. Lancet **353**:2129. https://doi.org/10.1016/s0140-6736(99)00214-7.

Shen, X., Ting, C.P., Xu, G., and Maimone, T.J. (2020). Programmable meroterpene synthesis. Nat. Commun. **11**:508. https://doi.org/10.1038/s41467-020-14354-5.

Shirakawa, M., Ueda, H., Shimada, T., and Hara-Nishimura, I. (2016). FAMA: a molecular link between stomata and myrosin cells. Trends Plant Sci. **21**:861–871.

Skalkos, D., Tatsis, E., Gerothanassis, I.P., and Troganis, A. (2002). Towards a consensus structure of hypericin in solution: direct evidence for a single tautomer and different ionization states in protic and nonprotic solvents by the use of variable temperature gradient 1H NMR. Tetrahedron **58**:4925–4929. https://doi.org/10.1016/s0040-4020(02)00419-2.

Soelberg, J., Jørgensen, L.B., and Jäger, A.K. (2007). Hyperforin accumulates in the translucent glands of Hypericum perforatum. Ann. Bot. **99**:1097–1100. https://doi.org/10.1093/aob/mcm057.

Srivastava, A., Malik, L., Smith, T., Sudbery, I., and Patro, R. (2019). Alevin efficiently estimates accurate gene abundances from dscRNA-seq data. Genome Biol. **20**:65. https://doi.org/10.1186/s13059-019-1670-y.

Stanke, M., Diekhans, M., Baertsch, R., and Haussler, D. (2008). Using native and syntenically mapped cDNA alignments to improve de novo gene finding. Bioinformatics **24**:637–644. https://doi.org/10.1093/bioinformatics/btn013.

Sun, S., Shen, X., Li, Y., Li, Y., Wang, S., Li, R., Zhang, H., Shen, G., Guo, B., Wei, J., et al. (2023). Single-cell RNA sequencing provides a high-resolution roadmap for understanding the multicellular compartment of specialized metabolism. Nat. Plants **9**:179–190. https://doi.org/10.1038/s41477-022-01291-y.

Tang, H., Bowers, J.E., Wang, X., Ming, R., Alam, M., and Paterson, A.H. (2008). Synteny and collinearity in plant genomes. Science **320**:486–488. https://doi.org/10.1126/science.1153917.

Tatsis, E.C., Exarchou, V., Troganis, A.N., and Gerothanassis, I.P. (2008). 1H NMR determination of hypericin and pseudohypericin in complex natural mixtures by the use of strongly deshielded OH groups. Anal. Chim. Acta **607**:219–226. https://doi.org/10.1016/j.aca.2007.11.040.

Tatsis, E.C., Boeren, S., Exarchou, V., Troganis, A.N., Vervoort, J., and Gerothanassis, I.P. (2007). Identification of the major constituents of Hypericum perforatum by LC/SPE/NMR and/or LC/MS. Phytochemistry **68**:383–393. https://doi.org/10.1016/j.phytochem.2006.11.026.

Tatsis, E.C., Carqueijeiro, I., Dugé de Bernonville, T., Franke, J., Dang, T.T.T., Oudin, A., Lanoue, A., Lafontaine, F., Stavrinides, A.K., Clastre, M., et al. (2017). A three enzyme system to generate the Strychnos alkaloid scaffold from a central biosynthetic intermediate. Nat. Commun. **8**:316. https://doi.org/10.1038/s41467-017-00154-x.

Tekel'ová, D., Repcák, M., Zemková, E., and Tóth, J. (2000). Quantitative changes of dianthrones, hyperforin and flavonoids content in the flower ontogenesis of Hypericum perforatum. Planta Med. **66**:778–780.

Tello, M., Kuzuyama, T., Heide, L., Noel, J.P., and Richard, S.B. (2008). The ABBA family of aromatic prenyltransferases: broadening natural product diversity. Cell. Mol. Life Sci. **65**:1459–1463.

Thulasiram, H.V., Erickson, H.K., and Poulter, C.D. (2007). Chimeras of Two Isoprenoid Synthases Catalyze All Four Coupling Reactions in Isoprenoid Biosynthesis. Science **316**:73–76. https://doi.org/10.1126/science.1137786.

Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., and Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat. Biotechnol. **28**:511–515. https://doi.org/10.1038/nbt.1621.

Vickery, C.R., La Clair, J.J., Burkart, M.D., and Noel, J.P. (2016). Harvesting the biosynthetic machineries that cultivate a variety of indispensable plant natural products. Curr. Opin. Chem. Biol. **31**:66–73. https://doi.org/10.1016/j.cbpa.2016.01.008.

Walsh, C.T., and Tang, Y. (2017). Natural Product Biosynthesis (Royal Society of Chemistry).

Wang, Y., Malaco Morotti, A.L., Xiao, Y., Wang, Z., Wu, S., Chen, J., and Tatsis, E.C. (2022a). Decoding the Cytochrome P450 Catalytic Activity in Divergence of Benzophenone and Xanthone Biosynthetic Pathways. ACS Catal. **12**:13630–13637. https://doi.org/10.1021/acscatal.2c03405.

Wang, Z., Xiao, Y., Wu, S., Chen, J., Li, A., and Tatsis, E.C. (2022b). Deciphering and reprogramming the cyclization regioselectivity in bifurcation of indole alkaloid biosynthesis. Chem. Sci. **13**:12389–12395. https://doi.org/10.1039/d2sc03612f.

Weng, J.-K., Lynch, J.H., Matos, J.O., and Dudareva, N. (2021). Adaptive mechanisms of plant specialized metabolism connecting chemistry to function. Nat. Chem. Biol. **17**:1037–1045.

Weng, J.K. (2014). The evolutionary paths towards complexity: a metabolic perspective. New Phytol. **201**:1141–1149. https://doi.org/10.1111/nph.12416.

Wu, S., Malaco Morotti, A.L., Wang, S., Wang, Y., Xu, X., Chen, J., Wang, G., and Tatsis, E.C. (2022). Convergent gene clusters underpin hyperforin biosynthesis in St John's wort. New Phytol. **235**:646–661. https://doi.org/10.1111/nph.18138.

Xu, X., and Jackson, D. (2023). Single-cell analysis opens a goldmine for plant functional studies. Curr. Opin. Biotechnol. **79**:102858. https://doi.org/10.1016/j.copbio.2022.102858.

Xu, Z., and Wang, H. (2007). LTR_FINDER: an efficient tool for the prediction of full-length LTR retrotransposons. Nucleic Acids Res. **35**:W265–W268. https://doi.org/10.1093/nar/gkm286.

Yang, X.W., Grossman, R.B., and Xu, G. (2018). Research Progress of Polycyclic Polyprenylated Acylphloroglucinols. Chem. Rev. **118**:3508–3558. https://doi.org/10.1021/acs.chemrev.7b00551.

Zappia, L., and Oshlack, A. (2018). Clustering trees: a visualization for evaluating clusterings at multiple resolutions. GigaScience **7**:giy083. https://doi.org/10.1093/gigascience/giy083.

Zeng, L., Zhang, N., Zhang, Q., Endress, P.K., Huang, J., and Ma, H. (2017). Resolution of deep eudicot phylogeny and their temporal diversification using nuclear genes from transcriptomic and genomic datasets. New Phytol. **214**:1338–1354. https://doi.org/10.1111/nph.14503.

Zhan, X., Qiu, T., Zhang, H., Hou, K., Liang, X., Chen, C., Wang, Z., Wu, Q., Wang, X., Li, X.L., et al. (2023). Mass spectrometry imaging and single-cell transcriptional profiling reveal the tissue-specific regulation of bioactive ingredient biosynthesis in Taxus leaves. Plant Commun. **4**:100630. https://doi.org/10.1016/j.xplc.2023.100630.

Zhao, Q., Yang, J., Cui, M.Y., Liu, J., Fang, Y., Yan, M., Qiu, W., Shang, H., Xu, Z., Yidiresi, R., et al. (2019). The Reference Genome Sequence of Scutellaria baicalensis Provides Insights into the Evolution of Wogonin Biosynthesis. Mol. Plant **12**:935–950. https://doi.org/10.1016/j.molp.2019.04.002.

Zheng, G.X.Y., Terry, J.M., Belgrader, P., Ryvkin, P., Bent, Z.W., Wilson, R., Ziraldo, S.B., Wheeler, T.D., McDermott, G.P., Zhu, J., et al. (2017). Massively parallel digital transcriptional profiling of single cells. Nat. Commun. **8**:14049. https://doi.org/10.1038/ncomms14049.

Zhou, W., Wang, Y., Li, B., Petijová, L., Hu, S., Zhang, Q., Niu, J., Wang, D., Wang, S., Dong, Y., et al. (2021). Whole-genome sequence data of Hypericum perforatum and functional characterization of melatonin biosynthesis by N-acetylserotonin O-methyltransferase. J. Pineal Res. **70**:e12709. https://doi.org/10.1111/jpi.12709.