Improved production of the antidiabetic metabolite montbretin A in Nicotiana benthamiana: discovery, characterization, and use of Crocosmia shikimate shunt genes

Krise, L.H.; Sunstrum, F.G.; Garcia, D.; López Pérez, G.; Jancsik, S.; Bohlmann, J.; Irmisch, S.

the plant journal

SEB SOCIETY FOR EXPERIMENTAL BIOLOGY

The Plant Journal (2024) **117**, 766–785 doi: 10.1111/tpj.16528

Improved production of the antidiabetic metabolite montbretin A in *Nicotiana benthamiana*: discovery, characterization, and use of *Crocosmia* shikimate shunt genes

Lars H. Kruse[1], Frederick G. Sunstrum[1], Daniela Garcia[1], Guillermo López Pérez[1], Sharon Jancsik[1], Joerg Bohlmann[1,2,3,*] and Sandra Irmisch[1,4,*]

[1]*Michael Smith Laboratories, University of British Columbia, Vancouver, British Columbia V6T 1Z4, Canada,*
[2]*Department of Botany, University of British Columbia, Vancouver, British Columbia V6T 1Z4, Canada,*
[3]*Department of Forest and Conservation Science, University of British Columbia, Vancouver, British Columbia V6T 1Z4, Canada, and*
[4]*Plant Sciences, Institute of Biology, Leiden University, Leiden 2333 BE, Netherlands*

Received 26 July 2023; revised 20 September 2023; accepted 23 October 2023; published online 14 November 2023.
*For correspondence (e-mail bohlmann@msl.ubc.ca; s.irmisch@biology.leidenuniv.nl).

SUMMARY

The plant-specialized metabolite montbretin A (MbA) is being developed as a new treatment option for type-2 diabetes, which is among the ten leading causes of premature death and disability worldwide. MbA is a complex acylated flavonoid glycoside produced in small amounts in below-ground organs of the perennial plant Montbretia (*Crocosmia* × *crocosmiiflora*). The lack of a scalable production system limits the development and potential application of MbA as a pharmaceutical or nutraceutical. Previous efforts to reconstruct montbretin biosynthesis in *Nicotiana benthamiana* (Nb) resulted in low yields of MbA and higher levels of montbretin B (MbB) and montbretin C (MbC). MbA, MbB, and MbC are nearly identical metabolites differing only in their acyl moieties, derived from caffeoyl-CoA, coumaroyl-CoA, and feruloyl-CoA, respectively. In contrast to MbA, MbB and MbC are not pharmaceutically active. To utilize the montbretia caffeoyl-CoA biosynthesis for improved MbA engineering in Nb, we cloned and characterized enzymes of the shikimate shunt of the general phenylpropanoid pathway, specifically hydroxycinnamoyl-CoA: shikimate hydroxycinnamoyl transferase (CcHCT), *p*-coumaroylshikimate 3′-hydroxylase (CcC3′H), and caffeoylshikimate esterase (CcCSE). Gene expression patterns suggest that CcCSE enables the predominant formation of MbA, relative to MbB and MbC, in montbretia. This observation is supported by results from *in vitro* characterization of CcCSE and reconstruction of the shikimate shunt in yeast. Using CcHCT together with montbretin biosynthetic genes in multigene constructs resulted in a 30-fold increase of MbA in Nb. This work advances our understanding of the phenylpropanoid pathway and features a critical step towards improved MbA production in bioengineered Nb.

Keywords: montbretia, *Crocosmia* × *crocosmiiflora*, type-2 diabetes, obesity, bioengineering, flavonoid biosynthesis, phenylpropanoid pathway, hydroxycinnamoyl-CoA: Shikimate hydroxycinnamoyl transferase (HCT), caffeoyl shikimate esterase (CSE), *p*-coumaroylshikimate 3′-hydroxylase (C3′H).

INTRODUCTION

Type-2 diabetes (T2D) is a growing epidemic and one of the leading causes of death and disability worldwide (The Lancet, 2023). The disease affects over half a billion people globally, with healthcare expenditures exceeding US$900 billion in 2021. The number of people affected by T2D is expected to rise to 1.3 billion by 2050, making it 'a defining disease of the 21st century' (Ong et al., 2023). T2D is often linked to obesity or being overweight, unhealthy lifestyle conditions, and food insecurity across low-, middle- and high-income societies (Diabetes Canada, 2020; FAO et al., 2021). The growing case numbers of T2D, as well as side effects caused by current T2D medications, require new or improved treatment options, including easily accessible pharmaceuticals and nutraceuticals, to control blood glucose levels in T2D patients and prediabetics. The

plant-specialized metabolite montbretin A (MbA) has been approved for Phase-1 clinical trials as a T2D treatment, after having previously passed efficacy and toxicity studies in animals (Yuen et al., 2016). MbA lowers the glycemic index of starch-containing food by highly efficient (K_i of 8 nM) and specific inhibition of the human pancreatic α-amylase (HPA), a key enzyme in starch digestion (Andersen et al., 2009; Asada et al., 1988; Tarling et al., 2008; Williams et al., 2015).

MbA is a complex acylated flavonoid glycoside (Figure 1a,b), which according to current phytochemical and taxonomic knowledge is unique to plants of the genus Montbretia (*Crocosmia*; family Iridaceae) that are native to eight different species in Southern and Eastern Africa (Royal Botanical Garden, Kew, 2023). The ornamental montbretia hybrid *C.* × *crocosmiiflora* is widely grown around the world as a perennial garden plant. *C.* × *crocosmiiflora* produces MbA in small amounts in bulb-like below-ground organs, called corms, during a short time of the growing season (Irmisch et al., 2018). The limited taxonomic occurrence of MbA, its low abundance in montbretia corms, and its lack of scalable horticultural production make it unfeasible to obtain sufficient amounts of the metabolite from its natural source. Due to its complex chemical structure (Figure 1b), the MbA supply problem also cannot be resolved by chemical or chemo-enzymatic synthesis. However, synthetic biology and bioengineering of a heterologous host have the potential for the development of a scalable production system for MbA. To this end, all of the required montbretia MbA biosynthetic genes and enzymes, as well as several precursor pathway genes, have been cloned and characterized, and proof of concept of metabolically engineered MbA production in *Nicotiana benthamiana* (Nb) has been established (Irmisch et al., 2018, 2020; Irmisch, Jancsik, et al., 2019; Irmisch, Ruebsam, et al., 2019). MbA is assembled from seven building blocks: the flavonoid myricetin, two units of UDP-glucose (UDP-Glc), two units of UDP-rhamnose (UDP-Rha), one UDP-xylose (UDP-Xyl), and the CoA-activated hydroxycinnamic acid caffeoyl-CoA (Irmisch et al., 2020). The assembly requires the activity of five UDP-dependent glycosyltransferases (CcUGT1-5) and a BAHD acyltransferase (CcAT1 or CcAT2) (Figure 1a).

Nb is widely used as a heterologous transient expression system for the production of a variety of different proteins and discovery of biosynthetic pathways, due to the ease of *Agrobacterium*-mediated expression of single or multiple genes (Bally et al., 2018). In the plant biotechnology industry, Nb is used for the production of therapeutic and diagnostic proteins, enzymes, vaccines, and antibodies (Schillberg & Finnern, 2021; Tschofen et al., 2016). Biosynthetic systems of several plant-specialized metabolites, including terpenoids, alkaloids, and flavonoids from a diverse range of plant species, have been elucidated and reconstructed using heterologous transient gene expression in Nb; and this system is also being explored for scalable of production of high-value plant metabolites (Calgaro-Kozina et al., 2020; De La Peña et al., 2023; Dudley et al., 2022; Irmisch et al., 2020; Kim et al., 2022; Lau & Sattely, 2015; Nett & Sattely, 2021; Reed & Osbourn, 2018; Schultz et al., 2019). However, one major challenge inherent to the bioengineering and heterologous production of specific specialized metabolites in Nb is the occurrence of unwanted side-products due to endogenous Nb enzyme activities (Dudley et al., 2022; Irmisch et al., 2020). Another common problem is the lack, or low abundance, of specific metabolic precursors in Nb, which may result in low yields of the target product or biosynthesis of non-target metabolites (De La Peña & Sattely, 2021; Irmisch, Ruebsam, et al., 2019; Kim et al., 2022; Reed & Osbourn, 2018).

In previous work, we successfully reconstructed MbA biosynthesis in Nb by transiently expressing the six montbretia MbA assembly genes (CcUGT1-5 and CcAT1), as well as three additional montbretia genes to enhance myricetin biosynthesis, specifically a MYB transcription factor (CcMYB4), flavonol synthase (CcFLS), and flavonol 3′5′-hydroxylase (CcCYP2) (Irmisch et al., 2020; Irmisch, Ruebsam, et al., 2019). However, levels of MbA were low, and two related compounds, montbretin B (MbB) and montbretin C (MbC), were produced in higher concentrations than MbA (Irmisch et al., 2020). MbA, MbB, and MbC are nearly identical, but they differ in their acyl moieties, which are derived from caffeoyl-CoA, coumaroyl-CoA, and feruloyl-CoA, respectively (Figure 1b). In contrast to MbA, MbB, and MbC are not effective as HPA inhibitors (K_i of 3600 and 6100 nM, respectively), highlighting the importance of the caffeoyl meta-hydroxy group (Tarling et al., 2008; Williams et al., 2015). Formation of MbA, MbB, and MbC by the same set of montbretin pathway assembly enzymes is likely due to lack of substrate specificity of CcAT1 and CcAT2, which accept caffeoyl-CoA, coumaroyl-CoA, and feruloyl-CoA (Irmisch et al., 2018). It is not known what mechanisms regulate the preferential accumulation of MbA, relative to MbB and MbC, in montbretia corms, and how the same might be accomplished in Nb. We hypothesize that differences in the availability of caffeoyl-CoA in montbretia corms and Nb leaves may be a factor influencing the different montbretin profiles in the native and heterologous system.

Caffeoyl-CoA, coumaroyl-CoA, and feruloyl-CoA are all derived from the general phenylpropanoid pathway (Figure 1c), which plays a central role in plant growth and development and is tightly regulated by a complex network of transcription factors and signaling pathways (Chezem & Clay, 2016; de Vries, Fürst-Jansen, et al., 2021; Douglas et al., 1992; Renault et al., 2019; Weng & Chapple, 2010; Xie et al., 2018; Zhong & Ye, 2009). The phenylpropanoid pathway begins with phenylalanine ammonia-lyase (PAL) converting phenylalanine to cinnamic acid, followed

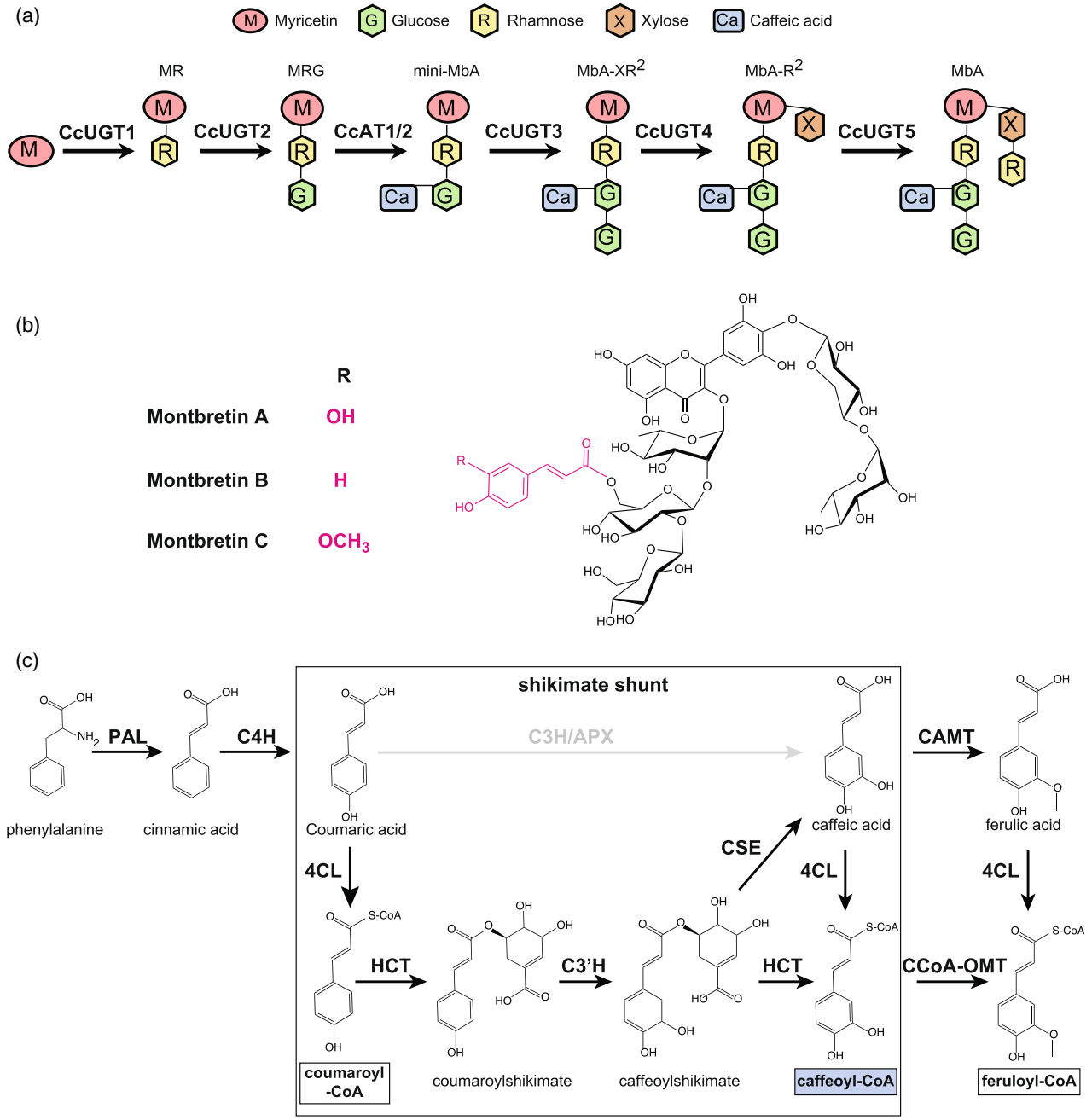

Figure 1. Biosynthesis of montbretin A and hydroxycinnamic acid-CoA esters. (a) Schematic of the six-step assembly pathway of montbretin A (MbA) in montbretia (Irmisch et al., 2018, 2020; Irmisch, Jancsik, et al., 2019; Irmisch, Ruebsam, et al., 2019). MbA is assembled from seven different building blocks, myricetin, two units of UDP-glucose, two units of UDP-rhamnose, UDP-xylose, and caffeoyl-CoA. The assembly pathway involves five different UDP-glycosyltransferases (CcUGT1-5) and a BAHD acyltransferase (CcAT1 or CcAT2). The different building blocks of the MbA molecule are shown as colored symbols. (b) Structure of montbretin A (MbA), montbretin B (MbB), and montbretin C (MbC). The different acyl moieties are highlighted in purple and are derived from either caffeoyl-CoA (MbA), coumaroyl-CoA (MbB), or feruloyl-CoA (MbC). (c) Illustration of the general phenylpropanoid pathway from phenylalanine to the hydroxycinnamic acid-CoA esters, caffeoyl-CoA, coumaroyl-CoA, and feruloyl-CoA, with focus on the shikimate shunt shown with a box. The shikimate shunt may not be conserved across all land plants, and additional enzymes may exist in different plant species. Arrows depict one possible direction of pathway flux. PAL, phenylalanine ammonia-lyase; C4H, cinnamate 4-hydroxylase; 4CL, p-coumaroyl ligase; HCT, hydroxycinnamoyl-CoA:shikimate hydroxycinnamoyltransferases; C3H, *p*-coumarate 3-hydroxylase; C3′H, *p*-coumaroylshikimate 3′-hydroxylase; CSE, caffeoylshikimate esterase; CAMT, caffeic acid O-methyltransferase; CCoA-OMT, caffeoyl-CoA O-methyltransferase; C3H/APX, bifunctional p-coumarate 3-hydroxylase.

by cinnamic acid 4-hydroxylase (C4H) to yield coumaric acid (Fraser & Chapple, 2011; Kriegshauser et al., 2021). Activation of coumaric acid to coumaroyl-CoA is catalyzed by 4-coumaroyl ligase (4CL) (Costa et al., 2005). It is important to note that coumaroyl-CoA provides the acyl moiety in MbB, and is also the precursor for the biosynthesis of

myricetin, the central flavonol building block in MbA, MbB, and MbC (Irmisch, Ruebsam, et al., 2019). Caffeoyl-CoA and feruloyl-CoA are produced from coumaroyl-CoA via the formation of shikimate esters, known as the shikimate shunt (Hoffmann et al., 2003, 2005; Moghe et al., 2023). Esterification of coumaroyl-CoA with shikimate by hydroxycinnamoyl-CoA: shikimate hydroxycinnamoyl transferases (HCT), gives rise to coumaroylshikimate (Hoffmann et al., 2003). Coumaroylshikimate can be hydroxylated by *p*-coumaroylshikimate 3′-hydroxylase (C3′H) to produce caffeoylshikimate (Alber et al., 2019; Schoch et al., 2001). In some plant species, HCT may catalyze the reaction of caffeoylshikimate to caffeoyl-CoA and free shikimate (Hoffmann et al., 2005; Vanholme et al., 2013). Alternatively, caffeoylshikimate esterase (CSE) catalyzes the hydrolysis of caffeoylshikimate into shikimate and caffeic acid, the latter can then be activated by 4CL to caffeoyl-CoA (Vanholme et al., 2013). Caffeoyl-CoA can be methylated by caffeoyl-CoA *O*-methyltransferase to produce feruloyl-CoA (CCoA-OMT) (Zhong et al., 1998). Routes that bypass the shikimate shunt involve direct conversion of coumaric acid into caffeic acid. For example, in *Brachypodium distachyon* and *Arabidopsis thaliana* a bifunctional ascorbate peroxidase, coumarate 3-hydroxylase (C3H)/APX, can hydroxylate coumaric acid (Barros et al., 2019). In poplar (*Populus trichocarpa*), a C4H/C3H complex has been reported for this function (Chen et al., 2011), but has recently been put in question (Gou et al., 2018). Additionally, older literature suggested that phenolases may oxidize coumaric acid to caffeic acid (Duke & Vaughn, 1982; Kojima & Takeuchi, 1989; Stafford & Dresler, 1972). Since 4CL enzymes are involved in the formation of caffeoyl-CoA, coumaroyl-CoA, and feruloyl-CoA (Figure 1b), we recently cloned and characterized montbretia 4CLs. We tested the two identified Cc4CLs for a possible preference for caffeic acid, relative to coumaric acid, as a possible mechanism for higher levels of MbA production in montbretia. However, both Cc4CLs preferred coumaric acid as their substrate and did not enhance MbA levels when co-expressed with the MbA assembly pathway in Nb (Sunstrum et al., 2021).

Here, we further investigate the shikimate-shunt of the phenylpropanoid pathway in montbretia for the biosynthesis of the different acyl-CoA precursors for MbA, MbB and MbC biosynthesis. We describe the discovery of CcHCT (AT8), CcC3′H, and CcCSE (EST2), the *in planta* characterization, the *in vitro* characterization of CcHCT (AT8) and CcCSE (EST2), as well as the reconstruction of the entire shikimate shunt pathway in yeast. We used both engineered yeast and Nb to shed light on the role of this pathway in montbretia. Heterologous expression of montbretia phenylpropanoid pathway enzymes resulted in increased MbA levels in Nb. This work highlights the importance of the phenylpropanoid pathway for MbA bioengineering in Nb and the metabolic plasticity of this host for production of rare and valuable specialized metabolites.

RESULTS

Identification of montbretia HCT candidates

To identify montbretia enzymes involved in the biosynthesis of caffeic acid and caffeoyl-CoA, we used BLASTP searches against a translated transcriptome database that covers a time course of corm development, previously described in Irmisch et al. (2020). Known plant HCT sequences typically belong to Clade 5 of the BAHD acyltransferase (AT) family (Kruse et al., 2022, 2023; Moghe et al., 2023). A search of the montbretia transcriptome using *A. thaliana* HCT (NP_199704.1) revealed three candidate genes, for which we cloned the coding sequences (CDS) from young corms. We named these acyltransferases *CcAT8*, *CcAT9*, and *CcAT10*, prior to testing them for HCT function. The three translated amino acid sequences showed the typical BAHD motifs, HxxxD and DFGWG (Figure S1), and fell into Clade 5 of the BAHD-AT family (Figure 2a; Figure S2, File S1). CcAT8 grouped with known monocot HCTs within the Sub-Clade 5a/b, while CcAT9 and CcAT10 appear on the margins of this subclade (Figure 2a, Figure S2). The latter two grouped together with OsHCT4 (XM_015786263.1) and PvHCT1-like (AFY17066.1) involved in the formation of glycerolesters and chlorogenic acid biosynthesis, respectively (Escamilla-Treviño et al., 2014; Kim et al., 2012).

Previously characterized genes involved in the MbA assembly pathway, *CcUGT1-5* and *CcAT1-2*, are co-expressed with the highest transcript levels during the early seasonal development of young corms coinciding with MbA accumulation (Irmisch et al., 2018, 2020). Analysis of the transcriptome data that cover the six different time points of corm development (Irmisch et al., 2020) did not reveal a strong correlation (>0.8) of expression patterns of *CcAT8*, *CcAT9* or *CcAT10* with the MbA assembly genes (Figure S3, Table S1).

Identification of montbretia C3′H candidates

A BLASTP search of the montbretia transcriptome for putative C3′H sequences using the C3H (CYP98A3) sequence from *A. thaliana* (NP_850337.1, CYP98A3) as bait identified a single montbretia transcript, *CcC3′H* (CYP98A216), for which we cloned the full-length CDS. The translated amino acid sequence shared 75% identity with *A. thaliana* CYP98A3 (Figure S4). The time course transcript expression pattern of *CcC3′H* did not strongly correlate (>0.8) with the expression of previously characterized MbA biosynthetic genes (Figure S3, Table S1).

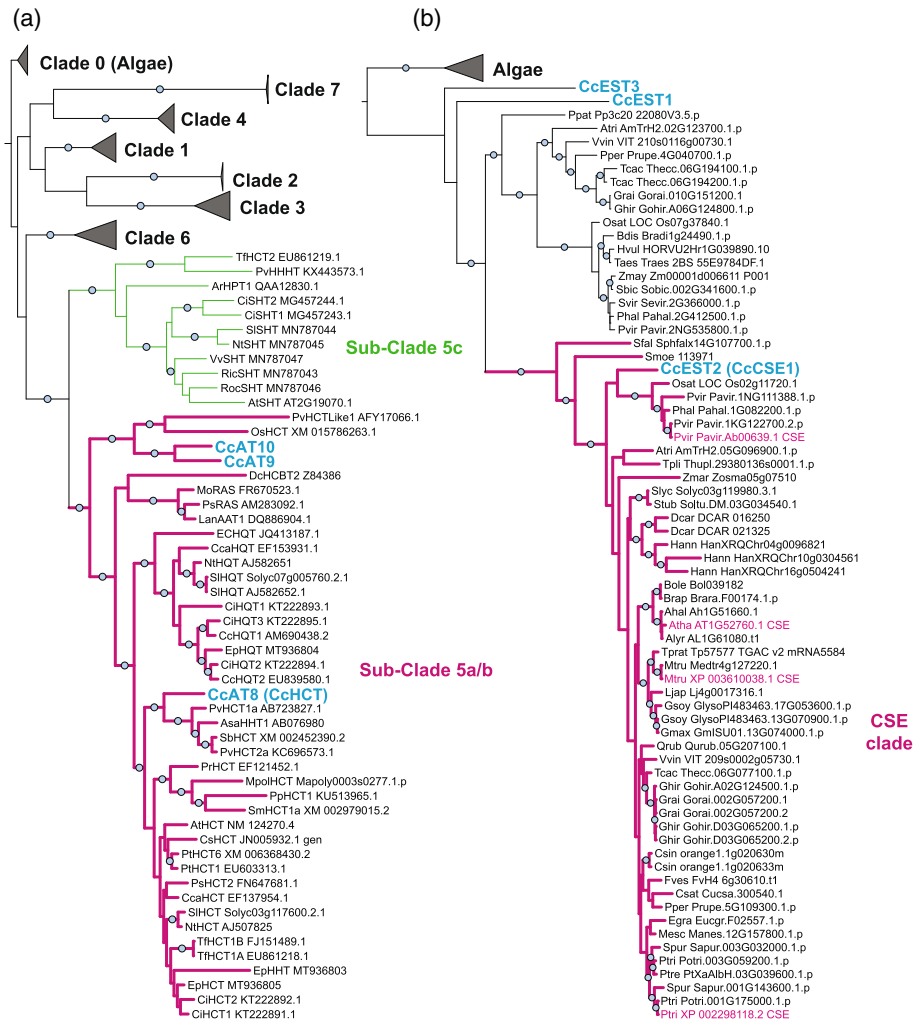

Figure 2. Position of montbretia candidate HCTs and CSEs within the phylogeny of the plant BAHD acyltransferase enzyme family (a) and plant esterases of the alpha/beta hydrolase enzyme family (b), respectively. (a) Phylogenetic tree of characterized BAHD acyltransferases (Transferase, Pfam: PF02458) (Kruse et al., 2022) including the montbretia HCT candidates, CcAT8, CcAT9 and CcAT10. The tree was rooted using the algal clade (Clade 0). Clade 5, with sub-Clade 5 a/b containing characterized HCTs, is shown in full. Sub-Clade 5c contains enzymes involved in phenolamide biosynthesis. Other clades are collapsed (see Figure S2 for the expanded version of the tree). (b) Phylogenetic tree of characterized and uncharacterized plant esterases of the alpha/beta hydrolase family (Hydrolase_4, Pfam: PF12146). Montbretia candidates and sequences from other plant species were identified using a BLASTP search with the Arabidopsis thaliana caffeoylshikimate esterases (AtCSE, AT1G52760.1) from the Uniprot database. Four previously functionally characterized CSEs from other species are shown with names in purple. The algae sequence clade was used for rooting the tree and is shown collapsed (see Figure S2 for the expanded version of the tree). The trees in (a) and (b) were calculated using IQ-TREE v1.6.10 and illustrated in ITOL v5.6.2 (Letunic & Bork, 2021). Light blue circles on branches indicate bootstrap support values >70. Names highlighted in blue represent the montbretia genes identified in this study. The clades of interest are shown with branches drawn as thicker lines in purple. The sequences used to calculate the trees are given in File S1 (BAHDs) and File S2 (ESTs). BAHD sequences are from Kruse et al., 2022 and further description can be found there. Species abbreviations for (b): Ahal, *Arabidopsis halleri*; Alyr, *Arabidopsis lyrata*, Atha, *Arabidopsis thaliana*; Atri, *Amborella trichopoda*; Bdis, *Brachypodium distachyon*; Bole, *Brassica oleracea*; Brap, *Brassica rapa*; Cbra, *Chara braunii*; Csat, *Cucumis sativus*; Csin, *Citrus sinensis*; Dcar, *Daucus carota*; Egra, *Eucalyptus grandis*; Fves, *Fragaria vesca*; Ghir, *Gossypium hirsutum*; Gmax, *Glycine max*; Grai, *Gossypium raimondii*; Gsoy, *Glycine soja*; Hann, *Helianthus annuus*; Hvul, *Hordeum vulgare*; Ljap, *Ionicera japonica*; Mesc, *Manihot esculenta*; Mtru, *Medicago truncatula*; Osat, *Oryza sativa*; Phal, *Panicum hallii*; Ppat, *Physcomitrium patens*; Pper, *Prunus persica*; Ptre, *Populus tremuloides*; Ptri, *Populus trichocarpa*; Pvir, *Panicum virgatum*; Qrub, *Quercus rubra*; Sbic, *Sorghum bicolor*; Sfal, *Sphagnum fallax*; Slyc, *Solanum lycopersicum*; Smoe, *Selaginella moellendorffii*; Spur, *Salix purpurea*; Stub, *Solanum tuberosum*; Svir, *Setaria viridis*; Tcac, *Theobroma cacao*; Teas, *Triticum aestivum*; Tpli, *Thuja plicata*; Tprat, *Trifolium pratense*; Vvin, *Vitis vinifera*; Zmar, *Zostera marina*; Zmay, *Zea mays*.

Identification of montbretia CSE candidates

A BLASTP search of the montbretia transcriptome for putative CSE sequences using the *A. thaliana* CSE (AtCSE, NP_175685.1) as query revealed three montbretia esterase-like transcripts, *CcEST1*, *CcEST2* and *CcEST3*, which were cloned as full-length CDS. Their translated amino acid sequences shared 37%, 67%, and 33% identity with AtCSE, respectively (Figure S5). A phylogeny with characterized CSEs from other

species and uncharacterized homologs from the UniProt database (The UniProt Consortium, 2023) placed CcEST2 into a large clade that contained functionally characterized CSEs from *A. thaliana* (AtCSE, AT1G52760.1), switchgrass (*Panicum virgatum*; PvCSE, Pavir.Ab00639.1), poplar (PtCSE, XP_002298118.2), and barrel clover (*Medicago truncatula*; MtCSE, XP_003610038.1) (Figure 2b; Figure S2b, File S2). In contrast, CcEST1 and CcEST3 clustered outside of the CSE-containing clade, suggesting they encode esterases that are functionally different from CSEs. All three transcripts showed strong co-expression (correlation >0.8) with MbA biosynthetic genes during the time course of corm development (Figure S3, Table S1).

Expression of individual montbretia phenylpropanoid pathway genes increases MbA formation in *N. benthamiana*

To test if MbA yields in Nb could be affected by the expression of *CcAT8-10*, *CcEST1-3*, or *CcC3′H*, we transiently co-expressed one of each of these seven genes (hereafter referred to as phenylpropanoid genes) together with a set of ten genes previously described as required for heterologous MbA production in Nb (Irmisch et al., 2020). These ten genes are the six montbretia genes, *CcUGT1-5* and *CcAT1*, required for the assembly of MbA from its different building blocks (Figure 1); a set of three montbretia genes to enhance myricetin biosynthesis, *CcMYB4*, *CcCYP2*, and *CcFLS*; and the p19 silencer of repression. Hereafter, we will refer to these ten genes as the '10-gene set', and use this term in the description of co-expression experiments with the montbretia candidate phenylpropanoid pathway genes described here.

CcAT8-10, *CcEST1-3*, and *CcC3′H* were (Irmisch et al., 2018; Tarling et al., 2008) individually cloned into pCAMBIA230035Su and plasmids were individually transformed into *Agrobacterium tumefaciens*. All genes were expressed in Nb under the control of the *35S* promoter. For each experiment, Nb leaves were co-infiltrated with the ten different *Agrobacterium* strains comprising the 10-gene set and one strain carrying one of the putative phenylpropanoid pathway genes. Experiments were performed with four biological replicates. As a control, leaves were infiltrated only with the 10-gene set. Leaf extracts were analyzed by liquid chromatography-mass spectrometry (LC–MS), and MbA, MbB, and MbC were quantified (Figure S6a). Controls, transformed with the 10-gene set only, yielded approximately 5 μg MbA × g^{-1} leaf fresh weight (FW), which accounted for 1.3% of the total of the three different montbretins (MbA plus MbB plus MbC) (Figure S6b,c), confirming previous results (Irmisch et al., 2020). In comparison, in montbretia corms, it was previously reported that MbA is the most abundant montbretin (Irmisch et al., 2018; Tarling et al., 2008) and makes up approximately 80% of the produced montbretins (File S3). Of the seven different putative phenylpropanoid pathway genes tested here, co-expression with *CcAT8* and *CcEST2*, resulted in significant increases of MbA compared to the control (Figure S6b). Co-expression with *CcAT8* resulted in a fivefold increase of MbA (*P*-value ≤0.001, 29 μg MbA × g^{-1} FW, 7.4% of total montbretins); and co-expression with *CcEST2* resulted in an eightfold increase of MbA (*P*-value ≤0.001, 45 μg MbA × g^{-1} FW, 7% of total montbretins). Co-expression with *CcC3′H* increased MbA levels slightly, but not significantly (*P*-value = 0.13, 15 μg MbA × g^{-1} FW, 1.8% total montbretins, Figure S6b). Co-expression with all other genes did not yield changes in MbA levels. The increase of MbA levels effected by *CcAT8* or *CcEST2* co-expression was confirmed with a second independent set of experiments, also performed with four replicates each (Figure 3). Here, control plants produced 4 μg MbA × g^{-1} FW (1.2% of total montbretins). Co-expression with *CcAT8* resulted in a 15.2-fold (*P*-value = 0.018) increase of MbA yielding 64 μg MbA × g^{-1} FW (13.5% of total montbretins) (Figure 3a,b). Co-expression with *CcEST2* resulted in a 9.4-fold (*P*-value ≤0.001) increase of MbA yielding 40 μg MbA × g^{-1} FW (16% of total montbretins) (Figure 3a,b).

In addition to MbA, MbC levels also increased significantly by threefold (*P*-value ≤0.001) in co-expression experiments with *CcAT8*, resulting in 185 μg MbC × g^{-1} FW in comparison to 60 μg MbC × g^{-1} FW in the controls. MbC levels also increased (but not significant) in comparison to the control in co-expression experiments with *CcEST2* to 127 μg MbC × g^{-1} FW. MbB levels decreased significantly by 0.57-fold (*P*-value = 0.011) to 150 μg MbB × g^{-1} FW and 0.51-fold (*P*-value = 0.005) to 134 μg MbB × g^{-1} FW, in co-expression experiments with *CcAT8* and *CcEST2*, respectively, in comparison to ~260 μg MbB × g^{-1} FW in controls. Co-expression of *CcC3′H* resulted in a significant twofold increase of MbB in comparison to the control (*P*-value = 0.0047).

Beyond the three montbretins, we also attempted to measure intermediates within the shikimate shunt (Figure 1c) as well as chlorogenic acid (i.e. caffeoylquinate), an important sink for caffeic acid in some plant species (Cardenas et al., 2021; Niggeweg et al., 2004). We did not observe significant changes in the levels of caffeic acid and chlorogenic acid relative to the controls (Figure S7a,c). Levels of caffeoylshikimate significantly increased 34.4-fold (*P*-value <0.001) as an effect of *CcAT8* co-expression (Figure S7b). Efforts to measure coumaroylshikimate and coumaroyl/caffeoyl-CoA were not successful.

Expression of different combinations of montbretia phenylpropanoid pathway genes has no obvious additive effect on MbA formation in *N. benthamiana*

In parallel with testing individual genes in the set of co-expression experiments shown in Figure S6, we also tested

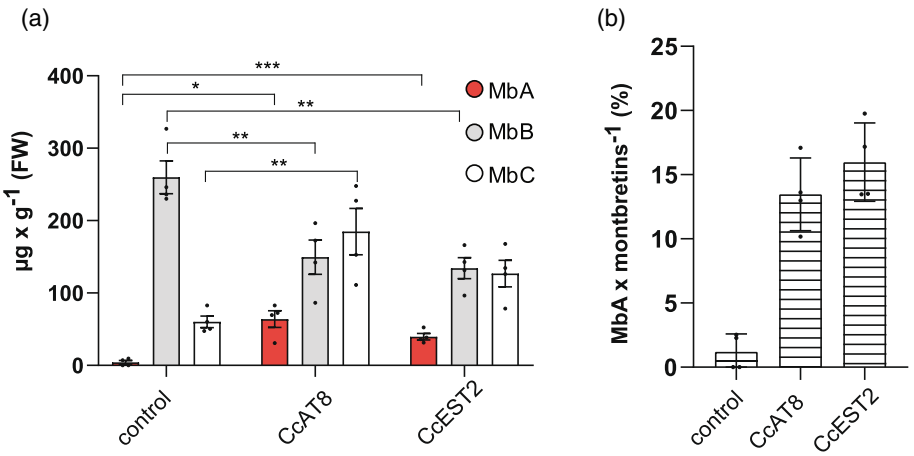

Figure 3. Levels of MbA, MbB, and MbC in *N. benthamiana* affected by co-expression of CcAT8 or CcEST2. (a) Levels of MbA, MbB, and MbC, and (b) percent of MbA of total montbretins (MbA + MbB + MbC). *N. benthamiana* leaves were transiently transformed using *A. tumefaciens* strains carrying plasmids with a set of ten genes previously shown to be required for montbretin production in Nb serving as the control (Irmisch et al., 2020), plus the additional transient transformation with CcAT8 or CcEST2. Leaves were collected 5 days post-infiltration. Metabolites were extracted with 50% MeOH and analyzed by LC–MS. Montbretin concentrations were calculated based on an external calibration curve of MbA. Error bars represent the standard error of the mean for four biological replicates (n = 4). Individual data points are shown as black dots. Asterisk symbols indicate significant differences (*P-value <0.05; **P-value <0.01; ***P-value <0.001). Also, see Figure S6 for results with the other phenylpropanoid pathway genes tested.

the effect of co-expression with different combinations of the candidate phenylpropanoid/shikimate shunt genes, specifically *CcAT8*, *CcEST2*, *CcC3′H* as well as the previously described *Cc4CL2* (identified in Sunstrum et al., 2021) in combination with the 10-gene set (Figure S8). Here we tested initially six additional combinations and quantified MbA, MbB, and MbC levels (Figure S8a). None of these combinations resulted in significantly increased MbA levels in comparison to co-expressing *CcAT8* or *CcEST2* alone with the 10-gene set (Figure S8a). To validate these results and test other possible combinations, we performed another set of similar experiments with a total of 13 combinations of *CcAT8*, *CcEST2*, *CcC3′H*, and *Cc4CL* with the 10-gene set (Figure S8d,e). Results from these experiments confirmed the individual effects of *CcAT8* and *CcEST2*, where individual co-expression of *CcAT8* or *CcEST* resulted in 3.0-fold (P-value <0.001, 9.9% of total montbretins) and 4.2-fold (P-value <0.001, 11.2% of total montbretins) increases of MbA levels, respectively (Figure S8d,e). The additional combinations tested here did not yield further increases in MbA (Figure S8e).

In vitro enzyme activities support CcAT8 as bona fide CcHCT and CcEST2 as bona fide CcCSE in the montbretia phenylpropanoid pathway

Given that *CcAT8* and *CcEST2* had the strongest significant *in vivo* effect on increasing MbA levels in Nb, we performed *in vitro* enzyme assays to verify the biochemical functions of CcAT8 and CcEST2. The full-length CDS of *CcAT8* and *CcEST2* were cloned into pASK-IBA37 for heterologous expression in *Escherichia coli*. Ni-affinity purified proteins (Figure S9) were used for enzyme assays, and products were identified based on accurate mass and authentic standards. CcAT8 was assayed with the acyl acceptor shikimate or quinate, and one of the acyl donors caffeoyl-, coumaroyl-, or feruloyl-CoA, which are substrates typically used by HCTs from other plant species (Kruse et al., 2022, 2023; Moghe et al., 2023). With shikimate as the acyl acceptor, all three acyl-CoA donors were converted into the respective shikimate esters, caffeoylshikimate (*m/z* 335.0893), coumaroylshikimate (*m/z* 319.0807) and feruloylshikimate (*m/z* 349.0932) (Figure 4). With quinate as the acceptor, CcAT8 was active only with coumaroyl-CoA as the acyl donor resulting in coumaroylquinate (*m/z* 353.0893) production. Protein preparations from *E. coli* carrying an empty vector were used as control (Figure S10). While most substrate combinations did not yield unspecific product formation in the control assays, trace amounts of product formation were observed when shikimate and coumaroyl-CoA were used in control assays (Figure S10), possibly due to spontaneous, non-enzymatic reaction or unspecific *E. coli* enzyme activity, similar to what has been observed before (Kruse et al., 2022).

The substrate for CSE (Vanholme et al., 2013), caffeoylshikimate, is not commercially available. Thus, we initially tested CcEST2 for activity with the surrogate substrate chlorogenic acid (i.e., caffeoylquinate). Consistent with esterase activity, samples taken at different time points of incubation with CcEST2 showed an increase in free caffeic acid and quinate, as well as a decrease in the substrate chlorogenic acid (Figure S11). To enable testing of caffeoylshikimate as well as coumaroylshikimate as substrates, we performed combined enzyme assays with CcAT8 and CcEST2. First, CcAT8 was incubated with

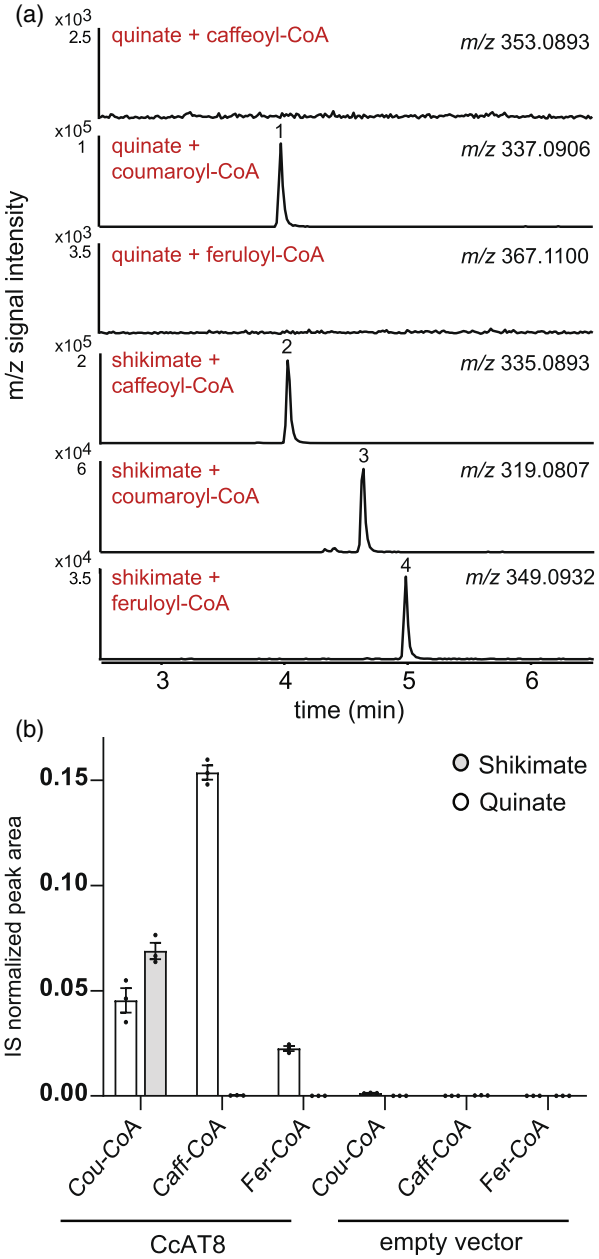

Figure 4. Enzyme activity of recombinant CcAT8. CcAT8 was heterologously expressed in *E. coli* and Ni-affinity purified protein was assayed with shikimate or quinate and either coumaroyl-CoA, caffeoyl-CoA or feruloyl-CoA (indicated in red). Controls were performed with a protein preparation from *E. coli* transformed with the empty vector (see Figure S10). (a) Products were analyzed by LC–MS and extracted ion chromatograms (EICs) with peaks for the detected products are shown. The mass-to-charge ratio (*m/z*) in negative ionization mode is given for each expected product. Peak 1, coumaroylquinate (*m/z* 337.0906); peak 2, caffeoylshikimate (*m/z* 335.0893); peak 3, coumaroylshikimate (*m/z* 319.0807); peak 4, feruloylshikimate (*m/z* 349.0932). (b) Bar graphs show the observed relative peak area of the detected products and black dots show individual data points. Three independent enzyme assay replicates were performed, and the peak area was normalized using the internal standard (IS, chloramphenicol). Error bars show the standard error of the mean ($n = 3$) and black dots indicate individual data points.

shikimate and either caffeoyl- or coumaroyl-CoA to produce the corresponding shikimate esters. Then CcEST2 was added and samples were analyzed for product formation over a 60 min incubation time course, showing an increase of free caffeic acid and a decrease of caffeoylshikimate over time (Figure 5). Negative control assays, performed with empty vector preparations instead of CcEST2, did not show this change in product formation and substrate depletion (Figure S11c,d), nor did changes occur in samples where coumaroylshikimate served as substrate (Figure S12).

Together with the *in vivo* results, the results from *in vitro* enzyme assays support a role of CcAT8 as bona fide HCT and a role of CcEST2 as bona fide CSE involved in caffeic acid and caffeoyl-CoA biosynthesis in montbretia. We are therefore annotating CcAT8 as CcHCT and CcEST2 as CcCSE. Hereafter, we will use the identifiers CcAT8/CcHCT and CcEST2/CcCSE, interchangeably depending on context.

Validation of the montbretia shikimate shunt pathway and its functional reconstruction in *Saccharomyces cerevisiae*

After having identified the four genes involved in the montbretia shikimate-shunt (Figure 1), *Cc4CL* (specifically Cc4CL2, Sunstrum et al., 2021), *CcHCT* (CcAT8, this work), a candidate *CcC′3H* (this work), and *CcCSE* (*CcEST2*, this work), we reconstructed pathway in yeast (*Saccharomyces cerevisiae*) to confirm the activity of CcC′3H *in vivo* and to assess the functionality of the complete pathway. Stepwise reconstruction in yeast provides a means to identify potential bottlenecks and assess flux through the pathway. Plasmids for heterologous expression of combinations of two, three, or four of these genes under the control of constitutive promotors were assembled using the yeast modular cloning toolkit (Lee et al., 2015) (Figure 6a; Figure S13) and transformed into the yeast strain BY4741. Single gene constructs, a non-transformed wild-type BY4741 strain, as well as a GFP expressing strain served as controls (Figure 6a; Figure S13). Yeast cultures were fed with coumaric acid as substrate, and the formation of phenylpropanoid shikimate shunt intermediates and products, specifically coumaroyl-CoA, coumaroylshikimate, caffeoylshikimate, caffeic acid and caffeoyl-CoA was analyzed by LC–MS.

Expression of Cc4CL alone did not yield detectable levels of the expected product 4-coumaroyl-CoA, despite depletion of the coumaric acid substrate, which was only observed when expressing *Cc4CL* alone or together with different combination of *CcHCT*, *CcC3′H* and *CcCSE* (Figure S14a). However, previous *in vitro* characterization of Cc4CL showed that this enzyme efficiently produces 4-coumaroyl-CoA (Sunstrum et al., 2021). It is possible that

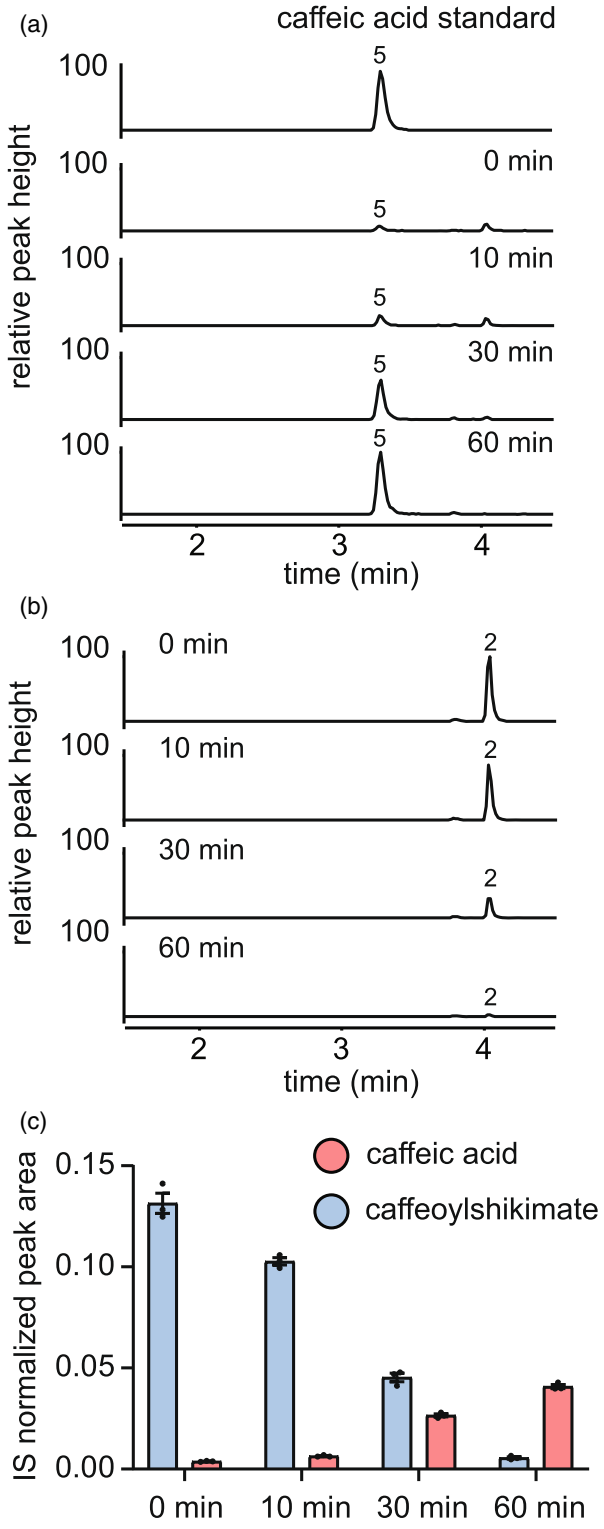

Figure 5. Activity of recombinant CcEST2 with caffeoylshikimate detected in coupled assays with CcAT. CcEST2 and CcAT8 enzymes were heterologously expressed in *E. coli* and Ni-affinity purified protein was used in assays. Substrates for CcEST2 were produced by incubating CcAT8 (CcHCT) for 1 h with shikimate and caffeoyl-CoA to produce caffeoylshikimate. After 1 h, CcEST2 was added and samples were taken at four different time points (0, 10, 30, and 60 min) over a 1-h time course and analyzed using LC–MS. (a) Representative extracted ion chromatograms (EICs) for the product caffeic acid (*m/z* 179.0443) and (b) the substrate caffeoylshikimate (*m/z* 335.0753). Peak height is normalized to the highest peak across all shown chromatograms (relative peak height). (c) Bar graph showing the observed relative peak area for caffeic acid and caffeoylshikimate. The peak area was normalized using the internal standard (IS, chloramphenicol). Coupled assays revealed a time-dependent increase in the product caffeic acid (a) and a decrease in the substrate caffeoylshikimate (b). Error bars represent the standard error of the mean for three independent assay replicates ($n = 3$) and black dots indicate individual data points. Peak 2, caffeoylshikimate; Peak 5, caffeic acid.

coumaroyl-CoA produced in these experiments was converted by endogenous yeast metabolism to avoid toxicity, which may be caused by coumaroyl-CoA as well as caffeoyl-CoAs (Incha et al., 2019; Liu et al., 2022; Parke & Ornston, 2004). Co-expression of Cc4CL and CcHCT led to the formation of the expected coumaroylshikimate intermediate in the shikimate shunt, confirming the functions of both Cc4CL and CcHCT (Figure 6b,c; Figure S14b). The additional expression of CcC3′H (CYP98A216) led to the formation of caffeoylshikimate, the next intermediate in the shikimate shunt, which validated the role of CcC3′H in hydroxylating coumaroylshikimate (Figure 6b,c; Figure S14b). While caffeoylshikimate accumulated, only traces of caffeic acid could be detected (Figure 6b; Figure S14b). Finally, with the addition of CcCSE, the co-expression of all four shikimate shunt genes in yeast led to the conversion of coumaric acid to caffeic acid, which validated the function of CcCSE (Figure 6b,c; Figure S14b). Here, only trace amounts of the intermediates coumaroylshikimate and caffeoylshikimate were detected, highlighting efficient metabolic flux from coumaric acid to caffeic acid via shikimate esters, and demonstrating the successful reconstruction of the shikimate shunt in yeast. Interestingly, yeast expressing the full shikimate shunt did not show depletion of coumaric acid as observed in other scenarios which expressed 4CL2, perhaps indicating a feedback regulation (Figure S14a). Competitive inhibition of 4CL by caffeic acid has previously been reported (Chen et al., 2013; Harding et al., 2002; Watanabe et al., 2018). As with coumaroyl-CoA, we were not able to detect caffeoyl-CoA.

In additional experiments, we tested if CcCSE could hydrolyze coumaroylshikimate. Yeast co-expressing Cc4CL, CcHCT, and CcCSE accumulated less coumaroylshikimate than yeast co-expressing Cc4CL and CcHCT, suggesting that coumaroylshikimate can be used as substrate by CcCSE as well (Figure 6b; Figure S14b). Yeast expressing CcC3′H alone also converted coumaric acid into caffeic acid (Figure 6c), which represents the shikimate-independent route of the phenylpropanoid pathway (Figure 1). This

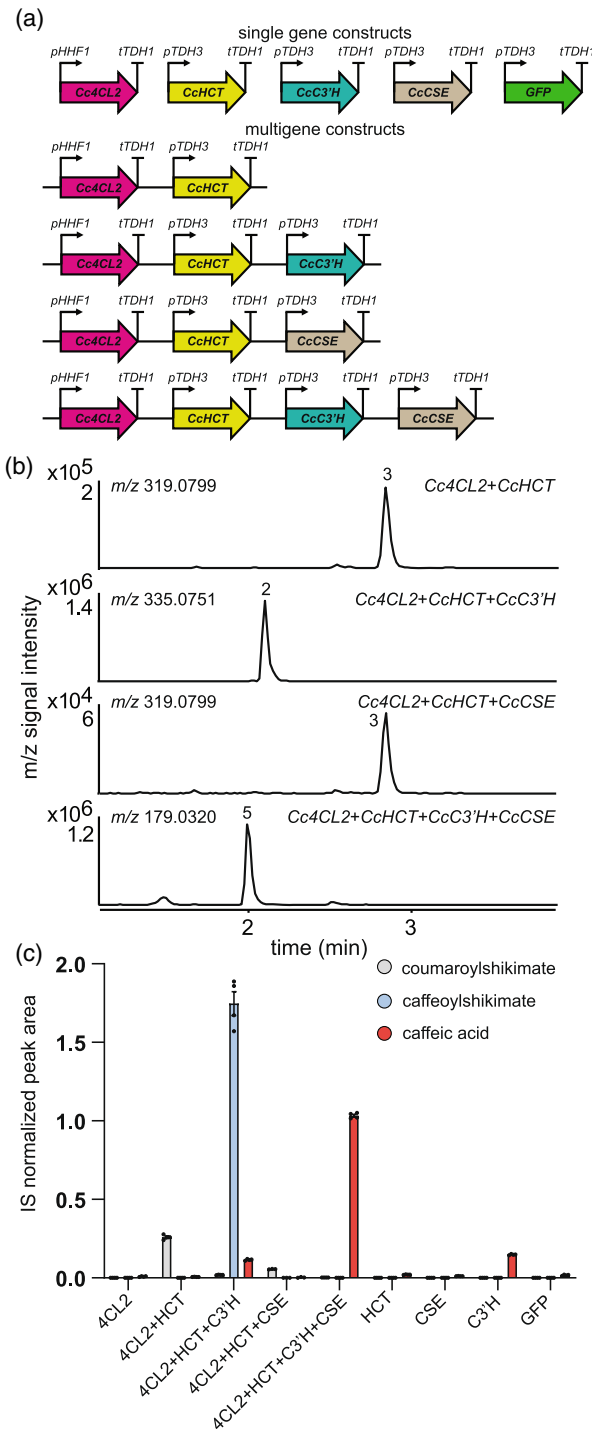

Figure 6. Functional validation and reconstruction of the montbretia shikimate shunt in yeast. (a) Four different single-gene and four different multigene constructs were made to constitutively express the montbretia shikimate shunt genes, Cc4CL2, CcHCT (CcAT8), CcC3′H, and CcCSE (CcEST2) in *Saccharomyces cerevisiae*. Shown for each construct are the montbretia gene(s), the promoter(s), and terminator(s) used. Details of the construct design are illustrated in Figure S13. (b) Conversion of coumaric acid, which was fed to transformed yeast strains, and product formation was measured after 3 days of incubation and detected by LC–MS analyses. Representative extracted ion chromatograms of the main product for each of the four multigene constructs are shown. (c) Integrated peak area for products detected in yeast carrying single-gene or multigene constructs. A yeast strain carrying a plasmid expressing GFP was used as control. Error bars represent the standard error of four independent replicates ($n = 4$). Dots indicate individual data points. Peak area was normalized to the internal standard (m/z 191.0275). A second experiment was conducted in the same manner and is shown in Figure S14(b). Peak 2, caffeoylshikimate; Peak 3, coumaroylshikimate; Peak 5, caffeic acid.

direct conversion produced sevenfold (P-value <0.001) less caffeic acid compared to yeast expressing the complete shikimate shunt (Figure 6c). In these experiments, the endogenous yeast cytochrome P450 reductase (CPR) was sufficient to support C3′H activity, not requiring heterologous CPR expression. As expected, individual expression of *CcHCT*, *CcCSE*, or *GFP* did not lead to accumulation of coumaroylshikimate, caffeoylshikimate, or caffeic acid (Figure 6c; Figure S14b).

Reconstruction of MbA biosynthesis in *N. benthamiana* using multigene constructs

Here we showed that MbA production in Nb can be improved by co-expression of the phenylpropanoid pathway shikimate shunt genes *CcHCT* (*CcAT8*) or *CcCSE* (CcEST2) in combination with the essential '10-gene set' for heterologous montbretin biosynthesis (Figure 3). The current laboratory process involves the transient transformation of Nb with 11 different and separately cultivated Agrobacterium strains, each containing a single gene on a separate plasmid. This approach would not be suitable for an industrial process due to the risk of process failure and being highly labor intensive. To streamline the system with fewer plasmids, we used the MoClo Tool Kit (Engler et al., 2014; Weber et al., 2011; Werner et al., 2012) to combine the 11 genes into two or three multigene constructs (MGC) (Figure 7a; Figure S15). As *CcHCT* and *CcCSE* effected MbA yield similarly (Figure 3), but no additive effect was observed with the two genes together (Figure S8), we chose to include *CcHCT* in the MGC. Eleven separate transcriptional units (level 1, Figure S15) were generated using the *p35S* promotor for constitutive expression and a set of different terminators (Engler et al., 2014) (Figure 7a; Figure S15). Level 1 units were used to assemble five different MGCs and each MGC contained different combinations of three to six different genes. Combinations of either two (two-plasmid system) or three (three-plasmid system) of these MGCs would cover the complete set of 11 genes for MbA biosynthesis. The two-plasmid transformation system comprised transformation of Nb leaves with MGC1 (18.8 kb) [$35S_{pro}$:(*CcUGT1-5* and *CcAT1*)] and MGC2 (15.3 kb) [$35S_{pro}$:(*CcFLS*, *CcCYP2*, *CcMYB4*, *CcHCT*, and *P19*)]; the three-plasmid transformation system comprised transformation of Nb leaves with MGC3 (14 kb) [$35S_{pro}$:

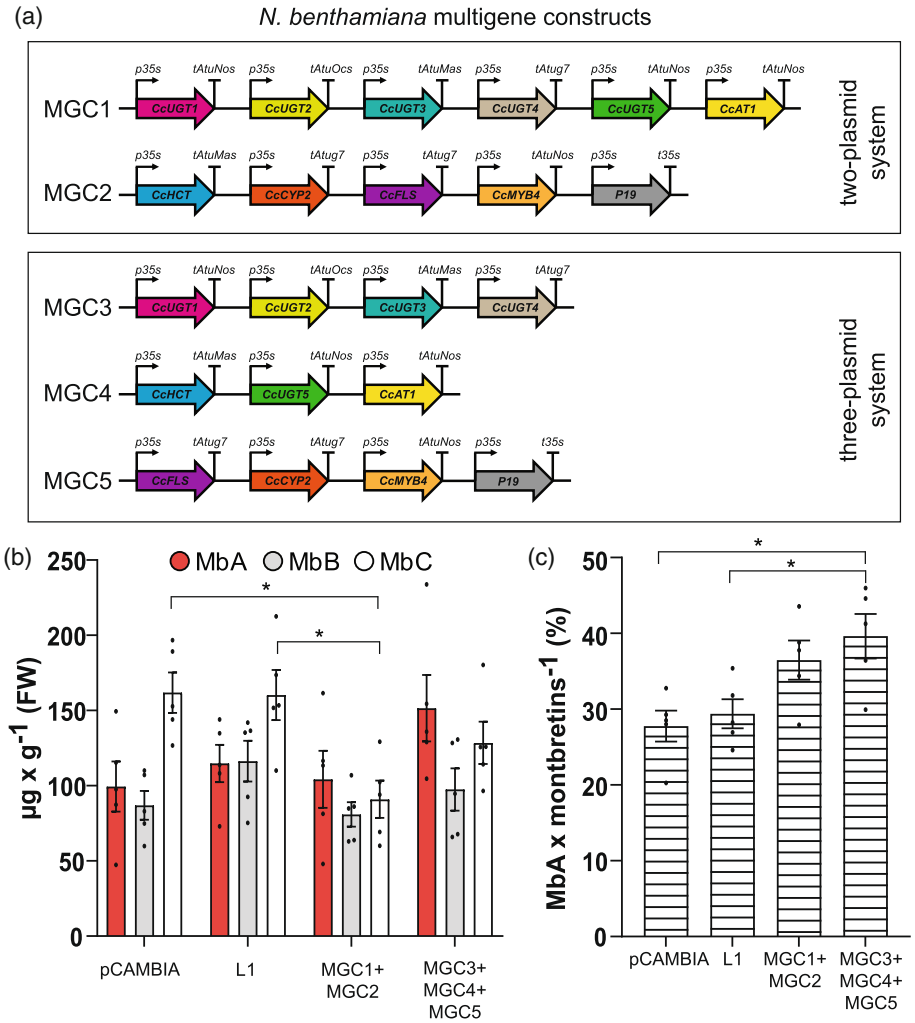

Figure 7. Montbretin A production in *N. benthamiana* using multigene constructs (MGCs). (a) Five different MGCs were designed to achieve the transient expression of 11 genes in *N. benthamiana* using a two-plasmid or three-plasmid system to streamline the transformation system. For each construct, the containing genes and the promoter and terminators controlling the expression are shown. Details on how constructs were made are shown in Figure S15. (b) *N. benthamiana* leaves were infiltrated with 11 different *Agrobacterium* strains carrying the 11 different genes on individual pCAMBIA230035Su plasmids (pCAMBIA); on individual level 1 plasmids (L1); or infiltrated with 2 or 3 *Agrobacterium* strains carrying the 11 genes on two plasmids MGC1 and MGC2 (MGC1 + MGC2); or on three plasmids MGC3, MGC4 and MGC5 (MGC3 + MGC4 + MGC5). Metabolites were extracted 5 days post-infiltration and montbretin concentrations were determined using an external MbA standard curve. (c) Percent proportion of MbA relative to the total of the three montbretins, MbA, MbB and MbC. Error bars represent the standard error of five biological replicates (n = 5). Individual datapoints are shown as black dots. Asterisk symbols indicate significant differences (*P-value <0.05).

(*CcUGT1-4*)], MGC4 (11.8 kb) [$35S_{pro}$:(*CcUGT5*, *CcHCT*, and *CcAT1*)], and MGC5 (13.2 kb) [$35S_{pro}$:(*CcFLS*, *CcMYB4*, *CcCYP2*, and *P19*)] (Figure 7a; Figure S15). For comparison, transformations were performed with a set of 11 plasmids representing the 11 different level 1 constructs, and a set of 11 different pCAMBiA constructs. Montbretin levels were quantified in Nb leaf extracts using LC–MS (Figure 7b; Figure S17). The use of the two-plasmid transformation system and the three-plasmid transformation system resulted in the accumulation of MbA at levels of 104.3 (±19.0 SE) µg MbA × g^{-1} FW and 151.6 (±22.1 SE) µg MbA × g^{-1} FW, respectively (Figure 7b). In comparison, 11 individual pCAMBIA and level 1 plasmids achieved 99.5 (±16.7 SE) µg MbA × g^{-1} FW and 114.8 (±12.4 SE) µg MbA × g^{-1} FW, respectively (Figure 7b). Overall, the amounts of MbA and MbB produced with these four different conditions were not found to be different. However, we found that the two-plasmid system produces less MbC in comparison to the 11 individual plasmid system, both for pCAMBIA and level 1 (P-value = 0.013 and 0.016, respectively, Figure 7b). Interestingly, when comparing the percent MbA relative to total montbretins between the different systems, we found that the three-plasmid system led to the production of a larger proportion of MbA (39.6%

MbA of total montbretins). In comparison percent MbA achieved with the two 11-plasmid systems, pCAMBIA (27.8% MbA of total montbretins, *P*-value = 0.015) or level 1 MbA at (29.4% MbA of total montbretins, *P*-value = 0.037) (Figure 7c), were lower. These results highlight that the process of transient transformation for MbA production in Nb was successfully streamlined with the use of either two or three MGCs, instead of 11 individual plasmids, without compromising on the amount of MbA produced and potentially increasing the ratio of MbA relative to other montbretins. Efforts to generate a single MGC with all 11 genes (35.4 kb) were not successful.

DISCUSSION

MbA is a promising new treatment option for T2D and clinical obesity (Andersen et al., 2009; Tarling et al., 2008; Williams et al., 2015; Yuen et al., 2016), but supply issues pose a challenge to its development as a pharmaceutical or nutraceutical. In previous work, we identified the genes and enzymes for MbA biosynthesis in montbretia and established proof-of-concept for bioengineering of MbA production in Nb (Irmisch et al., 2018, 2020; Irmisch, Jancsik, et al., 2019; Irmisch, Ruebsam, et al., 2019). Here, we investigated the biosynthesis of caffeic acid and caffeoyl-CoA, which are essential precursors for the formation of MbA in montbretia. While MbA is the predominant of the different montbretin metabolites that accumulates in montbretia corms, additional bioengineering tools are needed to improve MbA yields in Nb in absolute terms and relative to MbB and MbC. We identified and characterized a set of montbretia genes, specifically the shikimate shunt genes *CcHCT* (*CcAT8*), *CcC3′H* and *CcCSE* (*CcEST2*), and showed the utility of CcHCT and CcCSE to improve MbA yields in Nb. Using the extended set of genes for MbA bioengineering in Nb, we then improved the process of transient MbA pathway transformation by designing and testing variations of multigene constructs, instead of single-gene constructs, suitable for industrial process adaptation.

Montbretia shikimate shunt enzymes

It is not known which mechanisms control the preferential formation of MbA, relative to MbB and MbC, in montbretia. One possible mechanism, which we had previously tested, considered substrate specificity of the montbretin pathway branchpoint enzyme CcAT1 or its isoform CcAT2 (Irmisch et al., 2018, 2020). This enzyme introduces the different acyl sidechains, caffeoyl, coumaroyl or feruloyl, into MbA, MbB and MbC, respectively. Another hypothesized mechanism is the preferential biosynthesis of caffeoyl-CoA, relative to coumaroyl-CoA and feruloyl-CoA, by Cc4CLs. However, we found that CcAT is not specific for caffeoyl-CoA, relative to coumaroyl-CoA or feruloyl-CoA, and thus does not control selectivity for MbA formation (Irmisch et al., 2018, 2020). Similarly, montbretia 4CL enzymes are also not selective for favoring caffeoyl-CoA formation over coumaroyl-CoA and feruloyl-CoA (Sunstrum et al., 2021). We therefore set out to investigate other steps of the montbretia phenylpropanoid pathway, specifically for the formation of caffeic acid and caffeoyl-CoA (Figure 1). Within the phenylpropanoid pathway, coumaroyl-CoA is formed from coumaric acid by 4CL and caffeoyl- and feruloyl-CoA can then be produced from coumaroyl-CoA via shikimate ester intermediates, a pathway referred to as the shikimate shunt. We identified and characterized three montbretia genes, *CcHCT*, *CcC3′H*, and *CcCSE*, which we used together with *Cc4CL* to reconstitute the montbretia shikimate shunt for the conversion of coumaroyl-CoA into caffeic acid and caffeoyl-CoA in Nb and also in yeast.

In montbretia, CcHCT catalyzes the conjugation of coumaroyl-CoA and shikimate to coumaroylshikimate, the first step of the shikimate shunt. Like HCTs from other plant species, CcHCT also uses caffeoyl-CoA and feruloyl-CoA in addition to coumaroyl-CoA; however, the latter together with shikimate is considered the *bona fide* hydroxycinnamoyl-CoA HCT substrate (Hoffmann et al., 2003; Kruse et al., 2022; Moghe et al., 2023). CcHCT clustered most closely with characterized HCTs from other plant species (Figure 2), specifically with other characterized monocot HCTs from switchgrass and great millet (*Sorghum bicolor*) supporting its role as HCT in montbretia (Escamilla-Treviño et al., 2014; Walker et al., 2013).

The second step of the shikimate shunt is catalyzed by CcC3′H (CYP98A3) (Figure 1), a cytochrome P450 enzyme of the CYP98 family that hydroxylates coumaroylshikimate to caffeoylshikimate (Schoch et al., 2001). In addition, CcC3′H also converted coumaric acid into caffeic acid, although with lower efficiency compared to the conversion of the CoA esters. This is similar to what has been reported for C3H enzymes from other plant species (Franke et al., 2002; Schoch et al., 2001).

There are two possibilities for the third step from caffeoylshikimate to caffeoyl-CoA. One involves the reverse reaction of HCT; the other involves an ester hydrolysis catalyzed by CSE to yield caffeic acid which is then converted by 4CL to caffeoyl-CoA (Hoffmann et al., 2005; Kriegshauser et al., 2021; Vanholme et al., 2013). HCTs from *A. thaliana*, switchgrass, and poplar can catalyze the reverse reaction from caffeoylshikimate to caffeoyl-CoA and shikimate, but preferentially catalyze the forward-reaction (Escamilla-Treviño et al., 2014; Ha et al., 2016; Saleme et al., 2017; Wang et al., 2014). Interestingly, these three species also possess functional CSEs and downregulation or knock-out of CSE in these species resulted in severe growth and lignin phenotypes (de Vries, Brouckaert, et al., 2021; Ha et al., 2016; Saleme et al., 2017; Vanholme et al., 2013). CSE was first described in *A. thaliana* (Vanholme et al., 2013) and subsequently, CSEs have been characterized in switchgrass, barrel clover, and poplar (de

Vries, Brouckaert, et al., 2021; Ha et al., 2016; Saleme et al., 2017). *CSE* genes have been annotated in both monocots and dicots, but in some species (e.g., *Brachypodium distachyon*, *Zea mays*, *Beta vulgaris*, *Ricinus communis*, and *Physcomitrium patens*) *CSE* homologous genes seem to be missing (Ha et al., 2016). This pattern indicates that the CSE enzymatic function evolved independently in different species or *CSE*s were lost in several lineages (de Vries, Brouckaert, et al., 2021; Ha et al., 2016).

We could not find evidence supporting the HCT reverse reaction in montbretia, but we identified a functional CcCSE which clustered with other class I esterases (Ha et al., 2016) and displayed the canonical esterase activity. CcCSE is active with caffeoylshikimate and chlorogenic acid as substrates but did not show activity with coumaroylshikimate when assayed *in vitro*. However, the shikimate pathway reconstruction in yeast indicated that CcCSE also uses coumaroylshikimate, albeit with lower activity. This apparent preference of CcCSE toward caffeoylshikimate is similar to AtCSE, which strongly preferred caffeoylshikimate over coumaroylshikimate (Vanholme et al., 2013). Other characterized CSEs, such as the two CSEs from poplar, display broad substrate specificity, accepting caffeoylshikimate, feruloylshikimate, coumaroylshikimate, and sinapoylshikimate, although with highest affinity for caffeoylshikimate (de Vries, Brouckaert, et al., 2021). *CcCSE* showed a pattern of transcript expression matching that of previously characterized MbA biosynthetic genes and MbA accumulation during the early summer in developing young corms (Figure S4) (Irmisch et al., 2018, 2020; Irmisch, Jancsik, et al., 2019; Irmisch, Ruebsam, et al., 2019).

The results described here from the identification of montbretia shikimate shunt genes, and successful functional reconstruction of the shikimate shunt in yeast using these genes, support a model whereby caffeic acid and caffeoyl-CoA for MbA formation in montbretia are provided via the shikimate shunt. Of the different montbretia shikimate shunt genes, the transcript expression pattern of *CcCSE* in developing young corms matches that of other genes of the MbA biosynthetic system, suggesting that CcCSE may play a critical role in MbA precursor biosynthesis in the shikimate shunt.

The reconstruction of a functional shikimate shunt module in yeast expands the toolbox to metabolically engineer heterologous hosts for caffeic acid and caffeoyl-CoA production. While simpler biosynthetic pathways for caffeic acid production have been engineered in yeast (Liu et al., 2019; Zhou et al., 2021), the shikimate shunt may offer a system to produce high levels of caffeoyl-CoA and low levels of coumaroyl-CoA. The MbA biosynthetic system serves as an example where selective incorporation of caffeoyl-CoA is essential, and the same applies to other biosynthetic pathways that require caffeoyl moieties.

Co-expression of the shikimate shunt improves MbA formation in *N. benthamiana*

We had previously shown that transient expression of montbretia MbA pathway genes, together with montbretia myricetin precursor biosynthesis genes, in Nb resulted in approximately 10 μg MbA $\times$ g^{-1} FW, and more than 50 times higher levels of MbB (Irmisch et al., 2020). Here, we leverage the capacity of engineered Nb to produce montbretins (MbA, MbB, and MbC) and the metabolic plasticity of this system to tailor caffeoyl-CoA precursor biosynthesis, which resulted in the intended decrease of MbB and increase of MbA.

The previous predominant formation of MbB in Nb may have been due to the high endogenous availability of coumaroyl-CoA (Vanholme et al., 2019). Using co-expression of the shikimate shunt genes CcHCT or CcCSE we were able to significantly decrease MbB levels and increase MbA levels, presumably through improved availability of caffeoyl-CoA (Figure 3; Figure S6). MbC levels also increased, indicating conversion of caffeoyl-CoA by Nb endogenous O-methyltransferase activity to feruloyl-CoA (Figure 1b) (Zhong et al., 1998), which will be a target for downregulation in future work to further improve MbA levels and reduce MbC levels. Importantly, the overall amount of the three montbretins combined did not change, supporting our hypothesis that the availability of the different hydroxycinnamoyl-CoA esters, caffeoyl-CoA, coumaroyl-CoA, and feruloyl-CoA controls the ratio of MbA, MbB, and MbC in engineering Nb and possibly also in montbretia. It was not necessary to co-express the complete shikimate shunt in Nb along with other MbA biosynthetic pathway genes to achieve improved production of MbA, but we observed similar increases in MbA levels by co-expression of either CcHCT or CcCSE, implicating endogenous Nb enzymes in the precursor biosynthesis. An increase of MbA together with the accumulation of the pathway intermediate caffeoylshikimate upon co-expression of CcHCT (Figure S7b) suggests that coumaroylshikimate is efficiently transformed into caffeoylshikimate by NbC3′H and also suggests that Nb has a functional CSE. Indeed, while Nb CSE has not yet been reported, we found two predicted protein sequences (Niben261Chr08g0084016.1 and Niben261Chr01g1081002.1) with approximately 88% identity to Arabidopsis CSE (AT1G52760.1) in the Nb genome assembly (v2.6.1; https://solgenomics.net/organism/Nicotiana_benthamiana/genome).

The present work indicated remaining metabolic bottlenecks that will be explored in future work for further improvement of MbA production in Nb. The interplay and integration of heterologously expressed montbretia genes and endogenous Nb metabolism are of particular interest. For example, we observed accumulation of caffeoylshikimate in Nb when *CcHCT* was co-expressed with

the 10-gene set for MbA biosynthesis. This compound did not accumulate in the co-expression of the entire montbretia shikimate shunt in yeast (Figure 6), indicating efficient flux through the montbretia shunt pathway that was not yet fully achieved when utilizing both montbretia genes and endogenous metabolism in Nb. *CcCSE* co-expression with the 10-gene set for MbA biosynthesis led to increased MbA levels, suggesting *NbHCT* contributes to the pathway in this scenario. It also is possible that chlorogenic acid, which is abundant in Nb leaves, may be cleaved by CcCSE and contribute to increased caffeic acid and caffeoyl-CoA levels. However, we did not observe changes in chlorogenic acid levels affected by CcCSE expression (Figure S7). Chlorogenic acid appears to be closely linked with lignin biosynthesis (Escamilla-Treviño et al., 2014; Serrani-Yarce et al., 2021; Volpi e Silva et al., 2019), however recent analyses suggest that the pathways leading to either lignin or chlorogenic acid seem to occur independently and are not part of a metabolic network (Cardenas et al., 2021).

In our approach of engineering the biosynthesis of a complex molecule, such as MbA, into Nb, we are taking advantage of the host plant's endogenous metabolism (i.e., MbA building blocks), which would not be available in microbial hosts such as yeast or *E. coli*. However, the extent to which such metabolic precursors can be accessed and redirected may be limited. Precursor limitation has previously been observed for example in the engineering of etoposide (Lau & Sattely, 2015; Schultz et al., 2019) or triterpenoid biosynthesis (De La Peña & Sattely, 2021; Reed et al., 2017) in Nb. Expression of biosynthetic genes or transcriptional regulators that enhance individual precursor supply, or entire metabolic networks can alleviate bottlenecks and reshape endogenous precursor pathways (Butelli et al., 2008; Pandey et al., 2014; Wang et al., 2016). Previously, we showed that the expression of montbretia myricetin biosynthetic genes and a montbretia MYB transcription factor enhance myricetin availability for MbA biosynthesis in Nb (Irmisch, Ruebsam, et al., 2019). In the present study, we targeted caffeoyl-CoA availability to increase MbA yields. These two approaches generally represent the 'push' strategy in metabolic engineering (Yuan & Grotewold, 2015), enhancing precursor push toward MbA. In the present work, we found that individual co-expression of either CcHCT or CcCSE with the 10-gene set similarly increased MbA levels. CcHCT expression may deliver a 'push' effect into the shikimate shunt by producing more coumaroylshikimate, while the effect of CcCSE could be a 'pull' of metabolic flux towards caffeic acid by hydrolyzing caffeoylshikimate. In this case, both the 'push' and the 'pull' resulted in the same intended increase of MbA levels, which illustrates the metabolic plasticity of the Nb system. Combinations of 'push' and 'pull' strategies may result in the highest levels of desired end products (Yuan & Grotewold, 2015). However, co-expressing multiple montbretia shikimate shunt enzymes, including combinations of CcHCT and CcCSE, did not result in additive effects (Figure S8), indicating additional potential bottlenecks in Nb. Identifying these bottlenecks will open up new opportunities to improve MbA yields in Nb in the future.

Use of multigene constructs towards improved processes for MbA production

Transformation of Nb with multiple different *Agrobacterium* strains may reduce transformation efficiency (Montague et al., 2011; Stephenson et al., 2020). To address this general concern, and more specifically to streamline the transformation system for MbA production in Nb, we successfully designed and tested the stacking of 11 different genes onto two or three multigene constructs. As this approach requires the handling of fewer plasmids and fewer *Agrobacterium* strains for transient transformation, compared to individual-gene constructs, it reduces the risk of process failure in an industrial production system, which is being developed for MbA biosynthesis. The construction of plasmids carrying multigene expression cassettes was enabled by modular cloning. The number of genes per plasmid is, in theory, only limited by the final plasmid size, and constructs of up to 50 kb have been assembled and used for *Agrobacterium* transformation (Dudley et al., 2022; Grzech et al., 2022; Werner et al., 2012). In addition to the benefit of a streamlined process, we had speculated that the use of multigene constructs versus single-gene constructs may further improve the MbA yield. However, we found similar MbA levels when using either two or three Agrobacterium strains carrying multigene constructs compared to the use of 11 Agrobacterium strains carrying single-gene constructs (Figure 7). This observation is similar to another recent study of metabolic engineering with multigene constructs in Nb (Grzech et al., 2022).

Overall, the co-expression of *CcHCT* and the previously described 10-gene set using multigene constructs resulted in improved MbA yields in a streamlined transient Nb transformation system. We achieved reproducible levels of 0.1 mg MbA $\times$ g^{-1} FW, which is about 30 times higher than previously reported yields (Irmisch et al., 2020). For comparison, the level of MbA achieved here is in the range of other plant metabolites heterologously produced in Nb, such as strictosidine (0.23 mg $\times$ g^{-1} dry weight) (Dudley et al., 2022) or artemisinin (0.8 mg $\times$ g^{-1} dry weight) (Malhotra et al., 2016).

CONCLUSION

MbA is a promising new treatment option for T2D and clinical obesity, but development and application as a pharmaceutical or nutraceutical is limited by MbA supply from montbretia plants. Here we investigated the montbretia phenylpropanoid shikimate shunt and the use of its genes

to optimize caffeoyl-CoA availability in Nb for MbA production. Our findings suggest a model where CSE facilitates enhanced caffeic acid production and thus may enable preferential MbA formation in montbretia. Utilizing the newly identified montbretia genes improved MbA production in Nb and demonstrated how the endogenous metabolism of Nb can be tuned for the production of high-value small metabolites. This work highlights the potential to overcome limitations for MbA production in Nb, which may also apply to other metabolites derived from phenylpropanoid precursors.

EXPERIMENTAL PROCEDURES

Identification and cloning of montbretia candidate genes and phylogenetic sequence analyses

The databases and methods for the identification of montbretia (*Crocosmia* × *crocosmiiflora*) candidate transcripts and their open reading frames (ORF) and CDS were as described previously (Irmisch et al., 2018, 2020; Irmisch, Jancsik, et al., 2019; Irmisch, Ruebsam, et al., 2019). In brief, a BLASTP search against the translated transcriptome database that covers a time course of montbretia corm development (Irmisch et al., 2020) was conducted using sequences from *A. thaliana* as query sequences (Table S2). Montbretia candidate sequences were filtered based on ≥40% identity to the query amino acid (aa) sequence and ≥300 aa length. In addition, for the identification of esterases, two sequences were identified with an e-value of 0 that returned the term esterase in a BLASTP (Camacho et al., 2009) search and showed >2-fold higher expression in June compared to October, the period of MbA biosynthesis in young corms as described in Irmisch et al. (2020) (Irmisch et al., 2020). This approach yielded three candidate acyltransferases (CcAT8-10), three candidate esterase (CcEST1-3), and one candidate *p*-coumaroylshikimate 3′-hydroxylase (CcC′3H). Full-length ORFs were obtained for all seven candidate genes by PCR from young corm cDNA attained from RNA (see Table S3 for primer information) (Irmisch et al., 2018). The ORFs were cloned into the pJET1.2/blunt vector (Thermo Fisher Scientific, Waltham, MA, USA, https://www.thermofisher.com/) and sequences were verified through sequencing. Phylogenetic sequence analyses were performed using methods described in Kruse et al., 2022, 2023 (Kruse et al., 2022, 2023). Additional sequences for the esterase phylogeny were obtained using a BLAST search of the *A. thaliana* AtCSE against the UNIPROT (The UniProt Consortium, 2023) database. Amino acid sequence alignments of montbretia sequences and sequences from other plants (for AT sequences see File S1; for CSE sequences see File S2) were constructed with default settings in the GENEIOUS v2022.2.2 (Biomatters, MA, USA, https://www.geneious.com) plugin of MAFFT v.7.453 (Katoh et al., 2002) and used to generate a maximum likelihood phylogenetic tree (non-parametric bootstrap, 1000 replicates) with IQ-Tree v1.6.10 (Nguyen et al., 2015). ITOL v5.6.2 (Letunic & Bork, 2021) was used to visualize the phylogenetic tree.

Candidate gene expression in *Escherichia coli*

The ORFs for *CcHCT* (*CcAT8*) and *CcCSE* (*CcEST2*) were subcloned into expression vectors, heterologously expressed in *E. coli*, and proteins purified following previously described methods (Irmisch et al., 2018; Kruse et al., 2022). Briefly, sequences covering complete open reading frames were amplified and cloned using *Bsa*I or *Bsp*MI into the pASK-IBA37 vector (IBA Lifesciences, Göttingen, Germany, https://www.iba-lifesciences.com/). The *E. coli* 5-alpha strain (New England Biolabs, Ipswich, MA, USA, https://international.neb.com/) was used for heterologous enzyme expression. For this, cultures of 250 ml were grown to OD_{600} = 0.5 at 37°C and 210 rpm, expression induced using 200 μg L^{-1} anhydrotetracycline (Sigma-Aldrich, St. Louis, MO, USA, https://www.sigmaaldrich.com/), and grown for another 16 h at 25°C. Cells were collected by centrifugation, resuspended in 8 ml lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole, 1 mg ml^{-1} lysozyme, pH 8.0), and incubated for 1 h at 4°C. Final cell lysis was achieved by using BUGBUSTER® (MilliporeSigma, Burlington, MA, USA, https://www.sigmaaldrich.com/) at half the recommended concentration (0.5×) according to the manufacturer's protocol. Cell debris and insoluble proteins were removed through centrifugation (30 min, 21 000 ***g***, 4°C) and the cell-free lysate was incubated with 200 μl of 50% Ni-NTA-Agarose matrix (Qiagen, Hilden, Germany, https://www.qiagen.com) for 30 min. After removing the lysate, the matrix was washed with 1 ml wash buffer (50 mM NaH_2PO_4, 300 mM NaCl, 20 mM imidazole, pH 8.0) and then transferred to a 2 ml spin column (BioRad Laboratories, Hercules, CA, USA, https://www.bio-rad.com/) fitted with a frit. The matrix was washed again with 500 μl wash buffer, after which proteins were eluted five times with 100 μl elution buffer (50 mM NaH_2PO_4, 50 mM NaCl, 250 mM imidazole, pH 8.0). The protein concentration of the eluate fractions was determined by measuring the absorbance at 280 nm and fractions with similar protein concentrations (~0.2–0.3 mg ml^{-1}) were combined. Subsequently, PD MINITRAP™ SEPHADEX™ G-25 columns (GE Healthcare, Chicago, IL, USA, https://www.gehealthcare.com/) were used according to the manufacturer's instructions to transfer the protein into assay buffer (10 mM TRIS–HCl, pH 7.2). Glycerol was added to the protein samples to a final concentration of 50% (v/v) and samples were stored at −20°C or used directly for enzyme assays. As a control, a protein preparation of *E. coli* harboring the empty pASK-IBA37 vector was obtained in the same way (empty vector control).

CcHCT and CcCSE enzyme assays

In vitro assays were performed with CcHCT (CcAT8) and CcCSE (CcEST2) to confirm enzyme functions. To test for CcHCT activity, 1 μg of Ni-affinity purified CcHCT was incubated with 100 μM acyl donor substrate (coumaroyl-CoA, caffeoyl-CoA, or feruloyl-CoA; all from TransMIT, Gießen, Germany, https://chemicals.transmit.shop/) and 200 μM acceptor substrate (shikimate or quinate). Reactions were carried out in assay buffer in a total volume of 25 μl for 1 h and were shaken at 30°C. Control assays contained 5 μl of empty vector enzyme preparation. Assays were stopped by adding an equal volume of 100% (v/v) methanol (LC–MS grade, Sigma-Aldrich) which contained 25 μM chloramphenicol as internal standard. After vortexing, the samples were centrifuged at 21 000 ***g*** for 20 min at 4°C and 25 μl of the supernatant was transferred into LC vials with glass inserts and stored at −80°C until LC–MS analysis. To test for CSE activity, two different types of assays were performed. One set of assays used chlorogenic acid as a surrogate substrate. Here, 200 μM chlorogenic acid was incubated with CcSE (CcEST2) in 25 μl reactions at 30°C for 0, 10, 30, and 60 min. The reactions were stopped by adding an equal volume of 100% (v/v) methanol (LC–MS grade, Sigma-Aldrich). As described above, after vortexing and centrifugation 25 μl was used for LC–MS analysis. In a second set of assays, coupled assays with CcHCT were performed to generate the substrates caffeoylshikimate and coumaroylshikimate to be tested with CcCSE. For the coupled assays, CcHCT reactions were set up as described above for HCT assays to produce caffeoylshikimate

(using shikimate and caffeoyl-CoA as substrates) or to produce coumaroylshikimate (using shikimate and coumaroyl-CoA as substrates) and incubated for 1 h. Reactions were then placed on ice and 1.5 μg Ni-affinity purified CcCSE (CcEST2) was added and the reactions were continued for up to 1 h at 30°C. Samples were taken at 0, 10, 30, and 60 min after the addition of CSE. The addition of the empty vector preparation (5 μl) instead of CcCSE served as a negative control. Assays were stopped by adding an equal volume of 100% (v/v) methanol (LC–MS grade, Sigma-Aldrich) which contained 25 μM chloramphenicol as internal standard. After vortexing, the samples were centrifuged at 21 000 ***g*** for 20 min at 4°C, and 25 μl of the supernatant was transferred into LC vials with glass inserts. Products of CcHCT and CcCSE assays were analyzed and quantified using LC–MS as described below.

Transformation and transient expression of candidate genes in *Nicotiana benthamiana*

Transient expression of *CcUGT1-5*, *CcAT1-2*, *CcMYB4*, *CcFLS*, *CcCYP2*, *CcAT8-10*, *CcEST1-3*, *CcC3′H*, and *Cc4CL* in *N. benthamiana* was performed as described previously (Irmisch et al., 2018, 2020; Irmisch, Ruebsam, et al., 2019). In brief, the complete ORFs for each gene were cloned into the pCAMBIA230035Su vector following the User Cloning method (Geu-Flores et al., 2007; Nour-Eldin et al., 2006). The resulting plasmids (Table S4) were transformed into *Agrobacterium tumefaciens* strain C58pMP90 (GV3101pMP90) (Koncz & Schell, 1986). Overnight cultures of 1 ml were used to inoculate 10 ml cultures (LB-media containing 50 μg ml^{-1} kanamycin, 25 μg ml^{-1} rifampicin, and 25 μg ml^{-1} gentamycin) and grown over-night at 28°C. Cells were harvested by centrifugation and resuspended in infiltration buffer (10 mM MES, pH 5.6, 10 mM $MgCl_2$, 100 μM acetosyringone) to a final OD_{600} of 0.5 and incubated at room temperature for 1–3 h. Before infiltration, equal volumes of each strain required for a particular combination of genes were combined and used to infiltrate leaves of 3–4-week-old *N. benthamiana* plants using a needleless syringe (Figure S17). An *Agrobacterium* strain carrying the *pBIN:p19 plasmid was included in each combination*. Infiltrated plants were kept in the growth chamber (8 h light/16 h dark, at 24°C) for 5 days, after which the infiltrated leaves were sampled using a 9 mm hole puncher to obtain leaf disks. For each infiltrated leaf, two leaf disks were cut and pooled (total of 4 disks per plant). Intact leaf disks were extracted with 1 ml of 50% MeOH (v/v; 12.5 μM chloramphenicol as internal standard) per 100 mg (fresh weight) plant material by shaking at 100 rpm for 4 h at room temperature. The same process was used for transient expression using multigene constructs, except kanamycin was replaced with 50 μg ml^{-1} carbenicillin for growing *Agrobacterium* strains transformed with level 1 plasmids.

Construction of multigene plasmids for transient expression in *Nicotiana benthamiana*

Multigene plasmids for transient expression in *N. benthamiana* were assembled using the MoClo Toolkit (Addgene Kit #1000000044) and the MoClo Plant Parts Kit (Addgene Kit #1000000047) (Engler et al., 2014; Weber et al., 2011; Werner et al., 2012). Site-directed mutagenesis was used to remove internal *Bsa*I and *Bpi*I restriction sites from the candidate sequences without altering the encoded protein sequence using the Q5 Site-Directed Mutagenesis Kit (New England Biolabs) according to the manufacturer protocol (for primers see Table S3). As template, the previously generated pJET1.2/blunt vector constructs were used and the modified coding sequences were then amplified using gene-specific primers (Table S3) and cloned as *Bpi*I fragments into the level 0 (L0) vector pICH41308. The L0 served as the basis for the assembly of level 1 plasmids and subsequently level 2 plasmids (Figure S15, Table S4), following the process described in Weber et al. (2011) and Engler et al. (2014) (Engler et al., 2014; Weber et al., 2011).

Construction of multigene plasmids for heterologous expression in yeast

The generation of constructs for heterologous gene expression in *Saccharomyces cerevisiae* was as previously described (Sunstrum et al., 2021). In short, the yeast modular cloning toolkit (MoClo YTK, Addgene Kit #1000000061, https://www.addgene.org/kits/moclo-ytk/) (Lee et al., 2015) was used and cds were cloned into entry vectors, transcriptional units and multigene constructs using Golden-Gate cloning and *Bsa*I and *Bsm*BI restriction sites (see Figure S13, Table S4). Plasmids were transformed into the *Saccharomyces cerevisiae* strain BY4741 (GE Life Sciences, Watertown, MA, USA, https://gelifesciences.com/).

Yeast *in vivo* assays

Yeast *in vivo* feeding assays were performed as described previously (Sunstrum et al., 2021) with minor modifications. Specifically, starter cultures were grown overnight at 28°C with shaking at 180 rpm in six-well plates (Greiner bio-one, CELLSTAR®) containing 2.5 ml SD-URA media and using a single yeast colony carrying one of the plasmids. The next day, starter cultures were used to inoculate 250 μl SD-URA media in 48-well plates (Greiner bio-one, CELLSTAR®) to OD_{600} = 0.1. Cultures were supplied with 1 mM coumaric acid as substrate and grown for 72 h (28°C, 180 rpm). Plates were covered with breathable film (AXYGEN®) and placed in a plastic container with high humidity during cultivation to prevent evaporation. For the cultivation of wild-type yeast, YPD media or SD media was used. Assays were stopped by transferring cultures to glass LC vials (Agilent Technologies, Santa Clara, CA, USA, https://www.agilent.com/), adding equal volumes of 100% (v/v) MeOH, and subsequent incubation at 42°C for 10 min. After centrifugation (15 min, 4°C, 4300 ***g***) 50 μl of the supernatant was transferred to glass inserts (Agilent Technologies) placed in 1.5 ml glass LC vials, and stored at −80°C until LC–MS analysis. As an internal standard for peak normalization, a peak consistently present across all samples with stable intensity (*m/z* 191.0275, retention time: 0.84 min) was used. For each multigene construct, cultures were grown with four replicates for each experiment and each experiment was replicated twice for a total of eight-fold replication.

Liquid chromatography-mass spectrometry (LC–MS) analyses

LC–MS analyses were performed as previously described (Irmisch et al., 2020). Samples from experiments with *N. benthamiana*, except for those using multigene constructs, were analyzed on an Agilent (Agilent Technologies) 1100 LC system coupled to an Agilent MSD Trap XCT-Plus mass spectrometer with electro-spray in negative ionization mode (capillary voltage, 4000 eV; temp, 350°C; nebulizing gas, 60 psi; dry gas 12 L min^{-1}). Analyses of products from the enzyme *in vitro* assays, yeast *in vivo* assays and *N. benthamiana* samples from experiments with multigene constructs were done on an Agilent 1290 Infinity LC system coupled to an Agilent 6530 Accurate Mass Q-TOF mass spectrometer (both Agilent Technologies) with an electrospray ion source operated in negative ionization mode (capillary voltage, 4000 eV; temp, 350°C;

nebulizing gas, 60 psi; dry gas 12 L × min^{-1}) and an Agilent 1290 DAD, detection 190–400 nm. As reference compounds for mass calibration, Hexakis (1H,1H,3H tetrafluoropropoxy)phosphazene, purine, and ammonium trifluoroacetate were used according to the manufacturer's (Agilent Technologies) instructions. For data analysis and processing, we used the LC/MSD Trap Software 5.2 (Bruker Daltonik, Billerica, MA, USA, https://www.bruker.com/) and the MASSHUNTER Workstation Software vB.07.00, 2015 (Agilent Technologies). The same separation methods and protocols were used with both LC–MS systems: For analyses of *N. benthamiana* extracts, LC was performed on an Agilent ZORBAX SB-C18 column (50 × 4.6 mm, 1.8 μm particle size) using water with 0.2%(v/v) formic acid (solvent A) against a gradient of acetonitrile with 0.2%(v/v) formic acid (solvent B). The gradient profile was: 0–0.5 min of 95% (v/v) solvent A, 0.5–5 min from 5 to 20% (v/v) solvent B, 5–7 min of 90% (v/v) of solvent B, followed by equilibration to 95% (v/v) solvent A for 3 min. For analyses of products from enzyme *in vitro* assays and yeast *in vivo* assays, the concentration of formic acid in both solvents was increased to 2% (v/v) while keeping all other LC conditions the same. The injection volumes of 1 μl and 5 μl were used for samples analyzed on the Agilent 6530 Accurate Mass Q-TOF and for samples analyzed on the Agilent MSD Trap XCT-Plus MS, respectively. For additional details see Table S5.

Statistical analyses

All statistical analyses were performed in GRAPHPAD v9.5.1 (Boston, MA, USA, https://www.graphpad.com). One-way analysis of variance (one-way ANOVA) was used after testing for normal distribution to compare the accumulation of metabolites between multiple samples and Tukey's HSD test was used to test for differences in individual treatments. For comparing two groups of samples the unpaired *t*-test was used.

AUTHOR CONTRIBUTIONS

The research described here was conceived by SI, LK, and JB as part of a larger project on MbA biosynthesis and bioengineering. LK and SI planned experiments with guidance from JB, LK, FS, DG, GLP, and SJ performed experiments. LK and SI analyzed data. LK, SI, and JB wrote the manuscript. All authors commented on and approved the final version. JB, LK, and SI secured funding.

ACKNOWLEDGEMENTS

We thank Dr. Carol Ritland for project management, Macaire Man Saint Yuen for bioinformatics support, and Dr. Jessica Garzke for helpful discussions on the statistics. The research was supported with funds from the GlycoNet NCE (Strategic Initiatives grant to J.B.; Advanced Training Opportunity Program grants to L.K. and J.B. to support G.L.P. and to S.I. and J.B. to support D.G.) and with funds from the Natural Sciences and Engineering Research Council of Canada (NSERC) to J.B. and S.I. was supported by the Alexander von Humboldt Foundation through a Feodor Lynen Research Fellowship. We thank Dr. Doug Cossar and Dr. Michael McLean from PlantForm Corporation for many insightful discussions.

CONFLICT OF INTEREST STATEMENT

The authors declare that the research was conducted in the absence of any commercial or financial relationship that could be construed as a potential conflict of interest.

DATA AVAILABILITY STATEMENT

All sequences newly described in this study have been deposited in NCBI GenBank with the following accession numbers: *CcHCT* (*CcAT8*), OR245368; *CcAT9*, OR245367; *CcAT10*, OR245366; *CcEST1*, OR245363; *CcCSE* (*CcEST2*), OR245364; *CcEST3*, OR245365; *CcC3′H* (CYP98A216), OR245362. Transcriptome data from *Crocosmia* × *crocosmiiflora* has been published previously in the NCBI/GenBank under the project number PRJNA389589 (SRP108844).

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Figure S1. Alignment of the three montbretia candidate HCTs, CcAT8, CcAT9, and CcAT10, with the *A. thaliana* HCT(AtHCT).

Figure S2. Expanded phylogenetic trees of the BAHD acyltransferases (a) and plant esterases of the alpha/beta hydrolase enzyme family (b).

Figure S3. Transcript expression patterns of putative caffeic acid biosynthetic genes and previously characterized MbA pathway genes.

Figure S4. Alignment of the montbretia candidate CcC3′H sequence with the C3′H from *A. thaliana* (AtC3′H).

Figure S5. Alignment of the montbretia candidate CcEST sequences with the CSE from *A. thaliana* (AtCSE).

Figure S6. Levels of MbA, MbB, and MbC in *N. benthamiana* affected by co-expression of phenylpropanoid pathway genes.

Figure S7. Levels of (a) caffeic acid, (b) caffeoylshikimate, and (c) chlorogenic acid (caffeoylquinate) in *N. benthamiana* affected by co-expression of phenylpropanoid pathway genes together with a set of 10 genes required for MbA production.

Figure S8. Montbretin formation in *N. bentamiana* co-expressing MbA genes and different combinations of phenylpropanoid pathway genes.

Figure S9. SDS-PAGE of heterologously expressed CcAT8 (CcHCT) and CcEST2 (CcCSE).

Figure S10. Empty vector controls for CcAT8 assays.

Figure S11. CcEST2 (CcCSE) activity assays and negative control reactions.

Figure S12. CcEST2 enzyme assay with coumaroylshikimate.

Figure S13. Design of multigene constructs for yeast expression.

Figure S14. Controls and additional experiments to functional validation and reconstruction of the montbretia shikimate shunt in yeast.

Figure S15. Design of level 1 (L1) and level 2 (L2) constructs for MGC transformation and transient expression in *Nicotiana benthamiana*.

Figure S16. Representative extracted ion chromatograms of MbA, MbB, and MbC of *N. benthamiana* plants infiltrated with MGCs or single gene plasmids.

Figure S17. Representative *Nicotiana benthamiana* plant used for transformation.

Table S1. Co-expression *analysis* of montbretia genes.

Table S2. A. thaliana sequences used for the identification of candidate genes.

Table S3. Oligonucleotide sequences used in this study.

Table S4. Plasmids created in this study.

Table S5. Compounds identified in this study.

File S1. BAHD acyltransferase sequences used for the phylogenetic analyses.

File S2. Esterase sequences used for the phylogenetic analyses.

File S3. Montbretin accumulation in montbretia corms.

REFERENCES

Alber, A.V., Renault, H., Basilio-Lopes, A., Bassard, J.-E., Liu, Z., Ullmann, P. *et al.* (2019) Evolution of coumaroyl conjugate 3-hydroxylases in land plants: lignin biosynthesis and defense. *The Plant Journal*, **99**, 924–936. Available from: https://doi.org/10.1111/tpj.14373

Andersen, R., Brayer, G., Tarling, A.C., Withers, S. & Woods, K. (2009) Alpha-amylase inhibitors: the montbretins and uses thereof. Available from: https://patents.google.com/patent/CA2738834A1/en?oq=Montbretin+A [Accessed 20th August 2021]

Asada, Y., Hirayama, Y. & Furuya, T. (1988) Acylated flavonols from *Crocosmia crocosmiiflora*. *Phytochemistry*, **27**, 1497–1501. Available from: https://doi.org/10.1016/0031-9422(88)80223-1

Bally, J., Jung, H., Mortimer, C., Naim, F., Philips, J.G., Hellens, R. *et al.* (2018) The rise and rise of *Nicotiana benthamiana*: a plant for all reasons. *Annual Review of Phytopathology*, **56**, 405–426. Available from: https://doi.org/10.1146/annurev-phyto-080417-050141

Barros, J., Escamilla-Trevino, L., Song, L., Rao, X., Serrani-Yarce, J.C., Palacios, M.D. *et al.* (2019) 4-coumarate 3-hydroxylase in the lignin biosynthesis pathway is a cytosolic ascorbate peroxidase. *Nature Communications*, **10**, 1994. Available from: https://doi.org/10.1038/s41467-019-10082-7

Butelli, E., Titta, L., Giorgio, M., Mock, H.-P., Matros, A., Peterek, S. *et al.* (2008) Enrichment of tomato fruit with health-promoting anthocyanins by expression of select transcription factors. *Nature Biotechnology*, **26**, 1301–1308. Available from: https://doi.org/10.1038/nbt.1506

Calgaro-Kozina, A., Vuu, K.M., Keasling, J.D., Loqué, D., Sattely, E.S. & Shih, P.M. (2020) Engineering plant synthetic pathways for the biosynthesis of novel antifungals. *ACS Central Science*, **6**, 1394–1400. Available from: https://doi.org/10.1021/acscentsci.0c00241

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K. *et al.* (2009) BLAST+: architecture and applications. *BMC Bioinformatics*, **10**, 421. Available from: https://doi.org/10.1186/1471-2105-10-421

Cardenas, C.L., Costa, M.A., Laskar, D.D., Moinuddin, S.G.A., Lee, C., Davin, L.B. *et al.* (2021) RNAi modulation of chlorogenic acid and lignin deposition in *Nicotiana tabacum* and insufficient compensatory metabolic cross-talk. *Journal of Natural Products*, **84**, 694–706. Available from: https://doi.org/10.1021/acs.jnatprod.1c00054

Chen, H.-C., Li, Q., Shuford, C.M., Liu, J., Muddiman, D.C., Sederoff, R.R. *et al.* (2011) Membrane protein complexes catalyze both 4- and 3-hydroxylation of cinnamic acid derivatives in monolignol biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 21253–21258. Available from: https://doi.org/10.1073/pnas.1116416109

Chen, H.-C., Song, J., Williams, C.M., Shuford, C.M., Liu, J., Wang, J.P. *et al.* (2013) Monolignol pathway 4-coumaric acid:coenzyme A ligases in *Populus trichocarpa*: novel specificity, metabolic regulation, and simulation of coenzyme A ligation fluxes. *Plant Physiology*, **161**, 1501–1516. Available from: https://doi.org/10.1104/pp.112.210971

Chezem, W.R. & Clay, N.K. (2016) Regulation of plant secondary metabolism and associated specialized cell development by MYBs and bHLHs. *Phytochemistry*, **131**, 26–43. Available from: https://doi.org/10.1016/j.phytochem.2016.08.006

Costa, M.A., Bedgar, D.L., Moinuddin, S.G.A., Kim, K.-W., Cardenas, C.L., Cochrane, F.C. *et al.* (2005) Characterization in vitro and in vivo of the putative multigene 4-coumarate:CoA ligase network in *Arabidopsis*: syringyl lignin and sinapate/sinapyl alcohol derivative formation. *Phytochemistry*, **66**, 2072–2091. Available from: https://doi.org/10.1016/j.phytochem.2005.06.022

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

De La Peña, R. & Sattely, E.S. (2021) Rerouting plant terpene biosynthesis enables momilactone pathway elucidation. *Nature Chemical Biology*, **17**, 205–212. Available from: https://doi.org/10.1038/s41589-020-00669-3

de Vries, L., Brouckaert, M., Chanoca, A., Kim, H., Regner, M.R., Timokhin, V.I. *et al.* (2021) CRISPR-Cas9 editing of CAFFEOYL SHIKIMATE ESTERASE 1 and 2 shows their importance and partial redundancy in lignification in *Populus tremula* × *P. alba*. *Plant Biotechnology Journal*, **19**, 2221–2234. Available from: https://doi.org/10.1111/pbi.13651

de Vries, S., Fürst-Jansen, J.M.R., Irisarri, I., Dhabalia Ashok, A., Ischebeck, T., Feussner, K. *et al.* (2021) The evolution of the phenylpropanoid pathway entailed pronounced radiations and divergences of enzyme families. *The Plant Journal*, **107**, 975–1002. Available from: https://doi.org/10.1111/tpj.15387

Diabetes Canada. (2020) *Food security and diabetes – Diabetes Canada*. Ottawa: Diabetes Canada. Available from: https://www.diabetes.ca/advocacy---policies/our-policy-positions/food-security-and-diabetes [Accessed 27th February 2023]

Douglas, C.J., Ellard, M., Hauffe, K.D., Molitor, E., de Sá, M.M., Reinold, S. *et al.* (1992) General phenylpropanoid metabolism: regulation by environmental and developmental signals. In: Stafford, H.A. & Ibrahim, R.K. (Eds.) *Phenolic metabolism in plants* Recent advances in phytochemistry. Boston, MA: Springer US, pp. 63–89. Available from: https://doi.org/10.1007/978-1-4615-3430-3_3

Dudley, Q.M., Jo, S., Guerrero, D.A.S., Chhetry, M., Smedley, M.A., Harwood, W.A. *et al.* (2022) Reconstitution of monoterpene indole alkaloid biosynthesis in genome engineered *Nicotiana benthamiana*. *Communications Biology*, **5**, 1–12. Available from: https://doi.org/10.1038/s42003-022-03904-w

Duke, S.O. & Vaughn, K.C. (1982) Lack of involvement of polyphenol oxidase in ortho-hydroxylation of phenolic compounds in mung bean seedlings. *Physiologia Plantarum*, **54**, 381–385. Available from: https://doi.org/10.1111/j.1399-3054.1982.tb00696.x

Engler, C., Youles, M., Gruetzner, R., Ehnert, T.-M., Werner, S., Jones, J.D.G. *et al.* (2014) A golden gate modular cloning toolbox for plants. *ACS Synthetic Biology*, **3**, 839–843. Available from: https://doi.org/10.1021/sb4001504

Escamilla-Treviño, L.L., Shen, H., Hernandez, T., Yin, Y., Xu, Y. & Dixon, R.A. (2014) Early lignin pathway enzymes and routes to chlorogenic acid in switchgrass (*Panicum virgatum* L.). *Plant Molecular Biology*, **84**, 565–576. Available from: https://doi.org/10.1007/s11103-013-0152-y

FAO, IFAD, UNICEF, WFP & WHO. (2021) *The State of Food Security and Nutrition in the World 2021*. Rome, Italy: FAO, IFAD, UNICEF, WFP and WHO. Available from: https://doi.org/10.4060/cb4474en

Franke, R., Humphreys, J.M., Hemm, M.R., Denault, J.W., Ruegger, M.O., Cusumano, J.C. *et al.* (2002) The *Arabidopsis* REF8 gene encodes the 3-hydroxylase of phenylpropanoid metabolism. *The Plant Journal*, **30**, 33–45. Available from: https://doi.org/10.1046/j.1365-313X.2002.01266.x

Fraser, C.M. & Chapple, C. (2011) The phenylpropanoid pathway in *Arabidopsis*. *The Arabidopsis Book/American Society of Plant Biologists*, **9**, e0152. Available from: https://doi.org/10.1199/tab.0152

Geu-Flores, F., Nour-Eldin, H.H., Nielsen, M.T. & Halkier, B.A. (2007) USER fusion: a rapid and efficient method for simultaneous fusion and cloning of multiple PCR products. *Nucleic Acids Research*, **35**, e55. Available from: https://doi.org/10.1093/nar/gkm106

Gou, M., Ran, X., Martin, D.W. & Liu, C.-J. (2018) The scaffold proteins of lignin biosynthetic cytochrome P450 enzymes. *Nature Plants*, **4**, 299–310. Available from: https://doi.org/10.1038/s41477-018-0142-9

Grzech, D., Hong, B., Caputi, L., Sonawane, P.D. & O'Connor, S.E. (2022) Engineering the biosynthesis of late-stage vinblastine precursors precondylocarpine acetate, catharanthine, tabersonine in *Nicotiana benthamiana*. *ACS Synthetic Biology*, **12**, 27–34. Available from: https://doi.org/10.1021/acssynbio.2c00434

Ha, C.M., Escamilla-Trevino, L., Yarce, J.C.S., Kim, H., Ralph, J., Chen, F. *et al.* (2016) An essential role of caffeoyl shikimate esterase in monolignol biosynthesis in *Medicago truncatula*. *The Plant Journal*, **86**, 363–375. Available from: https://doi.org/10.1111/tpj.13177

Harding, S.A., Leshkevich, J., Chiang, V.L. & Tsai, C.-J. (2002) Differential substrate inhibition couples kinetically distinct 4-coumarate:coenzyme A ligases with spatially distinct metabolic roles in quaking aspen. *Plant Physiology*, **128**, 428–438. Available from: https://doi.org/10.1104/pp.010603

Hoffmann, L., **Besseau, S.**, **Geoffroy, P.**, **Ritzenthaler, C.**, **Meyer, D.**, **Lapierre, C.** ***et al.*** (2005) Acyltransferase-catalysed p-coumarate ester formation is a committed step of lignin biosynthesis. *Plant Biosystems – An International Journal Dealing with all Aspects of Plant Biology*, **139**, 50–53.

Hoffmann, L., **Maury, S.**, **Martz, F.**, **Geoffroy, P.** & **Legrand, M.** (2003) Purification, cloning, and properties of an acyltransferase controlling shikimate and quinate ester intermediates in phenylpropanoid metabolism. *The Journal of Biological Chemistry*, **278**, 95–103. Available from: https://doi.org/10.1074/jbc.M209362200

Incha, M.R., **Thompson, M.G.**, **Blake-Hedges, J.M.**, **Liu, Y.**, **Pearson, A.N.**, **Schmidt, M.** ***et al.*** (2019) Leveraging host metabolism for bisdemethoxycurcumin production in *Pseudomonas putida*. *Metabolic Engineering Communications*, **10**, e00119. Available from: https://doi.org/10.1016/j.mec.2019.e00119

Irmisch, S., **Jancsik, S.**, **Man Saint Yuen, M.**, **Madilao, L.L.** & **Bohlmann, J.** (2020) Complete biosynthesis of the anti-diabetic plant metabolite montbretin A. *Plant Physiology*, **184**, 97–109. Available from: https://doi.org/10.1104/pp.20.00522

Irmisch, S., **Jancsik, S.**, **Yuen, M.M.S.**, **Madilao, L.L.** & **Bohlmann, J.** (2019) Biosynthesis of the anti-diabetic metabolite montbretin A: glucosylation of the central intermediate mini-MbA. *The Plant Journal*, **100**, 879–891. Available from: https://doi.org/10.1111/tpj.14493

Irmisch, S., **Jo, S.**, **Roach, C.R.**, **Jancsik, S.**, **Yuen, M.M.S.**, **Madilao, L.L.** ***et al.*** (2018) Discovery of UDP-glycosyltransferases and BAHD-acyltransferases involved in the biosynthesis of the antidiabetic plant metabolite montbretin A. *Plant Cell*, **30**, 1864–1886. Available from: https://doi.org/10.1105/tpc.18.00406

Irmisch, S., **Ruebsam, H.**, **Jancsik, S.**, **Man Saint Yuen, M.**, **Madilao, L.L.** & **Bohlmann, J.** (2019) Flavonol biosynthesis genes and their use in engineering the plant antidiabetic metabolite montbretin A. *Plant Physiology*, **180**, 1277–1290. Available from: https://doi.org/10.1104/pp.19.00254

Katoh, K., **Misawa, K.**, **Kuma, K.** & **Miyata, T.** (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*, **30**, 3059–3066.

Kim, I.A., **Kim, B.-G.**, **Kim, M.** & **Ahn, J.-H.** (2012) Characterization of hydroxycinnamoyltransferase from rice and its application for biological synthesis of hydroxycinnamoyl glycerols. *Phytochemistry*, **76**, 25–31. Available from: https://doi.org/10.1016/j.phytochem.2011.12.015

Kim, S.S., **Wengier, D.L.**, **Ragland, C.J.** & **Sattely, E.S.** (2022) Transcriptional reactivation of lignin biosynthesis for the heterologous production of etoposide aglycone in *Nicotiana benthamiana*. *ACS Synthetic Biology*, **11**, 3379–3387. Available from: https://doi.org/10.1021/acssynbio.2c00289

Kojima, M. & **Takeuchi, W.** (1989) Detection and characterization of p-coumaric acid hydroxylase in mung bean, *Vigna mungo*, seedlings. *Journal of Biochemistry*, **105**, 265–270. Available from: https://doi.org/10.1093/oxfordjournals.jbchem.a122651

Koncz, C. & **Schell, J.** (1986) The promoter of TL-DNA gene 5 controls the tissue-specific expression of chimaeric genes carried by a novel type of agrobacterium binary vector. *Molecular and General Genetics*, **204**, 383–396. Available from: https://doi.org/10.1007/BF00331014

Kriegshauser, L., **Knosp, S.**, **Grienenberger, E.**, **Tatsumi, K.**, **Gütle, D.D.**, **Sørensen, I.** ***et al.*** (2021) Function of the HYDROXYCINNAMOYL-CoA: SHIKIMATE HYDROXYCINNAMOYL TRANSFERASE is evolutionarily conserved in embryophytes. *Plant Cell*, **33**, 1472–1491. Available from: https://doi.org/10.1093/plcell/koab044

Kruse, L.H., **Fehr, B.**, **Chobirko, J.D.** & **Moghe, G.D.** (2023) Phylogenomic analyses across land plants reveals motifs and coexpression patterns useful for functional prediction in the BAHD acyltransferase family. *Frontiers in Plant Science*, **14**, 1067613. Available from: https://doi.org/10.3389/fpls.2023.1067613

Kruse, L.H., **Weigle, A.T.**, **Irfan, M.**, **Martínez-Gómez, J.**, **Chobirko, J.D.**, **Schaffer, J.E.** ***et al.*** (2022) Orthology-based analysis helps map evolutionary diversification and predict substrate class use of BAHD acyltransferases. *The Plant Journal*, **111**, 1453–1468. Available from: https://doi.org/10.1111/tpj.15902

Lau, W. & **Sattely, E.S.** (2015) Six enzymes from mayapple that complete the biosynthetic pathway to the etoposide aglycone. *Science*, **349**, 1224–1228. Available from: https://doi.org/10.1126/science.aac7202

Lee, M.E., **DeLoache, W.C.**, **Cervantes, B.** & **Dueber, J.E.** (2015) A highly characterized yeast toolkit for modular, multipart assembly. *ACS Synthetic Biology*, **4**, 975–986. Available from: https://doi.org/10.1021/sb500366v

Letunic, I. & **Bork, P.** (2021) Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Research*, **49**, W293–W296. Available from: https://doi.org/10.1093/nar/gkab301

Liu, D., **Sica, M.S.**, **Mao, J.**, **Chao, L.F.-I.** & **Siewers, V.** (2022) A p-coumaroyl-CoA biosensor for dynamic regulation of naringenin biosynthesis in *Saccharomyces cerevisiae*. *ACS Synthetic Biology*, **11**, 3228–3238. Available from: https://doi.org/10.1021/acssynbio.2c00111

Liu, L., **Liu, H.**, **Zhang, W.**, **Yao, M.**, **Li, B.**, **Liu, D.** ***et al.*** (2019) Engineering the biosynthesis of caffeic acid in *Saccharomyces cerevisiae* with heterologous enzyme combinations. *Engineering*, **5**, 287–295. Available from: https://doi.org/10.1016/j.eng.2018.11.029

Malhotra, K., **Subramaniyan, M.**, **Rawat, K.**, **Kalamuddin, M.**, **Qureshi, M.I.**, **Malhotra, P.** ***et al.*** (2016) Compartmentalized metabolic engineering for artemisinin biosynthesis and effective malaria treatment by oral delivery of plant cells. *Molecular Plant*, **9**, 1464–1477. Available from: https://doi.org/10.1016/j.molp.2016.09.013

Moghe, G., **Kruse, L.H.**, **Petersen, M.**, **Scossa, F.**, **Fernie, A.R.**, **Gaquerel, E.** ***et al.*** (2023) BAHD company: the ever-expanding roles of the BAHD acyltransferase gene family in plants. *Annual Review of Plant Biology*, **74**, null–194. Available from: https://doi.org/10.1146/annurev-arplant-062922-050122

Montague, N.P., **Thuenemann, E.C.**, **Saxena, P.**, **Saunders, K.**, **Lenzi, P.** & **Lomonossoff, G.P.** (2011) Recent advances of Cowpea mosaic virus-based particle technology. *Human Vaccines*, **7**, 383–390. Available from: https://doi.org/10.4161/hv.7.3.14989

Nett, R.S. & **Sattely, E.S.** (2021) Total biosynthesis of the tubulin-binding alkaloid colchicine. *Journal of the American Chemical Society*, **143**, 19454–19465. Available from: https://doi.org/10.1021/jacs.1c08659

Nguyen, L.-T., **Schmidt, H.A.**, **von Haeseler, A.** & **Minh, B.Q.** (2015) IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*, **32**, 268–274. Available from: https://doi.org/10.1093/molbev/msu300

Niggeweg, R., **Michael, A.J.** & **Martin, C.** (2004) Engineering plants with increased levels of the antioxidant chlorogenic acid. *Nature Biotechnology*, **22**, 746–754. Available from: https://doi.org/10.1038/nbt966

Nour-Eldin, H.H., **Hansen, B.G.**, **Nørholm, M.H.H.**, **Jensen, J.K.** & **Halkier, B.A.** (2006) Advancing uracil-excision based cloning towards an ideal technique for cloning PCR fragments. *Nucleic Acids Research*, **34**, e122. Available from: https://doi.org/10.1093/nar/gkl635

Ong, K.L., **Stafford, L.K.**, **McLaughlin, S.A.**, **Boyko, E.J.**, **Vollset, S.E.**, **Smith, A.E.** ***et al.*** (2023) Global, regional, and national burden of diabetes from 1990 to 2021, with projections of prevalence to 2050: a systematic analysis for the Global Burden of Disease Study 2021. *The Lancet*, **402**, 203–234. Available from: https://doi.org/10.1016/S0140-6736(23)01301-6

Pandey, A., **Misra, P.**, **Khan, M.P.**, **Swarnkar, G.**, **Tewari, M.C.**, **Bhambhani, S.** ***et al.*** (2014) Co-expression of *Arabidopsis* transcription factor, AtMYB12, and soybean isoflavone synthase, GmIFS1, genes in tobacco leads to enhanced biosynthesis of isoflavones and flavonols resulting in osteoprotective activity. *Plant Biotechnology Journal*, **12**, 69–80. Available from: https://doi.org/10.1111/pbi.12118

Parke, D. & **Ornston, L.N.** (2004) Toxicity caused by hydroxycinnamoyl-coenzyme A thioester accumulation in mutants of *Acinetobacter sp.* strain ADP1. *Applied and Environmental Microbiology*, **70**, 2974–2983. Available from: https://doi.org/10.1128/AEM.70.5.2974-2983.2004

Reed, J. & **Osbourn, A.** (2018) Engineering terpenoid production through transient expression in *Nicotiana benthamiana*. *Plant Cell Reports*, **37**, 1431–1441. Available from: https://doi.org/10.1007/s00299-018-2296-3

Reed, J., **Stephenson, M.J.**, **Miettinen, K.**, **Brouwer, B.**, **Leveau, A.**, **Brett, P.** ***et al.*** (2017) A translational synthetic biology platform for rapid access to gram-scale quantities of novel drug-like molecules. *Metabolic Engineering*, **42**, 185–193. Available from: https://doi.org/10.1016/j.ymben.2017.06.012

Renault, H., **Werck-Reichhart, D.** & **Weng, J.-K.** (2019) Harnessing lignin evolution for biotechnological applications. *Current Opinion in Biotechnology*, **56**, 105–111. Available from: https://doi.org/10.1016/j.copbio.2018.10.011

Royal Botanical Garden, Kew. (2023) The international plant names index and world checklist of vascular plants 2023. *Plants World Online*.

Available from: http://powo.science.kew.org/taxon/urn:lsid:ipni.org:names:20292-1 [Accessed 4th July 2023]

Saleme, M.d.L.S., **Cesarino, I.**, **Vargas, L.**, **Kim, H.**, **Vanholme, R.**, **Goeminne, G.** ***et al.*** (2017) Silencing CAFFEOYL SHIKIMATE ESTERASE affects lignification and improves saccharification in poplar. *Plant Physiology*, **175**, 1040–1057. Available from: https://doi.org/10.1104/pp.17.00920

Schillberg, S. & **Finnern, R.** (2021) Plant molecular farming for the production of valuable proteins – critical evaluation of achievements and future challenges. *Journal of Plant Physiology*, **258–259**, 153359. Available from: https://doi.org/10.1016/j.jplph.2020.153359

Schoch, G., **Goepfert, S.**, **Morant, M.**, **Hehn, A.**, **Meyer, D.**, **Ullmann, P.** ***et al.*** (2001) CYP98A3 from *Arabidopsis thaliana* is a 3′-hydroxylase of phenolic esters, a missing link in the phenylpropanoid pathway. *The Journal of Biological Chemistry*, **276**, 36566–36574. Available from: https://doi.org/10.1074/jbc.M104047200

Schultz, B.J., **Kim, S.Y.**, **Lau, W.** & **Sattely, E.S.** (2019) Total biosynthesis for milligram-scale production of etoposide intermediates in a plant chassis. *Journal of the American Chemical Society*, **141**, 19231–19235. Available from: https://doi.org/10.1021/jacs.9b10717

Serrani-Yarce, J.C., **Escamilla-Trevino, L.**, **Barros, J.**, **Gallego-Giraldo, L.**, **Pu, Y.**, **Ragauskas, A.** ***et al.*** (2021) Targeting hydroxycinnamoyl CoA: shikimate hydroxycinnamoyl transferase for lignin modification in *Brachypodium distachyon*. *Biotechnology for Biofuels*, **14**, 50. Available from: https://doi.org/10.1186/s13068-021-01905-1

Stafford, H.A. & **Dresler, S.** (1972) 4-hydroxycinnamic acid hydroxylase and polyphenolase activities in *Sorghum vulgare*. *Plant Physiology*, **49**, 590–595.

Stephenson, M.J., **Reed, J.**, **Patron, N.J.**, **Lomonossoff, G.P.** & **Osbourn, A.** (2020) 6.11 – Engineering tobacco for plant natural product production. In: (Ben) Liu, H.-W. & Begley, T.P. (Eds.) *Comprehensive Natural Products III*. Oxford: Elsevier, pp. 244–262. Available from: https://doi.org/10.1016/B978-0-12-409547-2.14724-9

Sunstrum, F.G., **Liu, H.L.**, **Jancsik, S.**, **Madilao, L.L.**, **Bohlmann, J.** & **Irmisch, S.** (2021) 4-coumaroyl-CoA ligases in the biosynthesis of the anti-diabetic metabolite montbretin A. *PLoS One*, **16**, e0257478. Available from: https://doi.org/10.1371/journal.pone.0257478

Tarling, C.A., **Woods, K.**, **Zhang, R.**, **Brastianos, H.C.**, **Brayer, G.D.**, **Andersen, R.J.** ***et al.*** (2008) The search for novel human pancreatic alpha-amylase inhibitors: high-throughput screening of terrestrial and marine natural product extracts. *ChemBioChem*, **9**, 433–438. Available from: https://doi.org/10.1002/cbic.200700470

The Lancet. (2023) Diabetes: a defining disease of the 21st century. *The Lancet*, **401**, 2087. Available from: https://doi.org/10.1016/S0140-6736(23)01296-5

The UniProt Consortium. (2023) UniProt: the universal protein knowledgebase in 2023. *Nucleic Acids Research*, **51**, D523–D531. Available from: https://doi.org/10.1093/nar/gkac1052

Tschofen, M., **Knopp, D.**, **Hood, E.** & **Stöger, E.** (2016) Plant molecular farming: much more than medicines. *Annual Review of Analytical Chemistry*, **9**, 271–294. Available from: https://doi.org/10.1146/annurev-anchem-071015-041706

Vanholme, R., **Cesarino, I.**, **Rataj, K.**, **Xiao, Y.**, **Sundin, L.**, **Goeminne, G.** ***et al.*** (2013) Caffeoyl shikimate esterase (CSE) is an enzyme in the lignin biosynthetic pathway in *Arabidopsis*. *Science*, **341**, 1103–1106. Available from: https://doi.org/10.1126/science.1241602

Vanholme, R., **De Meester, B.**, **Ralph, J.** & **Boerjan, W.** (2019) Lignin biosynthesis and its integration into metabolism. *Current Opinion in Biotechnology*, **56**, 230–239. Available from: https://doi.org/10.1016/j.copbio.2019.02.018

Volpi e Silva, N., **Mazzafera, P.** & **Cesarino, I.** (2019) Should I stay or should I go: are chlorogenic acids mobilized towards lignin biosynthesis? *Phytochemistry*, **166**, 112063. Available from: https://doi.org/10.1016/j.phytochem.2019.112063

Walker, A.M., **Hayes, R.P.**, **Youn, B.**, **Vermerris, W.**, **Sattler, S.E.** & **Kang, C.** (2013) Elucidation of the structure and reaction mechanism of *Sorghum* hydroxycinnamoyltransferase and its structural relationship to other coenzyme A-dependent transferases and synthases. *Plant Physiology*, **162**, 640–651. Available from: https://doi.org/10.1104/pp.113.217836

Wang, J.P., **Naik, P.P.**, **Chen, H.-C.**, **Shi, R.**, **Lin, C.-Y.**, **Liu, J.** ***et al.*** (2014) Complete proteomic-based enzyme reaction and inhibition kinetics reveal how monolignol biosynthetic enzyme families affect metabolic flux and lignin in *Populus trichocarpa*. *Plant Cell*, **26**, 894–914. Available from: https://doi.org/10.1105/tpc.113.120881

Wang, Q., **Reddy, V.A.**, **Panicker, D.**, **Mao, H.-Z.**, **Kumar, N.**, **Rajan, C.** ***et al.*** (2016) Metabolic engineering of terpene biosynthesis in plants using a trichome-specific transcription factor MsYABBY5 from spearmint (*Mentha spicata*). *Plant Biotechnology Journal*, **14**, 1619–1632. Available from: https://doi.org/10.1111/pbi.12525

Watanabe, B., **Kirikae, H.**, **Koeduka, T.**, **Takeuchi, Y.**, **Asai, T.**, **Naito, Y.** ***et al.*** (2018) Synthesis and inhibitory activity of mechanism-based 4-coumaroyl-CoA ligase inhibitors. *Bioorganic & Medicinal Chemistry*, **26**, 2466–2474. Available from: https://doi.org/10.1016/j.bmc.2018.04.006

Weber, E., **Engler, C.**, **Gruetzner, R.**, **Werner, S.** & **Marillonnet, S.** (2011) A modular cloning system for standardized assembly of multigene constructs. *PLoS One*, **6**, e16765. Available from: https://doi.org/10.1371/journal.pone.0016765

Weng, J.-K. & **Chapple, C.** (2010) The origin and evolution of lignin biosynthesis. *The New Phytologist*, **187**, 273–285. Available from: https://doi.org/10.1111/j.1469-8137.2010.03327.x

Werner, S., **Engler, C.**, **Weber, E.**, **Gruetzner, R.** & **Marillonnet, S.** (2012) Fast track assembly of multigene constructs using Golden Gate cloning and the MoClo system. *Bioengineered Bugs*, **3**, 38–43. Available from: https://doi.org/10.4161/bbug.3.1.18223

Williams, L.K., **Zhang, X.**, **Caner, S.**, **Tysoe, C.**, **Nguyen, N.T.**, **Wicki, J.** ***et al.*** (2015) The amylase inhibitor montbretin A reveals a new glycosidase inhibition motif. *Nature Chemical Biology*, **11**, 691–696. Available from: https://doi.org/10.1038/nchembio.1865

Xie, M., **Zhang, J.**, **Tschaplinski, T.J.**, **Tuskan, G.A.**, **Chen, J.-G.** & **Muchero, W.** (2018) Regulation of lignin biosynthesis and its role in growth-defense tradeoffs. *Frontiers in Plant Science*, **9**, 1427. Available from: https://doi.org/10.3389/fpls.2018.01427

Yuan, L. & **Grotewold, E.** (2015) Metabolic engineering to enhance the value of plants as green factories. *Metabolic Engineering*, **27**, 83–91. Available from: https://doi.org/10.1016/j.ymben.2014.11.005

Yuen, V.G., **Coleman, J.**, **Withers, S.G.**, **Andersen, R.J.**, **Brayer, G.D.**, **Mustafa, S.** ***et al.*** (2016) Glucose lowering effect of montbretin A in Zucker diabetic fatty rats. *Molecular and Cellular Biochemistry*, **411**, 373–381. Available from: https://doi.org/10.1007/s11010-015-2599-4

Zhong, R., **Morrison, W.H., III**, **Negrel, J.** & **Ye, Z.-H.** (1998) Dual methylation pathways in lignin biosynthesis. *Plant Cell*, **10**, 2033–2045. Available from: https://doi.org/10.1105/tpc.10.12.2033

Zhong, R. & **Ye, Z.-H.** (2009) Transcriptional regulation of lignin biosynthesis. *Plant Signaling & Behavior*, **4**, 1028–1034. Available from: https://doi.org/10.4161/psb.4.11.9875

Zhou, P., **Yue, C.**, **Shen, B.**, **Du, Y.**, **Xu, N.** & **Ye, L.** (2021) Metabolic engineering of *Saccharomyces cerevisiae* for enhanced production of caffeic acid. *Applied Microbiology and Biotechnology*, **105**, 5809–5819. Available from: https://doi.org/10.1007/s00253-021-11445-1