COSY catalyzes *trans-cis* isomerization and lactonization in the biosynthesis of coumarins

Ruben Vanholme[1,2,‡], Lisa Sundin[1,2,‡], Keletso Carol Seetso[1,2], Hoon Kim[3], Xinyu Liu[1,2], Jin Li[1,2], Barbara De Meester[1,2], Lennart Hoengenaert[1,2], Geert Goeminne[1,2,4], Kris Morreel[1,2], Jurgen Haustraete[5], Huei-Hsuan Tsai[6,7,8], Wolfgang Schmidt[6,7,9,10], Bartel Vanholme[1,2], John Ralph[3] and Wout Boerjan[1,2,*]

[1]Department of Plant Biotechnology and Bioinformatics, Ghent University, Technologiepark 71, 9052 Gent, Belgium.

[2]Center for Plant Systems Biology, VIB, Technologiepark 71, 9052 Gent, Belgium.

[3]Department of Biochemistry, and the DOE Great Lakes Bioenergy Research Center, the Wisconsin Energy Institute, University of Wisconsin, 1552 University Avenue, Madison, Wisconsin 53726, USA.

[4]VIB Metabolomics Core, Technologiepark 71, 9052 Ghent, Belgium.

[5]Protein Core, VIB-UGent Center for Inflammation Research, VIB, Ghent University, Technologiepark 927, 9052 Gent, Belgium.

[6]Institute of Plant and Microbial Biology, Academia Sinica, Taipei 11529, Taiwan

[7]Molecular and Biological Agricultural Sciences Program, Taiwan International Graduate Program, National Chung Hsing University and Academia Sinica, Taipei 11529, Taiwan

[8]Graduate Institute of Biotechnology, National Chung Hsing University, Taichung 40227, Taiwan

[9]Biotechnology Center, National Chung Hsing University, Taichung 40227, Taiwan

[10]Genome and Systems Biology Degree Program, College of Life Science, National Taiwan University, Taipei 10617, Taiwan

*Correspondence to: wout.boerjan@ugent.vib.be

‡ Contributed equally

Coumarins or 1,2-benzopyrones comprise a large class of secondary metabolites ubiquitously found in the plant kingdom. In many plant species, coumarins are particularly important for iron acquisition and plant defense. Here we show that COUMARIN SYNTHASE (COSY) is a key enzyme in the biosynthesis of coumarins. *Arabidopsis thaliana cosy* mutants are strongly reduced in coumarin levels and accumulate *o*-hydroxyphenylpropanoids instead. Accordingly, *cosy* mutants have reduced iron content and show growth defects when grown under conditions where iron availability is limited. Recombinant COSY is able to produce umbelliferone, esculetin and scopoletin from their respective *o*-hydroxycinnamoyl-CoA thioesters via two reaction steps: a *trans-cis* isomerization followed by a lactonization. This conversion happens partially spontaneously and is catalyzed by light, which explains why the need for an enzyme for this conversion has been overlooked. The combined results show that COSY has an essential role in the biosynthesis of coumarins in organs that are shielded from light, such as roots. These findings provide routes to improving coumarin production in crops or via microbial fermentation.

Introduction

Coumarins are broadly distributed across the plant kingdom where they are suggested to have phytoalexin and allelopathic properties[1-3]. More specifically, the coumarin scopoletin is excreted in the rhizosphere where it inhibits the growth of fungal pathogens, while growth promoting bacteria are tolerant to its antimicrobial effect[4]. In addition, several strategy I plants (i.e., dicots and non-grass monocots) excrete coumarins such as esculetin, scopoletin, fraxetin and sideretin into the rhizosphere to participate in iron acquisition[5-10].

The coumarin biosynthetic pathway branches off from the phenylpropanoid biosynthetic pathway at the level of the (hydroxy)cinnamoyl-CoAs. The first committed step in the biosynthesis of coumarins is the hydroxylation at the position *ortho* to the aromatic ring's aliphatic side-chain via 2-oxoglutarate-dependent dioxygenase activity (Fig. 1)[11-13]. Experiments with *p*-coumaroyl-CoA 2′-hydroxylase (C2′H) from *Ruta graveolen*s, feruloyl-CoA 6′-hydroxylase 1 (F6′H1) from *Arabidopsis* and two Ib enzymes from *Ipomoea batatas* suggested that the subsequent *trans-cis* isomerization and lactonization to form the coumarin core structures occur solely spontaneously[11-14]. Nevertheless, the lactonization of the coumarin core structures reveals a striking resemblance to an esterification-type reaction mediated by BAHD acyltransferases. BAHD family members utilize activated CoA thioesters as acyl-donors and an alcohol or amine as the acyl acceptor to form an ester or amide[15]. Here, we report the discovery of COUMARIN SYNTHASE (COSY), a BAHD acyltransferase responsible for catalyzing the *trans-cis* isomerization and lactonization in coumarin biosynthesis, and the first described BAHD acyltransferase to act on a single substrate.

Results

We selected *COSY* (AT1G28680) from a co-expression analysis designed to identify genes involved in lignin biosynthesis[16]. Of thirteen genes so identified, eleven are now established shikimate, general phenylpropanoid, and monolignol specific biosynthetic genes. The two remaining genes, *COSY* and another gene encoding a methyltransferase, had no known role[16]. *COSY* is also moderately co-expressed with genes involved in lignin biosynthesis in a set of general phenylpropanoid and monolignol specific pathway mutants (Supplementary Fig. 1)[17]. Notably, lignin and coumarin biosynthetic pathways share many genes and intermediates, and both pathways are thus largely intertwined. COSY is a member of the BAHD acyltransferase family

that has about 55 members in *Arabidopsis*[18]; COSY is a member of subclade V, a subclade in which several identified proteins act on benzoyl- or hydroxycinnamoyl-CoAs[15].

According to external datasets, *COSY* is mainly expressed in roots (Supplementary Fig. 2). This expression pattern was confirmed by our experiments; *Arabidopsis* lines transformed with a construct in which 1.88 kb of the *COSY* promoter was fused to *GUS* and *GFP*, revealed GUS activity preferentially in the roots in 8-day- and 3-month-old plants (Supplementary Fig. 3). In roots, GFP fluorescence was strongest in the endodermis and pericycle, but was also observed in epidermis, cortex, stele, columella and lateral root cap cells (Supplementary Fig. 3). To study the function of COSY, three *Arabidopsis* T-DNA insertion alleles were analyzed; *cosy-1* and *cosy-2* in ecotype Nössen, and *cosy-3* in ecotype Landsberg *erecta*. *cosy-1* and *cosy-2* have an insertion in the first exon, whereas that of *cosy-3* is in the 5′-UTR (Supplementary Fig. 4a). Expression analysis using RT-qPCR showed higher residual *COSY* expression in *cosy-3* than in *cosy-1* and *cosy-2* (Supplementary Fig. 4b).

To potentially reveal the substrate and product of COSY, metabolic extracts of the roots and inflorescence stems from *cosy-1* and *cosy-2* mutants were compared with those of wild-type plants. The methanol-soluble metabolites were analyzed via ultra-high-pressure liquid chromatography-mass spectrometry (UHPLC-MS) which had been optimized to detect low-molecular-weight phenolic metabolites such as phenylpropanoids, coumarins, benzenoids, flavonols, lignans, and oligolignols[16,17]. In the principal component analysis (PCA) plot based on root metabolic profiles, mutant samples were discriminated from those of the wild type (Supplementary Fig. 5). No discrimination was observed between *cosy* mutants and the wild type in the PCA plot based on inflorescence stem samples, where the gene is lowly expressed compared to the roots (Supplementary Fig. 5). Univariate analysis of the metabolic profiles of roots resulted in a list of 24 compounds that were at least 4-fold higher in abundance in the *cosy* mutants and 10 compounds that were reduced in abundance to below 25% of the level in the wild type (Supplementary Tables 1-3). Eighteen of the 24 compounds that were more abundant in the *cosy* mutants could be structurally characterized based on MS/MS fragmentation spectra. The majority were *o*-hydroxylated phenylpropanoids that might be derived from 6-hydroxyferuloyl-CoA (compounds **6**-**16**; Fig. 1, Supplementary Fig. 6 and Supplementary Table 2). Strikingly, all of the 9 structurally characterized compounds that showed a reduced abundance in *cosy* mutants (**25**-**33**) were coumarins (Fig. 1, Supplementary Fig. 7 and Supplementary Table 3). These compounds included various combinations of scopoletin and sideretin coupled to moieties such as pentose, hexose, acetylhexose, sulfate and malonate. A similar univariate analysis based on inflorescence stem metabolic profiles did not show any compounds with a 4-fold increase in abundance in the mutants, and only one compound, scopolin, had an abundance reduced to below 25% of wild-type levels (Supplementary Tables 4 and 5). A targeted search for other coumarins in the extracts showed evidence for reduced esculin levels in the mutant root extracts (Supplementary Fig. 8). Because coumarins are excreted in the rhizosphere[4-10], we also determined the coumarin levels in root exudates. Scopoletin and sideretin levels were strongly reduced in the root exudates of *cosy* mutants, as compared to those in wild-type plant exudates (Supplementary Fig. 9). However, the reduction in the levels of scopoletin and sideretin in the *cosy* mutant root exudates was less extreme as compared to those in the *f6′h1-1* mutant, which is defective in the conversion of feruloyl-CoA to 6-hydroxyferuloyl-CoA (Supplementary Fig. 9). Taken together, characterization of the most significantly differential compounds via MS methods revealed that the biosynthesis of coumarins was perturbed in *cosy* mutants (Fig. 1, Supplementary Fig. 10).

Coumarins fluoresce blue when exposed to 365 nm UV light[19,20]. In agreement with the reduced abundance of a set of coumarins, fluorescence was reduced in the roots of *cosy-1*, *cosy-2* and *cosy-3* mutants as compared to wild-type plants (Supplementary Fig. 11). Root extracts of *cosy-1*/*cosy-2* heterozygous F1 plants derived from crossing *cosy-1* and *cosy-2* homozygous mutants were similarly fluorescent as those of the parental homozygous mutants, proving that the two mutant alleles failed to complement each other and that the observed metabolic phenotype is caused by mutation of the *COSY* gene (Supplementary Fig. 12).

The combined observed changes in the phenolic pools of the *cosy* mutants are in agreement with a blockage in the conversion of *o*-hydroxycinnamoyl-CoAs to coumarins, a conversion that is considered to happen spontaneously. To verify the catalytic activity of COSY *in vitro*, COSY was prepared as a recombinant protein in *Escherichia coli*. Potential substrates were chosen based on the phenolic profiling of the *cosy* mutants in which derivatives of 6-hydroxyferulic acid accumulated (Supplementary Table 2). The fact that all characterized enzymes of the BAHD family utilize CoA thioesters as acyl donors and that the majority uses an alcohol as the acyl acceptor furthermore suggested that 6-hydroxyferuloyl-CoA, previously thought to be converted solely spontaneously into scopoletin, was the substrate for COSY. Because 6-hydroxyferuloyl-CoA was not commercially available, the corresponding acid, 6-hydroxyferulic acid, was chemically synthesized and the enzymatic assay was performed in combination with heterologously expressed tobacco 4-coumarate-CoA ligase 1 (Nt4CL1), to catalyze the formation of the corresponding CoA thioester[21]. During initial tests, we noted that 6-hydroxyferulic acid was converted to scopoletin by light (Supplementary Fig. 13). In addition, also 6-hydroxyferuloyl-CoA was partially converted to scopoletin by light and to 6-hydroxyferulic acid and scopoletin by heat (Supplementary Fig. 13). Therefore, all reactions to test COSY activity were conducted in the dark, and no heat but urea was used to denature the enzymes to stop the enzymatic activities. Because *COSY* is mainly expressed in roots, a plant organ that typically grows in soil and that is thereby buffered from extreme temperatures and shielded from the light, it is also biologically most relevant to test the activity of COSY in the dark at moderate temperature.

The compound 6-hydroxyferulic acid was stable in the dark at moderate temperature (Fig. 2a). When using Nt4CL1, about 9% of the 6-hydroxyferulic acid was converted into 6-hydroxyferuloyl-CoA and about 8% was converted into scopoletin, demonstrating that Nt4CL1 is able to attach the CoA moiety to 6-hydroxyferulic acid and that a small portion of the 6-hydroxyferuloyl-CoA converted spontaneously into scopoletin (Fig. 2a). However, when Nt4CL1 and COSY were added simultaneously, 6-hydroxyferuloyl-CoA remained under the detection limit and 65% of the 6-hydroxyferulic acid was converted into scopoletin (Fig. 2a), demonstrating that COSY significantly speeded up the reaction efficiency from 6-hydroxyferuloyl-CoA into scopoletin. Similarly, Nt4CL1 converted 2-hydroxy-*p*-coumaric acid (2,4-dihydroxycinnamic acid) and 6-hydroxycaffeic acid (2,4,5-trihydroxycinnamic acid) into 2-hydroxy-*p*-coumaroyl-CoA and 6-hydroxycaffeoyl-CoA, respectively. Similar to the reaction with 6-hydroxyferulic acid, spontaneous *trans*-*cis* isomerization and lactonization resulted in the partial conversion to umbelliferone and esculetin, respectively. However, by adding COSY to these reaction mixtures, 100% of the respective *o*-hydroxycinnamoyl-CoAs appeared to be converted to umbelliferone and esculetin (Fig. 2b and 2c), again demonstrating that COSY speeded up the reaction efficiency from *o*-hydroxycinnamoyl-CoAs into coumarins. Adding CoA and either scopoletin or umbelliferone to COSY, did not result in 6-hydroxyferuloyl-CoA or 2-hydroxy-*p*-coumaroyl-CoA, respectively, showing that COSY was not able to catalyze the reverse reaction (Supplementary Fig. 14).

To obtain insight into the catalytic reaction mechanism of COSY, we modeled its structure based on the crystal structure of hydroxycinnamoyl-CoA:shikimate hydroxycinnamoyltransferase (HCT) of *Coleus blumei* (Supplementary Fig. 15)[22]. After docking 6-hydroxyferuloyl-CoA in the COSY model, His161 and Trp371 appeared to be involved in the catalytic activity, similarly as described for other BAHD acyltransferases[23,24]; Trp371 stabilizes the carbonyl functionality, and His161 deprotonates the acyl acceptor, priming it for nucleophilic attack (Supplementary Fig. 15, Fig. 2e). In the case of the COSY reaction, the deprotonation could take place on the hydroxyl at the *ortho* position to the aromatic ring's aliphatic side-chain. Subsequent electron delocalization would then enable *trans-cis* isomerization upon which the *cis*-isomer could swiftly undergo lactonization to form scopoletin (Fig. 2e).

The proposed COSY reaction mechanism indicates that the *trans-cis* isomerization is the rate-limiting step in the conversion of *o*-hydroxycinnamoyl-CoA thioesters to coumarins. From this mechanism follows that the lactonization of dihydrohydroxycinnamoyl-CoA thioesters (lacking the aliphatic double bond that hinders the *trans*-to-*cis* rotation) to dihydrocoumarins would happen spontaneously. In an attempt to make dihydro-6-hydroxyferuloyl-CoA for testing this hypothesis, we found that Nt4CL1 was not active on dihydro-6-hydroxyferulic acid (Supplementary Fig. 16). However, Nt4CL1 converted about 95% of dihydro-2-hydroxy-*p*-coumaric acid to dihydro-umbelliferone, while dihydro-2-hydroxy-*p*-coumaroyl-CoA remained below the detection limit (Fig. 2d). This experiment showed that dihydro-2-hydroxy-*p*-coumaroyl-CoA was swiftly converted to umbelliferone, even in the absence of light and COSY, strongly supporting the proposed COSY reaction mechanism. The observed spontaneous lactonization of dihydrohydroxycinnamoyl-CoAs also explains the initially surprising observation that dihydroscopoletin-derived compounds **17** and **18** accumulated in the *cosy* mutants, whereas scopoletin-derived compounds **25-33** were reduced (Fig. 1, Supplementary Tables 2 and 3). The dihydroscopoletin-derived compounds **17** and **18** are most likely derived from the spontaneous lactonization of dihydro-6-hydroxyferuloyl-CoA, itself generated by a double bond reductase from 6-hydroxyferuloyl-CoA. Additional support for the proposed reaction mechanism is the observed partial conversion of *o*-hydroxycinnamic acids and *o*-hydroxycinnamoyl-CoA thioesters to their corresponding coumarins in the light (Supplementary Fig. 13), which indicates that *trans-cis* photoisomerization is sufficient to trigger the subsequent lactonization.

Because coumarins are important for iron uptake, mutants deficient in the biosynthesis of coumarins or in their excretion to the rhizosphere accumulate less iron and chlorophyll in their shoot when grown on media or soil with reduced iron bioavailability[6,7,25,26]. When *cosy* mutants were grown on agar-based medium with bioavailable iron (40 µM Fe(III)-EDTA, pH 5.5), *cosy* mutant shoots had similar iron and chlorophyll levels as compared to their respective wild-type controls (Fig. 3, Supplementary Fig. 17). However, when grown on medium with poorly bioavailable iron (40 µM $FeCl_3$, pH 7.0), *cosy* mutants had lower iron and chlorophyll levels as compared to their control lines (Fig. 3, Supplementary Fig. 17). When *cosy* mutants were co-cultivated with wild-type plants on medium with poorly bioavailable iron, their iron and chlorophyll levels were restored (Fig. 3d, Supplementary Fig. 18). On the other hand, co-cultivation of *cosy* mutants with mutants deficient in the biosynthesis of coumarins (*f6ʹh1-1* mutants) did not result in a restoration of the iron and chlorophyll levels (Fig. 3d, Supplementary Fig. 18). These combined observations indicate that the reduced coumarin levels in the root exudates of *cosy* mutants hamper iron uptake, and that the iron uptake can be restored by supplying coumarins to the roots via co-cultivation with wild-type plants.

In another series of experiments, *cosy* mutants were grown on mildly acidic soil (pH 5.5) and on soil where iron bioavailability was reduced by elevated pH (pH 8.0). On mildly acidic soil, all three mutants grew phenotypically similar to their respective control lines (Supplementary Fig. 19). On alkaline soil (pH 8.0), all three mutants were severely chlorotic, starting from the first pair of true leaves, whereas the wild type only showed minor growth retardation (Fig. 3e, Supplementary Fig. 20). By watering the mutant plants with the synthetic chelator complex Fe(III)-EDDHA, the *cosy* mutants did not develop the chlorotic phenotype and grew almost as well as the wild type on alkaline soil (Supplementary Fig. 20). On alkaline soil, the chlorosis phenotype of the *cosy* mutants was also observed in plants grown in low light intensities (30 µmol photons m^{-2} s^{-1}), indicating that photo-oxidation was not the cause for the chlorotic phenotype[27] (Supplementary Fig. 21). Altogether, these results show that *cosy* mutants fail to acquire sufficient iron to sustain photosynthesis. The mutant phenotype of *cosy-3* was complemented in stable transgenic lines in which expression of *COSY* was driven by the Cauliflower Mosaic Virus 35S promoter (Supplementary Fig. 22), verifying that mutation of the *COSY* gene is the cause of the observed phenotypes.

Discussion

We discovered that the conversion of *o*-hydroxycinnamoyl-CoA thioesters into coumarins is catalyzed by COSY. Enzyme modeling and *in vitro* tests showed that COSY lowers the activation energy of the conversion of *trans*-*o*-hydroxycinnamoyl-CoA thioesters to *cis*-*o*-hydroxycinnamoyl-CoA thioesters (Fig. 2). This conversion appears as the rate-limiting step in coumarin biosynthesis, because the subsequent lactonization of *cis*-*o*-hydroxycinnamoyl-CoA thioesters to coumarins went swiftly, a hypothesis that was further supported by the observation that dihydrohydroxycinnamoyl-CoA thioesters do not need COSY to be converted to their corresponding dihydrocoumarins (Fig. 2d). The results further showed that the reaction also occurred upon photoisomerisation (Supplementary Fig. 13) and even partially spontaneously at moderate temperature (Fig. 2). However, these COSY-independent conversions of *o*-hydroxycinnamoyl-CoA thioesters into coumarins appeared marginal *in planta*. Profiling of phenolic metabolites showed that roots of *Arabidopsis cosy* mutants had severely reduced coumarin levels, whereas *o*-hydroxylated phenylpropanoids accumulated instead (Fig. 1), the latter most likely being derived from the COSY substrate, 6-hydroxyferuloyl-CoA. In addition to *cosy* roots, also root exudates of *cosy* mutants had reduced coumarin levels. In accordance with the role of coumarins in iron uptake, *cosy* mutants had reduced iron and chlorophyll levels when grown in conditions with limited iron bioavailability (Fig. 3). The phenotype of *cosy* mutants was successfully restored via genetic complementation by over-expression of *COSY* (Supplementary Fig. 22), via chemical complementation with bioavailable iron (Fig. 3), and via chemical complementation with root exudates from wild-type plants (Fig. 3). Furthermore, we showed that the phenotype of *cosy* mutants was similar to that of plants mutated in *F6'H*, the gene coding for the first committed step in the biosynthesis of coumarins in *Arabidopsis* [6,7,13]. Collectively, we disprove the generally accepted model that the *trans-cis* isomerization and lactonization of *o*-hydroxycinnamoyl-CoAs to coumarins occur solely spontaneously[11-14]. Consequently, these data necessitate revision of the currently accepted coumarin biosynthetic pathway.

Because of its importance in iron uptake and plant defense, the discovery of COSY will potentially facilitate the development of crops with increased iron uptake and pest resistance, for instance via biotechnological engineering or marker-assisted breeding. Consistent with a conserved role in coumarin biosynthesis, orthologues of *COSY* are present in a wide range of plant

species (Supplementary Fig. 23). In addition, coumarins are valuable for society as flavors, fragrances, and medicines. Because the conversion of cinnamoyl-CoAs into coumarins has been reported to be the main bottleneck in the production of coumarins through white biotechnology, COSY might also enhance current manufacturing systems for valuable coumarins[28,29].

Methods

Plant material. *Arabidopsis thaliana* accessions Nossen (No-0), Landsberg *erecta* (L*er*-0) and Columbia (Col-0) were used. Two T-DNA insertion lines are in the No-0 ecotype and were acquired from the RIKEN Arabidopsis Ds transposon mutant collection: *cosy-1* (15-5543-1) and *cosy*-2 (13-3585-1)[30,31]. The third T-DNA insertion line, *cosy-3* (GT_3.9907), is in the L*er*-0 ecotype and was derived from the EXON Trapping Insert Consortium (EXOTIC)[32]. The sequences of the primers used to verify the insertions (Supplementary Fig. 4) are given in Supplementary Table 6. The *f6′h1-1* mutant line (SALK_132418C, in the Col-0 ecotype) has been described before[6,7] and is a kind gift of Stephan Clemens and Nicole Schmid.

Vector construction and production of transgenic lines. Constructs were prepared using the Gateway Cloning Technology (Invitrogen, Carlsbad, CA). 1880 bp of the 5′ region directly upstream of the *COSY* start codon was amplified via PCR from genomic Col-0 DNA using the primers A and B (Supplementary Table 6), introducing the restriction sites *Bam*HI and *Xho*I. The PCR product was captured into the Gateway *pEN-R4L1* vector using T4 DNA Ligase (Invitrogen). The promoter region was confirmed by sequencing. The *COSY* promoter was then transferred into the destination vector *pMK7S*NFm14GW* with an LR reaction, resulting in *pCOSY:GUS-GFP-NLS* expression clone. The expression clones were introduced into *Agrobacterium tumefaciens* strain C58C1 (pMP90) by electroporation. The construct was introduced into wild-type *Arabidopsis* (Col-0) through the floral dip method. Single-locus lines were subsequently identified by segregation analysis of kanamycin resistance. For complementation with *pro35S:COSY*, the *COSY* gene was amplified via PCR from Col-0 cDNA using the primers C and D (Supplementary Table 6), introducing the attB1 and attB2 sites, respectively. The PCR product was captured into the Gateway vector pDONR221. The pDONR221 vector containing the *COSY* gene and the gateway vector pEN-L4-2-R1 containing the CaMV 35S promoter (flanked by attL4 and attR1 sites) were then transferred into the destination vector pK7m24GW-FAST. The expression clones were introduced into *Agrobacterium tumefaciens* strain C58C1 (pMP90) by electroporation. Next, the construct was introduced into *cosy-3* through the floral dip method. Single-locus lines were subsequently identified by segregation analysis via the FAST system[33].

Reporter gene analysis. Six homozygous *pCOSY:GUS-GFP-NLS* single-locus lines were investigated for GUS activity through histochemical staining. For analysis of 8-day-old seedlings, seeds were vapor-phase sterilized and sown on sterile plates containing 1/2 Murashige and Skoog (½ MS) medium (pH 5.7, 2.15 g MS basal salt mixture powder (Duchefa), 10 g sucrose, 0.5 g MES monohydrate and 8 g plant tissue culture agar [Agar No 4 (Lab M)] per L). After sowing, seeds were incubated at 4°C for at least 2 d, whereupon plates were placed in a vertical orientation in the tissue culture chamber room under a 16-h light/8-h dark photoperiod at 21°C. For analysis of 5-week-old rosettes, the plants were grown on soil in long-day conditions (16-h light/8-h dark, 22°C). For analysis of 3-month-old plants, the plants were grown on soil in short-day photoperiod (9-h light/15-h dark, 22°C) for eight weeks followed by long-day conditions (16-h light/8-h dark, 22°C) for four additional weeks. Seedlings, rosettes, inflorescence and mature roots of the six homozygous line were subjected to histochemical GUS staining at 37°C in a staining buffer

containing 1 mM 5-bromo-4-chloro-3-indolyl-β-D-glucopyranoside sodium salt (X-Gluc), 0.5 % Triton X-100, 1 mM ethylenediaminetetraacetic acid (EDTA) pH 8.0, 0.5 mM potassium ferricyanide [$K_3Fe(CN)_6$], 0.5 potassium ferrocyanide [$K_4Fe(CN)_6$] and 500 mM sodium phosphate buffer pH 7.0. Seedlings were incubated for 30 min, rosettes for 24 h, inflorescences for 48 h and mature roots overnight. The reaction was terminated by replacing the staining buffer with 70 % ethanol. Next, the samples were incubated in a destaining solution (50% glycerol, 25% lactic acid and 25% water) to remove the chlorophyll that otherwise interferes with the blue GUS-stain. The stem sections were photographed with an Olympus BX51 microscope and a 10x objective equipped with a Nikon Digital Sight DS-SM camera. The seedlings were photographed with a Leica S6D stereomicroscope. For GFP analysis, three homozygous *pCOSY:GUS-GFP-NLS* single-locus lines were grown for 14 d on half-strength ½ MS plates. Root cell walls were stained with propidium iodide (PI, 30 µM) at the onset of the experiment. The excitation energy of 488 nm was from an argon laser. The fluorescence emission of PI was collected between 550 and 650 nm and that of GFP between 500 and 550 nm. All images were captured with an inverted LSM 710 META confocal microscope equipped with 20X-Air objectives (Zeiss).

Expression analysis via RT-qPCR. Expression analysis for the characterization of the *cosy* mutants was performed on three pools for each genotype, each of the pools consisting out of two roots. Seeds were vapor-phase sterilized and sown on sterile plates containing ½ MS medium containing 40 µM Fe(III)-EDTA [pH 7.2, 2.135 g MS basal salt mixture powder (Duchefa), 10 g sucrose, 0.5 g MES monohydrate, 14.682 mg FeNaEDTA and 8 g Agar No 4 (Lab M) per L]. After sowing, seeds were incubated at 4°C for 2 d, whereupon plates were placed in a vertical orientation in the tissue culture room under a 16-h light/8-h dark photoperiod. Roots were harvested 8 d after germination. RNA was extracted using RNeasy kit (Qiagen, Valencia, CA) and a DNase treatment was performed using DNA-free™ (Ambion, Life Technologies, Carlsbad, California, U.S.). The extracted RNA was quantified using the nanodrop® ND-1000 spectrophotometer (Thermo Scientific, Wilmington, DE, USA). Next, 1000 ng RNA was used for cDNA synthesis using the First Strand cDNA Synthesis Kit (Thermo Scientific, Thermo Fisher Scientific, Waltham, MA, USA). A ten times diluted cDNA sample was used for RT-qPCR using a sensiFAST SYBR No-ROX-kit (Bioline, London, UK) on a Lightcycler 480 (Roche, Basel, Switzerland). Samples were run in three technical triplicates on the Lightcycler 480 with following protocol: one activation cycle of 10 min at 95°C; 45 amplification cycles of 10 s at 95°C, 10 s at 60 °C and 10 s at 72 °C; melting curve cycle measuring from 65 to 95 °C. Fluorescence values were exported from the Lightcycler 480 program whereupon Ct values, normalization factors and primer efficiencies were calculated according to Ramakers *et al.*[34] using three reference genes: Ser/Thr protein phosphatase 2A (*PP2A*; AT1G13320), SAND family protein (AT2G28390), and AT2G32170[35] (Supplementary Table 6). Two regions of the *COSY* cDNA were amplified using primer 6 and 7 in one reaction and primer 8 and 9 in the other (Supplementary Table 6).

Phenolic profiling. Plants that were used for the phenolic profiling of their roots and stems were grown in soil in a short-day photoperiod (9-h light/15-h dark), 22°C for eight weeks, followed by a long-day photoperiod (16-h light/8-h dark), 22°C for an additional four weeks. Then, roots and stems were harvested. Roots were washed to remove soil particles. A subsample of about 150 mg fresh weight was used for metabolite extraction from root samples. A basal 9-cm inflorescence stem sample from 1-10 cm above the rosette was used for phenolic profiling. The plant tissue was immediately flash-frozen in liquid nitrogen. Plant material was ground in a 2-mL Eppendorf tube with a 4-mm iron bead using a Retch ball mill (20 Hz). The phenolic metabolites were extracted from root and stem samples with 1 mL methanol for 15 min at 70°C, while shaking in a

thermomixer. After drying 800 µL of each methanol extract in a SpeedVac, a liquid-liquid extraction was performed with 100 µL cyclohexane and 100 µL H_2O. The lower water phase was injected. UHPLC-MS analysis was performed as previously described[16]. In short, 15 µL of the lower water phase was injected on a Waters Acquity UPLC® equipped with a BEH C18 column (2.1 x 150 mm, 1.7 µM, Waters) and coupled to a Synapt Q-Tof (Waters Corporation). A gradient of two buffers was used: buffer A (99/1/0.1 H_2O/acetonitrile/formic acid, pH 3), buffer B (99/1/0.1 acetonitrile/H_2O/formic acid, pH 3); 95% A for 0.1 min decreased to 50% A in 30 min (350 µL/min, column temperature 40°C). The flow was diverted to the MS equipped with an electrospray ionization source and lockspray interface for accurate mass measurements. The MS source parameters were capillary voltage, 1.5 kV; sampling cone, 40 V; extraction cone, 4 V; source temperature, 120°C; desolvation temperature, 350°C; cone gas flow, 50 L/h; and desolvation gas flow, 550 L/h. The collision energy for the trap and transfer cells was 6 V and 4 V, respectively. For data acquisition, the dynamic range enhancement mode was activated. Full-scan data were recorded in negative centroid V-mode; the mass range between m/z 100 and 1,200, with a scan speed of 0.2 s/scan, with Masslynx software (Waters). Leucin-enkephalin (250 pg/µL solubilized in water/acetonitrile 1:1 (vol:vol), with 0.1% formic acid) was used for the lock mass calibration, scanning every 10 s with a scan time of 0.5 s. The raw chromatograms were available from the MetaboLights database repository[36] as study MTBLS692. Data processing was done with Progenesis QI ver. 2.1 (Nonlinear Dynamics). Peak areas were normalized on total peak intensity. PCA analysis was performed in R on the mean-centered and unit variance-scaled peak abundances using the function prcomp. Peaks that were detected in less than five chromatograms were excluded from the PCA analysis. For the univariate analysis, peaks were further filtered based on (i) their detection in all replicates of at least one genotype and (ii) their average normalized peak intensity being higher than 5000 in at least one genotype. For statistical analysis, peak intensities were ArcSinh transformed. The following filters were used to select peaks that were significantly different in abundance between the mutant and wild-type samples: ANOVA p-value <0.001, post-hoc t-test p-value <0.01 for both *cosy-1* and *cosy-2,* and at least a four-fold difference (of the normalized but non-transformed peak intensities) between wild type and both *cosy-1* and *cosy-2* mutants. Post-hoc t-test between two genotypes was only performed for peaks that were detected in all replicates of both genotypes. Peaks having a zero value in all replicates of one genotype and a value different from zero in all samples of the other genotype were considered as 'discrete' and thus significantly different between the genotypes. Peaks having a zero value in all replicates of one genotype and in one or more replicates of the other genotype were considered as 'undetermined'. In root samples, these filters resulted in 56 peaks, of which 40 had a higher peak area in the mutants as compared to the wild type and 16 had a lower peak area in the mutants. The 40 peaks with a higher peak area could be assigned to 24 compounds, of which 18 could be (partially) structurally characterized based on MS/MS. The 16 peaks with a lower intensity in the mutant samples could be assigned to 10 compounds of which 9 could be (partially) structurally characterized. A similar analysis for the stem samples did not result in any differential peaks. We therefore weakened the filter (ii) of the average normalized peak intensity area to > 1000 (for the root extracts, this filter was set to > 4000). Using this filter, scopoletin was picked up as the only metabolite with reduced abundance in both *cosy-1* and *cosy-2* mutants, and no peaks were found to be higher in abundance.

Determination of root exudates. Seeds were vapor sterilized, then sown on sterile plates containing ½ MS medium[37]. The seeds were vernalized for 2 d at 4°C. Plates were placed vertically in the tissue culture room under a 16-h light/18-h dark regime at 21°C. Based on the method from

Schmid and coworkers[6], 10-day-old plants were transferred to Fe-sufficient (75 µM Fe(III)-EDTA) or Fe-deficient (no added Fe + 50 µM ferrozine) solid medium. After 4 d, the plants were transferred from the agar plates to a 12-well plate (8 plants per well) containing 3 mL of ultrapure water per well (Millipore). Plates were covered with tape on the sides to prevent dehydration, then placed on a shaker in the same chamber room. 1.8 mL of the exudates was collected after 6 h and concentrated in a CentriVap Concentrator (Labconco), then re-suspended in 100 µL water. Fifteen µL of the sample was injected for phenolic profiling through UPLC-MS as described above. In addition to the exudates, a calibration curve was made for scopoletin, fraxin, esculin, fraxetin and umbelliferone.

Fluorescence. Seeds were vapor-phase sterilized and sown on sterile plates containing ½ MS medium containing 40 µM Fe(III)-EDTA (pH 7.2, see Expression analysis via RT-qPCR). After sowing, seeds were incubated at 4°C for 2 d, whereupon plates were placed in a vertical orientation in the tissue culture chamber room under a 16-h light/8-h dark photoperiod. Fluorescence of 12-day-old plants was observed with a Gel Doc™ XR+ System (Biorad, California, USA), with an excitation wavelength of 365 nm.

Chemicals. Umbelliferone and esculetin were purchased from Sigma-Aldrich and scopoletin was purchased from Carbosynth Limited. The *o*-hydroxycinnamic acids were synthesized by saponification of either umbelliferone, esculetin or scopoletin with 2 M NaOH solution under slightly different conditions. To produce 2-hydroxy-*p*-coumaric acid, umbelliferone (50 mg) was stirred in 1.5 mL NaOH solution (2 M) at 80°C for 1 h, and acidified with 1 M HCl to pH 4-5. The product was collected by extraction with ethyl acetate (EtOAc). EtOAc was subsequently removed in a speedvac and the remaining pellet was resolved in 10 mL H_2O for further purification via a Waters AutoPurification™ System equipped with a Waters SunFire C18 analytical Column (100Å, 5 µm, 4.6 mm x 150 mm), a SunFire C18 Prep Column, (100Å, 5 µm, 10 mm x 150 mm) and an Acquity QDa Detector. A gradient of two buffers was used: buffer A (99/1/0.1 H_2O/acetonitrile/formic acid, pH 3), buffer B (99/1/0.1 acetonitrile/H_2O/formic acid, pH 3); 95% A for 0.1 min decreased to 50% A in 30 min. The flow was 2 mL/min and 10 mL/min for analytical and preparative separation, respectively. Columns were at room temperature during separation. The fraction collection was triggered by *m/z* 135.0 (i.e. the -CO_2 in-source fragment of 2-hydroxy-*p*-coumaric acid). After concentration of the collected fraction to 1.5 mL by use of a SpeedVac, the concentrate was stored overnight at -20°C upon which crystals were formed. Filtration of the crystals resulted in the purified 2-hydroxy-*p*-coumaric acid. Notably, solutions of 2-hydroxy-*p*-coumaric acid appeared instable and are recommended to be stored at -70°C. ^{1}H NMR (acetone-d_6) δ 6.42 (1H, dd, *J* = 8.5, 2.4 Hz, 5), 6.42 (1H, d, *J* = 16 Hz, 8), 6.47 (1H, d, *J* = 2.4 Hz, 3), 7.45 (1H, d, *J* = 8.5 Hz, 6), 7.94 (1H, d, *J* = 16 Hz, 7), 8.8 (1H, s, 4-OH), 9.05 (1H, s, 2-OH); ^{13}C NMR (acetone-d_6) δ 103.50 (3), 108.88 (5), 114.53 (1), 114.94 (8), 131.21 (6), 141.48 (7), 159.02 (2), 161.56 (4), 169.15 (9).

For 6-hydroxycaffeic acid, esculetin was acetylated with Ac_2O and pyridine first. The acetylated esculetin (500 mg, 2.02 mmol) was stirred in 2 M NaOH solution at 50°C for 1.5 h, and acidified with 1 M HCl to pH 4-5. The product solution was extracted with EtOAc to obtain the product (31 mg, 0.16 mmol, 7.8%). 6-hydroxycaffeic acid was further purified via a Waters AutoPurification™ System as described above, but with *m/z* 151.0 as trigger for fraction collection (i.e. the -CO_2 in-source fragment of 6-hydroxycaffeic acid). After concentration of the collected fraction to 1.5 mL by use of a SpeedVac, the concentrate was stored overnight at -20°C. In contrast to the procedure of 2-hydroxy-*p*-coumaric acid, no crystals were formed. The solvent was then simply further removed in a SpeedVac, resulting in purified 6-hydroxycaffeic acid. Solutions of

6-hydroxycaffeic acid appeared instable and are recommended to be stored at -70°C. ^{1}H NMR (acetone-d_6) δ 6.28 (1H, d, J = 16 Hz, 8), 6.49 (1H, s, 3), 7.03 (1H, s, 6), 7.93 (1H, d, J = 16 Hz, 7); ^{13}C NMR (acetone-d_6) δ 104.12 (3), 113.50 (1), 114.48 (6), 114.53 (8), 139.47 (5), 141.19 (7), 149.91 (4), 152.02 (2), 168.75 (9).

6-hydroxyferulic acid was obtained by stirring scopoletin (50 mg) for 1 h in 2 M NaOH solution at 80°C, after which the reaction mixture was acidified with 1 M HCl to pH 4-5. The crude product was further purified via a Waters AutoPurification™ System as described above for 2-hydroxy-*p*-coumaric acid, apart from the buffers used [buffer A (99/1/0.1 H_2O/acetonitrile/2 M ammonium acetate, pH 5) and buffer B (99/1/0.1 acetonitrile/H_2O/2 M ammonium acetate, pH 5)] and the trigger for fraction collection *m/z* 165.0 (i.e. the -CO_2 in-source fragment of 6-hydroxyferulic acid). Filtration of the crystals that appeared upon storing the 1.5 mL concentrate at -20°C, resulted in the purified 6-hydroxyferulic acid. ^{1}H NMR (acetone-d_6) δ 3.84 (3H, s, OMe), 6.39 (1H, d, J = 16 Hz, 8), 6.48 (1H, s, 3), 7.19 (1H, s, 6), 7.99 (1H, d, J = 16 Hz, 7); ^{13}C NMR (acetone-d_6) δ 56.82 (OMe), 104.10 (3), 111.58 (6), 113.20 (1), 114.71 (8), 141.11 (7), 142.53 (5), 151.16 (4), 152.95 (2), 168.90 (9).

Dihydro-6-hydroxyferulic acid and dihydro-2-hydroxy-*p*-coumaric acid were obtained by stirring scopoletin (80 mg) for 1 h in 2 M NaOH solution at 80°C. Similar to the synthesis of dihydrocinnamic acid (3-phenylpropanoic acid) from cinnamic acid described by Rao and coworkers[38], we then added Raney nickel (60 mg, W6 grade) and sodium borohydride (40 mg) to the reaction mixture, after which the reaction mixture was acidified to pH 5 with 1 M HCl and extracted in EtOAc. EtOAc was subsequently removed in a SpeedVac. The crude product was further purified via a Waters AutoPurification™ System as described above for 2-hydroxy-*p*-coumaric acid, apart from the trigger for fraction collection *m/z* 211.1 (i.e. the dihydro-6-hydroxyferulic acid). This fraction was concentrated in a SpeedVac and contained about 95% dihydro-6-hydroxyferulic acid and 5% dihydro-2-hydroxy-*p*-coumaric acid. This mixture was used to test the reactions with Nt4CL and COSY.

Stability assays in heat and light. To test the stability of 6-hydroxyferulic acid, the compound was dissolved in the buffer used for enzyme activity assays (see below); 100 µM 6-hydroxyferulic acid in a final volume of 40 µL water containing 100 mM Tris-HCL (pH 7.4), 5 mM MgCl, 5 mM ATP and 200 µM CoA. To test the stability in light, the mixtures (n=4) were placed for 1 h in a plant tissue culture growth room (21°C) equipped with fluorescent lighting (Radium Spectraluxplus, NL 58W/840, white, 50 µmol photons m^{-2} s^{-1}). Four control samples were placed for 1 h in the same plant tissue culture growth room, but wrapped in aluminum foil to shield them away from the light. To test the stability in heat, the mixtures were treated for 5 min at 95°C in the dark. As control, four samples were kept for 5 min at room temperature in the dark. The stability of 6-hydroxyferuloyl-CoA in light and heat was tested similarly. However, 6-hydroxyferuloyl-CoA could not be obtained pure. It was made in a 15-min reaction of 6-hydroxyferulic acid with Nt4CL (see below). The mixture also contained 6-hydroxyferulic acid and scopoletin.

Recombinant COSY preparation. Recombinant COSY was expressed in the *Escherichia coli* strain BL21codon + pICA2 after transformation with pLH36AT1G28680, in which expression is induced by isopropyl β-D-1-thiogalactopyranoside (IPTG) under control of a pL-promoter developed by the Protein Core of VIB. The pLH36 plasmid is provided with a His_6-tag followed by a murine caspase-3 site. The murine caspase-3 site can be used for the removal of the His_6-tag attached at the N-terminus of the protein of interest during purification. The transformed bacteria were grown in Luria Bertani medium supplemented with ampicillin (100 µg/mL) and kanamycin (50 µg/mL) overnight at 28°C before 1/100 inoculation in a 20-L fermenter provided with Luria

Bertani medium supplemented with ampicillin (100 µg/mL) and 1% glycerol. The initial stirring and airflow was 200 rpm and 1.5 L/min., respectively. This was further automatically adapted to keep the pO_2 at 30%. The temperature was kept at 28°C. The cells were grown to an optical density of 1.0 as measured by the absorbance at 600 nm, transferred to 20°C, and expression was induced by addition of 1 mM IPTG overnight. Cells were then harvested and frozen at -20°C. After thawing, the cells were resuspended at 3 mL/g in 20 mM Hepes pH 7.5, 500 mM NaCl, 20 mM imidazole, 1 mM PMSF. The cytoplasmic fraction was prepared by sonication of the cells and was isolated by centrifugation at 18,000 x g for 30 min. All steps were conducted at 4°C. The clear supernatant was applied to a 47-mL Ni-Sepharose 6 FF column (GE Healthcare), washed with 20 mM Hepes pH 7.5, 500 mM NaCl, 20 mM imidazole, 1 mM PMSF and 0.1 % Empigen over 20 column volumes followed by 5 column volumes without Empigen. The column was eluted with 20 mM Hepes pH 7.5, 20 mM NaCl, 400 mM imidazole, 1 mM PMSF after an extra wash step with 50 mM of imidazole in the same buffer. The elution fraction was diluted with 20 mM Tris pH 8.5 to a conductivity below 5 mS/cm and loaded on a 20-mL Source 15Q column (GE Healthcare) to remove contaminants. After equilibration, the protein of interest was eluted by a linear gradient over 10 column volumes of NaCl from 0 to 1 M in 20 mM Tris pH 8.5. After this ion-exchange chromatography step, the N-terminal His_6-tagged COSY was incubated for 1 h at 37°C with activated murine caspase-3 (1/50 m/m murine caspase-3/COSY) in the presence of 0.1 % CHAPS. The digested fraction was reloaded on a 20-mL Ni-Sepharose 6 FF column where the untagged mature COSY passed in the flow through of the column and the His_6 tag, uncleaved COSY and His_6-tagged murine caspase-3 bound to the column. The mature COSY was finally injected on Superdex 200 XK26/64 to PBS to remove minor contaminants.

Recombinant Nt4CL. The expression and purification of tobacco Nt4CL1 enzyme has been published previously[21].

Enzyme activity assays. The assay was performed in a final volume of 40 µL. The mixtures contained 100 mM Tris-HCL (pH 7.4, 8 µL of a 500-mM stock solution), 5 mM MgCl (8 µL of a 25-mM stock solution), 5 mM ATP (8 µL of a 25-mM stock solution), 200 µM CoA (4 µL of a 1-mM stock solution) and 100 µM substrate (8 µL of a 500-µM stock solution of 2-hydroxy-*p*-coumaric acid, 6-hydroxycaffeic acid, 6-hydroxyferulic acid, scopoletin, umbelliferone or the mixture of 95% dihydro-6-hydroxyferulic acid and 5% dihydro-2-hydroxy-*p*-coumaric acid). Other components were added as indicated: 1.94 µg Nt4CL1 (1 µL of a 1.94-mg/mL stock solution) and 1.94 µg COSY (1 µL of a 1.94-mg/mL stock solution). Each reaction was performed four times. After vortexing and 30 s of centrifugation, the reactions were carried for 15 min and except differently stated, the temperature was 20°C and the reactions were carried out in the dark (under red light). The reactions were stopped by adding 21 µL of 20 M urea. The reaction mixtures were measured on a UHPLC-MS. UHPLC-MS settings and data processing were the same as described for phenolic profiling, except for a shorter Acquity UPLC BEH C18 Column, (130Å, 1.7 µm, 2.1 mm x 50 mm) and the shortened gradient applied: 99% A for 0.1 min decreased to 50% A in 10 min (500 µL/min, column temperature 40°C). The respective acids (2-hydroxy-*p*-coumaric acid, 6-hydroxycaffeic acid, 6-hydroxyferulic acid, dihydro-2-hydroxy-*p*-coumaric acid, dihydro-6-hydroxyferulic acid) and esculetin were measured via buffers at pH 3; buffer A (99/1/0.1 H_2O/acetonitrile/formic acid) and buffer B (99/1/0.1 acetonitrile/H_2O/formic acid). 2-hydroxy-*p*-coumaric acid was detected as *m/z* 135.0404 at 2.12 min (i.e. the in-source -CO_2 fragment), 6-hydroxycaffeic acid as *m/z* 151.0379 at 1.35 min (i.e. the in-source -CO_2 fragment), 6-hydroxyferulic acid as *m/z* 165.0504 at 2.40 min (i.e. the in-source -CO_2 fragment), dihydro-2-hydroxy-*p*-coumaric acid as *m/z* 181.0458 at 1.84 min, dihydro-6-hydroxyferulic acid *m/z*

211.0569 at 2.08 min and esculetin as *m/z* 177.0169 at 1.93 min. The CoA thioesters (2-hydroxy-*p*-coumaroyl-CoA, 6-hydroxycaffeoyl-CoA, 6-hydroxyferuloyl-CoA, dihydro-6-hydroxyferuloyl-CoA), umbelliferone, scopoletin and dihydro-umbelliferone were measured via buffers at pH 7; buffer A (99/1/0.5 H_2O/acetonitrile/ammonium acetate) and buffer B (99/1/0.5 acetonitrile/H_2O/ammonium acetate). 2-hydroxy-*p*-coumaroyl-CoA was detected as *m/z* 928.1448 at 2.21 min, 6-hydroxycaffeoyl-CoA as *m/z* 470.5497 at 1.36 min, 6-hydroxyferuloyl-CoA as *m/z* 958.1577m/z at 2.30 min, dihydro-6-hydroxyferuloyl-CoA as *m/z* 181.0458 at 1.84 min, umbelliferone as *m/z* 161.0196 at 2.56 min, scopoletin as *m/z* 176.0066 at 2.60 min (i.e. the in-source -•CH_3 fragment) and dihydro-umbelliferone as *m/z* 163.0356 at 2.87 min.

Protein model. A model of COSY was built using SWISS-MODEL Homology Modelling[39]. As a first step in the process, the primary amino acid sequence of AT1G28680 was used as target in a BLAST search[40] against the primary amino acid sequence contained in the SWISS-MODEL Template Library (SMTL; last update: 2018-06-29, last included PDB release: 2018-06-22). The retained 14 templates were subsequently used to build an HHblits profile[41] that was used as target to search against all profiles of the SMTL. This resulted in 108 retained templates, with 5kjv.1 as top hit (HCT of *Coleus blumei*[22]; sequence identity 29.09; coverage 0.92). For each of the identified templates, the quality was predicted from features of the target-template alignment and the templates with the highest quality were selected for model building using ProMod3 Version 1.1.0.[42]. During the process, coordinates conserved between the target and the template were copied from the template to the model. Insertions and deletions were subsequently remodeled using a fragment library and side chains were rebuilt. Finally, the geometry of the resulting model was regularized by using a force field. In case loop modeling with ProMod3 failed, an alternative model was built with PROMOD-II. The obtained model of COSY has a GMQE score of 0.65 and a QMEAN score of -3.63[43]. 6-hydroxyferuloyl-CoA was drawn in ChemDraw and converted to its 3-dimensional structure that was saved as pdb format. Both ligand and protein were prepared for docking using MGLTtools-1.5.6. The docking was performed with AutoDock 4.2.1 and Vina 1.0.2[44] using default parameters unless stated otherwise. The grid spacing was set to 1 and the size of the docking grid was 22 Å×22 Å×24 Å, which encompassed the space occupied by *p*-coumaroyl-CoA in 5kjv.1. Final images were created with Swiss-PDB viewer[42] and rendered with POV-Ray (www.povray.org).

Quantitative analysis of shoot iron content. Total iron in shoots was determined by the BPDS method described by Tsai et al.[9], with slight modifications. Harvested shoot samples were washed with $CaCl_2$ and H_2O and dried in an oven at 70°C for at least 2 d before digestion. Dried samples were incubated in 225 µl of 65% (v/v) nitric acid (HNO_3) at 96°C for 6 h, followed by 150 µl of 30% (v/v) hydrogen peroxide (H_2O_2) at 56°C for 2 h. Dilution of each sample was made by the addition of 225 µl of sterile water. Samples were mixed in an assay solution that contained 1 mM BPDS, 0.6 M sodium acetate, and 0.48 M hydroxylamine hydrochloride. The concentration of Fe^{2+}-$BPDS_3$ complexes were measured by absorbance at 535 nm in a PowerWave XS2 microplate spectrophotometer (BioTek). Iron concentrations were determined against a standard curve made with $FeCl_3$ that was treated in the same way as the plant materials.

Growth conditions of medium-grown plants. *Arabidopsis* plants were grown directly on agar-based medium for 14 d under sterile conditions in a growth chamber. Seeds were surface-sterilized by immersing in 30% (v/v) commercial bleach containing 6% NaClO (Clorox Co., Oakland, CA) and 70% (v/v) absolute ethanol containing 0.1% (v/v) Tween 20 for 6 min, followed by three rinses with absolute ethanol. Seeds were sown on Petri dishes and kept for 2 d at 4°C in the dark before

the plates were transferred to a growth chamber with a 16-h light/8-h dark photoperiod at 21°C (light intensity 100 µmol m^{-2} s^{-2}). The growth medium was composed of 5 mM KNO_3, 2 mM $MgSO_4$, 2 mM $Ca(NO_3)_2$, 2.5 mM KH_2PO_4, 70 µM H_3BO_3, 14 µM $MnCl_2$, 1 µM $ZnSO_4$, 0.5 µM $CuSO_4$, 0.01 µM $CoCl_2$, and 0.2 µM Na_2MoO_4, supplemented with 0.5% (w/v) sucrose, and solidified with 0.4% Gelrite pure (Kelco). For bioavailable iron (control), 40 µM FeNa-EDTA and 1 g/L MES were added and the pH was adjusted to 5.5 with KOH (resulting in Estelle and Somerville media). For poorly bioavailable iron (low Fe) medium, 40 µM $FeCl_3$ and 1 g/L MOPS were added and pH was adjusted to 7.0 with KOH. The medium with low-bioavailability iron was also used for co-cultivation experiments of *cosy* mutants.

Growth conditions of soil-grown plants. For alkaline soil experiments, $CaCO_3$ (20 g kg^{-1}) and $NaHCO_3$ (12 g kg^{-1}) were added to the soil to obtain pH 8.0. As a control, regular soil was used with a pH of about 5.5. Depending on the growth conditions, each plant was watered with 1 mL 300 µM Fe(III)-EDDHA (Solufeed, Barnham, UK) 2 times, at 3 and 7 d after germination. Four plants were grown in a pot (containing about 68 cm^3 of soil), thus each pot received 4 mL 300 µM Fe(III)-EDDHA at day 3 and 4 mL at day 7. The plants were grown in short-day photoperiod (9-h light/15-h dark) at 22°C (n=21). Normal and low light conditions were 120 and 30 µmol photons m^{-2} s^{-1}, respectively.

Quantitative analysis of chlorophyll content. Rosettes, excluding the hypocotyl, of soil-grown plants were harvested and pooled two by two in a 2-mL Eppendorf tube. After determining the fresh weight of the rosettes, samples were frozen in liquid nitrogen and the tissue was homogenized using two precooled metal 4-mm beads in a Retch shaking mill. Minimum six biological replicates were harvested for each genotype. Leaf pigments were extracted by adding 500 µL of DMSO followed by 1 min of vortexing. After two min of centrifugation (20 000 g) the supernatant was transferred to a multi-well plate and the absorption was measured using a SPECTRAmax™ 250 microplate spectrophotometer (Molecular Devices Corporation) at 645 nm and 663 nm. ChlA and ChlB: were quantified via following equations ChlA (µg mL^{-1}) = 12.70 A663 - 2.69 A645; ChlB (µg mL^{-1}) = 22.90 A645 - 4.68 A663[45].

Phylogenetic tree. For the circular cladogram of COSY homologs, and analysis of the conserved protein residues, two independent BLASTP analyses with default settings were performed against the NCBI non-redundant protein database using the AT1G28680 protein sequence as query. One was restricted towards monocotyledons, whereas the other was restricted to eudicotyledons. For each organism in the output file, the top hit was retained. The total number of sequences was subsequently manually reduced to 54 by deleting sequences of closely related organisms. Following protein sequences were selected: *Sorghum bicolor*, XP_002437857; *Hordeum vulgare*, BAJ95032; *Brachypodium distachyon*, XP_003561089; *Musa acuminata*, XP_009409409; *Ananas comosus*, XP_020081186; *Asparagus officinalis*, XP_020268888; *Zea mays*, XP_020399988; *Oryza sativa*, BAX24637; *Populus trichocarpa*, XP_002305089; *Vitis vinifera*, XP_002271913; *Ricinus communis*, XP_002522535; *Glycine max*, XP_003541712; *Cucumis sativus*, XP_004134935; *Solanum lycopersicum*, XP_004238204; *Fragaria vesca*, XP_004295571; *Solanum tuberosum*, XP_006354146; *Capsella rubella*, XP_006307488; *Phaseolus vulgaris*, XP_007148072; *Prunus persica*, XP_007211625; *Citrus sinensis*, KDO77799; *Malus domestica*, XP_008383652; *Coffea canephora*, CDP17659; *Eucalyptus grandis*, XP_010052232; *Sesamum indicum*, XP_011089230; *Jatropha curcas*, XP_012091734; *Medicago truncatula*, XP_003593719; *Brassica oleracea*, XP_013638333; *Brassica napus*, XP_013718258; *Cynara cardunculus*, KVH97037; *Arachis duranensis*, XP_015943149;

Capsicum annuum, XP_016550069; *Nicotiana tabacum*, NP_001312097; *Gossypium hirsutum*, XP_016686828; *Daucus carota*, XP_017220052; *Theobroma cacao*, XP_007025453; *Juglans regia*, XP_018838506; *Nelumbo nucifera*, XP_019055767; *Lupinus angustifolius*, XP_019426044; *Cephalotus follicularis*, GAV61508; *Corchorus olitorius*, OMO65309; *Manihot esculenta*, XP_021597940; *Hevea brasiliensis*, XP_021692452; *Chenopodium quinoa*, XP_021749332; *Spinacia oleracea*, XP_021852800; *Helianthus annuus*, XP_021992790; *Vigna radiata*, XP_014517062; *Olea europaea*, XP_022868027; *Cucurbita maxima*, XP_022978699; *Trifolium pratense*, PNX98537; *Trema orientalis*, PON99106; *Lactuca sativa*, XP_023736798; *Quercus suber*, XP_023907465; *Morus notabilis*, XP_024029268. In addition, the four closest homologs of AT1G28680 were identified via BLASTP at TAIR (The Arabidopsis Information Resource) and these were added to the sequence list. The final 58 protein sequences were aligned using MUSCLE[46] with default parameters. After trimming, the alignment was used to select an optimal model for Maximum Likelihood (ML) analysis in MEGA7[47]. According to the Akaike Information Criteria, the best model for phylogenetic analysis was the Jones-Taylor-Thornton (JTT) model, Gamma distributed with Invariant sites (+G +I). Based on this information, an ML analysis was performed using MEGA-CC after defining the parameters with MEGA-Proto. Bootstrap probabilities of each node were calculated using 500 replicates. The consensus tree was visualized by Dendroscope[48] as cladogram with the *Arabidopsis* homologs as outgroup. Final modifications were made in Inkscape.

Data availability

The raw chromatograms of the phenolic profiling experiment of *cosy-1*, *cosy-2* and wild-type roots are available from the MetaboLights database repository[36] as study MTBLS692. The data that support the findings of this study are available from the corresponding author upon request.

References

1. Sun, H. *et al.* Scopoletin is a phytoalexin against *Alternaria alternata* in wild tobacco dependent on jasmonate signalling. *J. Exp. Bot.* **65,** 4305-4315 (2014).
2. Gnonlonfin, G. J. B., Sanni, A. & Brimer, L. Review scopoletin - a coumarin phytoalexin with medicinal properties. *Crit. Rev. Plant Sci.* **31,** 47-56 (2012).
3. Chong, J. *et al.* Downregulation of a pathogen-responsive tobacco UDP-Glc:phenylpropanoid glucosyltransferase reduces scopoletin glucoside accumulation, enhances oxidative stress, and weakens virus resistance. *Plant Cell* **14,** 1093-1107 (2002).
4. Stringlis, I. A. *et al.* MYB72-dependent coumarin exudation shapes root microbiome assembly to promote plant health. *Proc. Natl. Acad. Sci. USA* **115,** E5213-E5222 (2018).
5. Rodríguez-Celma, J. & Schmidt, W. Reduction-based iron uptake revisited: on the role of secreted iron-binding compounds. *Plant Signal. Behav.* **8,** e26116 (2013).
6. Schmid, N. B. *et al.* Feruloyl-CoA 6 '-hydroxylase1-dependent coumarins mediate iron acquisition from alkaline substrates in Arabidopsis. *Plant Physiol.* **164,** 160-172 (2014).
7. Schmidt, H. *et al.* Metabolome analysis of *Arabidopsis thaliana* roots identifies a key metabolic pathway for iron acquisition. *PLoS ONE* **9,** e102444 (2014).
8. Siwinska, J. *et al.* Scopoletin 8-hydroxylase: a novel enzyme involved in coumarin biosynthesis and iron-deficiency responses in Arabidopsis. *J. Exp. Bot.* **69,** 1735-1748 (2018).
9. Tsai, H.-H. *et al.* Scopoletin 8-hydroxylase-mediated fraxetin production is crucial for iron mobilization. *Plant Physiol.* **177,** 194-207 (2018).

10. Rajniak, J. *et al.* Biosynthesis of redox-active metabolites in response to iron deficiency in plants. *Nat. Chem. Biol.* **14,** 442-450 (2018).
11. Matsumoto, S., Mizutani, M., Sakata, K. & Shimizu, B.-I. Molecular cloning and functional analysis of the *ortho*-hydroxylases of *p*-coumaroyl coenzyme A/feruloyl coenzyme A involved in formation of umbelliferone and scopoletin in sweet potato, *Ipomoea batatas* (L.) Lam. *Phytochemistry* **74,** 49-57 (2012).
12. Vialart, G. *et al.* A 2-oxoglutarate-dependent dioxygenase from *Ruta graveolens* L. exhibits *p*-coumaroyl CoA 2'-hydroxylase activity (C2'H): a missing step in the synthesis of umbelliferone in plants. *Plant J.* **70,** 460-470 (2012).
13. Kai, K. *et al.* Scopoletin is biosynthesized via *ortho*-hydroxylation of feruloyl CoA by a 2-oxoglutarate-dependent dioxygenase in *Arabidopsis thaliana*. *Plant J.* **55,** 989-999 (2008).
14. Karamat, F. *et al.* CYP98A22, a phenolic ester 3 '-hydroxylase specialized in the synthesis of chlorogenic acid, as a new tool for enhancing the furanocoumarin concentration in *Ruta graveolens*. *BMC Plant Biol.* **12,** 152 (2012).
15. Tuominen, L. K., Johnson, V. E. & Tsai, C.-J. Differential phylogenetic expansions in BAHD acyltransferases across five angiosperm taxa and evidence of divergent expression among *Populus* paralogues. *BMC Genomics* **12,** 236 (2011).
16. Vanholme, R. *et al.* Caffeoyl shikimate esterase (CSE) is an enzyme in the lignin biosynthetic pathway in *Arabidopsis*. *Science* **341,** 1103-1106 (2013).
17. Vanholme, R. *et al.* A systems biology view of responses to lignin biosynthesis perturbations in *Arabidopsis*. *Plant Cell* **24,** 3506-3529 (2012).
18. Yu, X.-H., Gou, J.-Y. & Liu, C.-J. BAHD superfamily of acyl-CoA dependent acyltransferases in *Populus* and *Arabidopsis*: bioinformatics and gene expression. *Plant Mol. Biol.* **70,** 421-442 (2009).
19. Ahn, Y. O. *et al.* Scopolin-hydrolyzing β-glucosidases in roots of Arabidopsis. *Plant and Cell Physiology* **51,** 132-143 (2010).
20. Dorey, S. *et al.* Spatial and temporal induction of cell death, defense genes, and accumulation of salicylic acid in tobacco leaves reacting hypersensitively to a fungal glycoprotein elicitor. *Mol. Plant-Microbe Interact.* **10,** 646-655 (1997).
21. Beuerle, T. & Pichersky, E. Enzymatic synthesis and purification of aromatic coenzyme A esters. *Anal. Biochem.* **302,** 305-312 (2002).
22. Levsh, O. *et al.* Dynamic conformational states dictate selectivity toward the native substrate in a substrate-permissive acyltransferase. *Biochemistry* **55,** 6314-6326 (2016).
23. Walker, A. M. *et al.* Elucidation of the structure and reaction mechanism of sorghum hydroxycinnamoyltransferase and its structural relationship to other coenzyme A-dependent transferases and synthases. *Plant Physiol.* **162,** 640-651 (2013).
24. Ma, X., Koepke, J., Panjikar, S., Fritzsch, G. & Stöckigt, J. Crystal structure of vinorine synthase, the first representative of the BAHD superfamily. *J. Biol. Chem.* **280,** 13576-13583 (2005).
25. Rodríguez-Celma, J. *et al.* Mutually exclusive alterations in secondary metabolism are critical for the uptake of insoluble iron compounds by Arabidopsis and *Medicago truncatula*. *Plant Physiol.* **162,** 1473-1485 (2013).
26. Fourcroy, P. *et al.* Involvement of the ABCG37 transporter in secretion of scopoletin and derivatives by *Arabidopsis* roots in response to iron deficiency. *New Phytol.* **201,** 155-167 (2014).

27. Rizhsky, L., Liang, H. & Mittler, R. The water-water cycle is essential for chloroplast protection in the absence of stress. *J. Biol. Chem.* **278,** 38921-38925 (2003).
28. Yang, S.-M., Shim, G. Y., Kim, B.-G. & Ahn, J.-H. Biological synthesis of coumarins in *Escherichia coli*. *Microb. Cell Fact.* **14,** 65 (2015).
29. Lin, Y., Sun, X., Yuan, Q. & Yan, Y. Combinatorial biosynthesis of plant-specific coumarins in bacteria. *Metab. Eng.* **18,** 69-77 (2013).
30. Kuromori, T. *et al.* A collection of 11 800 single-copy *Ds* transposon insertion lines in *Arabidopsis*. *Plant J.* **37,** 897-905 (2004).
31. Ito, T. *et al.* A new resource of locally transposed *Dissociation* elements for screening gene-knockout lines in silico on the Arabidopsis genome. *Plant Physiol.* **129,** 1695-1699 (2002).
32. Sundaresan, V. *et al.* Patterns of gene action in plant development revealed by enhancer trap and gene trap transposable elements. *Genes Dev.* **9,** 1797-1810 (1995).
33. Shimada, T. L., Shimada, T. & Hara-Nishimura, I. A rapid and non-destructive screenable marker, FAST, for identifying transformed seeds of *Arabidopsis thaliana*. *Plant J.* **61,** 519-528 (2010).
34. Ramakers, C., Ruijter, J. M., Deprez, R. H. L. & Moorman, A. F. M. Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neurosci. Lett.* **339,** 62-66 (2003).
35. Czechowski, T., Stitt, M., Altmann, T., Udvardi, M. K. & Scheible, W.-R. Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. *Plant Physiol.* **139,** 5-17 (2005).
36. Kale, N. S. *et al.* MetaboLights: an open-access database repository for metabolomics data. *Curr. Protoc. Bioinform.* **53,** 14.13.11-14.13.18 (2016).
37. Murashige, T. & Skoog, F. A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Physiol. Plant.* **15,** 473-497 (1962).
38. Rao, G. K., Gowda, N. B. & Ramakrishna, R. A. Raney nickel–catalyzed hydrogenation of unsaturated carboxylic acids with sodium borohydride in water. *Synth. Commun.* **42,** 893-904 (2012).
39. Waterhouse, A. *et al.* SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* **46,** W296-W303 (2018).
40. Camacho, C. *et al.* BLAST+ : architecture and applications. *BMC Bioinformatics* **10,** 241 (2009).
41. Remmert, M., Biegert, A., Hauser, A. & Söding, J. HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. *Nat. Methods* **9,** 173-175 (2012).
42. Guex, N., Peitsch, M. C. & Schwede, T. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: a historical perspective. *Electrophoresis* **30,** S162-S173 (2009).
43. Benkert, P., Biasini, M. & Schwede, T. Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics* **27,** 343-350 (2011).
44. Trott, O. & Olson, A. J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **31,** 455-461 (2010).
45. Hiscox, J. D. & Israelstam, G. F. A method for the extraction of chlorophyll from leaf tissue without maceration. *Can. J. Bot.* **57,** 1332-1334 (1979).

46. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32,** 1792-1797 (2004).
47. Kumar, S., Stecher, G. & Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* **33,** 1870-1874 (2016).
48. Huson, D. H. *et al.* Dendroscope: an interactive viewer for large phylogenetic trees. *BMC Bioinformatics* **8,** 460 (2007).

Correspondence and requests for materials should be addressed to W.B.

Acknowledgements

The authors thank Annick Bleys for help in preparing the manuscript. We gratefully acknowledge funding through the framework of the IWT-SBO (project ARBOREF), the Hercules program of Ghent University for the Synapt Q-Tof (grant no. AUGE/014), the "Bijzonder Onderzoeksfonds-Zware Apparatuur" of Ghent University for the FT-ICR-MS instrument (174PZA05) and the Multidisciplinary Research Partnership 'Biotechnology for a Sustainable Economy' (01MRB510W) of Ghent University. R.V. is indebted to the Research Foundation-Flanders for a postdoctoral fellowship, L.S., L.H. and B.D.M. to the Agency for Innovation by Science and Technology, Flanders (IWT-Vlaanderen) and X.L. and J.L. to the China Scholarship Council (CSC). H.K. and J.R. were funded by the DOE Great Lakes Bioenergy Research Center (DOE BER Office of Science DE-FC02-07ER64494). Research in the Schmidt laboratory was supported by MoST.

Author contributions

R.V., L.S. W.S., and W.B. designed the experiments; R.V., L.S., K.C.S, H.K., X.L., B.D.M, L.H, J.L., G.G., H.H.T., and K.M. performed experiments; R.V., L.S., G.G., K.M., J.H., B.V., H.H.T., and J.R. collected and analyzed data; R.V. and W.B. wrote the manuscript with the help of all authors.

Additional information

Supplementary information is available for this paper. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to W.B.

Competing interests

A patent application for the use of *COSY* for the production of coumarins in plants and microorganisms has been filed.

Figure 1 | Revised coumarin biosynthetic pathway in *Arabidopsis* displaying metabolic shifts in the *cosy* mutants. The general phenylpropanoid pathway produces feruloyl-CoA that is metabolized via F6'H1 and COSY activity into scopoletin and various downstream products. COSY catalyzes the enzymatic step from 6-hydroxyferuloyl-CoA into scopoletin. The conversion of 6-hydroxyferuloyl-CoA to scopoletin also happens partially spontaneously (Fig. 2) and is catalyzed by light (Supplementary Fig. 13). The compounds in the red boxes show a significantly elevated abundance in roots of *cosy* mutants. Compounds in the blue box are significantly lower in abundance. Dashed arrows represent unknown but suggested/implicated conversions and are explained in more detail in the caption of Supplementary Fig. 10.

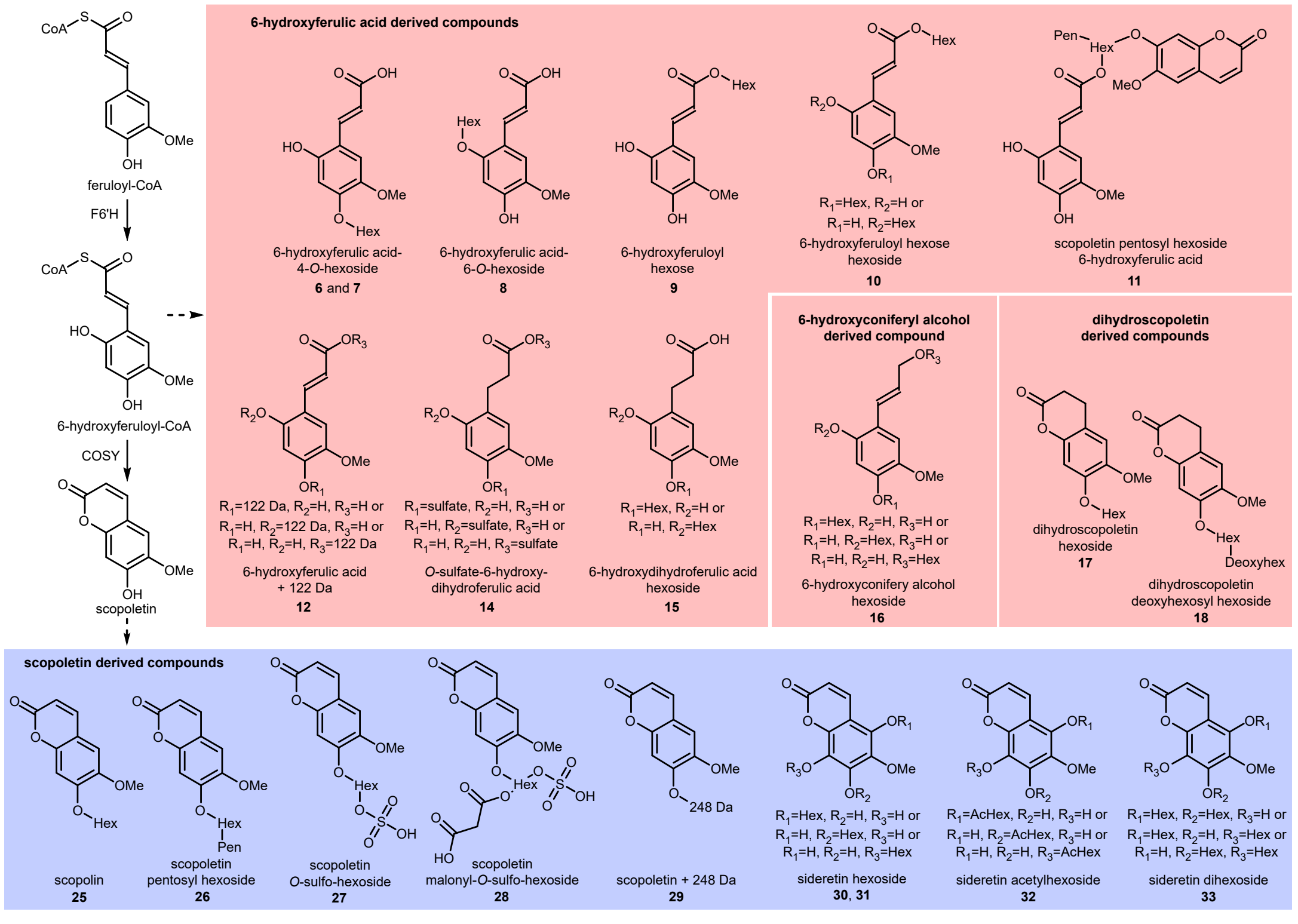
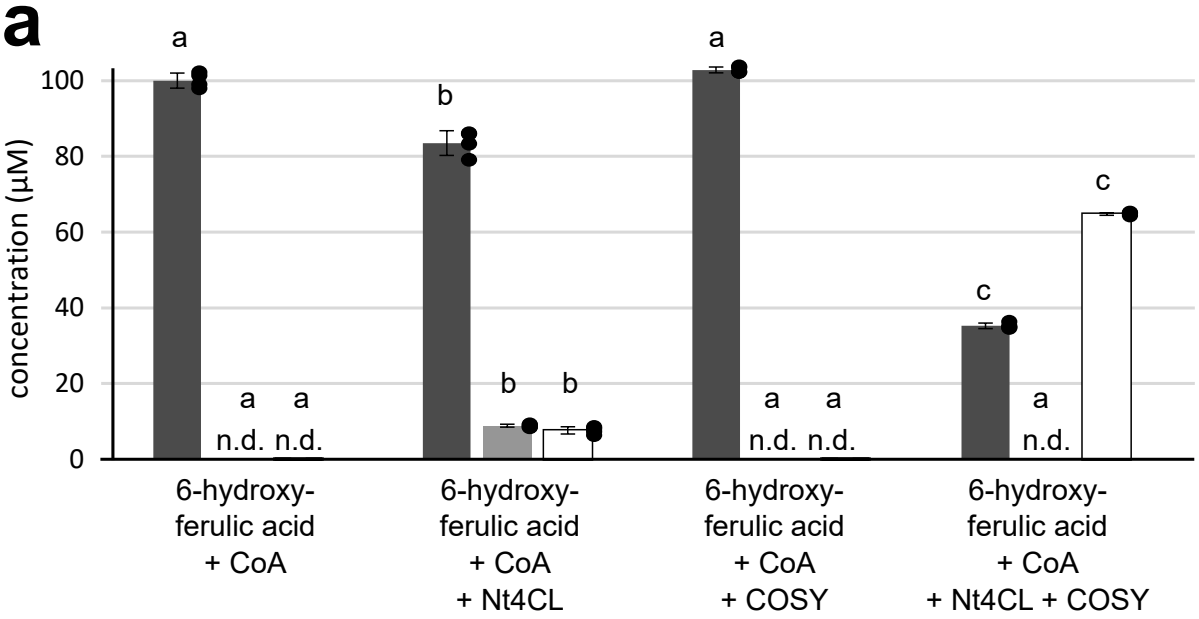
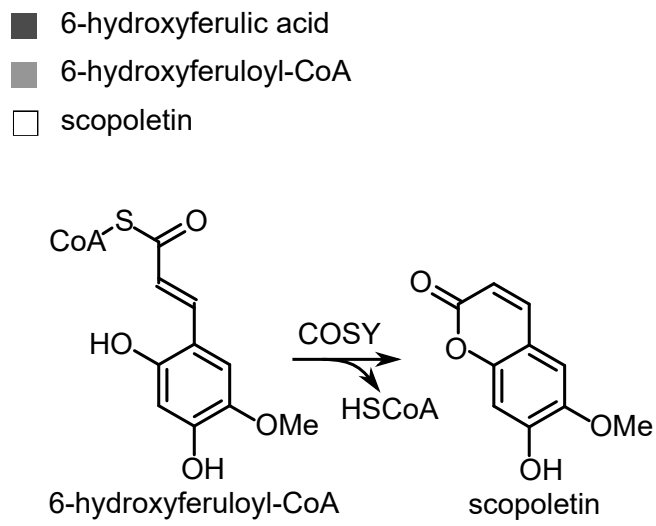

feruloyl-CoA
F6'H
6-hydroxyferuloyl-CoA
COSY
scopoletin
6-hydroxyferulic acid derived compounds
6-hydroxyferulic acid-4-O-hexoside
6 and 7
6-hydroxyferulic acid-6-O-hexoside
8
6-hydroxyferuloyl hexose
9
R1=Hex, R2=H or R1=H, R2=Hex
6-hydroxyferuloyl hexose hexoside
10
scopoletin pentosyl hexoside 6-hydroxyferulic acid
11
R1=122 Da, R2=H, R3=H or R1=H, R2=122 Da, R3=H or R1=H, R2=H, R3=122 Da
6-hydroxyferulic acid + 122 Da
12
R1=sulfate, R2=H, R3=H or R1=H, R2=sulfate, R3=H or R1=H, R2=H, R3=sulfate
O-sulfate-6-hydroxy-dihydroferulic acid
14
R1=Hex, R2=H or R1=H, R2=Hex
6-hydroxydihydroferulic acid hexoside
15
6-hydroxyconiferyl alcohol derived compound
R1=Hex, R2=H, R3=H or R1=H, R2=Hex, R3=H or R1=H, R2=H, R3=Hex
6-hydroxyconifery alcohol hexoside
16
dihydroscopoletin derived compounds
dihydroscopoletin hexoside
17
dihydroscopoletin deoxyhexosyl hexoside
18
scopoletin derived compounds
scopolin
25
scopoletin pentosyl hexoside
26
scopoletin O-sulfo-hexoside
27
scopoletin malonyl-O-sulfo-hexoside
28
scopoletin + 248 Da
29
R1=Hex, R2=H, R3=H or R1=H, R2=Hex, R3=H or R1=H, R2=H, R3=Hex
sideretin hexoside
30, 31
R1=AcHex, R2=H, R3=H or R1=H, R2=AcHex, R3=H or R1=H, R2=H, R3=AcHex
sideretin acetylhexoside
32
R1=Hex, R2=Hex, R3=H or R1=Hex, R2=H, R3=Hex or R1=H, R2=Hex, R3=Hex
sideretin dihexoside
33

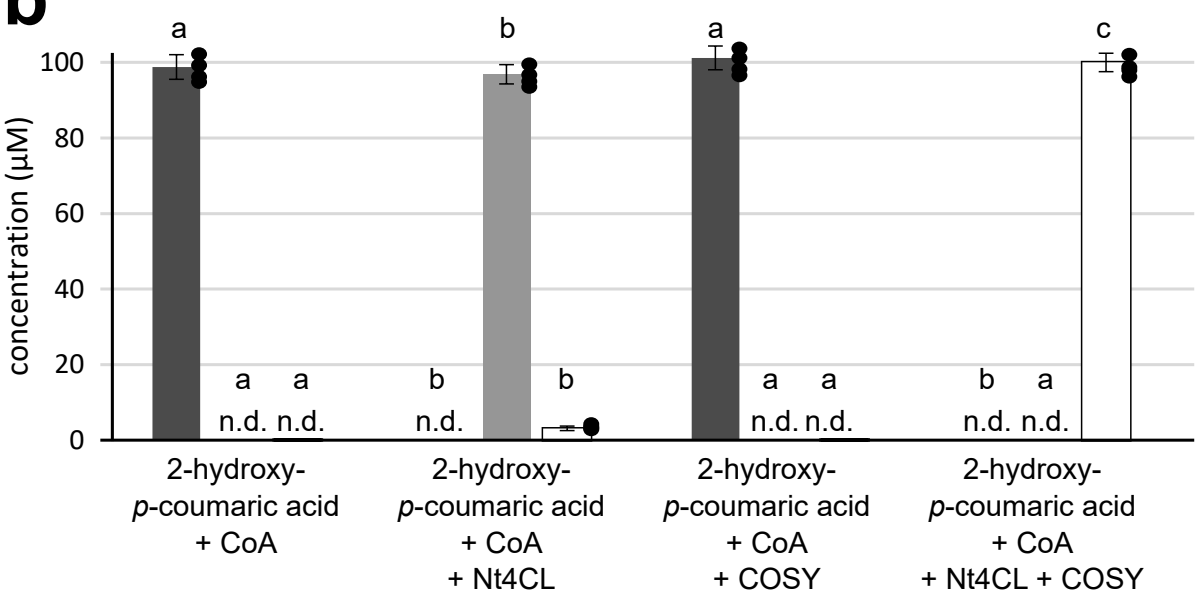
Figure 2 | Enzymatic activity of COSY. COSY and Nt4CL1 were supplemented with either 6-hydroxyferulic acid **(a)**, 2-hydroxy-*p*-coumaric acid **(b)** or 6-hydroxycaffeic acid **(c)**. The respective hydroxycinnamoyl-CoAs were formed by addition of Nt4CL, CoA and ATP. The *o*-hydroxycinnamoyl-CoAs were partly converted to scopoletin **(a)**, umbelliferone **(b)** and esculetin **(b)**, respectively. However, these coumarins were formed more efficiently from their respective *o*-hydroxycinnamoyl-CoAs in the presence of COSY. When Nt4CL1 was supplemented with dihydro-2-hydroxy-*p*-coumaric acid **(d)**, dihydro-2-hydroxy-*p*-coumaroyl-CoA remained below the detection limit, but a 95% conversion to dihydro-umbelliferone was detected. The height of the bars represent the average abundance of 4 technical replicates each, error bars indicate ± SD, dots indicate individual samples. Bars with the same letter do not differ significantly from each other ($p>0.05$; ANOVA, post-hoc Scheffe test). n.d., not detected (i.e. below detection limit). **(e)** Proposed molecular mechanism of COSY activity. The catalytic activity starts with the deprotonation of the hydroxyl at the *ortho* position to the aromatic ring's aliphatic side-chain (i.e., the 6-position), whereby electron delocalization makes the *trans*-*cis* isomerization and subsequent lactonization possible.

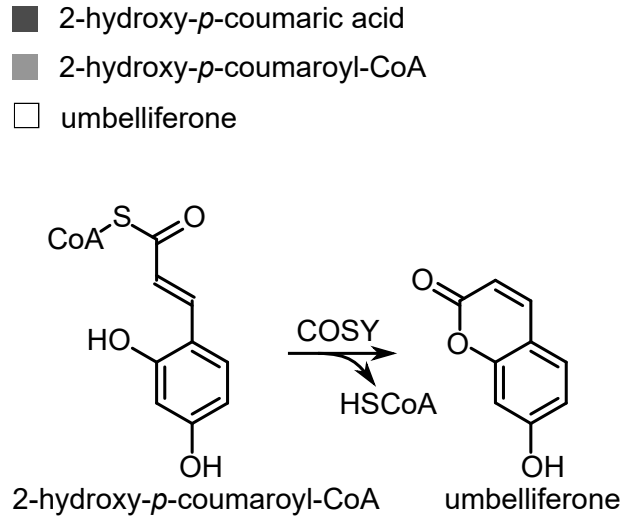
a


concentration (µM)
6-hydroxy-ferulic acid + CoA
6-hydroxy-ferulic acid + CoA + Nt4CL
6-hydroxy-ferulic acid + CoA + COSY
6-hydroxy-ferulic acid + CoA + Nt4CL + COSY
n.d.

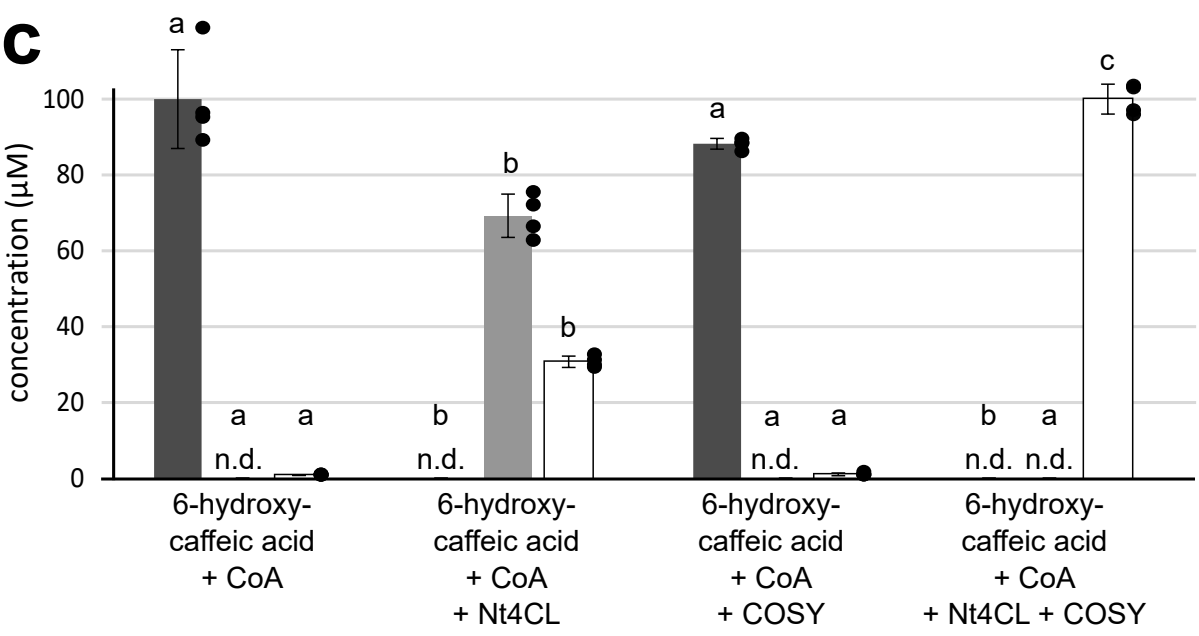

6-hydroxyferulic acid
6-hydroxyferuloyl-CoA
scopoletin
COSY
HSCoA
6-hydroxyferuloyl-CoA
scopoletin

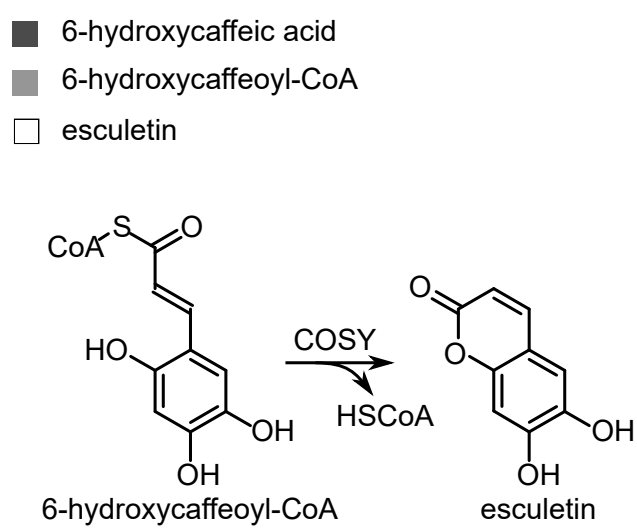
b


concentration (µM)
2-hydroxy-p-coumaric acid + CoA
2-hydroxy-p-coumaric acid + CoA + Nt4CL
2-hydroxy-p-coumaric acid + CoA + COSY
2-hydroxy-p-coumaric acid + CoA + Nt4CL + COSY
n.d.

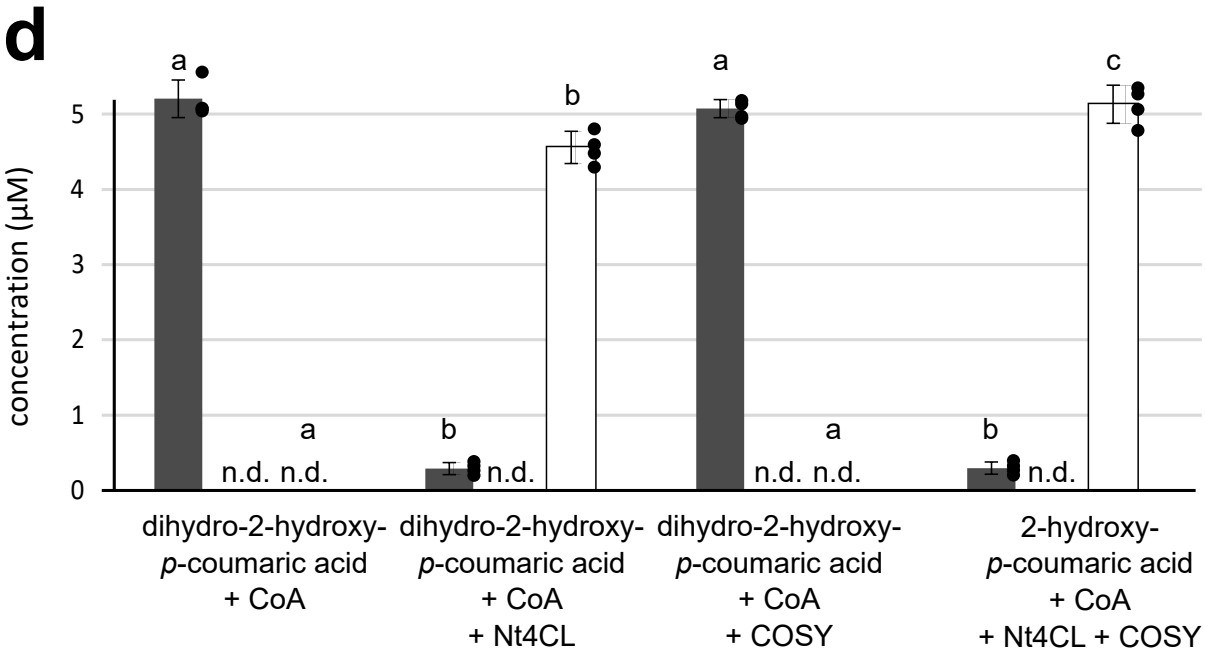

2-hydroxy-p-coumaric acid
2-hydroxy-p-coumaroyl-CoA
umbelliferone
COSY
HSCoA
2-hydroxy-p-coumaroyl-CoA
umbelliferone

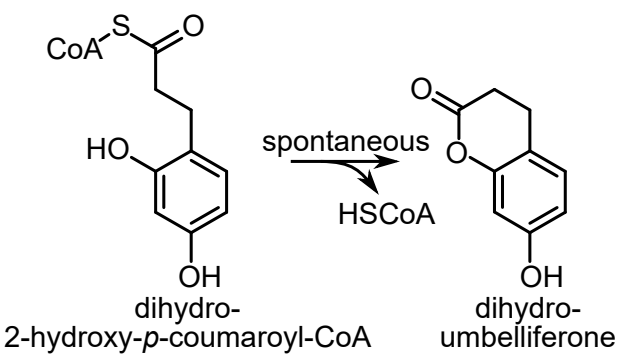
c


concentration (µM)
6-hydroxy-caffeic acid + CoA
6-hydroxy-caffeic acid + CoA + Nt4CL
6-hydroxy-caffeic acid + CoA + COSY
6-hydroxy-caffeic acid + CoA + Nt4CL + COSY
n.d.


6-hydroxycaffeic acid
6-hydroxycaffeoyl-CoA
esculetin
COSY
HSCoA
6-hydroxycaffeoyl-CoA
esculetin

d


concentration (µM)
dihydro-2-hydroxy-p-coumaric acid + CoA
dihydro-2-hydroxy-p-coumaric acid + CoA + Nt4CL
dihydro-2-hydroxy-p-coumaric acid + CoA + COSY
2-hydroxy-p-coumaric acid + CoA + Nt4CL + COSY
n.d.


dihydro-2-hydroxy-p-coumaric acid
dihydro-2-hydroxy-p-coumaroyl-CoA
dihydro-umbelliferone
spontaneous
HSCoA
dihydro-2-hydroxy-p-coumaroyl-CoA
dihydro-umbelliferone

e

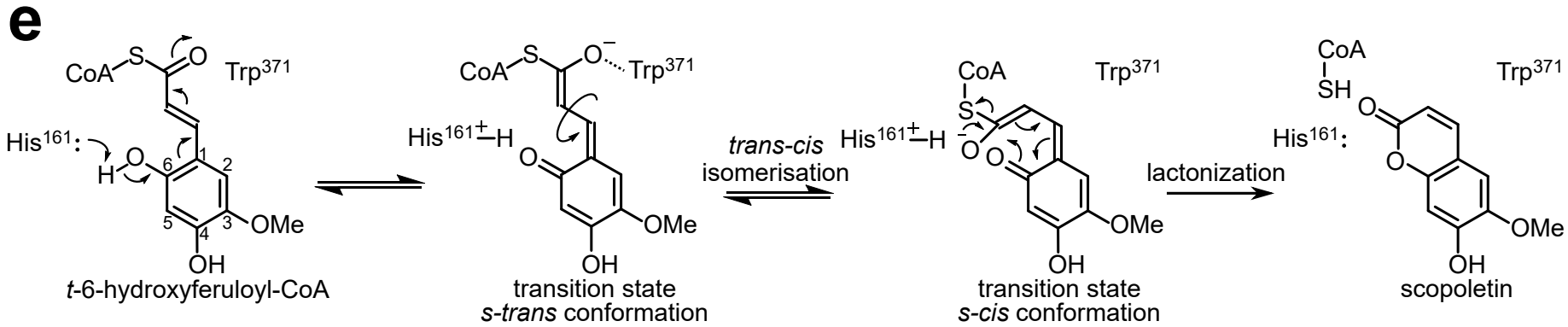
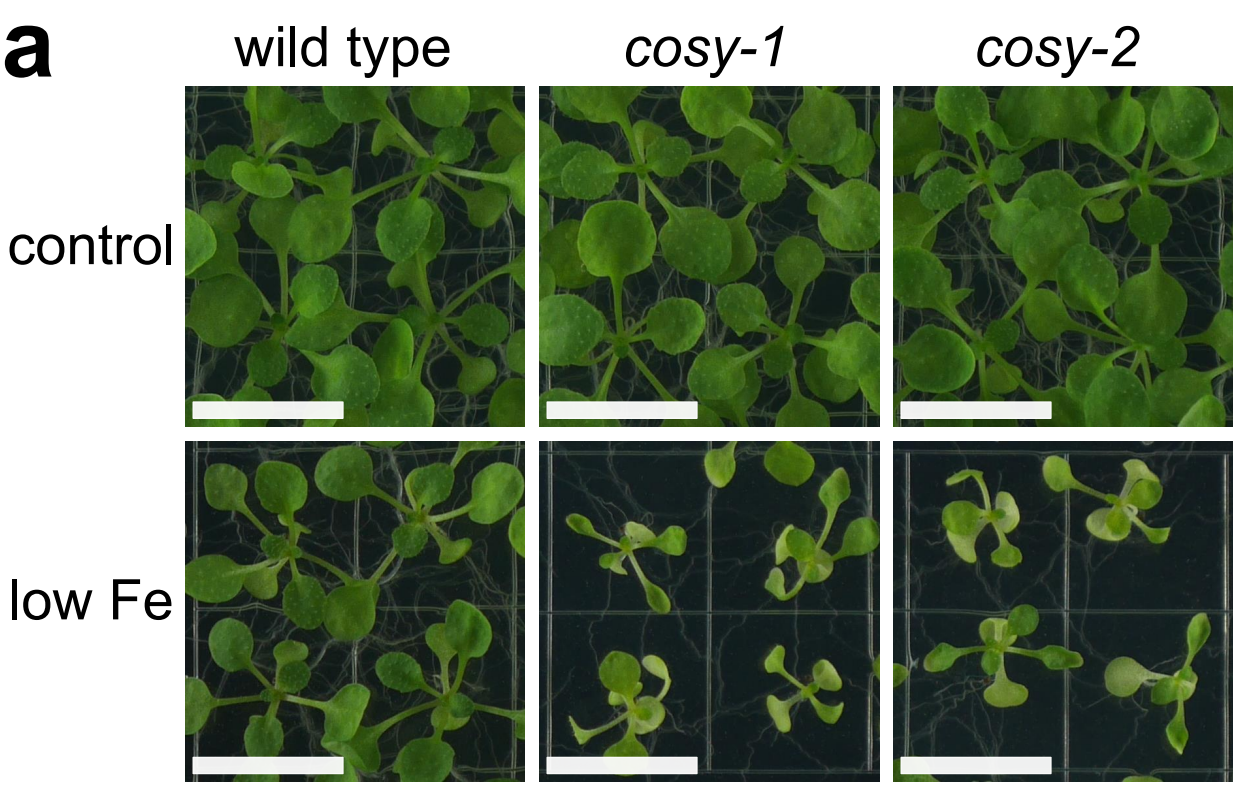
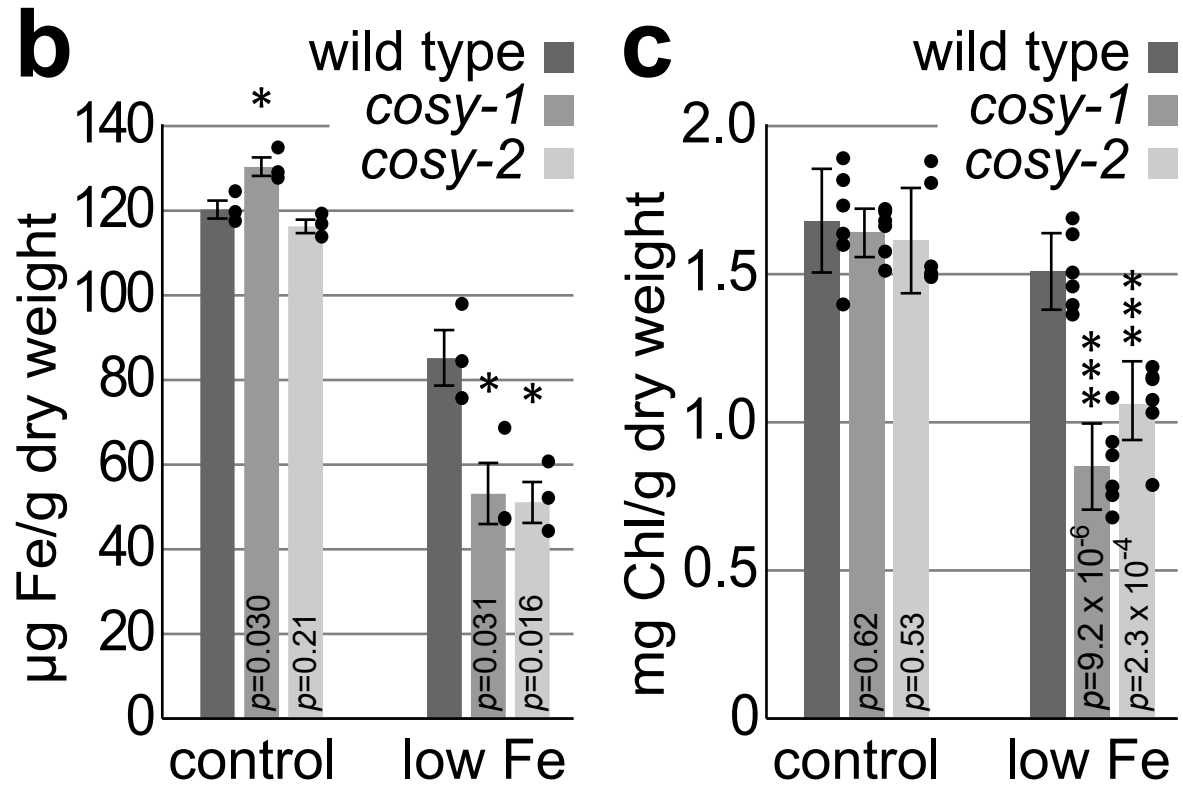
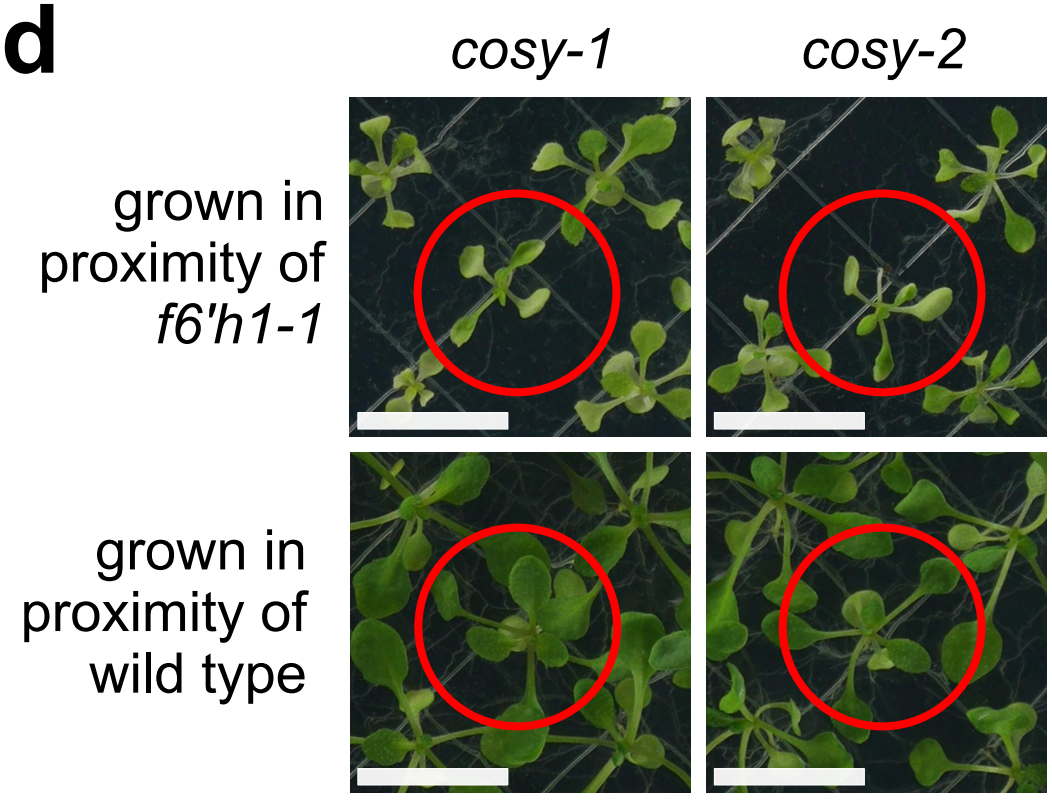
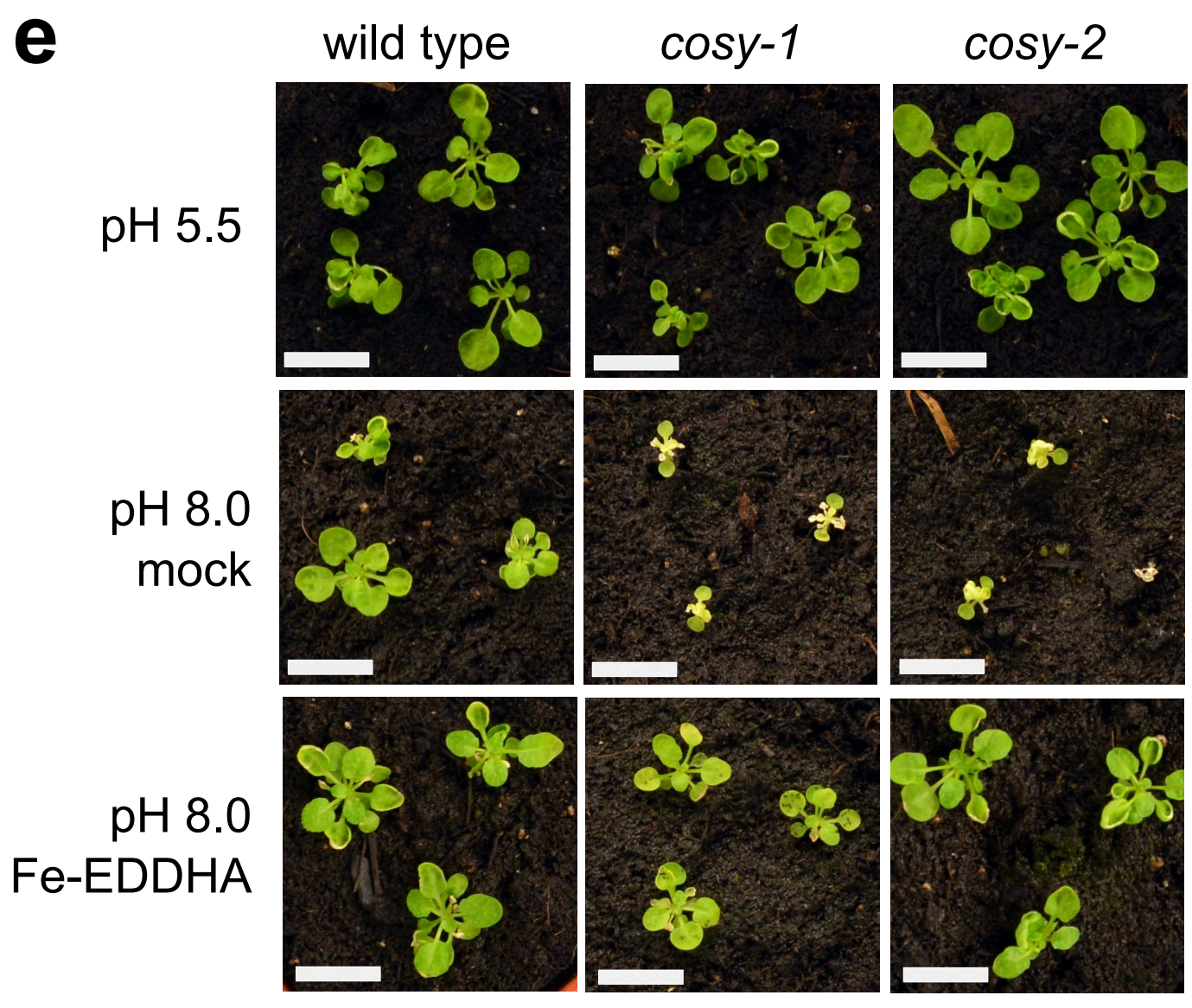

Trp371
His161
t-6-hydroxyferuloyl-CoA
transition state s-trans conformation
trans-cis isomerisation
transition state s-cis conformation
lactonization
scopoletin

Figure 3 | *cosy* mutants are sensitive to media or soil with reduced iron bioavailability. (a) Effects of iron bioavailability on the growth of *cosy* mutants. Phenotypes of plants grown for 14 days on medium with either bioavailable iron (control) or low-bioavailability iron (low Fe). Representative images of three independent experiments are shown. Scale bars = 1 cm. **(b)** Shoot iron concentration. The height of the bars represent the average, error bars indicate ± SEM and dots indicate individual biological replicates, n=3 pools of at least 20 plants. Asterisks indicate significant differences from the wild type (No-0) in each treatment, *0.05>*p*-value>0.01, exact *p*-values are given on the bar, unpaired two-sided Student's *t*-test. **(c)** Shoot chlorophyll (Chl) concentration. The height of the bars represent the average abundance, error bars indicate ± SEM and dots indicate individual biological replicates, n=6 pools of 2 plants. Asterisks indicate significant differences from the wild type (No-0) in each treatment, ***0.001>*p*-value, exact *p*-values are given on the bar, one-way ANOVA followed by post-hoc Scheffe test. **(d)** Rescue of *cosy* mutants by growing alongside wild-type seedlings. Phenotypes of *cosy* mutants (circled in red) grown for 14 d on media with low-bioavailability iron in close proximity of either *f6'h1-1* mutant or wild-type seedlings. Representative images from three independent experiments are shown. Scale bars = 1 cm. **(e)** *cosy* mutants are sensitive to alkaline soils. Three-week-old *cosy* seedlings developed chlorosis when grown on soil at pH 8.0. Watering with Fe(III)-EDDHA avoided the chlorosis development. The pictures show representative samples of 20 biological replicates per treatment. Scale bars = 1 cm.


a
wild type
cosy-1
cosy-2
control
low Fe


b
wild type
cosy-1
cosy-2
µg Fe/g dry weight
140
120
100
80
60
40
20
0
*
p=0.030
p=0.21
*
p=0.031
*
p=0.016
control
low Fe
c
wild type
cosy-1
cosy-2
mg Chl/g dry weight
2.0
1.5
1.0
0.5
0
p=0.62
p=0.53
****
p=9.2 x 10-6
***
p=2.3 x 10-4
control
low Fe


d
cosy-1
cosy-2
grown in proximity of f6'h1-1
grown in proximity of wild type


e
wild type
cosy-1
cosy-2
pH 5.5
pH 8.0 mock
pH 8.0 Fe-EDDHA