Methods

Plant material

Three commercial, high-morphine varieties (i.e. L, 11 and 40) of opium poppy (*Papaver somniferum*) and a mutant variety (i.e. T) that accumulates high levels of thebaine and oripavine, but lacks morphine and codeine[8], were cultivated in a growth chamber (Conviron, Winnipeg, Canada) at 20°C/18°C (light/dark) under high-intensity metal halide lights with a photoperiod of 16 h. Plant materials for gene expression and/or alkaloid analysis was harvested one day before the onset of flowering and stored at -80°C until further analysis. For virus-induced gene silencing (VIGS) experiments, the opium poppy variety Bea's Choice was cultivated under greenhouse conditions.

Chemicals

Morphine and codeine were gifts from Sanofi-Aventis (Paris, France; http://en.sanofi-aventis.com). Thebaine and oripavine were isolated from the latex of opium poppy variety T. Methanol extracts of T latex were concentrated under reduced pressure and spotted on thin layer chromatography (TLC) Silica gel 60 F_{254} plates (Merck, Whitehouse Station, NJ; http://www.merck.com). TLC was performed as described previously[6] and alkaloids were visualized under UV illumination at 254 nm. Silica was scraped off the plates from regions corresponding to the R_f values of authentic thebaine and oripavine and extracted three times with methanol. Pooled methanol extracts were concentrated under reduced pressure. The purity and identity of thebaine and oripavine were confirmed by electron-impact-mass spectrometry (EI-MS). The EI mass spectra were in agreement with those reported previously[15,26]. Noscapine and papaverine were purchased from Sigma-Aldrich (St. Louis, MO; http://www.sigmaaldrich.com). Other alkaloids were available as described previously[27,28]. [1-^{14}C]2-Oxoglutarate was purchased from American Radiolabeled Chemicals (St. Louis, MO; http://www.arcincusa.com). All other chemicals were purchased from Sigma-Aldrich.

Genetic inheritance

To determine the mode of inheritance underlying the alkaloid phenotype of variety T (i.e. high-thebaine/oripavine, morphine/codeine-free) reciprocal crosses were generated between T and the high-morphine cultivars 11 and 40. Flower buds were dissected one day prior to anthesis, immature stamens were excised and stigmas were pollinated with pollen from the appropriate variety or cross. Pollinated flowers were covered for 1-2 days to allow the development of seed capsules free from contaminating pollen. F_1 plants were either self-pollinated for the production of F_2 seed, or backcrossed with P_1 plants. The alkaloid phenotypes of individual F_2 and backcrossed plants were qualitatively scored as described in the metabolite profiling section. The frequency of plants from each cross that displayed the high-thebaine/oripavine, morphine/codeine-free phenotype (i.e. %T) is shown in Supplementary Table 1.

Nature Chemical Biology: doi:10.1038/nchembio.317

Phylogenetic analysis

Phylogeny (Supplementary Fig. 4) and amino acid alignments (Supplementary Fig. 5) were performed using ClustalX[29], and phylogenetic data were displayed using TREEVIEW[30]. Species and associated GenBank accession numbers are as follows: *Anisodus acutangulus* hyoscyamine 6β-hydroxylase AaH6H (ABM74185); *Arabidopsis thaliana* senescence-related gene 1 AtSRG1 (NP_173145); *Arabidopsis thaliana* 4-hydroxyphenylpyruvate dioxygenase At4HPPD (AAB58404); *Arabidopsis thaliana* anthocyanidin synthase AtAS (Q96323); *Atropa belladonna* hyoscyamine 6β-hydroxylase AbH6H (BAA78340); *Catharanthus roseus* desacetoxyvindoline 4-hydroxylase CrD4H (AAC49827); *Citrus unshiu* flavanol synthase CuFS (BAA36554); *Coptis japonica* norcoclaurine synthase-1 CjNCS (A2A1A0) *Cucurbita maxima* gibberellin 7-oxidase CmG7O (AAB64346); *Cucurbita maxima* gibberellin 20-oxidase CmG20O (AAB64345); *Citrus unshiu* flavanol synthase CuFS (BAA36554); *Papaver somniferum* thebaine 6-*O*-demethylase PsT6ODM (GQ500139); *Papaver somniferum* codeine *O*-demethylase PsCODM (GQ500141); *Papaver somniferum* PsDIOX2 (GQ500140); *Petunia hybrida* flavanone 3β-hydroxylase PhF3H (AAC49929); *Populus trichocarpa* Pt_dioxygenase-like (XP_002300453); *Ricinus communis* Rc_dioxygenase-like (EEF42734); *Solanum lycopersicum* ACC oxidase SlACCO (P24157); *Vitis vinifera* Vv_dioxygenase-like (CAO70478); *Hyoscyamus niger* hyoscyamine 6β-hydroxylase HnH6H (AAA3338); *Zea mays* flavonol synthase/flavanone 3-hydroxylase ZmFS/F3H (ACG44904).

Microarray construction

A custom-built opium poppy microarray was constructed based on expressed sequence tags (ESTs) derived from elicited cell culture[31] and stem cDNA libraries. A total of 22,752 spots were printed corresponding to 12,768 ESTs from cell culture and 9,984 ESTs from stem, which represented a total of 19,185 genetic elements and 14,355 unigenes. To construct the stem cDNA library, 10 cm of stem immediately below the flower buds of opium poppy variety L were harvested one day prior to anthesis. Total RNA isolation was performed as described previously[32] and poly(A)$^+$ RNA was selected by oligo(dT)-cellulose chromatography. A unidirectional cDNA library was constructed in λUni-ZAPII XR, according to the instructions of the manufacturer (Stratagene, Santa Clara, CA; http://www.stratagene.com). An amplified cDNA library derived from approximately 1 x 10^7 primary plaque-forming units was mass excised, and individual bacterial colonies were randomly isolated and cultured in 96-well microtiter plates. Plasmid DNA was prepared using the TempliPhi amplification kit (GE Healthcare Life Sciences, Piscataway, NJ; http://www.gelifesciences.com) and sequenced from the 5'-end using a 3730*xl* capillary electrophoresis DNA analyzer (Applied Biosystems, Foster City, CA; http://www.appliedbiosystems.com). Stem expressed sequence tags (ESTs) were analyzed as described previously[27,31]. Sequenced cDNAs from elicited cell culture and stem libraries were

Nature Chemical Biology: doi:10.1038/nchembio.317

amplified from pBluescript SK- using T3 and T7 primers in 0.2 ml capacity 96-well PCR plates (Corning, Corning, NY; http://www.corning.com). Agarose gel electrophoresis was used to ensure that each reaction produced sufficient amplicon abundance. PCR products were purified using Montage PCR_{96} plates (Millipore, Billerica, MA; http://www.millipore.com), recovered in 50 μl of water, transferred to polypropylene V-bottom 96-well plates (Corning), lyophilized to dryness, resuspended in 6 μl 3 x SSC buffer, and arrayed into 384-well polypropylene V-bottom plates (Axygen; Union City, CA; http://www.axygen.com) for printing (Microarray and Proteomics Facility, University of Alberta. Individual spots were printed using a BioMek FX (Beckman-Coulter, Fullerton, CA; http://www.beckman.com) onto SuperAmine Substrate (ArraIt, Sunnyvale, CA; http://www.arrayit.com) slides.

Microarray hybridization and analysis

RNA from opium poppy stem tissue was isolated using a previously described protocol[33] involving guanidinium thiocyanate-based extraction and cesium chloride-based density centrifugation. High quality RNA (100 μg) was reverse transcribed using BD PowerScript reverse transcriptase (BD Biosciences, Franklin Lakes, NJ; http://www.bdbiosciences.com) and labeled with Cy3- or Cy5-dCTP fluorescent dyes (GE Healthcare Life Sciences). Microarray slide preparation, probe-target hybridization, and washing steps were performed as described previously[31]. A total of 6 hybridization experiments were performed, including technical (i.e. dye-flip) replicates, in which the relative abundance of transcripts in variety T was compared with those of varieties L, 11 and 40, respectively (i.e. duplicate experiments for each of T vs L, T vs 11 and T vs 40). Fluorescence signatures were captured using a Scanarray 5000 scanner (PerkinElmer, Waltham, MA; http://www.perkinelmer.com) and analyzed using the TIGR TM4 suite of microarray tools[34]. Poor quality, low intensity or missing spots were flagged and excluded from further analysis. Florescence signals for each experiment were subjected to LOWESS normalization, followed by manual analysis. Transcripts that were potentially less or more abundant in T compared with the other varieties were identified using a signal intensity ratio cutoff of 1.8. Based on this criterion, eight genes were putatively underexpressed in T compared with at least two other varieties. Only one of these eight genes, represented on the microarray as the EST sequence 06_B04 (GenBank accession FE964517) originating from the cell culture cDNA library was putatively underepxressed in T compared with all three morphine-producing varieties.

Expression vector construction

The EST sequence 06_B04 (GenBank accession FE964517) identified using microarray analysis was used to query an EST database containing 10,148 sequences from elicitor-treated cell cultures[31] and 7,949 sequences from stems of opium poppy using the tBLASTn algorithm. Although the 06_B04 cDNA was incomplete, a full-length cDNA (named DIOX1) was identified in the cell culture EST database.

Nature Chemical Biology: doi:10.1038/nchembio.317

Two additional, full-length cDNAs with substantial nucleotide identity to DIOX1 were identified in the cell culture (DIOX2, 86%) and stem (DIOX3, 67%) cDNA libraries. Open reading frames (ORFs) encoding DIOX1, DIOX2 and DIOX3 were amplified from cDNA templates with *taq* polymerase (Invitrogen, Carlsbad, CA; http://www.invitrogen.com), using sense and antisense primers with flanking *Bam*HI and *Pst*I restriction sites, respectively (Supplementary Table 2). PCR products were individually ligated to pGEM-T (Promega, Madison, WI; http://www.promega.com), digested with *Bam*HI and *Pst*I, and ligated to pQE30 (Qiagen, Valencia, CA; http://www.qiagen.com) pre-cut with *Bam*HI and *Pst*I. DNA sequencing of cloned amplicons was performed following ligation into pGEM-T and pQE30 vectors (UCDNA Services, University of Calgary). Plasmid propagation was performed in the *Escherichia coli* strain XL1-Blue, and expression constructs were transformed into *E. coli* strain SG13009 to generate recombinant enzymes.

Protein expression and purification

Production of DIOX1, DIOX2 and DIOX3 proteins in *E. coli* cultures was induced with 0.3 mM isopropyl β-D-thiogalactopyranoside (IPTG) followed by incubation at either 20°C for 4 hours (DIOX1 and DIOX2) or 4°C for 24 hours (DIOX3). Cells were harvested and bacterial pellets were resuspended in Buffer A (100 mM Tris-HCl pH 7.4, 10% (v/v) glycerol and 14 mM 2-mercaptoethanol). Lysis was achieved using a French Press at 0.1 GPa (15,000 psi) and lysates were cleared by centrifugation at 10,000*g* for 10 min. Cleared lysates were loaded onto Talon cobalt affinity columns (Clontech, Mountain View, CA; http://www.clontech.com) and purified, His-tagged proteins were eluted according to the manufacturer's instructions. The columns with protein bound were washed with Buffer A, and recombinant proteins were eluted using Buffer A containing 100 mM imidazole. Desalting was performed using PD10 columns (GE Healthcare Life Sciences) and protein concentrations were determined using the Bio-Rad Protein Assay (Bio-Rad, Hercules, CA; http://www.bio-rad.com). The purity of recombinant proteins was determined using SDS-PAGE[35].

Enzyme assays

The direct enzyme assay for 2-oxoglutarate-dependent dioxygenase activity was performed using a reaction mixture of 100 mM Tris-HCl (pH 7.4), 10% (v/v) glycerol, 14 mM 2-mercaptoethanol, 1 mM alkaloid, 10 mM 2-oxoglutarate, 10 mM sodium ascorbate, 0.5 mM $FeSO_4$, and up to 100 μg of purified recombinant enzyme. Assays were carried out at 30°C for 1 or 4 hours, stopped by immersing the reaction tube in boiling water for 5 min, and subjected to LC-MS/MS analysis. 2-Oxoglutarate-dependent dioxygenase activity was also assayed using an indirect method based on the *O*-demethylation-coupled decarboxylation of [1-^{14}C]2-oxoglutarate[14]. Briefly, the standard assay contained 10 μM of a 10% mole/mole (n/n) solution of [1-^{14}C]2-oxoglutarate (specific activity 55 mCi/mmol) diluted with 90% n/n unlabeled 2-oxoglutarate, 10 μM unlabelled alkaloid substrate, 10

Nature Chemical Biology: doi:10.1038/nchembio.317

mM sodium ascorbate, 0.5 mM iron sulfate, and 5 μg purified enzyme in a 500 μl buffered (100 mM Tris-HCl, 10% [v/v] glycerol, 14 mM 2-mercaptoethanol, pH 7.4) reaction. Assays were initiated by the addition of enzyme, incubated for 45 min at 30°C, and stopped by removing the $^{14}CO_2$-trapping glass fiber filters (Whatman grade GF/D, pretreated with NCS-II tissue solubilizer, Amersham Biosciences) from the reaction vial. For enzyme kinetic analyses, 10 μM of a 1% (n/n) solution of [1-^{14}C]2-oxoglutarate (specific activity 55 mCi/mmol) diluted with 99% (n/n) unlabeled 2-oxoglutarate was used. Results from assays lacking an alkaloid substrate were subtracted from corresponding assays containing alkaloid substrates to account for the uncoupled consumption of 2-oxoglutarate. Kinetic data for T6ODM were obtained by varying the thebaine or oripavine concentrations in the reaction between 1 and 500 μM at a constant 2-oxoglutarate concentration of 500 μM. Conversely, the 2-oxoglutarate concentration was varied between 1 and 500 μM at a constant thebaine concentration of 30 μM, which produced the maximum reaction velocity (Supplementary Fig. 10). Kinetic data for CODM were obtained by varying the of 500 μM, and varying the 2-oxoglutarate concentration between 1 and 500 μM at a constant codeine concentration of 50 μM. Saturation curves and kinetic constants were calculated based on Michaelis-Menten kinetics using FigP v. 2.98 (BioSoft, Cambridge, UK; http://www.biosoft.com). The release of formaldehyde upon alkaloid *O*-demethylation was monitored using using a fluorescence-based modification of the Nash assay[21]. Nash reagent was prepared by adding 0.3 ml of glacial acetic acid and 0.2 ml acetyl acetone to 100 ml of 2 M ammonium acetate. Enzyme assays were performed as described above, except that unlabelled 2-oxoglutarate was used and the reactions were quenched by the addition of 2 volumes of Nash reagent, followed by a 10 min incubation period at 60°C to convert formaldehyde to diacetyldihydrolutidine (DDL). The fluorescence of DDL was recorded using a Cary Eclipse fluorescence spectrophotometer (Varian; www.varianinc.com) at λ_{ex} = 412 nm and λ_{em} = 505 nm. The acylcyclohexanediones, prohexadione calcium or trinexapac-ethyl, were tested as possible enzyme inhibitors at concentrations up to 500 μM using 100 μM 2-oxoglutarate in a standard assay.

Liquid chromatography-tandem mass spectrometry

Enzyme assays were diluted 1:10 with 0.1% (v/v) formic acid (morphinan alkaloids) or 0.2% acetic acid (protoberberine alkaloids) and analyzed using a 6410 Triple Quadrupole LC-MS/MS system (Agilent Technologies; Santa Clara, CA; http://www.agilent.com). Liquid chromatography was performed using a Zorbax Eclipse Plus C_{18} column (2.1 x 50 mm, 1.8 μm particle size; Agilent Technologies) at a flow rate of 0.4 ml/min. For morphinan alkaloids, the column was equilibrated in solvent A and alkaloids were eluted using a linear gradient to 100% solvent B over 4 min [solvent A: 20% (v/v) acetonitrile/0.1% (v/v) formic acid/79.9% (v/v) water; solvent B: 99.9% (v/v) acetonitrile/0.1% (v/v) formic acid]. For protoberberine alkaloids, the column was equilibrated in solvent C and alkaloids were eluted under the following conditions: 0-1 min 0% solvent D, 1-10 min to

Nature Chemical Biology: doi:10.1038/nchembio.317

35% solvent D, 10-11 min to 100% solvent D [solvent C: 2% (v/v) acetonitrile/0.2% (v/v) acetic acid/97.8% (v/v) water; solvent D: 98% (v/v) acetonitrile/0.2% (v/v) acetic acid/2% (v/v) water]. Injection into the mass analyzer was performed using an electrospray ionization (ESI) probe inlet. Ions were generated and focused using an ESI voltage of 4000 kV, 9 l/min gas flow, 40 psi nebulizing pressure, and gas temperature of 330°C. MS data acquisition was carried out in positive ion mode over 50-400 *m/z*. The collision-induced dissociation (CID) mass spectra were recorded using collision energies of -15.0 eV (thebaine and oripavine), -25.0 (codeinone, morphinone, scoulerine and 3-*O*-demethylscoulerine), -32.0 eV (morphine) and -30.0 eV (codeine). Argon collision gas was set at a pressure of 1.8 x 10^{-3} torr. Alkaloids were identified based on either previously published ESI mass spectra[36,37] or by comparison with ESI mass spectra of authentic standards.

Virus-induced gene silencing

Construction of vectors – Unique 3'-UTR sequences were used for the construction of virus-induced gene silencing (VIGS) vectors to specifically silence genes encoding DIOX1, DIOX2 and DIOX3, and to avoid the suppression of highly homologous genes: DIOX-b (*DIOX1*-specific; 221 bp), DIOX-c (*DIOX2*-specific; 152 bp) and DIOX-d (*DIOX3*-specific; 292 bp)[38]. In addition, a construct was assembled for the purpose of simultaneously silencing *DIOX1*, *DIOX2* and *DIOX3*. The design of this non-specific construct (DIOX-a; 342 bp) was based on a highly conserved ORF sequence in all three genes. Full-length cDNA clones were used as templates to amplify fragments shown in Supplementary Fig. 12. Forward and reverse PCR primers were designed with flanking *Bam*HI and *Xho*I restriction endonuclease sites, respectively (Supplementary Table 2). Amplicons were generated using *Taq* DNA polymerase (Invitrogen), ligated to pGEM-T (Promega), digested with *Bam*HI and *Xho*I, and isolated inserts ligated to pTRV2[39] pre-cut with *Bam*HI and *Xho*I. DNA sequencing of cloned products was performed after ligation into both vectors. Plasmid propagation was achieved in the *E. coli* strain XL1-Blue, and the pTRV2 constructs (i.e. DIOX-a, DIOX-b, DIOX-c, DIOX-d and the empty vector) and pTRV1[39] were independently mobilized in *Agrobacterium tumefaciens* strain GV3101.
Infiltration of plants – Bacteria were prepared for infiltration using a previously reported protocol[40]. Independent overnight liquid cultures of *A. tumefaciens* containing each construct were used to inoculate 500 ml of Luria-Bertani (LB) medium containing 10 mM MES, 20 μM acetosyringone and 50 μg/ml kanamycin. Cultures were maintained at 28°C for 24 hours, harvested by centrifugation at 3000*g* for 20 min, and resuspended in infiltration solution (10 mM MES, 200 μM acetosyringone, 10 mM $MgCl_2$) to an OD_{600} of 2.5. *Agrobacterium tumefaciens* harboring DIOX-a, DIOX-b, DIOX-c, DIOX-d and the pTRV2 empty vector were each mixed 1:1 (v/v) with *A. tumefaciens* containing pTRV1, and incubated for one hour at 20°C prior to infiltration. Opium poppy plants used for VIGS analysis were 2-3 weeks old with emerging first leaves. Infiltration of the *A. tumefaciens* inoculum to the emerging leaves was performed using a 1-cc syringe. Plants inoculated with pTRV1 and pTRV2, the latter

Nature Chemical Biology: doi:10.1038/nchembio.317

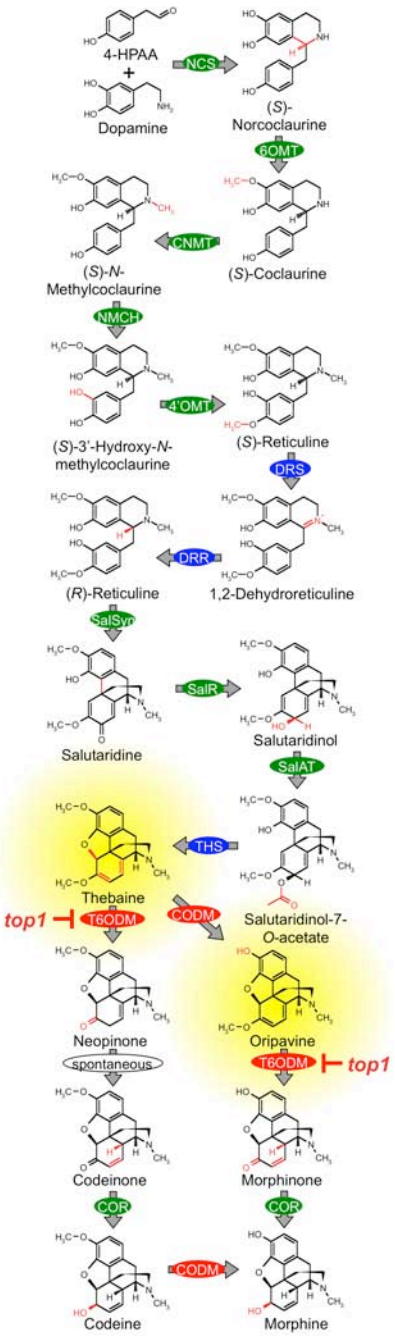
containing a fragment of the opium poppy gene encoding phytoene desaturase (PapsPDS)[40], displayed photobleaching and were used as a visual marker of VIGS efficiency, which was typically in the range of 20-25%.

Metabolite profiling – Infiltrated opium poppy plants were analyzed at maturity (i.e. the emergence of flower buds). Stems were cut immediately below the flower bud and 10 μl of exuding latex was collected. At the same time, three 1-cm segments of stem tissue directly below the flower bud were excised and flash frozen in liquid nitrogen for RT-qPCR analysis. Initial phenotypic screening was performed by thin layer chromatography (TLC). Latex samples were suspended in 30 μl methanol and 10 μl was spotted on TLC Silica gel 60 F_{254} plates (Merck). Separation was achieved using a previously described solvent system[6] and alkaloids were visualized by shadowing under 254 nm UV illumination. Major alkaloids were identified based on the comparison of R_f values with those of authentic standards. TLC results were confirmed using high performance liquid chromatography (HPLC) as described previously[8]. Fifteen microliters of methanol-latex suspension was diluted with 235 μl of methanol, vortexed and centrifuged for 10 min at 10,000*g* to remove insoluble debris and 100 μl of the supernatant was analyzed by HPLC.

Gene expression analysis

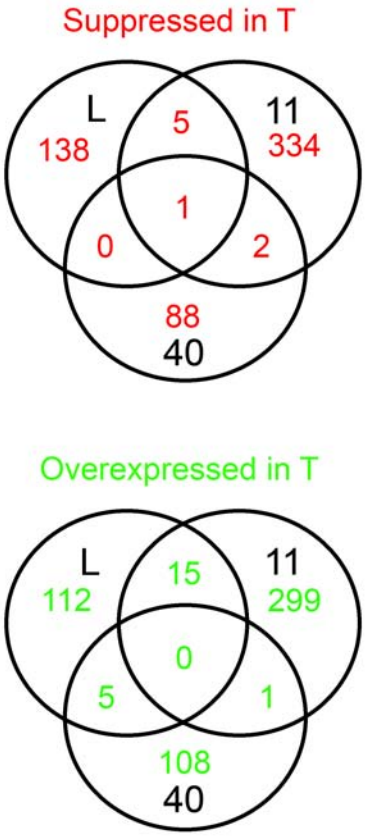
Total RNA was isolated with TRIzol (Invitrogen) according to the manufacturer's instructions. Reverse transcription was performed at 42°C for 60 min using 2.5 mM anchored oligo(dT) primer (dT20VN), 0.5 mM dNTP, 10 to 40 ng/μl RNA, and 5 microunits/ul reverse transcriptase (Fermentas, Burlington, Canada; http://www.fermentas.com) following denaturing of the RNA-primer mix at 70°C for 5 min. Real-time quantitative PCR using SYBR Green detection was performed using a 7300 Real-Time PCR system (Applied Biosystems). Each 10-μL PCR included 1 μL of cDNA (taken directly from the RT reaction in the case of stem, or diluted 50% [v/v] with water for bud, leaf and root), 300 nM forward and reverse primers, and 1x Power SYBR Green PCR Master Mix (Applied Biosystems). Primer sequences are listed in Supplementary Table 2. Reactions were subjected to 40 cycles of template denaturation, primer annealing and primer extension. To evaluate qPCR specificity, the amplicons of all primer pairs were subjected to melt-curve analysis using the dissociation method suggested by the instrument manufacturer (Applied Biosystems). Gene expression data for VIGS analysis were determined based on 27 independent values per plant line (i.e. 3 technical replicates performed on each of 3 stem segments taken from each of 3 individual plants). Organ-specific gene expression data (Supplementary Fig. 13) were based on nine independent values per plant line (i.e. 3 technical replicates on each of 3 individual plants). The $2^{-\Delta\Delta Ct}$ method was used for the analysis of relative gene expression[41] as described previously[8]. The gene encoding elongation factor 1a (elf1a) was used as the internal control and the plant line showing the highest expression level served as the calibrator for each target gene.

Nature Chemical Biology: doi:10.1038/nchembio.317



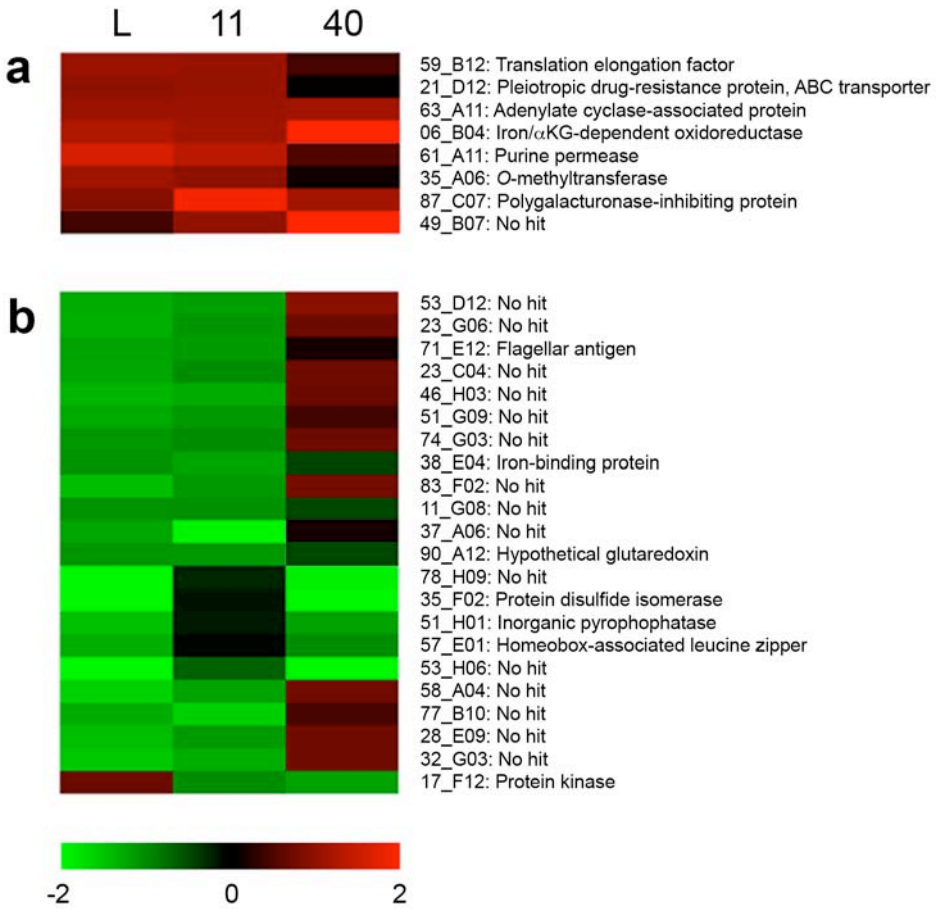
Supplementary Figure 1. Biosynthesis of morphine in opium poppy. The enzymes reported in this study, and the location of the metabolic block in the *top1* mutant[6], are shown in red. The *top1* variety of opium poppy accumulates thebaine and oripavine (highlighted in yellow) rather than codeine and morphine. The T variety displays a *top1* alkaloid profile[8] and exhibits a dramatic reduction in shoot *T6ODM* transcript levels (Supplementary Fig. 13). Enzymes for which corresponding cDNAs have previously been isolated are shown in green. Although activity has been detected in plant protein extracts for enzymes shown in blue, the corresponding genes have not yet been isolated. An enzyme capable of catalyzing thebaine biosynthesis has been suggested (THS)[42]. However, the formation of thebaine from salutaridinol 7-*O*-acetate occurs spontaneously at pH 8-9[43]. Enzyme-catalyzed and/or spontaneous structural rearrangements and functional group modulations are indicated on each molecule in red. Abbreviations: NCS, norcoclaurine synthase; 6OMT, norcoclaurine 6-*O*-methyltransferase; CNMT, coclaurine *N*-methyltransferase; NMCH, *N*-methylcoclaurine 3'-hydroxylase; 4'OMT, 3'-hydroxy-*N*-methylcoclaurine 4'-*O*-methyltransferase; DRS, 1,2-dehydroreticuline synthase; DRR, 1,2-dehydroreticuline reductase; SalSyn, salutaridine synthase; SalR, salutaridine reductase; SalAT, salutaridinol 7-*O*-acetyltransferase; THS, thebaine synthase; T6ODM, thebaine 6-*O*-demethylase; CODM, codeine *O*-demethylase; COR, codeinone reductase.

Nature Chemical Biology: doi:10.1038/nchembio.317



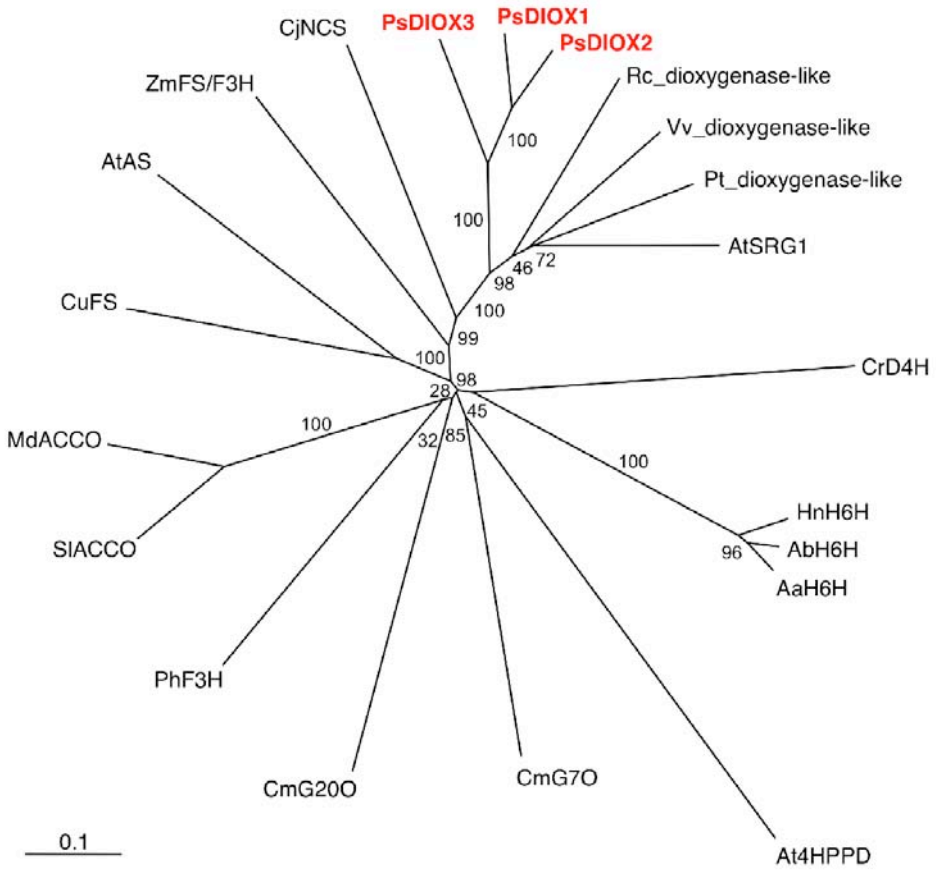
Supplementary Figure 2. Venn diagrams summarizing the results of microarray hybridization experiments that compared the abundance of transcripts in the stems of opium poppy variety T with varieties L, 11 and 40. The numbers of genetic elements on the microarray that showed decreased hybridization of RNA from variety T compared with varieties L, 11 and/or 40 are indicated in the upper panel (red). Conversely, the numbers of genetic elements that showed increased hybridization of RNA from variety T compared with varieties L, 11 and/or 40 are indicated in the lower panel (green). Decreased or increased hybridization was indicative of lower or higher transcript abundance, respectively, and thus revealed a relative suppression in the expression of specific genes. Genes were considered differentially expressed based on a signal intensity ratio cutoff of 1.8. Microarray hybridization experiments involving pair-wise comparisons of varieties T vs L, T vs 11, and T vs 40 revealed eight genes putatively suppressed in T compared with at least two other varieties, but only a single gene suppressed in T compared with all three varieties (upper panel).

Nature Chemical Biology: doi:10.1038/nchembio.317



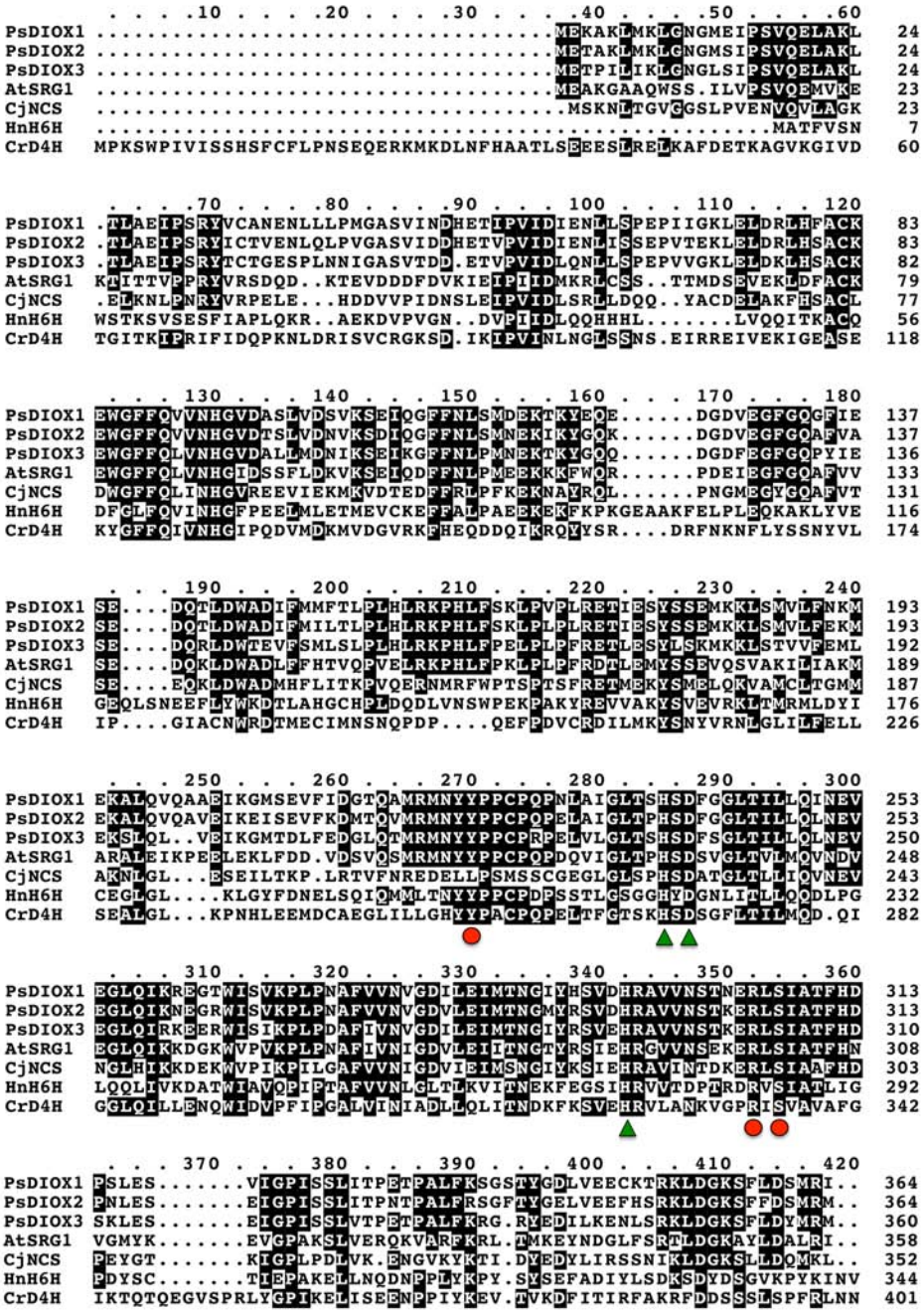
Supplementary Figure 3. Heat maps illustrating the relative abundance of transcripts in the stems of opium poppy variety T with varieties L, 11 and 40. Results are shown only for genes exhibiting low (**a**) or high (**b**) expression in T compared with at least two other varieties, based on a ratio cutoff of 1.8. Corresponding functional annotations and microarray coordinates are shown to the right of each diagram. Average signal intensity ratios from 6 independent microarray hybridization experiments were $\log_2$ normalized and plotted based on the indicated color scheme. Positive values (red color) indicate relatively lower transcript levels in variety T, whereas negative values (green color) indicate relatively higher transcript levels in variety T compared with varieties L, 11 and 40. Images were generated using MultiExperiment Viewer (TIGR TM4 Microarray Software Suite)[34]. Note that 06_B04 in (**a**) corresponds to DIOX1, which was renamed T6ODM after functional characterization.

Nature Chemical Biology: doi:10.1038/nchembio.317



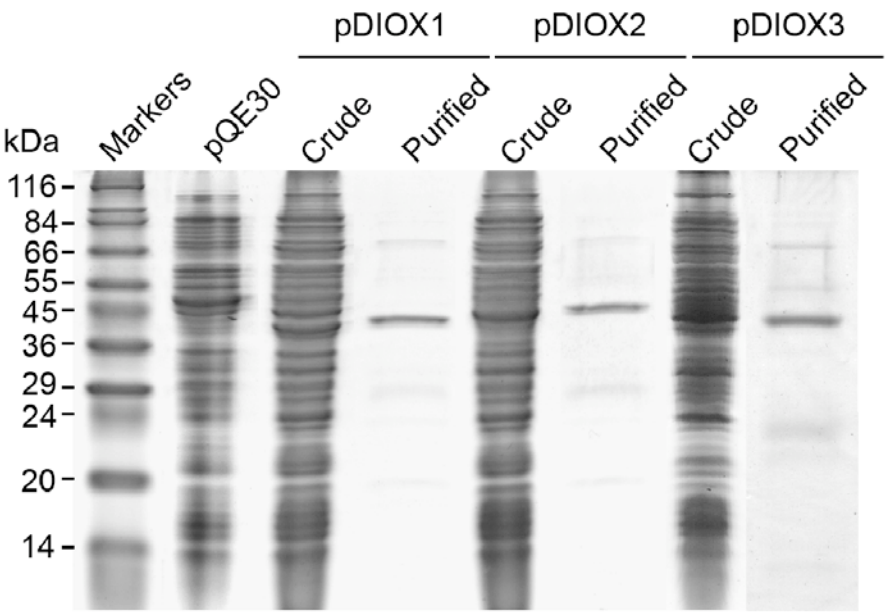
Supplementary Figure 4. Unrooted neighbor-joining phylogenetic tree for selected plant 2-oxoglutarate (2OG)/Fe(II)-dependent dioxygenases. Boostrap frequencies for each clade are percentages of 1,000 iterations. Bootstrapping provides several alternative versions of the branch points that separate each of the different clades. The bootstrap value indicates the frequency of a branch occurring at the indicated position in alternative phylogenetic analyses. Abbreviations and GenBank accession numbers are listed in the phylogenetic analysis section of the Methods. Note that DIOX1 and DIOX3 were renamed thebaine 6-*O*-demethylase (T6ODM) and codeine *O*-demethylase (CODM), respectively, after functional characterization.

Nature Chemical Biology: doi:10.1038/nchembio.317



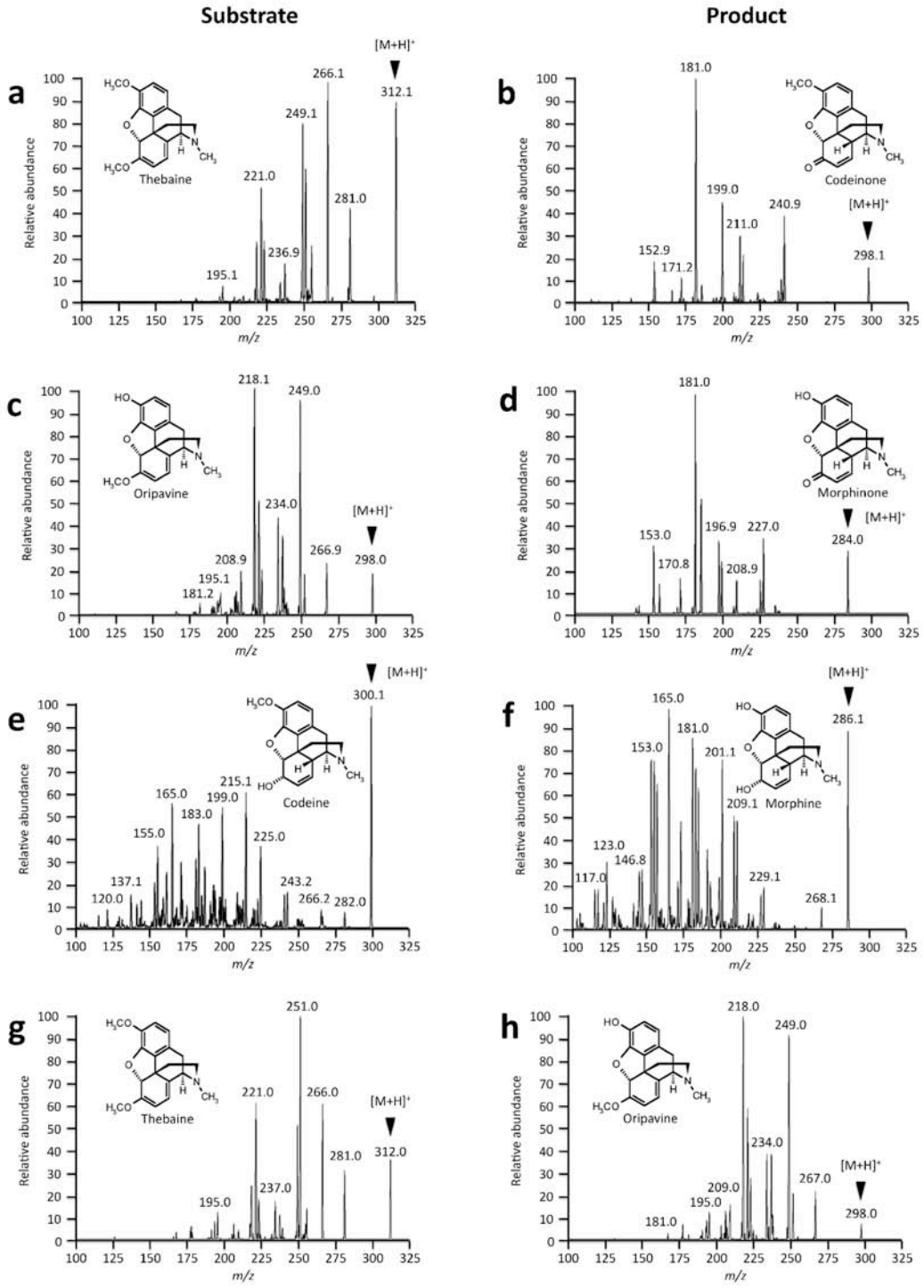
Supplementary Figure 5. Alignment of the deduced amino acid sequences of opium poppy DIOX1, DIOX2, and DIOX3 with other plant 2-oxoglutarate (2OG)/Fe(II)-dependent dioxygenases. Sequences were aligned using ClustalX[29]. Shaded boxes indicate residues that are identical in at least 40% of the aligned proteins. Dots represent introduced gaps into sequences to maximize the alignment. Abbreviations: AtSRG1, *Arabidopsis thaliana* senescence-related gene 1; CjNCS, *Coptis japonica* norcoclaurine synthase; HnH6H, *Hyoscyamus niger* hyoscyamine 6β-hydroxylase, CrD4H, *Catharanthus roseus* desacetoxyvindoline 4-hydroxylase. Residues corresponding to a canonical $HXDX_nH$ catalytic triad (green triangles) required for coordinating Fe(II)[19] and a YX_nRXS motif (red circles) implicated in 2OG binding[10] are indicated. Note that DIOX1 and DIOX3 were renamed thebaine 6-*O*-demethylase (T6ODM) and codeine *O*-demethylase (CODM), respectively, after functional characterization.

Nature Chemical Biology: doi:10.1038/nchembio.317



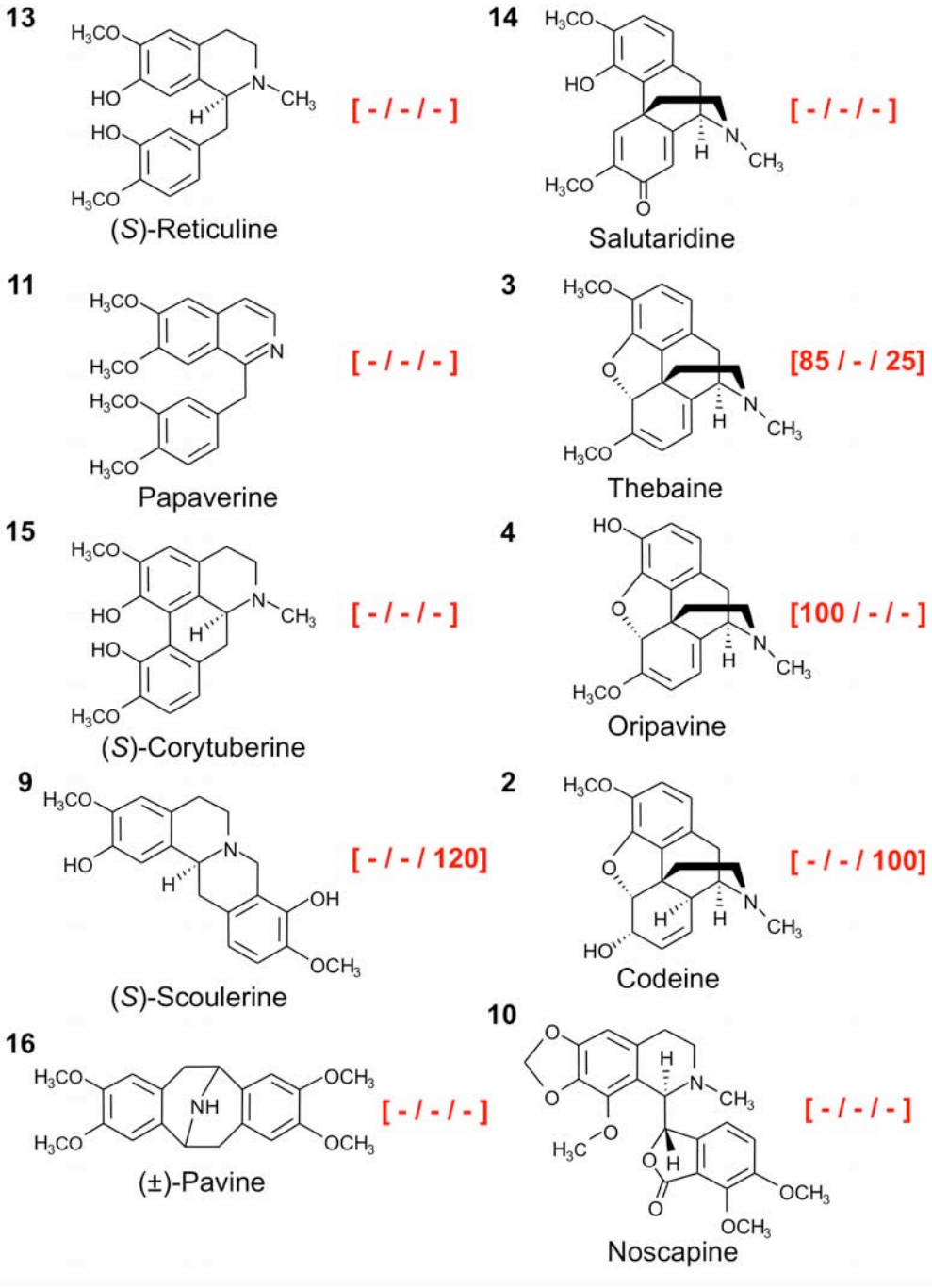
Supplementary Figure 6. SDS-PAGE of recombinant proteins produced by pDIOX1 (thebaine 6-*O*-demethylase, T6ODM), pDIOX2, and pDIOX3 (codeine *O*-demethylase, CODM) in *Escherichia coli*. The left lane contains molecular weight protein markers and corresponding sizes are indicated to the left of the panel. All other lanes feature total (crude) or purified protein from *E. coli* strain SG13009 cells induced with IPTG. Purification of polyhistadine-tagged recombinant proteins was achieved using a cobalt-affinity column. Bacteria harboring the empty pQE30 vector were included as a negative control. Visualization was achieved using Coomassie blue staining.

Nature Chemical Biology: doi:10.1038/nchembio.317



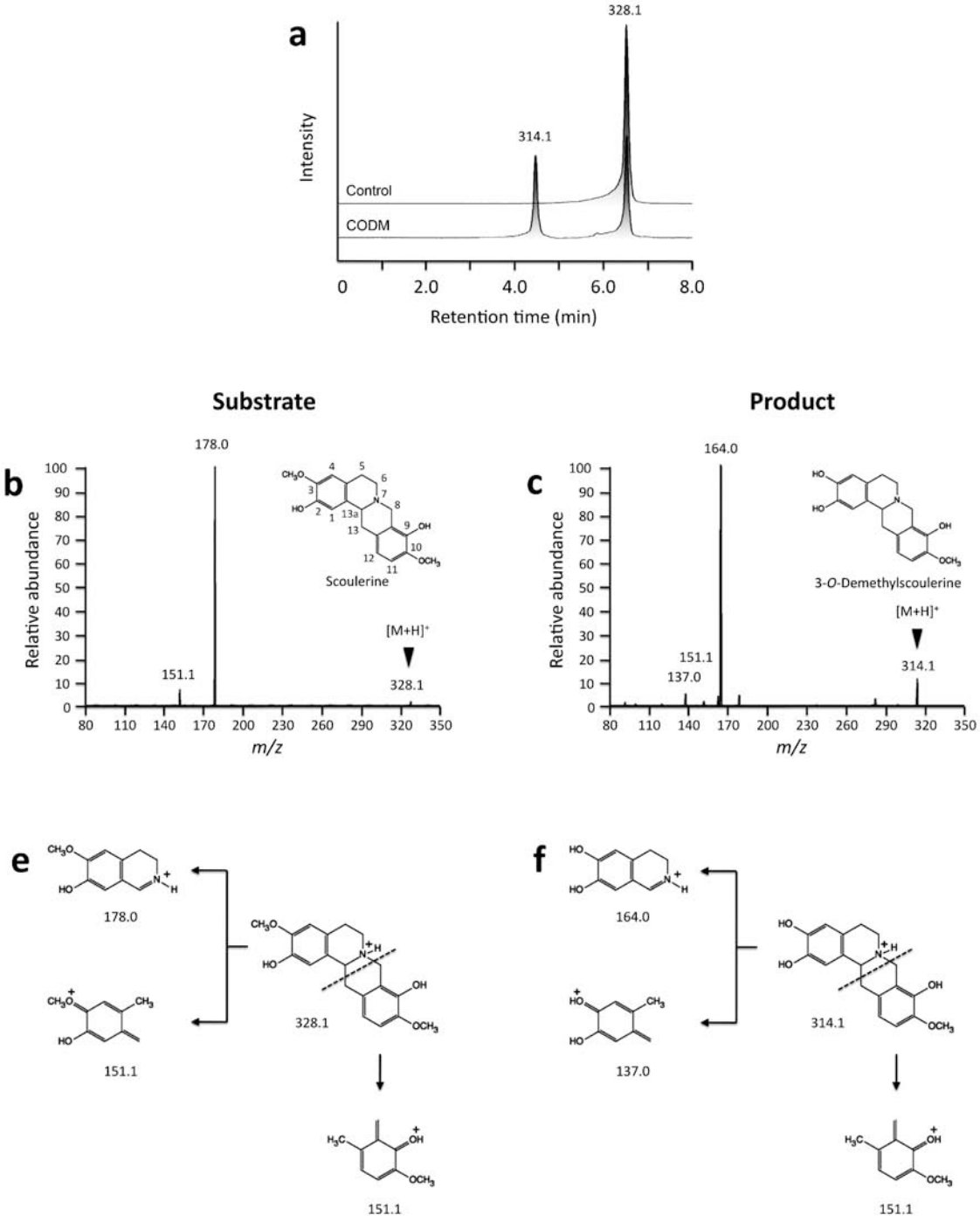
Supplementary Figure 7. Collision-induced dissociation (CID) mass spectra for substrates (left panels) and products (right panels) of thebaine 6-*O*-demethylase (T6ODM, a-d) and codeine *O*-demethylase (CODM, e-h) enzyme assays. Following liquid chromatography (LC), molecular parent ions (arrowheads) were generated and focused using electrospray ionization (ESI), and subjected to mass spectrometry. To unambiguously identify reaction components, daughter ions were generated using argon gas collision at the following energies: -15.0 eV (thebaine and oripavine), -25.0 (codeinone and morphinone), -32.0 eV (morphine) and -30.0 eV (codeine). The observed ESI mass spectra were in agreement with previously published ESI spectra[36] and with those acquired for authentic standards. Structures corresponding to the parent molecules are shown.

Nature Chemical Biology: doi:10.1038/nchembio.317



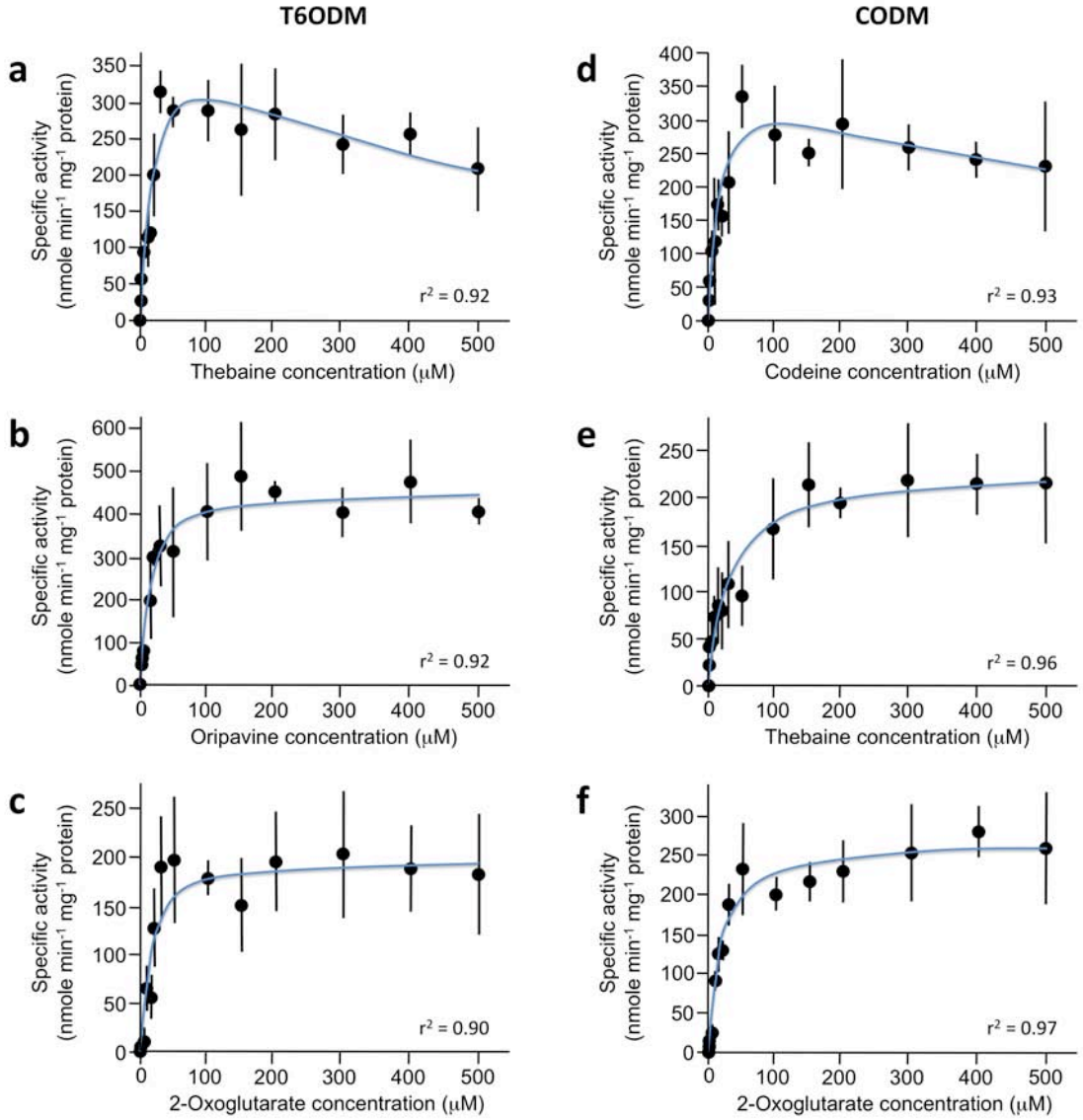
Supplementary Figure 8. Substrate specificities of recombinant thebaine 6-*O*-demethylase (T6ODM), DIOX2 and codeine *O*-demethylase (CODM). Enzyme assays were based on the decarboxylation of [1-^{14}C]2-oxoglutarate coupled with the *O*-demethylation of a benzylisoquinoline alkaloid co-substrate. The incubation time (45 min), protein concentration (10 ng/μl) and other assay parameters were optimized prior to enzyme kinetic analyses. The structures of compounds tested as potential enzymatic substrates are shown adjacent to values indicating percent relative activities for T6ODM, DIOX2, and CODM, respectively. Hyphens indicate that enzyme activity was not detected.

Nature Chemical Biology: doi:10.1038/nchembio.317



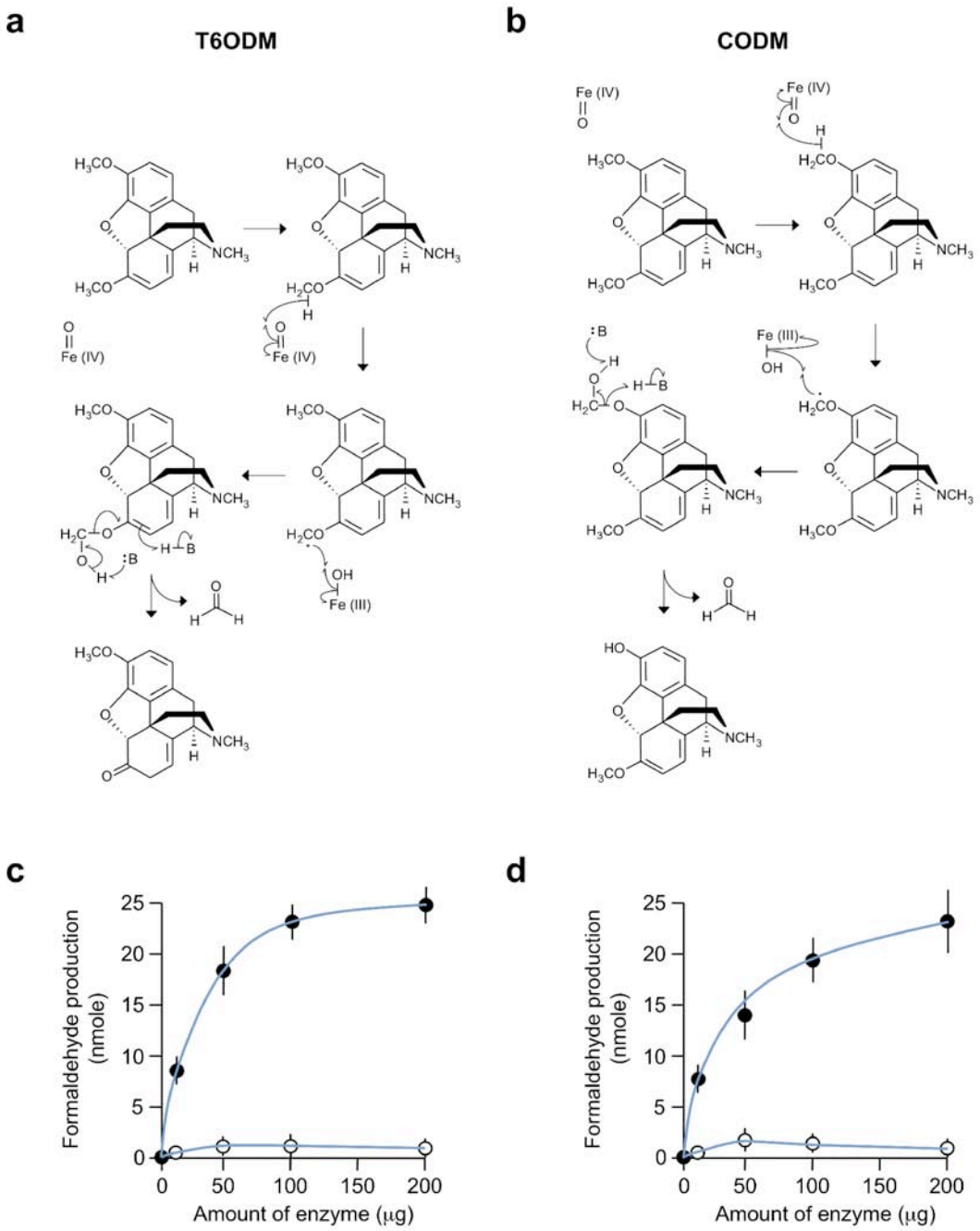
Supplementary Figure 9. Extracted ion chromatograms (EICs) and collision-induced dissociation (CID) mass spectra showing the 3-*O*-demethylation of (*S*)-scoulerine by codiene *O*-demethylase (CODM). a, The upper (control) EIC corresponds to an assay performed with boiled enzyme, whereas the lower (CODM) EIC shows an assay performed with native enzyme. Incubation of native CODM with (*S*)-scoulerine (*m/z* 328.1) as the substrate yielded a reaction product of *m/z* 314.1, consistent with the loss of a methyl group. CID analysis at a collision energy of -25.0 eV supported *O*-demethylation of (*S*)-scoulerine at position 3 (**b**) yielding 3-*O*-demethylscoulerine (**c**). Proposed mass spectral fragmentation of (*S*)-scoulerine (**e**) and 3-*O*-demethylscoulerine (**f**) under electrospray CID conditions[37].

Nature Chemical Biology: doi:10.1038/nchembio.317



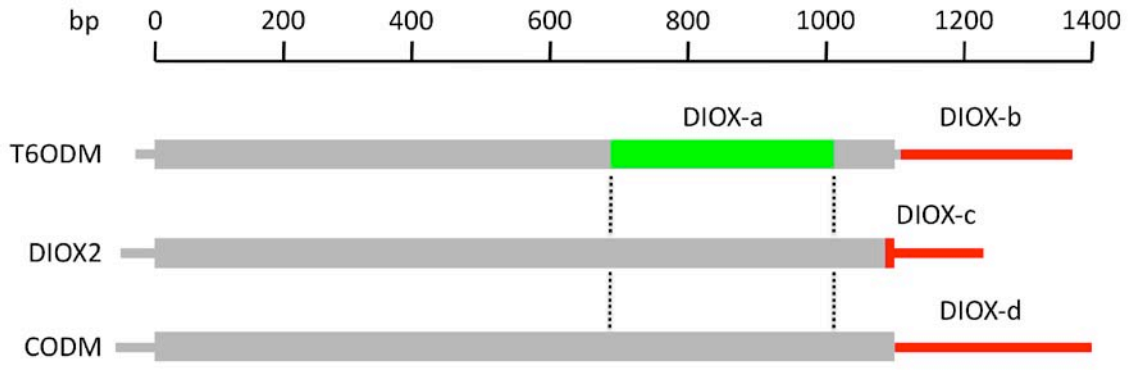
Supplementary Figure 10. Steady-state enzyme kinetics of purified recombinant thebaine 6-*O*-demethylase (T6ODM, left panels) and codeine *O*-demethylase (CODM, right panels) with varying different substrate concentrations. Enzyme assays were based on the decarboxylation of [1-^{14}C]2-oxoglutarate coupled with the *O*-demethylation of a benzylisoquinoline alkaloid co-substrate. The incubation time (45 min), protein concentration (10.0 ng/μl) and other assay parameters were optimized prior to enzyme kinetic analyses. Values represent the mean specific activity ± standard deviation monitored as a function of substrate concentration for three independent replicates. Data was subjected to further analysis using FigP v. 2.98 (BioSoft, Cambridge, UK), generating maximum velocity (V_m) and substrate affinity (K_m) constants based on Michaelis-Menten kinetics. Curve-fitting for data shown in (**a**) and (**b**) revealed moderate substrate inhibition; thus, optimal velocity (V_{opt}) and inhibition (K_i) constants were also calculated (Supplementary Table 3). Corresponding r^2 values are displayed in the right-hand corners of each panel.

Nature Chemical Biology: doi:10.1038/nchembio.317



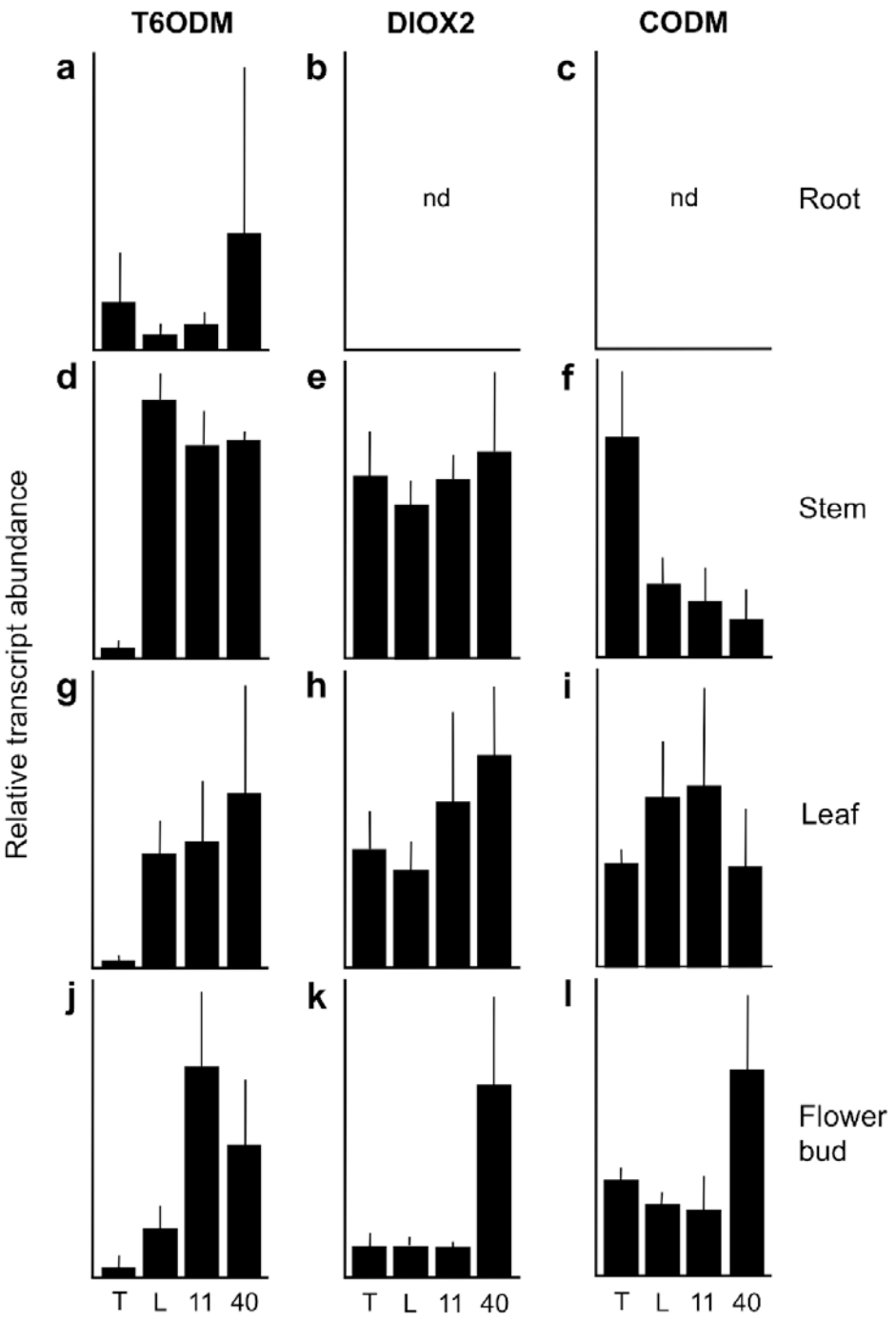
Supplementary Figure 11. Proposed reaction mechanisms for the 6-*O*-demethylation of thebaine by T6ODM (a) and the 3-*O*-demethylation of codeine by CODM (b). Identical mechanisms are proposed for the 6-*O*-demethylation of oripavine and the 3-*O*-demethylation of thebaine by T6ODM and CODM, respectively. In all cases, alkyl hydroxylation proceeds through a radical mechanism involving an iron-oxo intermediate, followed by the elimination of formaldehyde. The production of formaldehyde was demonstrated in all cases using the Nash assay[21]. The quantification of formaldehyde release after the incubation of native (closed circles) or denatured (open circles) T6ODM with thebaine (**c**) or CODM with codeine (**d**) is shown.

Nature Chemical Biology: doi:10.1038/nchembio.317



Supplementary Figure 12. Regions of T6ODM, DIOX2 and CODM cDNAs used to construct virus-induced gene silencing (VIGS) vectors in pTRV2. Coding regions of each cDNA are shown as thicker line elements, whereas non-coding 5' and 3'-untranslated regions (UTRs) are shown as thinner line elements. The green segment and dotted lines show a region of the T6ODM cDNA that is highly conserved in DIOX2 and CODM, and was used to construct the DIOX-a VIGS vector. The red segments represent unique regions in each cDNA that were used to construct gene-specific VIGS vectors, DIOX-b, DIOX-c and DIOX-d. DNA sequence length is shown in base pairs (bp) with respect to the start codon in each cDNA.

Nature Chemical Biology: doi:10.1038/nchembio.317



Supplementary Figure 13. Relative abundance of transcripts encoding thebaine 6-*O*-demethylase (T6ODM), DIOX2 and codeine *O*-demethylase (CODM) in opium poppy plant organs. Real-time quantitative PCR was used to quantify the relative transcript abundance in roots, stems, leaves and flower buds of opium poppy varieties T, L, 11 and 40. Data were calculated using nine independent trials per plant line (i.e. 3 technical replicates on each of 3 individual plants). Normalization was performed using elongation factor 1a (*elf1a*) as the internal control, and the plant line exhibiting the highest expression level served as the calibrator for each target gene. *DIOX2* and *CODM* transcripts were below detection limits in root (**a** and **c**, respectively). Abbreviation: nd, not detected.

Nature Chemical Biology: doi:10.1038/nchembio.317

Supplementary Table 1: Genetic inheritance of the alkaloid phenotype in opium poppy variety T.

Generation	Cross	Phenotype (%T, n=48)
F_1	♂40 x ♀T	0
	♀40 x ♂T	0
F_2	♂40 x ♀T	25
	♀40 x ♂T	23
	♂11 x ♀T	29
	♀11 x ♂T	23
Backcross	♂40 x [♂40x ♀T]	0
	♀40 x [♀Tx ♂40]	0
	♀T x [♀40x ♂T]	48
	♂T x [♀Tx ♂40]	56

Nature Chemical Biology: doi:10.1038/nchembio.317

Supplementary Table 2: PCR primers used for the assembly of expression vectors and virus-induced gene-silencing constructs, and for real-time quantitative PCR (RT-qPCR) analysis.

Protein Expression
Primers used to amplify ORFs for ligation to pQE30

ORF	Forward primer	Reverse primer
DIOX1 (*T6ODM*)	GCGCGGATCCATGGAGAAAGCAAAACTT	GCGCCTGCAGCACAACGCACTTTCGAGA
DIOX2	GCGCGGATCCATGGAGACAGCAAAACTT	GCGCCTGCAGAGAGTCAAAAAGCAATGA
DIOX3 (*CODM*)	GCGCGGATCCATGGAGACACCAATACTT	GCGCCTGCAGGCACCATATGAATTCTTC

Virus-induced gene silencing
Primers used in sequence amplification for ligation to TRV2

Sequence	Forward primer	Reverse primer
DIOX-a	GCGCGGATCCCCTTGTCCTCAACCAAAT	GCGCCTCGAGTCCACTTTTAAACAAAGC
DIOX-b	GCGCGGATCCCGACGTGATTGCATGTCA	GCGCCTCGAGCACAACGCACTTTCGAGA
DIOX-c	GCGCGGATCCTTGATTCGATGAGGATG	GCGCCTCGAGCTTGAGAAAAGTTTTATT
DIOX-d	GCGCGGATCCGTGAGAAAGTGTGAACAT	GCGCCTCGAGCAATCCAATACATTATTT

RT-qPCR
Primers used to examine gene-specific transcript abundance

Gene	Forward primer	Reverse primer
DIOX1 (*T6ODM*)	TTGAGGCACAAATGAGAAAATTGA	CACAACGCACTTTCGAGAAATTAC
DIOX2	TGTGAGAAACTGAAGAACACCCAAT	AAGGACTCAAACCACTGAAAGACG
DIOX3 (*CODM*)	TTGTGCTTAAATTTCGTGGATGAC	TGATTACATCACTTGACCCAAACAG

Nature Chemical Biology: doi:10.1038/nchembio.317

Supplementary Table 3: Kinetic data for T6ODM and CODM.

Enzyme	Substrate	K_m (μM)	K_i (μM)	V_{max} (pkat)	k_{cat} (s^{-1})	k_{cat}/K_m (s^{-1} M^{-1})
T6ODM	Thebaine	20.3±7.1	518±237	2.09±0.33	0.0170±0.0027	837.1
T6ODM	Oripavine	15.4±2.7	-	2.35±0.09	0.0191±0.0007	1242.6
T6ODM	2-Oxoglutarate	16.4±5.3	-	1.00±0.07	0.0081±0.0006	492.9
CODM	Codeine	20.5±8.0	642±409	1.97±0.35	0.0161±0.0029	785.4
CODM	Thebaine	41.9±8.0	-	1.21±0.06	0.0099±0.0005	235.2
CODM	2-Oxoglutarate	19.0±3.3	-	1.35±0.05	0.0110±0.0004	578.9

Nature Chemical Biology: doi:10.1038/nchembio.317

References

26. Wheeler, D. M. S., Kinstle, T. H. & Rinehart, K. L. J. Mass spectral studies of alkaloids related to morphine. *J. Am. Chem. Soc.* **89**, 4494-4501 (1967).
27. Liscombe, D. K. *et al.* Targeted metabolite and transcript profiling for elucidating enzyme function: Isolation of novel *N*-methyltransferases from three benzylisoquinoline alkaloid-producing species. *Plant J.* doi: 10.1111/j.1365-313X.2009.03980.x (2009).
28. Ziegler, J., Brandt, W., Geissler, R. & Facchini, P. J. Removal of substrate inhibition and increase in maximal velocity in the short chain dehydrogenase/reductase salutaridine reductase involved in morphine biosynthesis. *J. Biol. Chem.* **284**, 26758-26767 (2009).
29. Chenna, R. *et al.* Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res.* **31**, 3497–3500 (2003).
30. Page, R.D. TREEVIEW: an application to display phylogenetic trees on personal computers. *Comput. Appl. Biosci.* **12**, 357–358 (1996).
31. Zulak, K. G. *et al*. Gene transcript and metabolite profiling of elicitor-induced opium poppy cell cultures reveals the coordinate regulation of primary and secondary metabolism. *Planta* **225**, 1085–1106 (2007).
32. Logemann, J., Schell, J. & Willmitzer, L. Improved method for the isolation of RNA from plant tissues. *Anal. Biochem.* **163**, 16-20 (1987).
33. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry* **18**, 5294-5299 (1979).
34. Saeed, A. I. *et al.* TM4: a free, open-source system for microarray data management and analysis. *Biotechniques* **34**, 374–378 (2003).
35. Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685 (1970).
36. Raith K. *et al.* Electrospray tandem mass spectrometric investigations of morphinans. *J. Am. Soc. Mass Spectrom.* **14**, 1262-1269 (2003).
37. Schmidt, J., Boettcher, C., Kuhnt, C., Kutchan, T.M., Zenk, M.H. Poppy alkaloid profiling by electrospray tandem mass spectrometry and electrospray FT-ICR mass spectrometry after [ring-$^{13}C_6$]-tyramine feeding. *Phytochemistry* **68**, 189-202 (2007).
38. Dinesh-Kumar, S. P. *et al.* Virus-induced gene silencing. *Methods Mol. Biol*. **236**, 287-294 (2003).
39. Liu, Y. L., Schiff, M., Marathe, R. & Dinesh-Kumar, S. P. Tobacco Rar1, *EDS1* and *NPR1/NIM1* like genes are required for *N*-mediated resistance to tobacco mosaic virus. *Plant J.* **30**, 415–429 (2002).
40. Hileman, L.C., Drea, S., de Martino, G., Litt, A. & Irish, V. Virus-induced gene silencing is an effective tool for assaying gene function in the basal eudicot species *Papaver somniferum* (opium poppy). *Plant J*. **44**, 334–341 (2005).

Nature Chemical Biology: doi:10.1038/nchembio.317

41. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta Ct}$ method. *Methods* **25**, 402–408 (2001).

42. Fisinger, U., Grobe, N. & Zenk, M.H. Thebaine synthase: a new enzyme in the morphine pathway in *Papaver somniferum*. *Nat. Prod. Commun.* **2**, 249-253 (2007).

43. Lenz, R. & Zenk, M.H. Purification and properties of codeinone reductase (NADPH) from *Papaver somniferum* cell cultures and differentiated plants. *Eur. J. Biochem*. **233**, 132-139 (1995).

Nature Chemical Biology: doi:10.1038/nchembio.317