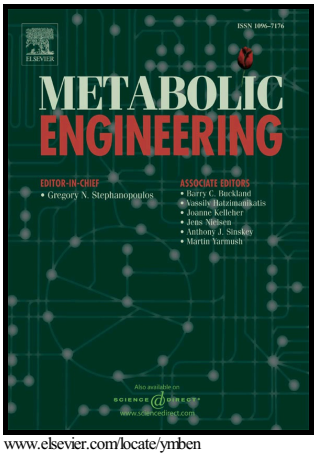

www.elsevier.com/locate/ymben

De novo synthesis of the sedative valerenic acid in *Saccharomyces cerevisiae*

Jeff Wong, Leo d'Espaux, Ishaan Dev, Cas van der Horst, Jay Keasling

PII: S1096-7176(17)30422-6
DOI: https://doi.org/10.1016/j.ymben.2018.03.005
Reference: YMBEN1359

To appear in: *Metabolic Engineering*

Received date: 8 November 2017
Revised date: 3 March 2018
Accepted date: 5 March 2018

Cite this article as: Jeff Wong, Leo d'Espaux, Ishaan Dev, Cas van der Horst and Jay Keasling, *De novo* synthesis of the sedative valerenic acid in *Saccharomyces cerevisiae*, *Metabolic Engineering,* https://doi.org/10.1016/j.ymben.2018.03.005

De novo synthesis of the sedative valerenic acid in *Saccharomyces cerevisiae*

Jeff Wong[a,b,c1], Leo d'Espaux[b,c1], Ishaan Dev[b,c,d], Cas van der Horst[b,c], Jay Keasling[b,c,d*]
[a]Department of Plant and Microbial Biology, University of California, Berkeley, CA 94720
[b]DOE Joint BioEnergy Institute, Emeryville, CA 94608, United States
[c]Biological Systems and Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, United States
[d]Department of Chemical Engineering and Bioengineering, University of California at Berkeley, Berkeley, CA, USA Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA Joint BioEnergy Institute, Emeryville, CA, USA keasling@berkeley.edu
*Corresponding author
^

Valeriana officinalis **(Valerian) root extracts have been used by European and Asian cultures for millennia for their anxiolytic and sedative properties. However, the efficacy of these extracts suffers from variable yields and composition, making these extracts a prime candidate for microbial production. Recently, valerenic acid, a C15 sesquiterpenoid, was identified as the active compound that modulates the $GABA_A$ channel. Although the first committed step, valerena-4,7(11)-diene synthase, has been identified and described, the complete valerenic acid biosynthetic pathway remains to be elucidated. Sequence homology and tissue-specific expression profiles of *V. officinalis* putative P450s led to the discovery of a *V. officinalis* valerena-4,7(11)-diene oxidase, VoCYP71DJ1, which required coexpression with a *V. officinalis* alcohol dehydrogenase and aldehyde dehydrogenase to complete valerenic acid biosynthesis in yeast. Further, we demonstrated the stable integration of all pathway enzymes in yeast, resulting in the production of 140mg/L of valerena-4,7(11)-diene and 4mg/L of valerenic acid in millititer plates. These findings showcase *Saccharomyces cerevisiae*'s potential as an expression platform for facilitating multiply-oxidized medicinal terpenoid pathway discovery, possibly paving the way for scale up and FDA approval of valerenic acid and other active compounds from plant-derived herbal medicines.**

1. Introduction

The popularity of herbal medicines is exemplified by a $60 billion industry worldwide, and 20% of the total drug market (Kirk and Dunker 2014). Valerian — a medicinal preparation of *Valeriana officinalis* rhizome, root, and stolon extracts — is one such herbal medicine, and has been valued for its anxiolytic and sedative properties for millenia (Houghton 1988; Eadie 2004). Valerian is currently 'generally recognized as safe' by the US Food and Drug Administration, and has been approved as a natural sleep aid in several countries (Kumar 2006). While *V. officinalis* extracts contain many bioactive compounds such as phenolic acids, valepotriates, lignans, flavonoids, amino acids, alkaloids, and tannins, the compound responsible for valerian's well-known activities is the sesquiterpenoid valerenic acid (Hritcu and Cioanca 2016; Trauner et al. 2008; Takemoto et al. 2009) (Fig. 1a). Interestingly, valerenic acid has been shown to be produced in only two related genera in the *Caprifoliaceae* family, *Valeriana* and *Centranthus*, primarily in the root and inflorescence tissues (Hassan et al. 2008). Until recently, the mechanism for valerenic acid activity was unknown. *In vivo* studies in mice showed that valerenic acid allosterically modulates $GABA_A$ receptor activity leading

[1] Present address: Demetrix, Inc., Emeryville, CA 94608, United States

to sedative or anxiolytic effects (Becker et al. 2014; Benke et al. 2009; Trauner et al. 2008). Interestingly, studies showed that *V. officinalis* root extracts contain another naturally occurring valerenic acid-derived sesquiterpenoid, acetoxyvalerenic acid, that diminishes the anxiolytic effects of valerenic acid by binding to identical sites (Felgentreff et al. 2012; Khom et al. 2007; Benke et al. 2009) (Fig. 1a). Indeed, studies have shown that extracts with high valerenic acid:acetoxyvalerenic acid ratios have more pronounced sedative effects (Felgentreff et al. 2012). Effective use of valerian is hampered by inaccurate dosage guidelines and highly variable acetoxyvalerenic acid content, which often makes up a significant amount of crude root extract (Becker et al. 2014; Felgentreff et al. 2012). Indeed, synthesis of valerenic acid is a source of isolated valerenic acid but suffers from low yields (ca. 6%) after many chemical conversion steps starting from (*R*)-pulegone (Kopp et al. 2009). Microbial production of valerenic acid promises isolation of valerian's active compound in the absence of antagonistic acetoxyvalerenic acid, lowers production costs, and does not suffer from inconsistent composition nor variable yields of plant derived valerenic acid — the current source of this herbal medicine.

The biosynthetic pathway for valerenic acid has not been fully identified, but is thought to proceed from central carbon metabolism through the mevalonate pathway and farnesyl pyrophosphate (FPP), then cyclized by a sesquiterpene synthase, valerena-4,7(11)-diene synthase (VDS), to form valerena-4,7(11)-diene, and likely decorated by one or more P450s, acyltransferases, and other modifying enzymes (Fig. S1). Previous studies have found that the highest concentration of valerenic acid is localized to *V. officinalis* root tissues, and an expression study showed *VDS* transcripts were almost exclusively expressed in the root relative to other tissues (Yeo et al. 2013).

A diverse set of plant P450s has been shown to oxidize a primary carbon on a sesquiterpene hydrocarbon substrate; however, several P450s in the CYP71D clade from a related plant family, *Asteraceae*, have been shown to oxidize substrates highly similar to valerenadiene. These include the *Artemisia annua* AaCYP71AV1 (which converts amorpha-4,11-diene to artemisinic acid, an important precursor to the anti-malarial drug artemisinin), various germacrene A oxidases (GAOs) (which convert germacrene A to germacra-1(10),4,11(13)-trien-12-oic acid, the precursor to the anti-cancer compound costunolide (Nguyen et al. 2010; Ro et al. 2006; Ikezawa et al. 2011), and *Thapsia garganica* TgCYP76AE2 (which catalyzes analogous oxidations on epikunzeaol to form epidihydrocostunolide (Andersen et al. 2017)). Due to the high similarity of *Asteraceae* sesquiterpenoid and the valerenic acid biosynthetic pathways, previous studies have postulated that a single P450 may catalyze conversion of valerenadiene into valerenic acid (Fig. 1a).

Thus far, attempts at identifying the full biosynthetic pathway for producing the important drug valerenic acid or reconstituting this activity in a heterologous host have failed (Ricigliano et al. 2016). We decided to use yeast as an expression platform, as our preliminary studies and other studies have encountered difficulties using *N. benthamiana* as a heterologous host for oxidized sesquiterpene production likely due to endogenous activities such as glycosylation (Fig. S2) (van Herpen et al. 2010). Here, we engineered a yeast chassis for the production of the valerenic acid precursor valerenadiene at a titer of ~140 mg/L. Then, we used phylogenetic and expression analysis to identify a root-upregulated *V. officinalis* valerenadiene oxidase, VoCYP71DJ1, and integrated dehydrogenases to produce a yeast strain capable of generating valerenic acid at 4 mg/L.

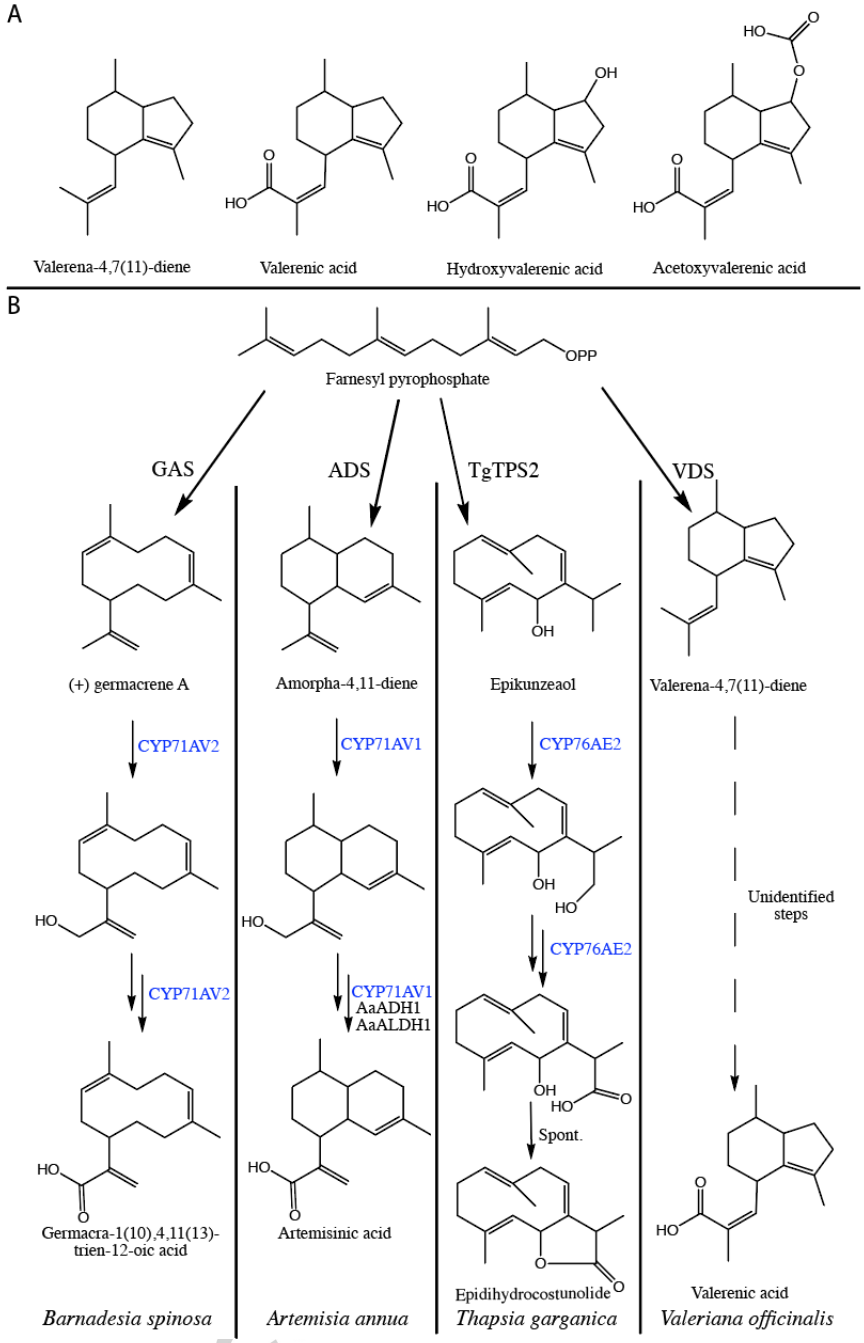

Fig. 1 (A) Important valerenic acid derivatives and related compounds. Valerenic acid, hydroxyvalerenic acid and acetoxyvalerenic acid are the three most abundant valerenic acids in *V. officinalis*. Valerenic acid and valerena-4,7(11)-diene are reported to have sedative effects, while acetoxyvalerenic acid has antagonistic activities. (B) P450 activity in *Asteraceae*, *Apiaceae* and *Caprifoliaceae* sesquiterpenoid drug pathways. Sesquiterpenes are produced in all plant species from MEV pathway-derived farnesyl pyrophosphate (FPP) by sesquiterpene synthases (GAS, ADS, TgTPS2, VDS). Subsequently, P450 (CYP) activity is responsible for sesquiterpene acid formation in characterized biosynthetic pathways. *Artemisia annua* alcohol dehydrogenase (AaADH) and *A. annua* aldehyde dehydrogenase (AaALDH) have been shown to catalyze the conversion of artemisinic alcohol to artemisinic acid alongside P450 activity in the artemisinin pathway. Functionally characterized P450s are shown in blue. Terpenoid pathways in *V. officinalis* are shown in more detail in Figure S1.

Table 1 Yeast strains used in this study

Strain	Parent (+ additional genetic changes)	References
GTy23	erg9::KanMX_P_{CTR3}-ERG9 leu2-3, 112::His3MX6_P_{GAL1}-ERG19/P_{GAL10}-ERG8 ura3-52::URA3_P_{GAL1}-mvaS(A110G)/P_{GAL10}-mvaE(CO) his3Δ1::hphMX4_P_{GAL1}-ERG12/P_{GAL10}-IDI1	
JWy601	GTy23 (ura3-52 prototrophy removed for use of Cas9 system)	This study
JWy602	JWy601 (ARS1622b::P_{GAL1}-VDS-GFP)	This study
JWy603	JWy601 (ARS1622b::P_{GAL1}-VDS)	This study
JWy604	JWy601 (ARS1622b::P_{GAL1}-nMBP-VDS)	This study
JWy605	JWy601 (ARS1622b::P_{GAL1}-nMBP-VDS-GFP)	This study
JWy606	JWy601 (ARS1622b::P_{GAL1}-nMBP-VDS-ERG20)	This study
JWy607	JWy601 (ARS1622b::P_{GAL1}-nMBP-VDS-ERG20-GFP)	This study
JWy608	JWy605 (ARS1014a::P_{GAL1}-nMBP-VDS-ERG20 ARS1114a::P_{GAL1}-nMBP-VDS-ERG20 ARS308a::P_{GAL1}-nMBP-VDS-ERG20)	This study
JWy609	JWy608 (ARS911b::P_{GAL10}-AaCPR, P_{GAL1}-VoCYP714A33)	This study
JWy610	JWy608 (ARS911b::P_{GAL10}-AaCPR, P_{GAL1}-VoCYP81Q107)	This study
JWy611	JWy608 (ARS911b::P_{GAL10}-AaCPR, P_{GAL1}-VoCYP71D510)	This study
JWy612	JWy608 (ARS911b::P_{GAL10}-AaCPR, P_{GAL1}-VoCYP71D511)	This study
JWy613	JWy608 (ARS911b::P_{GAL10}-AaCPR, P_{GAL1}-VoCYP71BE87)	This study
JWy614	JWy608 (ARS911b::P_{GAL10}-AaCPR, P_{GAL1}-VoCYP71DJ1)	This study
JWy615	JWy614(ARS1021b::P_{GAL10}-VoADH1, P_{GAL1}-VoALDH1)	This study
JWy616	JWy615 (ARS607b::P_{GAL1}-VoCYP714A33)	This study
JWy617	JWy615 (ARS607b::P_{GAL1}-VoCYP81Q107)	This study
JWy618	JWy615 (ARS607b::P_{GAL1}-VoCYP71D510)	This study
JWy619	JWy615 (ARS607b::P_{GAL1}-VoCYP71D511)	This study
JWy620	JWy615 (ARS607b::P_{GAL1}-VoCYP71BE87)	This study
JWy621	JWy615(ARS511b::P_{GAL1}-nMBP-VoADH1)	This study
JWy622	JWy615(ARS511b::P_{GAL1}-nMBP-VoALDH1)	This study
JWy623	JWy615(ARS511b::P_{GAL1}-nMBP-VDS-ERG20)	This study
JWy624	JWy614(ARS1021b::P_{GAL10}-AaADH1, P_{GAL1}-AaALDH1)	This study
JWy625	JWy608 (ARS911b::P_{GAL10}-VoCPR1, P_{GAL1}-VoCYP71DJ1)	This study
JWy626	JWy626(ARS1021b::P_{GAL10}-AaADH1, P_{GAL1}-AaALDH1)	This study

2. Materials and methods

2.1. Strain construction

The parent *Saccharomyces cerevisiae* strain used for all engineering was GTy23 {erg9::KanMX_PCTR3-ERG9 leu2-3,112::His3MX6_PGAL1-ERG19/PGAL10-ERG8 ura3-52::URA3_PGAL1-mvaS(A110G)/PGAL10-mvaE(CO) his3Δ1::hphMX4_PGAL1-ERG12/PGAL10-IDI1} previously used by our lab (Reider Apel* et al. 2016). The integration cassettes for all subsequent strains (Table 1) were created using the software tools CASdesigner (casdesigner.jbei.org) and DIVA (diva.jbei.org) and integrated using the previously reported, cloning-free methodology via Cas9-aided homologous recombination (Reider Apel et al. 2017). Integration cassettes containing 1-kb flanking homology regions targeting a chosen genomic locus were constructed by PCR amplifying donor DNA fragments using primers generated by CASdesigner, then co-transformed with a Cas9-

gRNA plasmid (pCut) targeting the chosen genomic locus. CASdesigner primers provide 30–60 nt of inter-fragment homology allowing 1–5 separate fragments to assemble via homologous recombination *in vivo*. pCuts targeting genomic loci were assembled in vivo from a linear backbone and a linear PCR fragment containing the new gRNA sequence, as described previously (Reider Apel et al. 2017). The new gRNA sequence for the URA3 locus was chosen using DNA2.0 (www.dna20.com/eCommerce/cas9/input). To generate donor DNA fragments, native sequences—e.g., chromosomal homology regions, promoters—were amplified from CEN.PK2-1C genomic DNA, while heterologous sequences—e.g., P450 coding sequences (Fig. S9)—were amplified from synthetic gene blocks codon-optimized for expression in *S. cerevisiae*. All genes integrated in this study were expressed under galactose inducible promoters.

All PCRs used Phusion Hot Start II DNA polymerase (www.thermofisher.com, cat. F549L). The following touchdown PCR cycling conditions were used for all PCRs: 1 cycle of 98°C for 15 sec; 25 cycles of 98°C for 10 sec, 65°C for 30 sec (dropping 1 degree each cycle after the first cycle), 72°C for 30 sec, and then 25 cycles of 98°C for 10 sec, 50°C for 30 sec, 72°C for 30 sec. Transformations were performed via heat-shock using ~ 200 ng pCut, ~ 1 μg donor DNA per sample, and 20 min heat shock at 42°C, then plated all cells on selective agarose plates (Gietz and Woods 2002). For assembling a pCut targeting a new site by homologous recombination, we used 200 ng linear pCut backbone and 500 ng of a 1- kb fragment containing the gRNA sequence, as described (Reider Apel et al. 2017). For multi-site integrations, we used 200 ng total linear pCut backbone, and the same amounts of gRNA fragment and donor DNA for each site as we would have for a single integration. Colonies were screened by PCR directed at the target locus, and for integrations, one representative colony was sequenced. Three to four biological replicates were analyzed for each strain.

2.3. Synthetic genes and oligonucleotides.

Oligonucleotides and synthetic genes were commercially synthesized (Integrated DNA Technologies, Inc.). All codon optimized sequences were designed based on the IDT online tool. Sequences of synthetic genes can be found in Figure S9. P450 ORF predictions for gene synthesis were selected from the publically available transcriptome at Medicinal Plant Genomics Resource (http://medicinalplantgenomics.msu.edu). Previously described synthetic, yeast codon optimized *Artemisia annua* ADH1 (JQ582842.1) and ALDH1 (JQ609276.1) were used in this study (Paddon et al. 2013).

2.4. Culture and fermentation conditions

Selective agar plates used for transformations were purchased from Teknova (www.teknova.com, cat. C3080). Liquid selective medium used to grow transformants contained 0.2% (w/v) complete supplement mixture (CSM) lacking uracil (www.sunrisescience.com, cat. 1004-100), 0.67% yeast nitrogen base (www.difco.com, cat. 291920), and 2% dextrose. Nonselective medium contained 1% yeast extract, 2% peptone (Difco cat. 288620 and 211677, respectively), and either 2% dextrose (YPD) or 2% galactose and 0.2% dextrose (YPG). Nonselective agar YPD plates were purchased from Teknova (cat. Y100). Cultures were grown in plastic 96-deep well plates (www.vwr.com, cat. 29445-166) and glass test tubes for strain maintenance, while 2 ml of medium in 24-deep well plastic plates (CWR cat. 89080-534) were used for all production runs. Production cultures were spiked with 100 mg/L trans-caryophyllene (sigma cat. C9653) as an internal standard. Plastic plates were covered with aeraseal film (www.excelscientific.com, cat. BS-25) and shaken at 800 rpm in a Multitron shaker (www.infors-ht.com, model AJ185). Production runs were cultured for

48 hr in 2 ml of YPG before terpenoid extraction for analysis. Glass tubes were shaken at 200 rpm. All strains were grown at 30°C.

2.5. Confocal microscopy

To visualize GFP expression of tagged VDS variants in yeast strains, strains were grown in 5 ml YPD overnight, then back-diluted 1:100 into the same medium and grown 3–6 h at 200 rpm and 30 °C. Then, 1 ml of culture volume was centrifuged at 21,952 x g on a table-top centrifuge, washed with 1x water, and 1ul of the cell pellet was imaged using a Zeiss LSM 710 confocal system mounted on a Zeiss inverted microscope (www.zeiss.com) with a 63Å~ objective and processed using Zeiss Zen software.

2.6. Identification of differentially expressed P450s

Previously reported assembled transcripts and expression abundance estimations were retrieved from the Medicinal Plant Genomics Resource (http://medicinalplantgenomics.msu.edu) and Plantrans DB (http://lifecenter.sgst.cn/plantransdb). To select P450s for functional testing, published transcriptomes of *V. officinalis* were mined using a profile hidden Markov model search (HMMER) (Finn et al. 2011). Correlation of expression profiles was calculated by the Pearson product-moment correlation coefficient of log2 FPKM, with all log2 FPKM values less than zero were set to zero. Subsequent heatmaps were generated using Multiple Experiment Viewer Software (MeV) version 4.5 (Howe et al. 2010). DNA sequences identified in this work are deposited in GenBank™.

2.7. Phylogenetic analysis

Phylogenetic analysis was performed using the entire predicted amino acid sequences of *V. officinalis* P450 family proteins and related terpene-modifying P450 proteins from the GenBank database. Fourteen P450s with homology to *V. officinalis* candidates with known roles in terpenoid biosynthetic pathways were used in the analysis. The following accessions were used: *Valeriana officinalis* VoCYP71D442 (ALU63882.1), *Santalum album* SaCYP76F40 (AHB33947.1), *Santalum album* SaCYP76F39 (AHB33940.1), *Catharanthus roseus* CrCYP76B6 (CAC80883), *Thapsia garganica* TgCYP76AE2 (AQY76213.1), *Lactuca sativa* LsCYP71BL2 (AEI59780.1), *Barnadesia spinosa* BsCYP71AV7 (D5JBX1.1), *Tanacetum cinerariifolium* TcCYP71AV2 (AGO03789.1), *Artemisia annua* AaCYP71AV1 (Q1PS23.1), *Cynara cardunculus var. Scolymus* CcCYP71BL5 (AIA09038.1), *Cynara cardunculus var. Scolymus* CcCYP71AV9 (AIA09035.1), *Cichorium intybus* CiCYP71AV8 (E1B2Z9.1), *Arabidopsis thaliana* AtCYP714A1 (NP_001332750.1), and *Arabidopsis thaliana* AtCYP714A2 (NP_001331286.1). Sequence alignments were generated on the basis of comparison of the amino acid sequences using the MAFFT L-INS-i algorithm with default parameters. Alignments for each partition were generated using the default settings (gap opening penalty = 1.53 and offset value = 0.00) (Katoh et al. 2002). A consistent alignment was selected using TrimAl, with the parameter automated1 (Capella-Gutiérrez et al. 2009). Maximum likelihood analyses were conducted with RAxML v.7.2.8 (Stamatakis 2014). Twenty randomized starting trees were generated with which the initial rearrangement setting and the number of distinct rate categories were determined. The best-known likelihood tree was found by performing 1000 repetitions for each of the amino acid datasets. One thousand non-parametric bootstrap replications were then performed using the bootstrap algorithm. The resulting tree was visualized using FigTree. The scale bar of 0.2

indicates a 20% change and each number shown next to the branches is the number of replicate trees in which the related taxa clustered in the bootstrap test. Gblocks curated alignment for *V. officinalis* P450s and related CYP71 clade P450s is shown in Figure S4.

2.8. Metabolite quantification using GC-MS

Yeast cultures were grown in 2 mL of YPgal with 0.25μM $CuSO_4$ (Sigma cat. 209201) in 24-deep well plastic plates for 48 hours. For GC-MS analysis of valerenadiene, the cultures were overlaid with 0.4mL dodecane (spiked with internal standard), which was pipetted into 1.5mL Eppendorf tubes and spun at 21,952 x g for 1 minute. The resulting organic phase was removed and transferred to GC vials and mixed with EtOAc. For GC-MS analysis of valerenic acid, the cultures were extracted 1:1 with EtOAc spiked with 50 mg/L trans-caryophyllene by shaking with 100μL of glass beads for 8 minutes at a frequency of 28 Hz in a Retsch mixer mill MM 400, then centrifuged at 21,952 x g for 1 minute. The resulting organic phase was removed and dried down at 54°C in vacuum, resuspended in 41 μL EtOAc, 4 μL of 40% v/v tetrabutylammonium hydroxide (TBAH) solution (Sigma cat. 86854). Then, 5 μL of iodomethane (Sigma cat. I8504) was added to the sample, and the mixture was agitated by vortex for 10 s. An aliquot of the sample (1 μL) was injected into a cyclosil B column (J&W Scientific) operating at a He flow rate of 1 mL/min on GC–MS (GC model 6890, MS model 5973 Inert, Agilent). An initial temperature of 120°C was held for 3 min, followed by ramping to 250°C at a rate of 20°C/min, and then held at 250°C for another 3 min. The total flow was set to 8.3 mL/min and helium flow was set to 1 mL/min. All production measurements were performed in biological triplicates or quadruplicates. A caryophyllene standard (Sigma cat. C9653) and a valerenic acid standard (Sigma cat. 51964) containing known concentrations of the internal standard were used to determine titers of valerenadiene and valerenic acid, respectively.

3. Results and discussion

3.1 Engineering production of valerenadiene precursor in yeast

First, we engineered a yeast strain to produce the valerenic acid precursor, valerenadiene, by expressing chromosomally integrated copies of valerenadiene synthase (*VDS*), which converts FPP to valerenadiene. Previous studies have produced low titers of valerenadiene in yeast (~1 mg/L) from unstable, high-copy plasmids (Yeo et al. 2013; Pyle et al. 2012). We used a previously developed mevalonate overproducing strain, GTy23, in which all enzymes converting acetyl-CoA to GPP are overexpressed from galactose-inducible promoters. After, we integrated a single copy of *VDS*, but noted low production, about 3 mg/L, consistent with other non-modified sesquiterpene synthases (Fig. 2) (Son et al. 2014; Beekwilder et al. 2014). Many studies have shown diterpene and sesquiterpene synthases suffer from cytosolic insolubility or instability when expressed in yeast (Ignea et al. 2015; Reider Apel et al. 2017). Thus, we employed protein tagging strategies to visualize insoluble protein and to improve protein solubility. A single integration of a VDS-GFP tagged variant showed severe protein aggregation (Fig. 2b, Fig. S3, left). We proceeded to test other recently described protein tag variations. We tested several combinations of an N-terminal MBP tag, a C-terminal GFP tag, and a C-terminal ERG20 tag (ERG20 being the enzyme catalyzing the formation of FPP). The best VDS variant was MBP-VDS-ERG20, which resulted in substantial titer improvements to ~30 mg/L, almost an order of magnitude increase over untagged VDS. Integrating three additional copies of MBP-VDS-ERG20 led to strain JWy608, which produced a titer of ~140 mg/L valerenadiene (Fig. 2). This strain was used for all P450 functional testing.

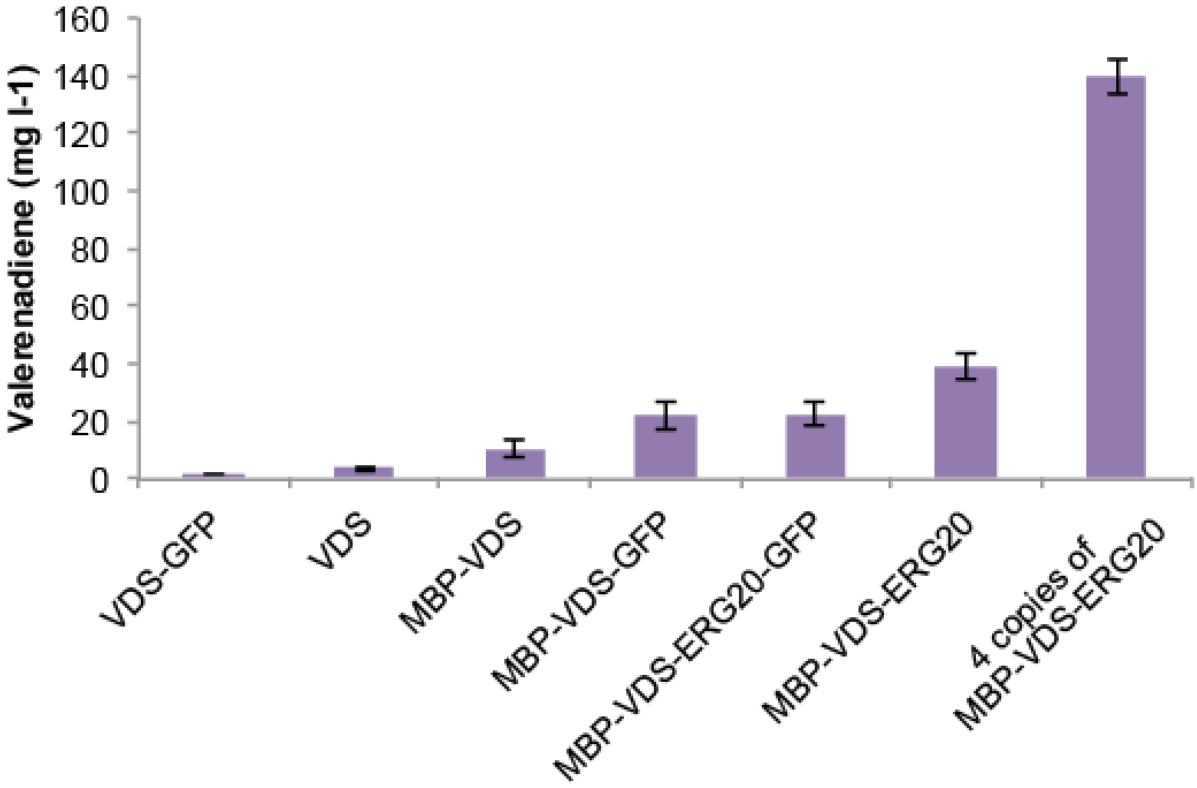

Fig. 2. Titer improved with VDS tagging strategies. Confocal microscopy of yeast cells expressing VDS-GFP shows GFP aggregation. ERG20 appeared to improve VDS solubility and provide increased flux through the pathway by increasing the supply of FPP. Cultures were extracted after 48 hours and analyzed for valerenadiene production by GC-MS. MBP-VDS-ERG20 narrowly outperformed both MBP-VDS-GFP and MBP-VDS-ERG20-GFP in production of valerenadiene. This VDS variant was integrated two additional times to produce the high titer valerenadiene strain, JWy608. Data represent the averages of three replicate cultures; error bars show s.d.

3.2 Identifying putative valerenadiene oxidase genes from V. officinalis

To select P450s for functional testing, we mined the published transcriptomes of *V. officinalis* using a profile hidden Markov model to search for all P450s (Yeo et al. 2013; Pyle et al. 2012; Finn et al. 2011). Subsequently, we used isoform expression data generated by a previous study to identify likely candidates involved in valerenic acid biosynthesis based on root tissue preferential expression, and their correlation with the expression pattern of *VDS* in the tissue types (Yeo et al. 2013). However, we also identified two P450s, *VoCYP81Q107* and *VoCYP71D510*, that had constitutive expression over all tissue types, similar to the expression of upstream isoprenoid biosynthetic gene homologs for *HMGR*, *IDI*, and *FPPS* (Fig. 3). Of the candidates, the expression profile of *VoCYP71DJ1* shared the highest similarity to the expression profile of *VDS*, with no expression in the leaf or callus, low expression in the stem, and the highest expression in the root (Fig. 3).

We made a phylogenetic tree of all putative *V. officinalis* P450s and included other functionally tested P450s that oxidize sesquiterpenes (Fig. 4). We noticed that all of these upregulated P450s had homology to known terpene modifying enzymes, with the exception of VoCYP81Q107. VoCYP714A33 has homology to gibberellin oxidases (Nomura et al. 2013), such as AtCYP714A1 and AtCYP714A1, while VoCYP71BE87 shares homology with the *Vitis vinifera* P450 VvCYP71BE5 responsible for the formation of the sesquiterpenoid (-)-rotundone, an important

component of wine flavor (Takase et al. 2016). VoCYP81Q107 has homology to P450s involved in the sesamin biosynthetic pathway (Hata et al. 2010). Four of the six identified P450 candidates were classified as CYP71D P450s, consistent with the classification of related *Asteraceae* and *Apiaceae* sesquiterpene oxidases that have been found to catalyze the oxidation of a primary carbon on sesquiterpenes forming the respective acids (Nguyen et al. 2010; Ro et al. 2006; Andersen et al. 2017; Ikezawa et al. 2011; Nelson and Werck-Reichhart 2011). Because amorphadiene oxidase AaCYP71AV1 catalyzes three successive oxidations on amorphadiene, similar to the oxidations seen in our target pathway, we expected high homology of AaCYP71AV1 to our candidates. Of the selected putative *V. officinalis* P450s in the CYP71D family, VoCYP71D510 and VoCYP71D511 share approximately 51% homology with AaCYP71AV1, while VoCYP71DJ1 and VoCYP71BE87 share approximately 44% homology with AaCYP71AV1 at the protein sequence level (Fig. S4). VoCYP71D510 and VoCYP71D511 share high homology to VoCYP71D442, previously identified as a possible *V. officinalis* P450 candidate involved in valerenic acid biosynthesis, but this enzyme was not coexpressed with *VDS* or upstream enzymes and did not produce oxidized sesquiterpenes (Ricigliano et al. 2016).

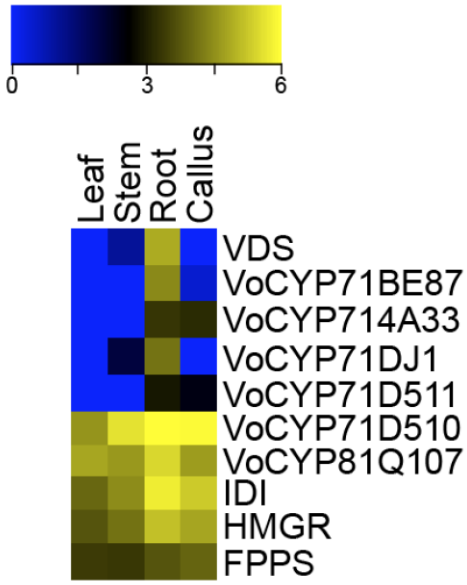

Fig. 3. Heat map of candidate P450s involved in valerenic acid biosynthesis upregulated in the root. Sesquiterpene biosynthetic precursor enzymes upstream of valerenadiene are highly expressed in all tissues, while VDS is almost exclusively expressed in root tissue. P450s were selected by high expression in the root of *V. officinalis*; note, VoCYP71DJ1 (valerenadiene oxidase) shares a similar expression profile among tissue types with VDS. Expression values in log2 FPKM (fragments per Kilobase of transcript per million fragments mapped) were used, negative values were set to zero. Expression values shown represent the different developmental tissues.

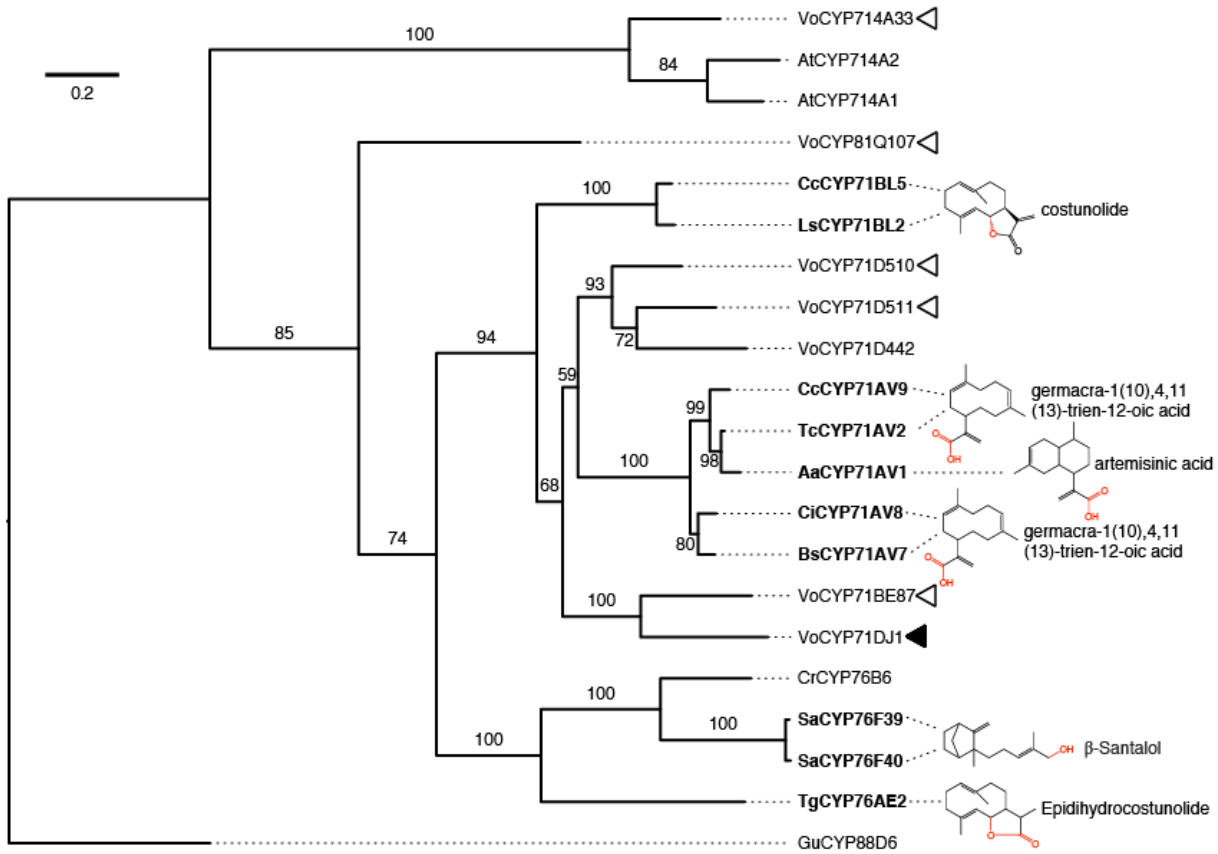

Fig. 4. Phylogenetic tree of *V. officinalis* P450s and related terpene-modifying P450s. Note, Germacrene A oxidase P450s from *Asteraceae* species clade with many of the *V. officinalis* candidates. Respective oxidized sesquiterpene products of functionally tested P450s with oxidations accented in red are shown on left. The neighbor-joining tree was generated using MAFFT and RAxML. The numbers indicate the bootstrap value (%) from 100 replications. The scale bar shows the amino acid substitution ratio. GuCYP88D6, *Glycyrrhiza uralensis* β-amyrin-11-oxidase (AB433179), was used as the outgroup. *V. officinalis* CYP proteins identified in this study are marked with arrowheads. The closed arrowhead indicates VoCYP71DJ1 having the activity of valerenic alcohol synthesis and the open arrowheads indicate CYPs incapable of producing oxidized valerenadiene in this study.

3.3 Functional identification of V. officinalis P450s acting on valerenadiene

To test the candidate P450s for activity, we cointegrated a single copy of each candidate with *A. annua* cytochrome P450 reductase, *AaCPR1*, into our high valerenadiene producer strain, JWy608. Unfortunately, no valerenic acid was produced by any of the candidate P450s. However, we detected the production of trace amounts of hydroxylated valerenadiene by one of our candidate P450s, VoCYP71DJ1 (Fig. 1a, Fig. S6). We were surprised to find that none of the other candidates functioned as a valerenadiene oxidase, despite sequence similarity and upregulated expression in the root.

3.3 Expression of alcohol and aldehyde dehydrogenases forms valerenic acid

Many plant sesquiterpenoid pathways produce respective sesquiterpenoid acids in two steps, requiring only a synthase and a P450. However, several of these pathways and those of other terpenoids show increased production of the final product by overexpressing alcohol and aldehyde dehydrogenases, including the artemisinic acid, jolkinol C, zerumbone, and germacra-1(10),4,11(13)-

trien-12-oic acid pathways (Okamoto et al. 2011; de Kraker et al. 2001; Luo et al. 2016; Paddon et al. 2013).

We mined the *V. officinalis* transcriptome for genes with homology to *A. annua* dehydrogenases and found an alcohol dehydrogenase and an aldehyde dehydrogenase *VoADH1* and *VoALDH1,* respectively. *VoADH1* had poor expression in the root transcriptome, while a *VoALDH1* was highly expressed in the root and stem tissues (Fig. S5). *VoADH* and *VoALDH* were cointegrated into our yeast strain expressing *VoCYP71DJ1*, JWy614, resulting in trace amounts of valerenic acid, as determined by the mass spectrum and retention time relative to an authentic valerenic acid standard (Fig. 5).

Due to the low titer of valerenic acid in this strain, we surmised that another P450 candidate could be responsible for the conversion of valerenic alcohol to valerenic acid. We individually integrated all other P450 candidates into the valerenic acid base strain, JWy615, but failed to see improved valerenic acid titer (Fig. S7). Additionally, hydroxyvalerenic acid (Fig. 1a), a valerenic acid derivative, was also not detected in these strains expressing additional P450 candidates.

To further improve titer, we integrated the previously published *A. annua* dehydrogenases *AaADH* and *AaALDH*, used to improve artemisinic acid titer in yeast, surmising that they might have activity on valerenic alcohol and valerenic aldehyde substrates. Cointegration of *A. annua AaADH1* and *AaALDH1* in JWy614 increased valerenic acid titer (Fig. 5); thus, although the VoADH1 and VoALDH1 identified in this study formed valerenic acid in yeast, it is possible that one or both of these enzymes are not actually responsible for the formation of valerenic acid in the native plant, which is consistent with their expression profiles. Our final strain, JWy627, produces 4mg/L of valerenic acid. These results show that the techniques that led to the successful microbial production of artemisinin can be applied to other terpenoid pathways.

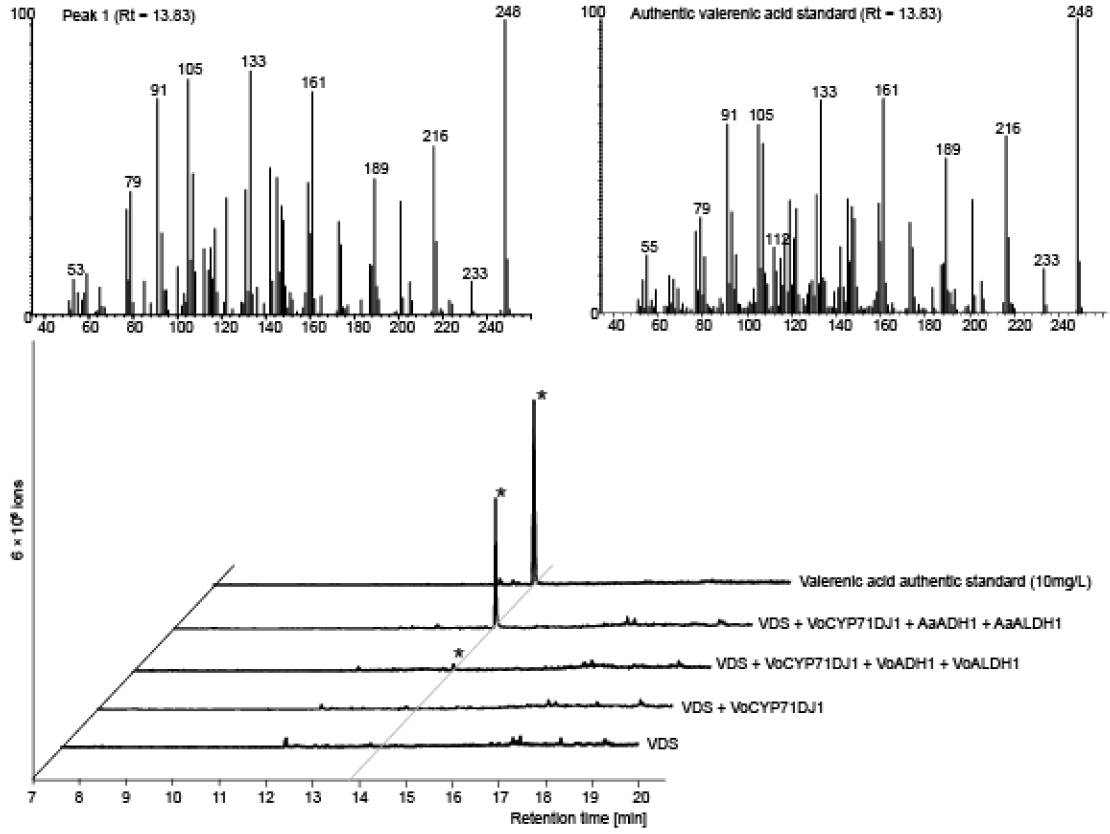

Fig. 5. *In vivo* production of valerenic acid by *V. officinalis* P450s and dehydrogenases in yeast. GC analysis identifies sesquiterpenoid products of extracts of yeast cultures with integrated candidate genes.
Coexpression of *A. annua* dehydrogenases improved production of valerenic acid to ~4 mg/L. VoCYP71DJ1 was cointegrated with *A. annua* CPR. Ethyl acetate-extractable fractions were derivatized and analyzed by GC–MS in extracted ion mode (*m*/*z* 248). Mass spectra of valerenic acid (retention time 13.83 min, detected as methyl ester) relative to authentic standard is shown. Valerenic acid denoted with asterisk.

4. Conclusion

We have engineered a yeast chassis for the production of the sedative valerenic acid by engineering the production of the backbone valerenadiene, identified novel P450 oxidizing valerenadiene, and overexpressed these two genes, and completed the valerenic acid biosynthetic pathway in yeast. Valerenadiene and valerenic acid were produced at 140 mg/L and 4 mg/L, respectively. Phylogenetic and expression analyses were necessary to identify a valerenadiene oxidase, VoCYP71DJ1. Further, expression of an ADH and ALDH were required to produce a yeast strain capable of generating valerenic acid. Microbially produced valerenic acid may allow for more accurate studies of this drug, as plant derived material contains many bioactive compounds, including acetoxyvalerenic acid, a compound with antagonistic effects of valerenic acid. The gene testing strategy used in this study could prove valuable for gene discovery in other medicinally important sesquiterpenoid pathways. These findings also illustrate that closely related P450s have been fine-tuned by evolutionary pressure for specialized metabolism. Additional strain engineering will be necessary to improve valerenic acid titers for industrial applications.

Acknowledgements

This work was supported by the US Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231. The authors thank Dr. David Nelson (Department of Molecular Sciences, University of Tennessee) for helpful suggestions about P450 nomenclature and classification, and Patrick Shih, Solomon Stonebloom, and Henrik Scheller for their support and guidance during the initial stages of this project.

References

Andersen, T.B., Martinez-Swatson, K.A., Rasmussen, S.A., Boughton, B.A., Jørgensen, K., Andersen-Ranberg, J., Nyberg, N., Christensen, S.B. and Simonsen, H.T. 2017. Localization and in-Vivo Characterization of Thapsia garganica CYP76AE2 Indicates a Role in Thapsigargin Biosynthesis. *Plant Physiology* 174(1), pp. 56–72.
Becker, A., Felgentreff, F., Schröder, H., Meier, B. and Brattström, A. 2014. The anxiolytic effects of a Valerian extract is based on valerenic acid. *BMC complementary and alternative medicine* 14, p. 267.
Beekwilder, J., van Houwelingen, A., Cankar, K., van Dijk, A.D.J., de Jong, R.M., Stoopen, G., Bouwmeester, H., Achkar, J., Sonke, T. and Bosch, D. 2014. Valencene synthase from the heartwood of Nootka cypress (Callitropsis nootkatensis) for biotechnological production of valencene. *Plant*

Biotechnology Journal 12(2), pp. 174–182.
Benke, D., Barberis, A., Kopp, S., Altmann, K.-H., Schubiger, M., Vogt, K.E., Rudolph, U. and Möhler, H. 2009. GABA A receptors as in vivo substrate for the anxiolytic action of valerenic acid, a major constituent of valerian root extracts. *Neuropharmacology* 56(1), pp. 174–181.
Capella-Gutiérrez, S., Silla-Martínez, J.M. and Gabaldón, T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25(15), pp. 1972–1973.
Eadie, M.J. 2004. Could valerian have been the first anticonvulsant? *Epilepsia* 45(11), pp. 1338–1343.
Felgentreff, F., Becker, A., Meier, B. and Brattström, A. 2012. Valerian extract characterized by high valerenic acid and low acetoxy valerenic acid contents demonstrates anxiolytic activity. *Phytomedicine: International Journal of Phytotherapy and Phytopharmacology* 19(13), pp. 1216–1222.
Finn, R.D., Clements, J. and Eddy, S.R. 2011. HMMER web server: interactive sequence similarity searching. *Nucleic Acids Research* 39(Web Server issue), pp. W29–37.
Gietz, R.D. and Woods, R.A. 2002. Screening for protein-protein interactions in the yeast two-hybrid system. *Methods in Molecular Biology* 185, pp. 471–486.
Hassan, E., Tayebeh, R., Tousi Samaneh, E., Sadr Zeinalabedin, B., Vahid, N. and Mehdi, Z. 2008. Quantification of Valerenic Acid and its Derivatives in Some Species of Valeriana L. and Centranthus longiflorus Stev. *Asian Journal of Plant Sciences* 7(2), pp. 195–200.
Hata, N., Hayashi, Y., Okazawa, A., Ono, E., Satake, H. and Kobayashi, A. 2010. Comparison of sesamin contents and CYP81Q1 gene expressions in aboveground vegetative organs between two Japanese sesame (Sesamum indicum L.) varieties differing in seed sesamin contents. *Plant Science* 178(6), pp. 510–516.
van Herpen, T.W.J.M., Cankar, K., Nogueira, M., Bosch, D., Bouwmeester, H.J. and Beekwilder, J. 2010. Nicotiana benthamiana as a production platform for artemisinin precursors. *Plos One* 5(12), p. e14222.
Houghton, P.J. 1988. The biological activity of valerian and related plants. *Journal of Ethnopharmacology* 22(2), pp. 121–142.
Howe, E., Holton, K., Nair, S., Schlauch, D., Sinha, R. and Quackenbush, J. 2010. MeV: MultiExperiment Viewer. In: Ochs, M. F., Casagrande, J. T., and Davuluri, R. V. eds. *Biomedical informatics for cancer research*. Boston, MA: Springer US, pp. 267–277.
Hritcu, L. and Cioanca, O. 2016. Prevalence of use of herbal medicines and complementary and alternative medicine in europe. In: Grosso, C. ed. *Herbal medicine in depression*. Cham: Springer International Publishing, pp. 135–181.
Ignea, C., Trikka, F.A., Nikolaidis, A.K., Georgantea, P., Ioannou, E., Loupassaki, S., Kefalas, P., Kanellis, A.K., Roussis, V., Makris, A.M. and Kampranis, S.C. 2015. Efficient diterpene production in yeast by engineering Erg20p into a geranylgeranyl diphosphate synthase. *Metabolic Engineering* 27, pp. 65–75.
Ikezawa, N., Göpfert, J.C., Nguyen, D.T., Kim, S.-U., O'Maille, P.E., Spring, O. and Ro, D.-K. 2011. Lettuce costunolide synthase (CYP71BL2) and its homolog (CYP71BL1) from sunflower catalyze distinct regio- and stereoselective hydroxylations in sesquiterpene lactone metabolism. *The Journal of Biological Chemistry* 286(24), pp. 21601–21611.
Katoh, K., Misawa, K., Kuma, K. and Miyata, T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research* 30(14), pp. 3059–3066.
Khom, S., Baburin, I., Timin, E., Hohaus, A., Trauner, G., Kopp, B. and Hering, S. 2007. Valerenic acid potentiates and inhibits GABA(A) receptors: molecular mechanism and subunit specificity. *Neuropharmacology* 53(1), pp. 178–187.
Kirk, J. and Dunker, K.S. 2014. Dietary counseling: the ingredient for successfully addressing the use

of herbal supplements and probiotics in chronic kidney disease. *Advances in chronic kidney disease* 21(4), pp. 377–384.
Kopp, S., Schweizer, W. and Altmann, K.-H. 2009. Total synthesis of valerenic acid. *Synlett : accounts and rapid communications in synthetic organic chemistry* 2009(11), pp. 1769–1772.
de Kraker, J.W., Franssen, M.C., Dalm, M.C., de Groot, A. and Bouwmeester, H.J. 2001. Biosynthesis of germacrene A carboxylic acid in chicory roots. Demonstration of a cytochrome P450 (+)-germacrene a hydroxylase and NADP+-dependent sesquiterpenoid dehydrogenase(s) involved in sesquiterpene lactone biosynthesis. *Plant Physiology* 125(4), pp. 1930–1940.
Kumar, V. 2006. Potential medicinal plants for CNS disorders: an overview. *Phytotherapy Research* 20(12), pp. 1023–1035.
Luo, D., Callari, R., Hamberger, B., Wubshet, S.G., Nielsen, M.T., Andersen-Ranberg, J., Hallström, B.M., Cozzi, F., Heider, H., Lindberg Møller, B., Staerk, D. and Hamberger, B. 2016. Oxidation and cyclization of casbene in the biosynthesis of Euphorbia factors from mature seeds of Euphorbia lathyris L. *Proceedings of the National Academy of Sciences of the United States of America* 113(34), pp. 5082–5089.
Nelson, D. and Werck-Reichhart, D. 2011. A P450-centric view of plant evolution. *The Plant Journal: for Cell and Molecular Biology* 66(1), pp. 194–211.
Nguyen, D.T., Göpfert, J.C., Ikezawa, N., Macnevin, G., Kathiresan, M., Conrad, J., Spring, O. and Ro, D.-K. 2010. Biochemical conservation and evolution of germacrene A oxidase in asteraceae. *The Journal of Biological Chemistry* 285(22), pp. 16588–16598.
Nomura, T., Magome, H., Hanada, A., Takeda-Kamiya, N., Mander, L.N., Kamiya, Y. and Yamaguchi, S. 2013. Functional analysis of Arabidopsis CYP714A1 and CYP714A2 reveals that they are distinct gibberellin modification enzymes. *Plant & Cell Physiology* 54(11), pp. 1837–1851.
Okamoto, S., Yu, F., Harada, H., Okajima, T., Hattan, J., Misawa, N. and Utsumi, R. 2011. A short-chain dehydrogenase involved in terpene metabolism from Zingiber zerumbet. *The FEBS Journal* 278(16), pp. 2892–2900.
Paddon, C.J., Westfall, P.J., Pitera, D.J., Benjamin, K., Fisher, K., McPhee, D., Leavell, M.D., Tai, A., Main, A., Eng, D., Polichuk, D.R., Teoh, K.H., Reed, D.W., Treynor, T., Lenihan, J., Fleck, M., Bajad, S., Dang, G., Dengrove, D., Diola, D., Dorin, G., Ellens, K.W., Fickes, S., Galazzo, J., Gaucher, S.P., Geistlinger, T., Henry, R., Hepp, M., Horning, T., Iqbal, T., Jiang, H., Kizer, L., Lieu, B., Melis, D., Moss, N., Regentin, R., Secrest, S., Tsuruta, H., Vazquez, R., Westblade, L.F., Xu, L., Yu, M., Zhang, Y., Zhao, L., Lievense, J., Covello, P.S., Keasling, J.D., Reiling, K.K., Renninger, N.S. and Newman, J.D. 2013. High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature* 496(7446), pp. 528–532.
Pyle, B.W., Tran, H.T., Pickel, B., Haslam, T.M., Gao, Z., MacNevin, G., Vederas, J.C., Kim, S.-U. and Ro, D.-K. 2012. Enzymatic synthesis of valerena-4,7(11)-diene by a unique sesquiterpene synthase from the valerian plant (Valeriana officinalis). *The FEBS Journal* 279(17), pp. 3136–3146.
Reider Apel*, A., Espaux*, L. d', Wehrs*, M., Sachs, D., Li, R.A., Tong, G.J., Garber, M., Nnadi, O., Zhuang, W., Hillson, N.J., Keasling, J.D. and Aindrila Mukhopadhyay 2016. A Cas9-based toolkit to program gene expression in Saccharomyces cerevisiae.
Reider Apel, A., Espaux, L. d', Wehrs, M., Sachs, D., Li, R.A., Tong, G.J., Garber, M., Nnadi, O., Zhuang, W., Hillson, N.J., Keasling, J.D. and Mukhopadhyay, A. 2017. A Cas9-based toolkit to program gene expression in Saccharomyces cerevisiae. *Nucleic Acids Research* 45(1), pp. 496–508.
Ricigliano, V., Kumar, S., Kinison, S., Brooks, C., Nybo, S.E., Chappell, J. and Howarth, D.G. 2016. Regulation of sesquiterpenoid metabolism in recombinant and elicited Valeriana officinalis hairy roots. *Phytochemistry* 125, pp. 43–53.
Ro, D.-K., Paradise, E.M., Ouellet, M., Fisher, K.J., Newman, K.L., Ndungu, J.M., Ho, K.A., Eachus, R.A., Ham, T.S., Kirby, J., Chang, M.C.Y., Withers, S.T., Shiba, Y., Sarpong, R. and Keasling, J.D.

2006. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* 440(7086), pp. 940–943.

Son, Y.-J., Kwon, M., Ro, D.-K. and Kim, S.-U. 2014. Enantioselective microbial synthesis of the indigenous natural product (-)-α-bisabolol by a sesquiterpene synthase from chamomile (Matricaria recutita). *The Biochemical Journal* 463(2), pp. 239–248.

Stamatakis, A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30(9), pp. 1312–1313.

Takase, H., Sasaki, K., Shinmori, H., Shinohara, A., Mochizuki, C., Kobayashi, H., Ikoma, G., Saito, H., Matsuo, H., Suzuki, S. and Takata, R. 2016. Cytochrome P450 CYP71BE5 in grapevine (Vitis vinifera) catalyzes the formation of the spicy aroma compound (-)-rotundone. *Journal of Experimental Botany* 67(3), pp. 787–798.

Takemoto, H., Yagura, T. and Ito, M. 2009. Evaluation of volatile components from spikenard: valerena-4,7(11)-diene is a highly active sedative compound. *Journal of natural medicines* 63(4), pp. 380–385.

Trauner, G., Khom, S., Baburin, I., Benedek, B., Hering, S. and Kopp, B. 2008. Modulation of GABAA receptors by valerian extracts is related to the content of valerenic acid. *Planta Medica* 74(1), pp. 19–24.

Yeo, Y.-S., Nybo, S.E., Chittiboyina, A.G., Weerasooriya, A.D., Wang, Y.-H., Góngora-Castillo, E., Vaillancourt, B., Buell, C.R., DellaPenna, D., Celiz, M.D., Jones, A.D., Wurtele, E.S., Ransom, N., Dudareva, N., Shaaban, K.A., Tibrewal, N., Chandra, S., Smillie, T., Khan, I.A., Coates, R.M., Watt, D.S. and Chappell, J. 2013. Functional identification of valerena-1,10-diene synthase, a terpene synthase catalyzing a unique chemical cascade in the biosynthesis of biologically active sesquiterpenes in Valeriana officinalis. *The Journal of Biological Chemistry* 288(5), pp. 3163–3173.