MR JIA WEI ZHOU (Orcid ID : 0000-0001-6628-3056)

Article type : Regular Article

Friedelane-type triterpene cyclase in celastrol biosynthesis from *Tripterygium wilfordii* and its application for triterpenes biosynthesis in yeast

Jiawei Zhou[1,2], Tianyuan Hu[2], Linhui Gao[1], Ping Su[4], Yifeng Zhang[2], Yujun Zhao[4], Shang Chen[1], Lichan Tu[2], Yadi Song[1], Xing Wang[1], Luqi Huang[4], Wei Gao[2, 1, 3]*

[1]School of Traditional Chinese Medicine, Capital Medical University, Beijing 100069, China; [2]School of Pharmaceutical Science, Capital Medical University, Beijing 100069, China; [3]Advanced Innovation Center for Human Brain Protection, Capital Medical University, Beijing 100069, China; [4]State Key Laboratory Breeding Base of Dao-di Herbs, National Resource Center for Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing 100700, China

* Correspondence: weigao@ccmu.edu.cn (W.G.); Tel: +86-10-8391-6572 (W.G.)

Received: 19 January 2019
Accepted: 14 March 2019

Summary

Celastrol is a promising bioactive compound isolated from *Tripterygium wilfordii* and has been proven to possess many encouraging preclinical applications. However, the celastrol biosynthetic pathway is poorly understood, especially the key oxidosqualene cyclase (OSC) responsible for cyclization of the main scaffold.

Here, we report on the isolation and characterization of three OSCs from *T. wilfordii*: *Tw*OSC1, *Tw*OSC2 and *Tw*OSC3. Both *Tw*OSC1 and *Tw*OSC3 were multi-product friedelin synthases, while *Tw*OSC2 was a β-amyrin synthase.

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/nph.15809

We further found that *Tw*OSC1 and *Tw*OSC3 were involved in the biosynthesis of celastrol and that their common product, friedelin, was a precursor of celastrol. We then reconstituted the biosynthetic pathway of friedelin in engineered yeast constructed by the CRISPR/Cas9 system, with protein modification and medium optimization, leading to heterologous production of friedelin at 37.07 mg L^{-1} in a shake flask culture.

Our study was the first to identify the genes responsible for biosynthesis of the main scaffold of celastrol and other triterpenes in *T. wilfordii*. As friedelin has been found in many plants, the results and approaches described here laid a solid foundation for further explaining the biosynthesis of celastrol and related triterpenoids. Moreover, our results provide insights for metabolic engineering of friedelane-type triterpenes.

Key words: biosynthesis, celastrol, oxidosqualene cyclases (OSC), *Tripterygium wilfordii*, triterpenes.

Introduction

Tripterygium wilfordii Hook.F is a medicinal plant known in traditional Chinese medicine as 'Lei gong teng' (also called thunder god vine), the debarked root of which has a long history of use in the treatment of rheumatism and inflammation (Jiang, 1994; Tao & Lipsky, 2000; Zhou, ZL *et al.*, 2012). Some listed drugs developed from *T. wilfordii* including Tripterygium glycoside tablets and Leigongteng tablets, have also been used to treat autoimmune and inflammatory diseases in China since the 1970s (Li, 2000; Tao *et al.*, 2001; Yang *et al.*, 2005). A lot of pharmacological studies have suggested that extracts of *T. wilfordii* can be used to treat tumours (Wong *et al.*, 2012), nephrotic syndrome (Xu *et al.*, 2009), HIV (Duan *et al.*, 2000a), Crohn's disease (Ren *et al.*, 2007), and a series of autoimmune and inflammatory diseases (Ma *et al.*, 2007). These remarkable bioactivities are mainly dependent on the active compounds isolated from *T. wilfordii,* such as sesquiterpenoid alkaloids, diterpenoids and triterpenes. Among these components, triptolide and celastrol are thought to be two of the natural products with the greatest potential to be developed into modern drugs (Corson &

Crews, 2007). Celastrol is a friedelane-type triterpene that was first isolated from the root of *T. wilfordii*, and was found to be effective in the treatment of inflammatory, autoimmune diseases (Astry *et al.*, 2015), Alzheimer's disease (Allison *et al.*, 2001) and cancer (Yang, H *et al.*, 2006; Pang *et al.*, 2010). Interestingly, researchers also found this compound to be a leptin sensitizer and a promising agent for the treatment of obesity (Liu *et al.*, 2015; Ma *et al.*, 2015).

Despite considerable pharmaceutical interest, the low celastrol levels, slow growth rate and restricted habitat of this plant have all limited the further study and application of celastrol. To address the above issues, methods have been developed to obtain a sustainable and reliable supply, such as cell culture systems and total synthesis (Su *et al.*, 2014; Camelio *et al.*, 2015). In recent years, metabolic engineering in microorganisms has been proven to be a promising method for the production of high-value natural products (Zhou, YJ *et al.*, 2012; Paddon *et al.*, 2013; Galanie *et al.*, 2015; Lau & Sattely, 2015). However, the biosynthetic pathway of celastrol remains unknown. The pharmaceutical potential of triptolide has led many researchers to investigate the biosynthetic pathway of this compound (Forman *et al.*, 2017; Hansen *et al.*, 2017; Inabuy *et al.*, 2017; Su *et al.*, 2018), but there have been few reports on the biosynthetic pathway of celastrol.

The common precursors of terpenoids are isopentenyl diphosphate (IPP) and its isomer dimethylallyl diphosphate (DMAPP) derived from the mevalonate (MVA) pathway and methylerythritol phosphate (MEP) pathway. Celastrol is a friedelane-type triterpene that is synthesized via the MVA pathway. Although the biosynthetic pathway of celastrol is unclear, a potential route (Fig. 1) has been proposed based on reported metabolomic data of triterpenes from *T. wilfordii* (Zhang *et al.*, 1991; Li *et al.*, 1997; Duan *et al.*, 2000b; Miao *et al.*, 2000; Yang, JH *et al.*, 2006). In this pathway, oxidosqualene cyclase (OSC) catalyses the complicated cyclization of 2, 3-oxidosqualene to friedelin, which is further decorated by cytochrome P450 to form polpunonic acid and wilforic acid C. After a series of oxidation reactions and rearrangements, these intermediates are finally transformed to celastrol. The cyclization of 2, 3-oxidosqualene is the first committed step in this unknown pathway and determines the scaffold of celastrol; thus it is important to determine the OSCs that are responsible for the biosynthesis of celastrol.

Most triterpene scaffolds are cyclized via the chair-chair-chair (CCC) conformation catalysed by OSC, while sterols are cyclized via chair-boat-chair (CBC) conformation (Fig.2). The cyclization of oxidosqualene involved four main steps:(a) binding of the substrate and predetermination the ring system, (b) initiation of the reaction by protonation of the epoxide, (c) cyclization and rearrangement of carbocation species, and (d) termination by deprotonation or water capture to yield a final terpene product (Thimmappa *et al.*, 2014). More than 200 different triterpene skeletons have been discovered from natural resources or enzymatic reactions (Xu *et al.*, 2004). Friedelin is one of the most highly rearranged triterpenes known in plants.

Here, we aimed to identify the OSCs responsible for the biosynthesis of celastrol. Three candidate OSC genes were screened from the transcriptome of *T. wilfordii*. *Tw*OSC1 and *Tw*OSC3 are friedelane-type triterpene cyclases responsible for the biosynthesis of celastrol, while *Tw*OSC2 is a *β*-amyrin synthase. Both *Tw*OSC1 and *Tw*OSC3 can cyclize 2, 3-oxidosqualene to friedelin as the major product, with *β*-amyrin and *α*-amyrin as the minor products. We also proved that friedelin is a precursor of celastrol. In addition, the yield of friedelin in the engineered yeast was also improved by using site-directed mutagenesis, the CRISPR/Cas9 system and medium optimization. Our findings laid the foundation for exploration of the biosynthetic pathway of celastrol and other friedelane-type triterpenes and provided insights into the biosynthetic pathways of friedelane-type, oleanane-type and ursane-type triterpenes in *T. wilfordii* and other plants.

Materials and Methods

Plant materials

T. wilfordii cell suspensions were initiated from calluses of *T. wilfordii* and cultured in Murashige and Skoog (MS) medium containing 30 g L^{-1} sucrose with 0.1 mg L^{-1} kinetin (KT), 0.5 mg L^{-1} 2,4-dichlorophenoxyacetic acid (2,4-D) and 0.5 mg L^{-1} indole-3-butyric acid (IBA). These cell suspensions were cultured in the dark at 25 °C with rotary shaking at 120 rpm as described previously (Guan *et al.*, 2017). Mature *T. wilfordii* plant tissues were collected from the Yongan national forest in Fujian Province, China. Samples of root, stem,

leaves and flowers were harvested from three individual plants, and ground in liquid nitrogen using a Retsch MM400 mixer mill (Retsch GmbH, Germany, https://www.retsch.com/) and then stored at -80 °C until use.

RNA isolation, cDNA synthesis and gene cloning

Total RNA from *T. wilfordii* suspension cells and plant tissues was isolated by the modified cetyltrimethylammonium bromide (CTAB) method as described previously (Zhou *et al.*, 2018). Total RNA was purified using an RNA Purification Kit (Tiangen Biotech, Beijing, China). First-strand cDNA was cloned using the SMARTer^{TM} RACE cDNA Amplification Kit (Clontech Laboratories,Cal., USA). To mine the candidates, tblastn analysis of OSCs in the *T. wilfordii* transcriptome was done by the BioEdit 7.0.9.0.(Altschul *et al.*, 1997). Friedelin synthase from *Kalanchoe daigremontiana* (*Kd*FRS, ADK35125) was used as query sequence. The full-length predicted cDNA sequences of the candidates were obtained from the transcriptome data of *T. wilfordii*. Special primers (Table S1) for cloning the *Tw*OSCs were designed based on the full-length predicted cDNA sequences. The amplification reaction was performed using Phusion high-fidelity PCR master mix (New England Biolabs) according to the manufacturer's instructions. The PCR products were purified and ligated into the pEASY-T3 vector, transformed into *E. coli* DH5*α* cells, and then cultured in Luria-Bertani medium at 37 □ in the dark. Positive colonies were verified by sequencing.

Sequence and phylogenetic analyses

The open reading frames (ORFs) and amino acid sequences of the *Tw*OSCs were identified using the online tool ORF Finder (http://www.ncbi.nlm.nih.gov/gorf/gorf.html). Multiple sequence alignments were conducted by DNAMAN. Interpro (www.ebi.ac.uk/Tools/InterProScan) was used to identify functional domains. The amino acid sequences of OSCs from other species were downloaded from the National Center for Biotechnology Information (NCBI) database and aligned by ClustalW. Then, a neighbour-joining tree was built by MEGA6 software via the bootstrap method with 1000 replications (Tamura *et al.*, 2013).

MeJA-mediated induction of *T. wilfordii* suspension cells

The *T. wilfordii* suspension cells were treated with methyl jasmonate (MeJA) at a final concentration of 50 μmol L^{-1} after subculturing. The control group was treated with the same volume of carrier solution. After 0, 4, 8, 12, 24, 48, 72, and 360 h, the suspension cells were harvested in liquid nitrogen and then stored at -80 °C before RNA extraction or UPLC analysis. Four biological duplicates were prepared for each time point in each group.

Quantitative real-time PCR

The total RNA of the treated suspension cells was extracted by the modified CTAB method as described above. First-stand cDNA for quantitative real-time PCR (qRT-PCR) was reverse transcribed from the total RNA according to manufacturer's instruction for the FastQuant RT Kit (Tiangen Biotech, Beijing, China). qRT-PCRs were performed with gene specific primers (Table S1) and the KAPA SYBR® FAST qPCR Kit (KAPA Biosystems, Massachusetts, USA) on an Applied Biosystems QuantStudio 5 real-time PCR System (Applied Biosystems, New York, USA). The PCR conditions were as follows: an initial incubation at 95 □ for 3 min, followed by 45 cycles of 95 □ for 3 s and 60 □ for 30 s and then by a melting curve cycle. Elongation factor 1 α (EF1α) was used as a reference gene (Miao *et al.*, 2015). Relative transcript abundance was evaluated using the $2^{-\Delta\Delta Ct}$ method (Livak & Schmittgen, 2001) with triplicate measurements from four biological replicates.

UPLC analysis of celastrol content in induced *T. wilfordii* suspension cells and plant tissues

The *T. wilfordii* suspension cells and powders of plant tissues were stored at -80 °C for at least 4 h prior to freeze drying for 48 h (Christ ALPHA 1-2, Germany). Approximately 50 mg of each sample was suspended in 1 mL of 80% (v/v) methanol overnight at room temperature (25 °C), and then extracted by sonication in an ultrasonic water bath for 60 min. After centrifugation for 2 min, 10000× g and room temperature, the supernatant was filtered through a 0.22-μm membrane filter (polytetrafluoroethylene) before UPLC analysis.

The analyses were conducted using an Agilent 1290 Infinity II Efficient UHPLC system quipped with a DAD detector. Chromatographic separation was conducted using a Waters ACQUITY UPLC HSS T3 analytical column (2.1×100 mm, 1.8 μm) maintained at 40 ℃. The mobile phase, consisting of a mixture of 0.1% (v/v) formic acid in water (A) and acetonitrile (C), was pumped at a flow rate of 0.4 mL/min. The gradient program was 30% C at 0 min, 35% C at 5-8 min, 70% C at 15 min, and 90% C at 21 min. The detection wavelength was 425 nm, and UV spectra from 190 to 500 nm were also recorded. The injection volume was 5 μL.

Functional identification in yeast

The ORFs of *Tw*OSCs were amplified and sub-cloned into the pYES2 expression vector (Invitrogen) separately. The ORFs of *Tw*OSC1 and *Tw*OSC2 were inserted between the *Kpn*I-*Not*I sites of pYES2 via digestion by the corresponding restriction endonucleases. *Tw*OSC3 was inserted between the *Kpn*I-*Bam*HI sites of pYES2 in the same manner. All primers are shown in Table S1. The constructed vectors were transformed into lanosterol synthase-deficient yeast (*Saccharomyces cerevisiae*, purchased from ATCC, cell line number: 4021900, -ERG7, -*ura*) using the Frozen-EZ Yeast Transformation II Kit (ZYMO RESEARCH). The pYES2 vector was transformed into yeast as a control.

The transformants were cultured on solid synthetic complete medium without uracil (SC-Ura) for selection. The positive transformant was then cultured in 20 mL of Sc-Ura medium containing 2% glucose and incubated with shaking at 220 rpm and 30 ℃ for 2 days. Then, the cells were collected and induced in 20 mL of Sc-Ura medium with 2% galactose in place of glucose and further cultured at 220 rpm and 30 ℃ for 12 h. After induction, the yeast cells were collected and re-suspended in 0.1 M potassium phosphate buffer (pH 7.0) with 2% glucose, and cultured at 220 rpm and 30 ℃ for 1 day. Finally, the yeast cells were collected and lysed with 10 mL of 20% KOH and 50% EtOH by ultrasonic extraction for 1 h or refluxed for 5 min with 10 mL of 20% KOH and 50% EtOH. The supernatant was extracted with 10 mL of hexane three times. All the extracts were combined and then evaporated by rotary evaporation.

The extracts were dissolved in pyridine and derivatized with N-methyl-N-(trimethylsilyl) trifluoroacetamide at 70 °C for 1 h. Then, the derived extracts were evaporated under N_2 and dissolved in 1 mL trichloromethane ($CHCl_3$) for analysis by gas chromatography and mass spectrometry (GC-MS). The analysis was performed on an Agilent 7000 gas chromatograph (split, 20:1; injector temperature, 250 □) with a DB-5ms (15 m×250 μm×0.1 μm) capillary column. One microliter of the concentrated organic phase was then injected under a He flow rate of 1 mL/min with a temperature program of 1 min at 50 °C, followed by a gradient from 50 to 260 °C at 50 °C min^{-1} and then, to 272 °C at 1 °C min^{-1} with a 4 min held. The ion trap temperature was 230 □. The electron energy was 70 eV. Spectra were recorded in the range of 10-550 m/z.

RNAi study

A specific 300-500 bp fragment was selected based on the blast analysis in *T. wilfordii* transcriptome database (Fig.S1 and Fig. S2) and amplified from the ORF of each *Tw*OSC by specific primers (Table S1) and then inserted into the vector pK7GWIWG2D (Invitrogen) by the Gateway cloning system (Invitrogen). The vectors containing each fragment were transformed into *E. coli* and then verified by PCR and sequencing. The resulting vectors were transformed into *T. wilfordii* suspension cells according to Zhao *et al.* (Zhao *et al.*, 2017). Suspension cells in the logarithmic growth phase were plated on MS solid medium (pH = 5.8) and grown for 7 days before transformation. The empty vector was transformed into *T. wilfordii* suspension cells as a control, and all samples had five biological duplicates. The transformed cell suspensions were incubated for 7 days on MS medium and then transferred onto selection medium (solid MS medium supplemented with 0.5 mg L^{-1} 2, 4-D, 0.1 mg L^{-1} KT, 0.5 mg L^{-1} IBA) supplemented with 100 mg L^{-1} kanamycin for selection of transformed cell suspensions. Positive transformed cell suspensions were further inoculated in the selection MS medium for culture growth with shaking. After three generations of screening, the transformed cell suspensions were harvested to measure gene expression and celastrol levels.

Feeding study

After subculturing, *T. wilfordii* suspension cells (2 g fresh weight of suspension cells in 40 mL of MS medium in 100-mL flasks were further cultured for 10 days, and then, friedelin (2.77 mg) dissolved in 300 μL ethanol/Tween 80 (1/1, v/v) was added into the culture. The control group was treated with the same volume of carrier solution alone, and each sample set had four biological duplicates. After treatment, the suspension cells were harvested at 0 and 7 days by negative-pressure filtration. Then, the suspension cells were stored at -80 °C for at least 4 h prior to freeze drying for 48 h. Approximately 50 mg of dried suspension cells were extracted with 1.0 mL of 80% (v/v) methanol as described above. The supernatant was filtered through a 0.22-μm membrane filter (polytetrafluoroethylene) for UPLC analysis.

Molecular docking and site-directed mutagenesis

The sequence logo and graphical representation of the DCTAE motif were generated with WebLogo 3 (http://weblogo.threeplusone.com/create.cgi) with default parameters (Crooks *et al.*, 2004), the alignment of the motif was visualized with DNAMAN. The three-dimensional protein structures of the *Tw*OSCs were created by MODELER 9v11 (Eswar *et al.*, 2006) based on the identified crystal structure of human OSC (RCSB Protein Data Bank ID: 1W6K). The three-dimensional protein structures of the *Tw*OSC models are shown in Fig. S3. The product friedelin or substrate 2, 3-oxidosqualene was docked with the model structures separately using SYBYL X-1.3. Molecular docking was performed as previously described (Su *et al.*, 2017). The docking results were visualized with PyMOL.

Site-directed mutagenesis of the *Tw*OSCs was conducted using the Fast Mutagenesis System (TransGen Biotech). The mutants of *Tw*OSC1, *Tw*OSC2 and *Tw*OSC3 were amplified from the templates using the specific primers (Table S1). The pYES2 vector containing each *Tw*OSC was used as a template. The constructed mutants were verified by PCR and sequencing. Positive mutants were transformed into lanosterol synthase-deficient yeast and cultured. The products were then extracted and analysed by GC-MS as described above.

Strain construction and fermentation

Specific gRNAs targeting *rox1*, *ypl062w* and *yjl064w* were designed by the open-source tool at http://yeastriction.tnw.tudelft.nl. Single and multiple gRNA expression vectors for *yjl064w*, *ypl062w* and *rox1* were constructed as described previously (Mans *et al.*, 2015). The 120 bp repair fragments used for effective repair by the homologous recombination repair machinery were obtained by two complementary single-stranded oligos using an annealing program: 5 min at 95 °C, followed by a gradient from 95 to 25 °C at -1 °C min^{-1}, then to 10 °C and held. For each transformation, the gRNA expression vector containing the gRNA sequence of choice, Cas9 encoding vector and the double-stranded DNA cassette for repair of the double strand break were combined and co-transformed into the BY4741 strain; this constructed strain was named as strain ZH0.

The squalene synthase (SQS) gene and the catalytic domain of HMG1 (tHMG1) were amplified from *S. cerevisiae* BY4741 genomic DNA and then inserted into the pESC-leu expression vector. All the primers used in this experiment are shown in Table S1. After verification by sequencing, this pESC-SQS+tHMG1 vector was transformed into the strain ZH0 to obtain strain ZH1. All of the strains used in this study are shown in Table S2. The vector pYES2-*Tw*OSC1^{T502E} was transformed into strain ZH1, strain BY4741 and a lanosterol synthase-deficient strain. Then, the positive strains were cultured in flasks (300 mL) containing 60 mL of YPD medium at 30 °C and 220 rpm for 2 days. Then, the cells were collected and induced in 60 mL of YPG medium at 30 °C and 200 rpm for 60 h. The products of the yeast were extracted and analysed as described above.

The high-yield strain containing pYES2-*Tw*OSC1^{T502E} was further cultured in flasks (300 mL) containing 30 mL of optimized YPD medium (5% glucose, 1% yeast extract, 3% peptone, 0.8% KH_2PO_4, and 0.6% $MgSO_4$) which was proven to effectively promote rapid propagation of yeast cells in a previous study (Song *et al.*, 2017). After culturing at 30 °C and 220 rpm for 2 days, the cells were collected and induced in 30 mL of optimized YPG medium (5% galactose, 1% yeast extract, 3% peptone, 0.8% KH_2PO_4, and 0.6% $MgSO_4$) at 30 °C and 200 rpm for 60 h. The products of the yeast were extracted and analysed as described above.

Results

Cloning and sequence analysis of the *Tw*OSCs

In our previous study, we sequenced and reported a transcriptomic library (SRA accession number: SRR6001265) of *T. wilfordii* (Su *et al.*, 2017). Transcripts of OSC candidates were screened via homology-based searches of the transcriptome. Three full-length cDNA sequences of OSCs (*Tw*OSC1, *Tw*OSC2 and *Tw*OSC3) were obtained by polymerase chain reaction and sequencing (GenBank accession numbers KY885467, KY885468 and KY885469). Protein domain analysis showed that all the proteins have a terpenoid cyclase/protein prenyltransferase alpha-alpha toroid domain, but only *Tw*OSC2 has the conserved site of terpene synthase (Fig. S4). Based on the multiple sequence alignments, all three *Tw*OSCs were homologous (average similarity identified as 76.46%) to OSCs in several other species (Fig. S5).The three *Tw*OSCs all have the conserved region DCTAE of the OSC family (Abe & Prestwich, 1995), which is associated with substrate binding. The conserved repetitive QW (QXXXXXW) motifs always appear as four to eight repeats in the OSC family (Haralampidis *et al.*, 2002) and appear as four repeats in the *Tw*OSCs. In addition, the *Tw*OSC2 and *Tw*OSC3 both have the MWCYCR motif, which is predicted to be a highly conserved motif of *β*-amyrin synthase (Tetsuo Kushiro *et al.*, 2000).

Phylogenetic tree construction

The results of the phylogenetic analysis showed that *Tw*OSC2 was different from *Tw*OSC1 and *Tw*OSC3, which clustered with the previously characterized *β*-amyrin synthase genes *Ks*bAS, *Hh*bAS and *Ae*bAS (Fig. 3). In contrast, *Tw*OSC1 and *Tw*OSC3 clustered with *Mi*FRSs and *Pd*FRS into a clade. As previously described (Souzamoreira *et al.*, 2016; Han *et al.*, 2018), both *Mi*FRSs and *Pd*FRS are monofunctional friedelin synthases, that can produce friedelin only. Interestingly, there is also a multi-product friedelin synthase, *Kd*FRS (Wang *et al.*, 2010), which can produce friedelin as a major product, taraxerol and *β*-amyrin as a minor product. This multi-product friedelin synthase did not cluster with other friedelin synthases but formed a branch with glutinol and lupeol synthases from *K. daigremontiana* (*Kd*GLS and

*Kd*LUS, respectively). These results indicated that *Tw*OSC1 and *Tw*OSC3 were friedelin synthases, while *Tw*OSC2 was a β-amyrin synthase.

The transcript levels of the *Tw*OSCs were significantly up-regulated after MeJA treatment

MeJA is a plant hormone used in plant defense and many diverse developmental pathways, which can increase the levels of plant secondary metabolites by inducing the expression of genes in the biosynthetic pathways of secondary metabolites (Martin *et al.*, 2002). In our previous work, we found that MeJA can induce MVA pathway gene expression and accumulation of celastrol in *T. wilfordii* suspension cells (Liu *et al.*, 2016). Here, we used real-time PCR to quantify *Tw*OSCs expression after MeJA treatment. We found that the relative expression levels of all three *Tw*OSCs were significantly up-regulated which correlated with the accumulation of celastrol (98.16 μg g^{-1} in 360 h) after MeJA treatment (Fig. 4a, b). This result showed that all three *Tw*OSCs may involve in the biosynthesis of triterpene in *T. wilfordii*.

The tissue expression patterns of the *Tw*OSCs were correlated with the celastrol distribution

To determine whether the gene expression levels of the *Tw*OSCs in *T. wilfordii* were correlated with celastrol accumulation, the relative transcript levels of the *Tw*OSCs and the celastrol levels were determined across all established tissues (Fig.4c, d and Fig. S6). The relative expression levels of *Tw*OSC2 and *Tw*OSC3 were highest in the root, which was also the tissue with the highest celastrol content (1.87 mg g^{-1} dry weight). The transcript level of *Tw*OSC1 was highest in the leaves, which also exhibited low celastrol levels.

Functional characterization of *Tw*OSCs in yeast

To identify the functions of the putative OSCs from *T. wilfordii*, the complete ORFs of all the *Tw*OSCs were sub-cloned into the vector pYES2 and then transformed into lanosterol synthase-deficient yeast. In parallel, the empty vector was transformed into the yeast mutant to serve as a negative control. After culturing and induction, the metabolite extracts were extracted and analysed by GC-MS. Based on the results of GC analysis (Fig. 5), the extracts from yeast harboring *Tw*OSC2 contained only one product, which was confirmed to be *β*-amyrin. Yeast harboring *Tw*OSC1 and *Tw*OSC3 contained friedelin, *β*-amyrin and *α*-amyrin, with friedelin as the major product. The structures of all these products were explained by comparison of mass spectral characteristics with those of authentic standards (Fig. S7, S8 and S9). This result from GC-MS analysis suggested that *Tw*OSC1 and *Tw*OSC3 were multi-product friedelin synthases and that *Tw*OSC2 was a *β*-amyrin synthase.

RNAi identified candidate *Tw*OSCs for biosynthesis of celastrol

To further investigate the *Tw*OSC responsible for the biosynthesis of celastrol, *Tw*OSC1, *Tw*OSC2 and *Tw*OSC3 were selected as candidate genes for the RNA interference (RNAi) study. In our early work, a stable and highly efficient transformation system suitable for *T. wilfordii* cell suspensions was developed based on particle bombardment (Zhao *et al.*, 2017). Here, we used this transformation system to transform the candidate genes. pK7GWIWG2D(II)-*Tw*OSC1, -*Tw*OSC2, -*Tw*OSC3 and the empty vector were separately transformed into *T. wilfordii* cells and screened on MS solid medium with kanamycin. *Tw*OSC1 expression was suppressed by 42.5% and celastrol levels decreased by 32.0% compared with the control. *Tw*OSC3 expression was suppressed by 37.3%, and celastrol levels decreased by 84.2%. In contrast, RNAi suppressed the transcription of *Tw*OSC2 by 47% and increased celastrol levels by 9.5% (Fig. 6a, b, c). RNAi combined with gene expression analysis and UPLC analysis of celastrol profiles showed that *Tw*OSC1 and *Tw*OSC3 were involved in the biosynthesis of celastrol and that *Tw*OSC3 may have a stronger effect than *Tw*OSC1 on celastrol biosynthesis.

Friedelin is a precursor of celastrol

Based on the results of the RNAi study and the observed tissue expression pattern, we determined that *Tw*OSC1 and *Tw*OSC3 were responsible for the biosynthesis of celastrol. The GC-MS analysis showed that both *Tw*OSC1 and *Tw*OSC3 can produce friedelin, which is also a friedelane-type triterpene, similar to celastrol. To further investigate the potential role of friedelin in celastrol biosynthesis, we fed friedelin to suspension cell cultures, which led to a significant increase in celastrol accumulation compared with the control groups after 7 days (Fig. 6d). This finding indicated that friedelin did serve as a precursor to celastrol.

Site-directed mutagenesis of *Tw*OSC proteins

The reaction catalysed by OSCs is one of the most complicated reactions, and can yield a wide array of triterpene skeletons derived from the same simple substrate 2, 3-oxidosqualene. However, the exact mechanism underlying this product diversity remains unknown. Enzymes with a single product and high yield will be helpful for exploration of the subsequent biosynthesis pathway and further commercial application. Thus, we wanted to identify the crucial amino acids that are involved in the production of friedelin. Recently, Souza-Moreira *et al.* found that the amino acid two positions upstream of the DCTAE site in friedelin synthase, usually a Leucine (L) residue, was required for friedelin production (Souzamoreira *et al.*, 2016). Multiple sequence alignment of the friedelin-producing OSCs and molecular docking of *Tw*OSC1 and *Tw*OSC3 showed that the L residue was unique to friedelin-producing OSCs and was near the active sites of these enzymes (Fig. 7a, b).

To further investigate this residue in friedelin synthase, we constructed additional mutants by site-directed mutagenesis. Substitution of the L residue with phenylalanine (F), histidine (H), proline (P) or arginine (R) in *Tw*OSC1 and substitution of the L residue with F, P, R, or alanine (A) in *Tw*OSC3 abolished all the products, but substitution of the L residue with valine (V) decreased the production of friedelin and increased the production of *β*-amyrin. Substitution of the L residue with Serine (S) in *Tw*OSC3 also abolished the production of friedelin and increased the production of *β*-amyrin (Fig. S10 and S11).

Moreover, substitution of the L residue with isoleucine (I) in *Tw*OSC1 and *Tw*OSC3 did not impact friedelin production much but increase the production of *β*-amyrin. In addition, we also constructed a mutant of *Tw*OSC2 to determine whether substitution of the V residue at the same position with L could change the product to friedelin. However, this substitution in *Tw*OSC2 did not lead to friedelin production and decreased the yield of *β*-amyrin. Fortunately, we found a residue in *Tw*OSC1 (Threonine (T) 502) that can increase the production of friedelin. The best mutant was T502E (Glutamic acid), which produced 10.63 mg L^{-1} friedelin in a shake flask culture. The levels of all the products of the *Tw*OSCs and mutants are shown in Fig. 7c and Table S3.

Improvement of the yield of friedelin in engineered *S. cerevisiae*

To improve the yield of friedelin, two key genes in the triterpene pathway were overexpressed to enhance the metabolic flux towards (3S)-2, 3-oxidosqualene. ERG9 and tHMG1 from *S. cerevisiae* were separately cloned into the pESC-leu vector then transformed into the yeast BY4741. In addition to overexpressing the key genes in the triterpene pathway, three genes were deleted to further improve the production of friedelin (Fig. 8a). These genes included a transcriptional regulator, *rox1* and two other genes, namely, *ypl062w* and *yjl064w. Rox1* was reported to repress genes in the MVA pathway and ergosterol biosynthesis while *ypl062w* and *yjl064w* could improve MVA pathway flux when knocked out and especially when combined (Özaydın *et al.*, 2013). When pYES2-*Tw*OSC1^{T502E} was transformed into strain ZH1, strain BY4741 and the lanosterol synthase-deficient strain cultured under the same conditions, the highest yield was 22.52 mg L^{-1} in strain ZH1, which was higher than yield in the lanosterol synthase-deficient strain and almost two-fold higher than the yield in the parent strain BY4741(Fig. 8b). After medium optimization, the highest yield of strain ZH1 could be improved to 37.07 mg L^{-1}. This is the first study on the biosynthesis of friedelin in engineered yeast with a high yield.

Discussion

T. wilfordii is known to contain a diversity of triterpenes, and celastrol was one of the valuable compounds isolated from *T. wilfordii*. Here, we focused on the OSCs that were responsible for the main scaffold cyclization of celastrol. Three OSCs were mined from the transcriptome of *T. wilfordii*. To identify the *Tw*OSC that was responsible for the biosynthesis of celastrol, we tested the expression levels of the OSCs after MeJA treatment and determined the tissue expression patterns of the OSCs in *T. wilfordii*. However, the results showed that all three *Tw*OSCs could be induced by MeJA and the tissue expression patterns of these OSCs, except that of *Tw*OSC1, were correlated with the distribution of celastrol. In a previous study, Corsino *et al.* found that triterpenes, once biosynthesized in the leaves, were translocated to the root bark and further transformed to quinonemethide triterpenoids in the Celastraceae and Hippocrateaceae families (Corsino *et al.*, 2000), which indicated that *Tw*OSC1 may also participate in the biosynthesis of triterpenes in the root. Subsequently, functional identification of *Tw*OSCs in lanosterol synthase-deficient yeast revealed that *Tw*OSC1 and *Tw*OSC3 were multi-product friedelin synthases that could produce friedelin as major product, and β-amyrin and α-amyrin as minor products. This is the first report of simultaneous production of friedelin, β-amyrin and α-amyrin by friedelin synthase. This study also showed that *Tw*OSC2 is a β-amyrin synthase with only one product, β-amyrin under the experimental conditions. This finding indicated that the common product of both *Tw*OSC1 and *Tw*OSC2 was β-amyrin, belonging to the oleanane-type triterpene family and not the friedelane-type triterpene family. To further identify the roles of the *Tw*OSCs, we performed an RNAi study to suppress the gene expression of the *Tw*OSCs. Based on the results of the RNAi study, we determined that decreased expression of *Tw*OSC1 and *Tw*OSC3 can cause a decrease in celastrol levels. This finding indicated that *Tw*OSC1 and *Tw*OSC3 were the genes responsible for the biosynthesis of celastrol. The feeding study further confirmed that the common product of *Tw*OSC1 and *Tw*OSC3, friedelane-type friedelin, is one precursor of celastrol, and can increase the production of celastrol.

The mechanistic diversity of OSCs remains intriguing. Although many attempts have been made to study the underlying mechanism (Hart *et al.*, 1999; Herrera *et al.*, 2000; Tetsuo Kushiro *et al.*, 2000; Liu *et al.*, 2012; Salmon *et al.*, 2016; Xue *et al.*, 2018), the mechanism of conversion between scaffolds remains poorly understood. Sequence analysis showed that all three *Tw*OSCs contained terpenoid cyclase/protein prenyltransferase alpha-alpha toroid domains, but only *Tw*OSC2 contained conserved sites of terpene synthase, which is rich in aromatic residues, in the C-terminal region. Multiple sequence alignments showed that all these protein contained the conserved DCTAE motif and four conserved repetitive QW motifs. Moreover, *Tw*OSC2 and *Tw*OSC3 both had the MWCYCR motif, which is important for *β*-amyrin synthase. However, we found that *Tw*OSC1 (containing the IWCYCR motif) as well as *Tw*OSC2 and *Tw*OSC3 could produce *β*-amyrin. Tetsuo *et al.* found that the Tyrptophan residue in the MWCYCR motif was responsible for the formation of pentacyclic triterpenes and that substitution of this residue with H resulted in the production of tetracyclic dammaradienols (Tetsuo Kushiro *et al.*, 2000). Here, our results suggested that the methionine residue in the MWCYCR motif was not necessary for the production of *β*-amyrin.

To identify the key residues involved in functional conversion between friedelin and *β*-amyrin, we performed site-directed mutagenesis. Our results showed that substitution of L residue (two positions upstream of the DCTAE motif) with F, H, P or R in *Tw*OSC1 and substitution of the L residue with F, P, R, or A in *Tw*OSC3 abolished all the products. From the docking result, we thought that the steric hindrance may be the main effect to prevent the reaction in these substitutions (Fig. S12). In contrast, substitution of the L residue with V in *Tw*OSC1 and *Tw*OSC3 decreased the production of friedelin and increased the production of *β*-amyrin, and substitution of the L residue with S in *Tw*OSC3 also abolished the production of friedelin but increased the production of *β*-amyrin. Substitution of the L residue with I in *Tw*OSC1 and *Tw*OSC3 did not impact friedelin production much but increase the production of *β*-amyrin. From the proposed oxidosqualene cyclization mechanisms, we thought that the product diversity was depending on the ability to stabilize the cations. Thus, we thought that these substitutions may make the enzyme stabilize oleanyl cation better and make the reaction blocked at oleanyl cation and turn to generate *β*-amyrin. The substitution of the V

residue with L in *Tw*OSC2 did not lead to friedelin production and decreased the yield of *β*-amyrin, which also confirmed this inference. This result strongly indicated that the L residue was important for the biosynthesis of friedelin but not the only factor influencing the production of friedelin. In the other site of *Tw*OSC1 (T502), we found that substitution of T with E, I or Lysine (K) increased the production of friedelin and the E residue led to the greatest increase. In a previous study, Thoma *et al.* found that human OSC consists of two (*α*/*α*) barrel domains that are connected by loops and three smaller *β*-structures (Thoma *et al.*, 2004). The substrate first entered the active site via channel 1 and then formed the product in the active site. Then, the product was separated from the enzyme by channel 2. The structural model of *Tw*OSC1 showed that residue T502 was located in channel 2 (Fig. S13); thus we speculated that T502 mutants increased the yield of friedelin by influencing the release of friedelin from the enzyme.

Biosynthesis of friedelin in yeast with high yields will help reveal the biosynthetic pathway of celastrol and further commercial application. To improve the yield of friedelin in yeast, we constructed an engineered strain by knocking out three down-regulated genes and overexpressing two key genes in the triterpene biosynthetic pathway. When the *Tw*OSC1 mutant T502E, with the highest yield, was transformed into the three different strains, strain ZH1 showed the best production capacity with a yield of 22.52 mg L^{-1} friedelin, which was almost two-fold higher than that obtained with the parent strain. The lanosterol synthase-deficient yeast was diploid yeast, which lacked one copy of the ERG1 gene so that it could increase the flux of triterpenoid productions by improving the accumulation of the common precursor, 2, 3-oxidosqualene. Here, we also tested the production of friedelin in this strain. The friedelin yield in this strain was 19.03 mg L^{-1}, which was lower than that in strain ZH1. Moreover, we further improved the yield of friedelin in strain ZH1 by optimizing the fermentation medium. By a combination of mutagenesis, strain construction and medium optimization, the highest yield of friedelin was determined to be 37.05 mg L^{-1} in a shake flask culture, which was more than five-fold the yield observed for wild-type of *Tw*OSC1 (7.35 mg L^{-1}).

Acknowledgements

This work was supported by the National Natural Science Foundation of China (81773830), Beijing Natural Science Foundation Program and Scientific Research Key Program of Beijing Municipal Commission of Education (KZ201710025022) and Support Project of High-level Teachers in Beijing Municipal Universities in the Period of 13th Five-year Plan (CIT&TCD20170324) and the National Program for Special Support of Eminent Professionals. We thank Yuru Tong from Shenyang Pharmaceutical University for the analysis of docking results. We also acknowledge Beijing University of Chinese Medicine for supporting SYBYL-based computing services.

Author contributions

J.Z., L.H. and W.G planned and designed the research. J.Z., T.H., L.G., Yifeng Zhang., L.T and S.C. performed the research. P.S., Yujun Zhao., X.W. and J.Z. collected and analysed the data. J.Z., Y.S. and T.H. wrote the manuscript.

ORCID

Jiawei Zhou iD: https://orcid.org/0000-0001-6628-3056

References

Abe I, Prestwich GD. 1995. Identification of the active site of vertebrate oxidosqualene cyclase. *Lipids* **30**: 231-234.

Allison AC, Cacabelos R, Lombardi VR, Alvarez XA, Vigo C. 2001. Celastrol, a potent antioxidant and anti-inflammatory drug, as a possible treatment for Alzheimer's disease. *Prog Neuropsychopharmacol Biol Psychiatry* **25**: 1341-1357.

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**: 3389-3402.

Astry B, Venkatesha SH, Laurence A, Christensen-Quick A, Garzino-Demo A, Frieman MB, O'Shea JJ, Moudgil KD. 2015. Celastrol, a Chinese herbal compound, controls autoimmune inflammation by altering the balance of pathogenic and regulatory T cells in the target organ. *Clin Immunol* **157**: 228-238.

Camelio AM, Johnson TC, Siegel D. 2015. Total Synthesis of Celastrol, Development of a Platform to Access Celastroid Natural Products. *J Am Chem Soc* **137**: 11864-11867.

Corsino J, de Carvalho PR, Kato MJ, Latorre LR, Oliveira OM, Araujo AR, Bolzani VD, Franca SC, Pereira AM, Furlan M. 2000. Biosynthesis of friedelane and quinonemethide triterpenoids is compartmentalized in *Maytenus aquifolium* and *Salacia campestris*. *Phytochemistry* **55**: 741-748.

Corson TW, Crews CM. 2007. Molecular understanding and modern application of traditional medicines: triumphs and trials. *Cell* **130**: 769-774.

Crooks GE, Hon G, Chandonia JM, Brenner SE. 2004. WebLogo: a sequence logo generator. *Genome Res* **14**: 1188-1190.

Duan H, Takaishi Y, Imakura Y, Jia Y, Li D, Cosentino LM, Lee KH. 2000a. Sesquiterpene alkaloids from *Tripterygium hypoglaucum* and *Tripterygium wilfordii*: a new class of potent anti-HIV agents. *J Nat Prod* **63**: 357-361.

Duan H, Takaishi Y, Momota H, Ohmoto Y, Taki T, Jia Y, Li D. 2000b. Triterpenoids from *Tripterygium wilfordii*. *Phytochemistry* **53**: 805-810.

Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, Eramian D, Shen MY, Pieper U, Sali A. 2006. Comparative protein structure modeling using Modeller. *Curr Protoc Bioinformatics* **Chapter 5**: Unit-5.6.

Forman V, Callari R, Folly C, Heider H, Hamberger B. 2017. Production of Putative Diterpene Carboxylic Acid Intermediates of Triptolide in Yeast. *Molecules* **22**: 981.

Galanie S, Thodey K, Trenchard IJ, Filsinger Interrante M, Smolke CD. 2015. Complete biosynthesis of opioids in yeast. *Science* **349**: 1095-1100.

Guan H, Zhao Y, Su P, Tong Y, Liu Y, Hu T, Zhang Y, Zhang X, Li J, Wu X, et al. 2017. Molecular cloning and functional identification of sterol C24-methyltransferase gene from *Tripterygium wilfordii*. *Acta Pharm Sin B* **7**: 603-609.

Han JY, Ahn C-H, Adhikari PB, Kondeti S, Choi YE. 2019. Functional characterization of an oxidosqualene cyclase (*PdFRS*) encoding a monofunctional friedelin synthase in *Populus davidiana*. *Planta* **249**: 95-111.

Hansen NL, Heskes AM, Hamberger B, Olsen CE, Hallstrom BM, Andersen-Ranberg J, Hamberger B. 2017. The terpene synthase gene family in *Tripterygium wilfordii* harbors a labdane-type diterpene synthase among the monoterpene synthase TPS-b subfamily. *Plant J* **89**: 429-441.

Haralampidis K, Trojanowska M, Osbourn AE. 2002. Biosynthesis of triterpenoid saponins in plants. *Adv Biochem Eng Biotechnol* **75**: 31-49.

Hart EA, Ling H, Darr LB, Wilson WK, Jihai Pang A, Matsuda SPT. 1999. Directed Evolution To Investigate Steric Control of Enzymatic Oxidosqualene Cyclization. An Isoleucine-to-Valine Mutation in Cycloartenol Synthase Allows Lanosterol and Parkeol Biosynthesis. *J.am.chem.soc* **121**: 9887-9888.

Herrera JBR, And WKW, Matsuda SPT. 2000. A Tyrosine-to-Threonine Mutation Converts Cycloartenol Synthase to an Oxidosqualene Cyclase that Forms Lanosterol as Its Major Product. *J Am Chem Soc* **122**: 6765-6766.

Inabuy FS, Fischedick JT, Lange I, Hartmann M, Srividya N, Parrish AN, Xu M, Peters RJ, Lange BM. 2017. Biosynthesis of Diterpenoids in Tripterygium Adventitious Root Cultures. *Plant Physiol* **175**: 92-103.

Jiang X. 1994. Clinical observations on the use of the Chinese herb *Tripterygium wilfordii* Hook for the treatment of nephrotic syndrome. *Pediatr Nephrol* **8**: 343-344.

Lau W, Sattely ES. 2015. Six enzymes from mayapple that complete the biosynthetic pathway to the etoposide aglycone. *Science* **349**: 1224-1228.

Li K, Duan H, Kawazoe K, Takaishi Y. 1997. Terpenoids from *Tripterygium wilfordii*. *Phytochemistry* **45**: 791-796.

Li LF. 2000. Treatment of pyoderma gangrenosum with oral *Tripterygium wilfordii* multiglycoside. *J Dermatol* **27**: 478-481.

Liu J, Lee J, Salazar Hernandez MA, Mazitschek R, Ozcan U. 2015. Treatment of obesity with celastrol. *Cell* **161**: 999-1011.

Liu YJ, Zhao YJ, Su P, Zhang M, Tong YR, Hu TY, Huang LQ, Gao W. 2016. The MVA pathway genes expressions and accumulation of celastrol in *Tripterygium wilfordii* suspension cells in response to methyl jasmonate treatment. *J Asian Nat Prod Res* **18**: 619-628.

Liu YT, Hu TC, Chang CH, Shie WS, Wu TK. 2012. Protein engineering of *Saccharomyces cerevisiae* oxidosqualene-lanosterol cyclase into parkeol synthase. *Org Lett* **14**: 5222-5225.

Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**: 402-408.

Ma J, Dey M, Yang H, Poulev A, Pouleva R, Dorn R, Lipsky PE, Kennelly EJ, Raskin I. 2007. Anti-inflammatory and immunosuppressive compounds from *Tripterygium wilfordii*. *Phytochemistry* **68**: 1172-1178.

Ma X, Xu L, Alberobello AT, Gavrilova O, Bagattin A, Skarulis M, Liu J, Finkel T, Mueller E. 2015. Celastrol Protects against Obesity and Metabolic Dysfunction through Activation of a HSF1-PGC1alpha Transcriptional Axis. *Cell Metab* **22**: 695-708.

Mans R, Rossum HMV, Wijsman M, Backx A, Kuijpers NGA, Broek MVD, Daran-Lapujade P, Pronk JT, Maris AJAV, Daran JMG. 2015. CRISPR/Cas9: a molecular Swiss army knife for simultaneous introduction of multiple genetic modifications in *Saccharomyces cerevisiae*. *FEMS Yeast Res* **15**: fov004.

Martin D, Tholl D, Gershenzon J, Bohlmann J. 2002. Methyl jasmonate induces traumatic resin ducts, terpenoid resin biosynthesis, and terpenoid accumulation in developing xylem of Norway spruce stems. *Plant Physiol* **129**: 1003-1018.

Miao GP, Li W, Zhang B, Zhang ZF, Ma ZQ, Feng JT, Zhang X, Zhu CS. 2015. Identification of Genes Involved in the Biosynthesis of *Tripterygium wilfordii* Hook.f. Secondary Metabolites by Suppression Subtractive Hybridization. *Plant Molecular Biology Reporter* **33**: 756-769.

Miao K, Zhang X, Dong Y. 2000. Studies on triterpenoids constituents of *Tripterygium wilfordii* Hook. F. *Natural Product Research & Development* **12**: 1-7.

Özaydın B, Burd H, Lee TS, Keasling JD. 2013. Carotenoid-based phenotypic screen of the yeast deletion collection reveals new genes with roles in isoprenoid production. *Metabolic Engineering* **15**: 174-183.

Paddon CJ, Westfall PJ, Pitera DJ, Benjamin K, Fisher K, McPhee D, Leavell MD, Tai A, Main A, Eng D, et al. 2013. High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature* **496**: 528-532.

Pang X, Yi Z, Zhang J, Lu B, Sung B, Qu W, Aggarwal BB, Liu M. 2010. Celastrol suppresses angiogenesis-mediated tumor growth through inhibition of AKT/mammalian target of rapamycin pathway. *Cancer Res* **70**: 1951-1959.

Ren J, Tao Q, Wang X, Wang Z, Li J. 2007. Efficacy of T2 in active Crohn's disease: a prospective study report. *Dig Dis Sci* **52**: 1790-1797.

Salmon M, Thimmappa RB, Minto RE, Melton RE, Hughes RK, O'Maille PE, Hemmings AM, Osbourn A. 2016. A conserved amino acid residue critical for product and substrate specificity in plant triterpene synthases. *Proc Natl Acad Sci U S A* **113**: E4407-4414.

Song TQ, Ding MZ, Zhai F, Liu D, Liu H, Xiao WH, Yuan YJ. 2017. Engineering *Saccharomyces cerevisiae* for geranylgeraniol overproduction by combinatorial design. *Sci Rep* **7**: 14991.

Souzamoreira TM, Alves TB, Pinheiro KA, Felippe LG, De Lima GM, Watanabe TF, Barbosa CC, Santos VA, Lopes NP, Valentini SR. 2016. Friedelin Synthase from *Maytenus ilicifolia*: Leucine 482 Plays an Essential Role in the Production of the Most Rearranged Pentacyclic Triterpene. *Sci Rep* **6**: 36858.

Su P, Cheng Q, Wang X, Cheng X, Zhang M, Tong Y, Li F, Gao W, Huang L. 2014. Characterization of eight terpenoids from tissue cultures of the Chinese herbal plant, *Tripterygium wilfordii*, by high-performance liquid chromatography coupled with electrospray ionization tandem mass spectrometry. *Biomed Chromatogr* **28**: 1183-1192.

Su P, Guan H, Zhang Y, Wang X, Gao L, Zhao Y, Hu T, Zhou J, Ma B, Tu L, et al. 2017. Probing the Single Key Amino Acid Responsible for the Novel Catalytic Function of ent-Kaurene Oxidase Supported by NADPH-Cytochrome P450 Reductases in *Tripterygium wilfordii*. *Front Plant Sci* **8**: 1756.

Su P, Guan H, Zhao Y, Tong Y, Xu M, Zhang Y, Hu T, Yang J, Cheng Q, Gao L, et al. 2018. Identification and functional characterization of diterpene synthases for triptolide biosynthesis from *Tripterygium wilfordii*. *Plant J* **93**: 50-65.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol* **30**: 2725-2729.

Tao X, Cush JJ, Garret M, Lipsky PE. 2001. A phase I study of ethyl acetate extract of the chinese antirheumatic herb *Tripterygium wilfordii* hook F in rheumatoid arthritis. *J Rheumatol* **28**: 2160-2167.

Tao X, Lipsky PE. 2000. The Chinese anti-inflammatory and immunosuppressive herbal remedy *Tripterygium wilfordii* Hook F. *Rheum Dis Clin North Am* **26**: 29-50, viii.

Tetsuo Kushiro, Masaaki Shibuya, Kazuo Masuda A, Yutaka Ebizuka. 2000. Mutational Studies on Triterpene Synthases: Engineering Lupeol Synthase into β-Amyrin Synthase. *J Am Chem Soc* **122**: 6816-6824.

Thimmappa R, Geisler K, Louveau T, O'Maille P, Osbourn A. 2014. Triterpene biosynthesis in plants. *Annu Rev Plant Biol* **65**: 225-257.

Thoma R, Schulz-Gasch T, D'Arcy B, Benz J, Aebi J, Dehmlow H, Hennig M, Stihle M, Ruf A. 2004. Insight into steroid scaffold formation from the structure of human oxidosqualene cyclase. *Nature* **432**: 118-122.

Wang Z, Yeats T, Han H, Jetter R. 2010. Cloning and characterization of oxidosqualene cyclases from *Kalanchoe daigremontiana*: enzymes catalyzing up to 10 rearrangement steps yielding friedelin and other triterpenoids. *Journal of Biological Chemistry* **285**: 29703-29712.

Wong KF, Yuan Y, Luk JM. 2012. *Tripterygium wilfordii* bioactive compounds as anticancer and anti-inflammatory agents. *Clin Exp Pharmacol Physiol* **39**: 311-320.

Xu G, Tu W, Jiang D, Xu C. 2009. *Tripterygium wilfordii* Hook F treatment for idiopathic refractory nephrotic syndrome in adults: a meta-analysis. *Nephron Clin Pract* **111**: c223-228.

Xu R, Fazio GC, Matsuda SP. 2004. On the origins of triterpenoid skeletal diversity. *Phytochemistry* **65**: 261-291.

Xue Z, Tan Z, Huang A, Zhou Y, Sun J, Wang X, Thimmappa RB, Stephenson MJ, Osbourn A, Qi X. 2018. Identification of key amino acid residues determining product specificity of 2,3-oxidosqualene cyclase in Oryza species. *New Phytol* **218**: 1076-1088.

Yang F, Dong XG, An ZM, Wu GQ, Zhang LJ, Gong NB, Gu TN. 2005. Retrospect and prospect of studies on *Tripterygium Wilfordii* Hook f. *Chin J Integr Med* **11**: 89-96.

Yang H, Chen D, Cui QC, Yuan X, Dou QP. 2006. Celastrol, a triterpene extracted from the Chinese "Thunder of God Vine," is a potent proteasome inhibitor and suppresses human prostate cancer growth in nude mice. *Cancer Res* **66**: 4758-4765.

Yang JH, Luo SD, Wang YS, Zhao JF, Zhang HB, Li L. 2006. Triterpenes from *Tripterygium wilfordii* Hook. *J Asian Nat Prod Res* **8**: 425-429.

Zhang DM, Yu DQ, Xie FZ. 1991. [The structure of tripterygone]. *Yao Xue Xue Bao* **26**: 341-344.

Zhao Y, Zhang Y, Su P, Yang J, Huang L, Gao W. 2017. Genetic Transformation System for Woody Plant *Tripterygium wilfordii* and Its Application to Product Natural Celastrol. *Front Plant Sci* **8**: 2221.

Zhou J, Zhang Y, Hu T, Su P, Zhang Y, Liu Y, Huang L, Gao W. 2018. Functional characterization of squalene epoxidase genes in the medicinal plant *Tripterygium wilfordii. Int J Biol Macromol* **120**: 203-212.

Zhou YJ, Gao W, Rong Q, Jin G, Chu H, Liu W, Yang W, Zhu Z, Li G, Zhu G, et al. 2012. Modular pathway engineering of diterpenoid synthases and the mevalonic acid pathway for miltiradiene production. *J Am Chem Soc* **134**: 3234-3241.

Zhou ZL, Yang YX, Ding J, Li YC, Miao ZH. 2012. Triptolide: structural modifications, structure-activity relationships, bioactivities, clinical development and mechanisms. *Nat Prod Rep* **29**: 457-475.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information tab for this article:

Figure Legends

Fig. 1 Proposed biogenetic pathway for celastrol in *Tripterygium wilfordii.* The red solid arrow indicates a biosynthetic reaction step identified in this work, and each dashed arrow indicates one or multiple proposed step reactions. All triterpenoid intermediates in the proposed pathway have been reported in *T. wilfordii.*

Fig. 2 The proposed oxidosqualene cyclization mechanisms. Friedelin is a most rearranged friedelane-type triterpene found in many plants including *Tripterygium wilfordii*.

Fig. 3 Phylogenetic tree of *Tw*OSCs and the characterized oxidosqualene cyclases (OSCs) from other species. The phylogenetic tree was created by MEGA 6.0and the neighbour-joining method. The bootstrap confidence values were obtained based on 1000 replicates. The *Tw*OSCs from *Tripterygium wilfordii* are highlighted with red diamond.

Fig. 4 Analysis of methyl jasmonate (MeJA)-induced and tissue expression of *Tw*OSCs combined with celastrol distribution. (a) Relative expression of *Tw*OSCs in the MeJA induced *Tripterygium wilfordii* suspension cells. CK: control group; MJ: MeJA treated group. (b) Celastrol content in MeJA-induced *T. wilfordii* suspension cells. CK: control group; MJ: MeJA treated group. (c) Relative expression of *Tw*OSCs in different *T. wilfordii* tissues. (d) *T. wilfordii* plant and distribution of celastrol in *T. wilfordii*. Error bars represent SDs based on triplicate measurements of at least four biological replicates.

Fig. 5 Gas chromatography-mass spectrometry (GC-MS) analysis of the products in yeast strains harboring the *Tw*OSCs from *Tripterygium wilfordii*. Standards: β-amyrin (1), α-amyrin (2), and friedelin (3).

Fig. 6 Analysis of RNA interference and feeding in *Tripterygium wilfordii* suspension cells. (a) The RNAi process. (b) Relative expression of *Tw*OSCs in the RNAi suspension cells and control suspension cells. (c) Celastrol content in the RNAi suspension cells and control suspension cells. (d) The content of celastrol in the friedelin-feeding group and control group

after 0 and 7days. Student's t-test was used to test statistically significant difference of increased celastrol levels between control group and feeding group. The data represent the average ± SD of at least four independent lines of suspension cells.

Fig. 7 Molecular docking and mutagenesis assay of *Tw*OSCs from *Tripterygium wilfordii*. (a) Sequence representation in the DCTAE motifs of *Tw*OSCs. Accession numbers and abbreviations are provided in Supporting Information Table S4. (b) Molecular docking of *Tw*OSC1 and *Tw*OSC3 with friedelin. (c) Product analysis of the mutants. The activities of wild-type and all the mutants are presented as the means ± SDs, $n = 5$.

Fig. 8 Strain construction and friedelin production using *Tw*OSC1^{T502E} from *Tripterygium wilfordii*. (a) A strategy to construct strain ZH1. (b) The levels of products in the three strains containing *Tw*OSC1^{T502E}. BY4741 stands for strain BY4741 containing *Tw*OSC1^{T502E}; GIL77 stands for the lanosterol synthase-deficient strain containing *Tw*OSC1^{T502E}; ZH1 stands for strain ZH1 containing *Tw*OSC1^{T502E}; and ZH1-MO stands for strain ZH1 containing *Tw*OSC1^{T502E} in the optimization medium. The data are presented as the means ± SDs, n = 3.

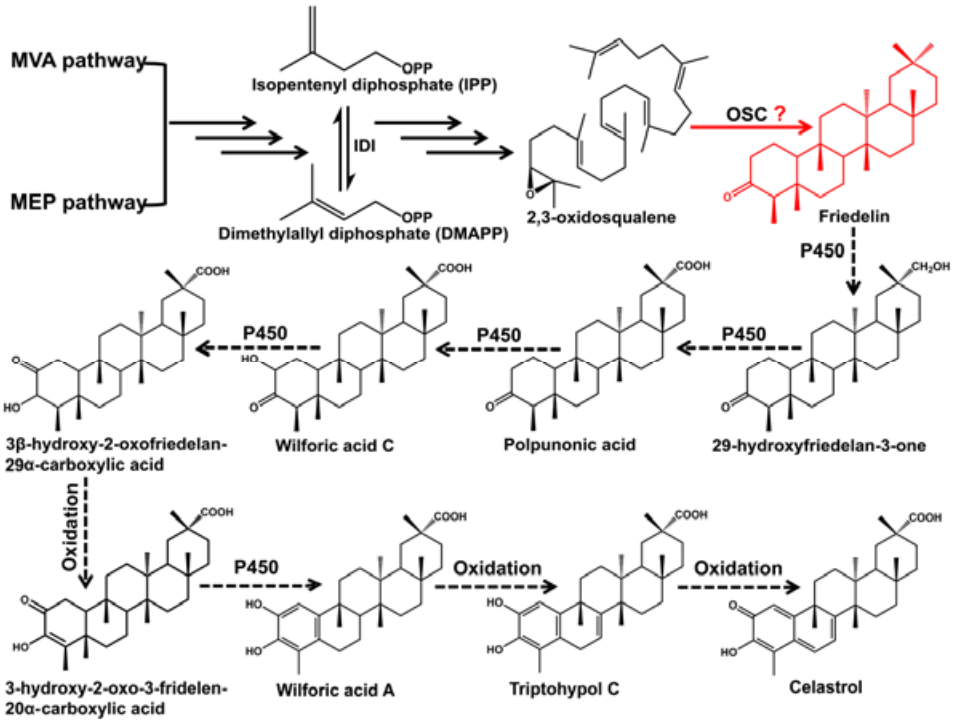

Fig. 1

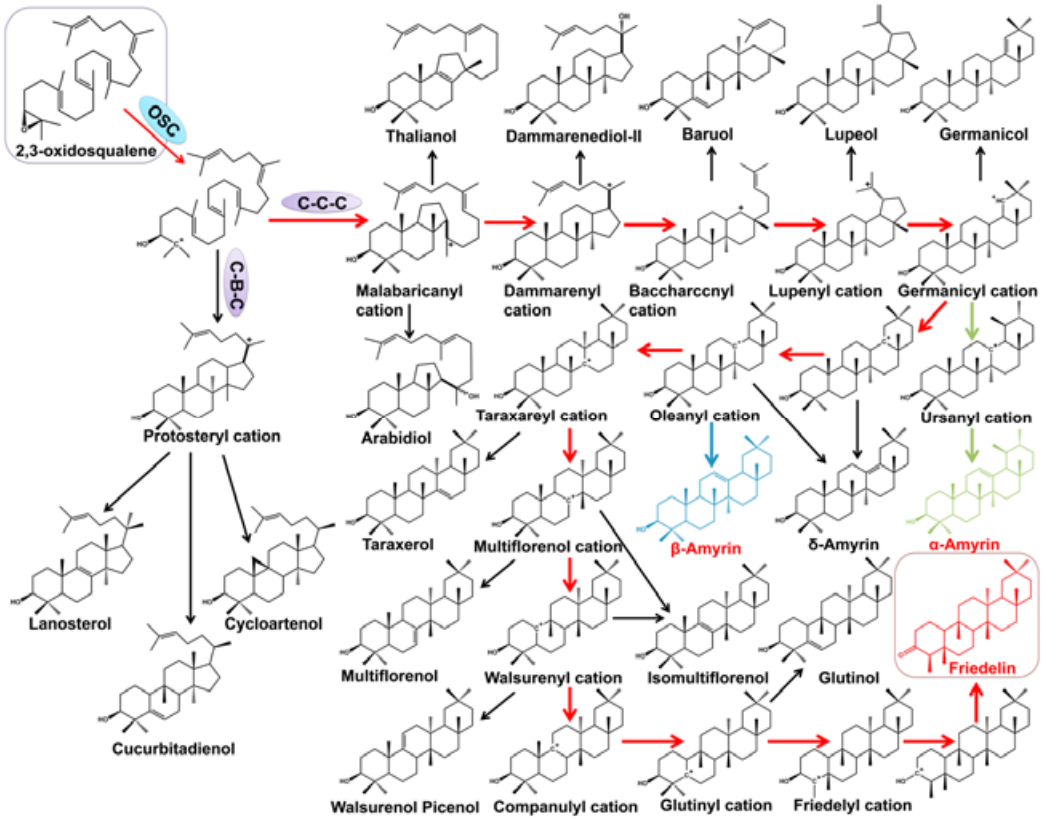

Fig. 2

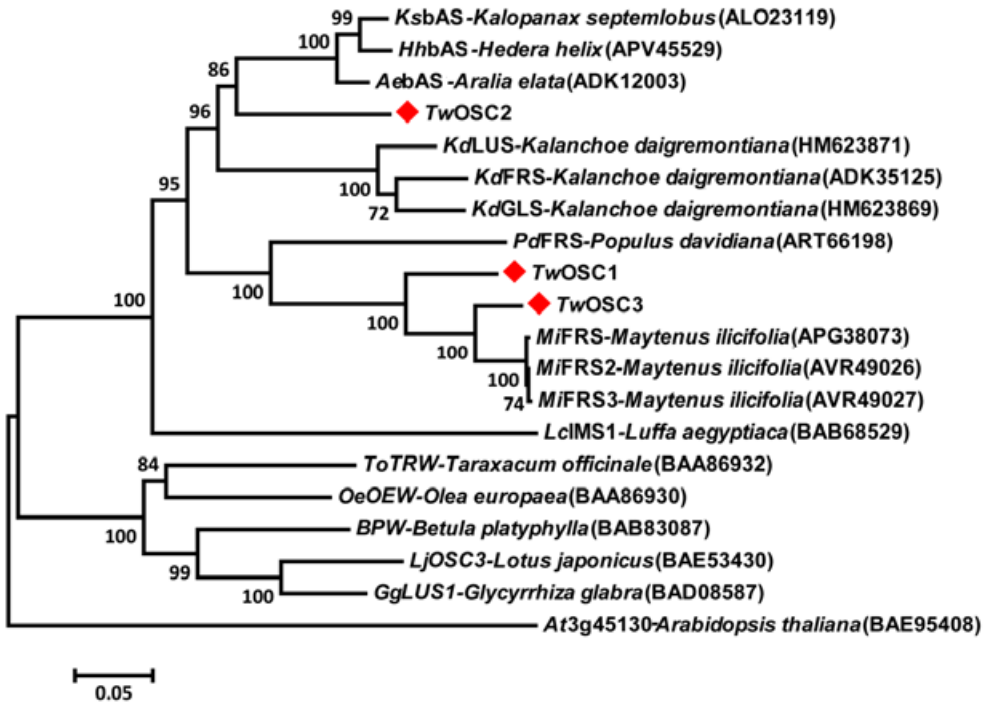

Fig. 3

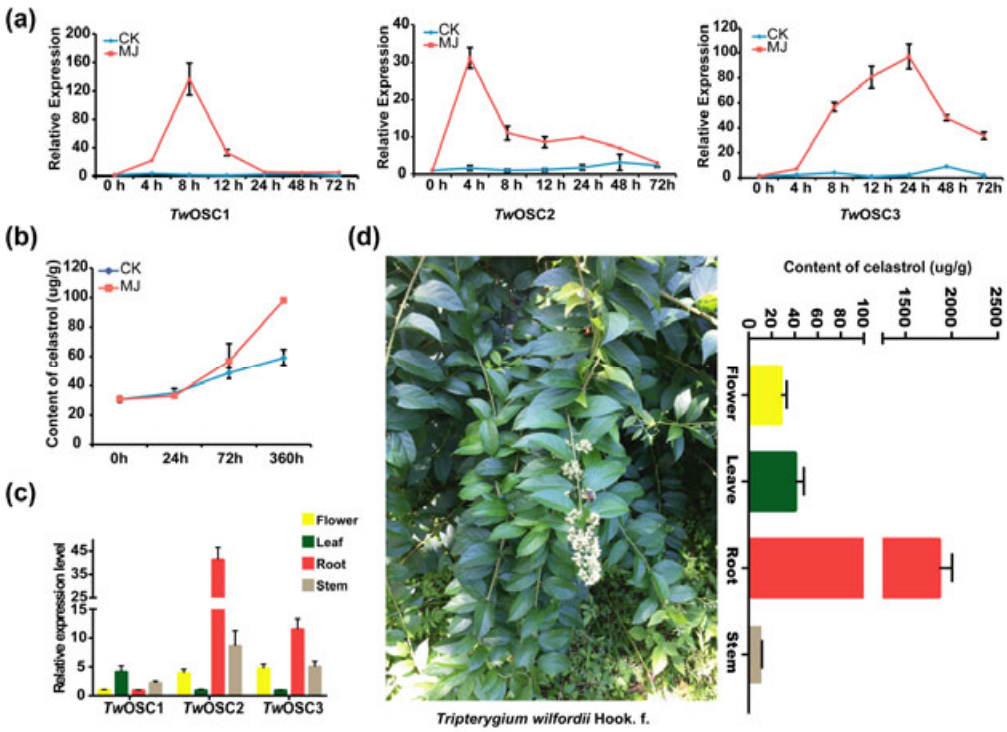

Fig. 4

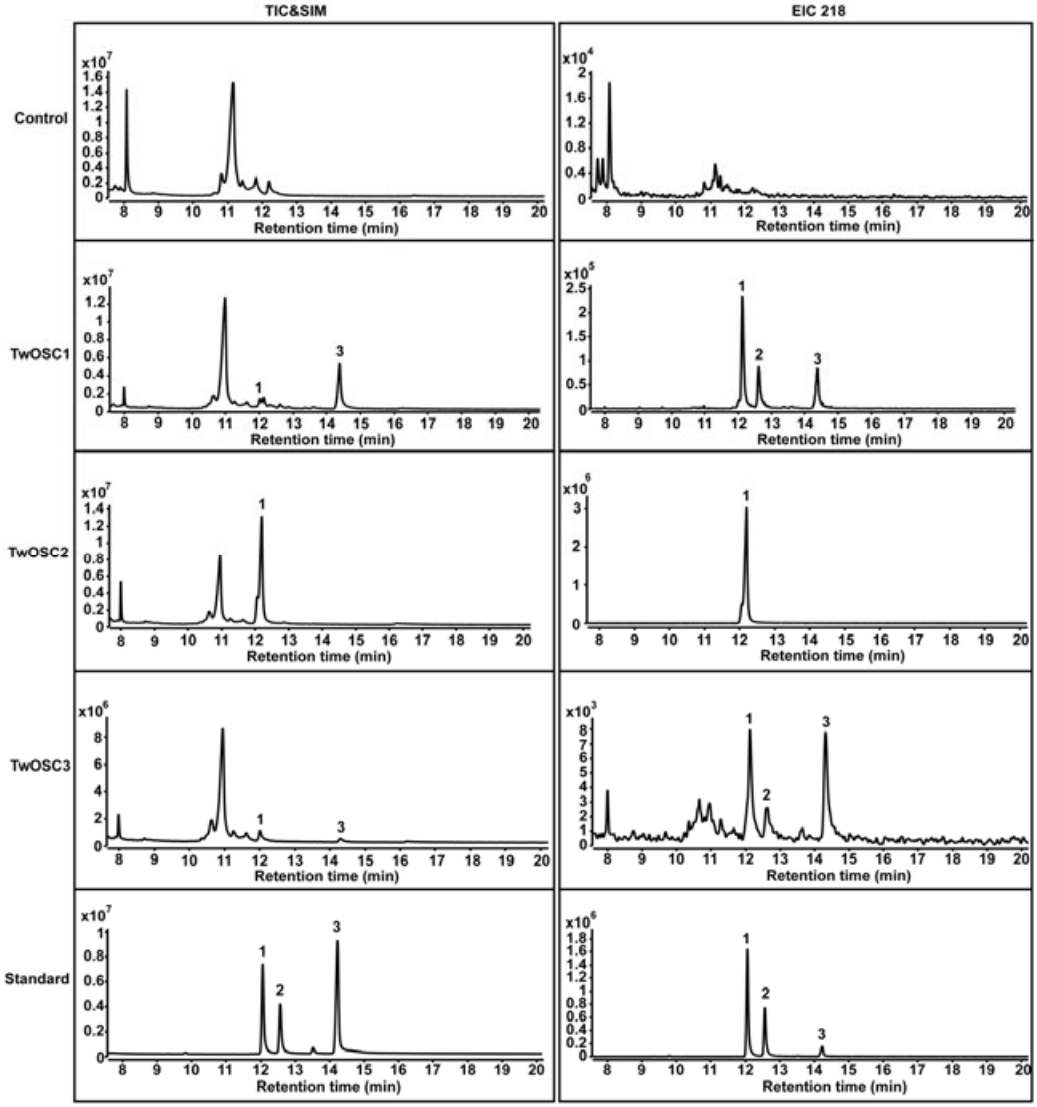

Fig. 5

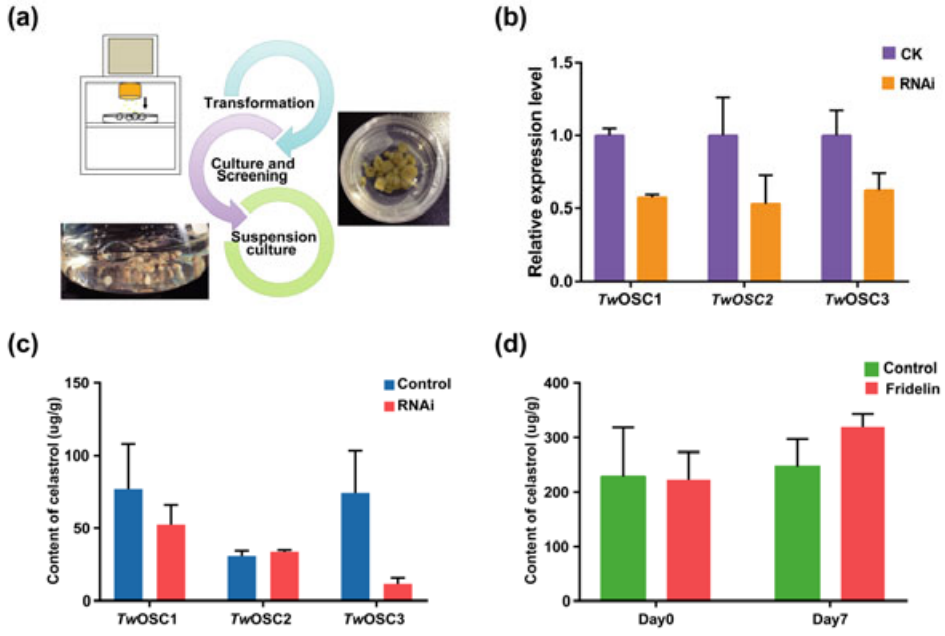

Fig. 6

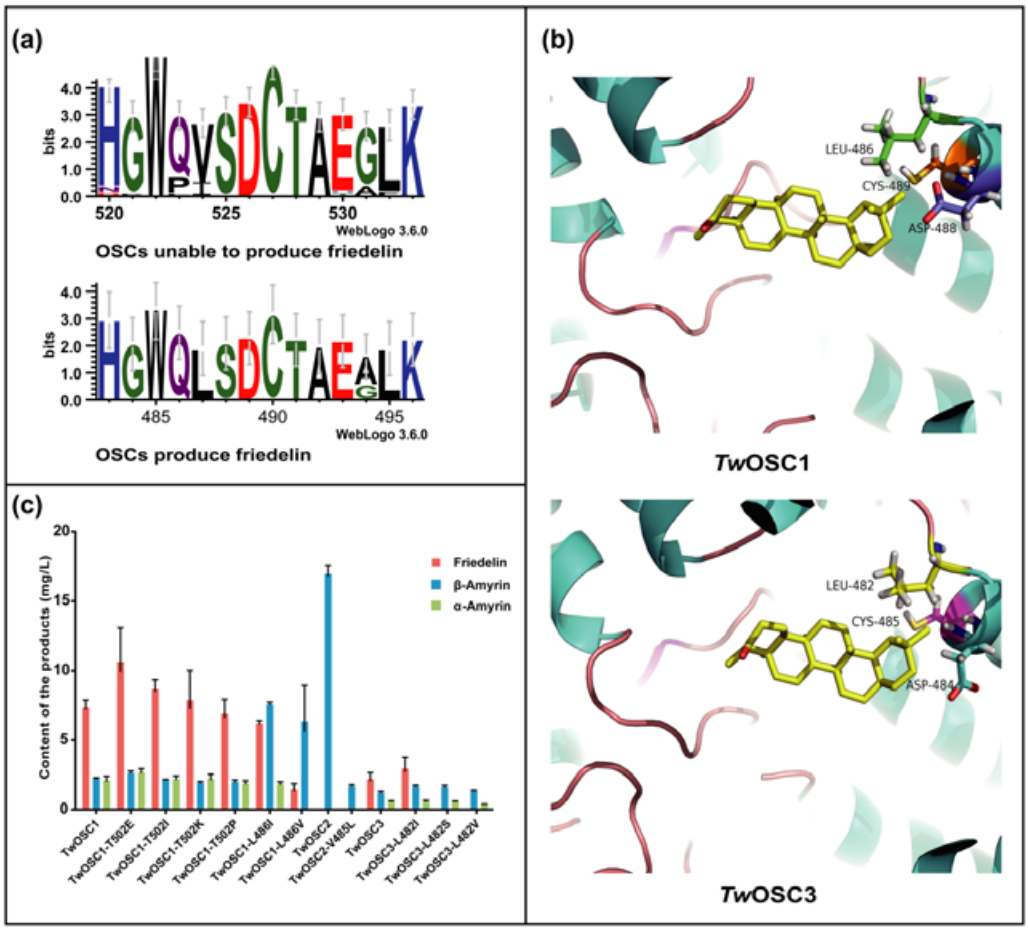

Fig. 7

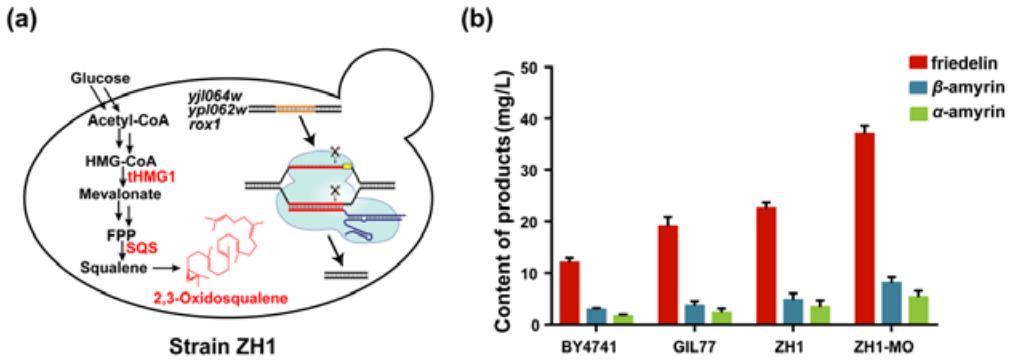

Fig. 8