

Two bi-functional cytochrome P450 CYP72 enzymes from olive (*Olea europaea*) catalyze the oxidative C-C bond cleavage in the biosynthesis of secoxy-iridoids – flavor and quality determinants in olive oil

Carlos E. Rodríguez-López¹ , Benke Hong¹, Christian Paetz² , Yoko Nakamura^{1,2}, Konstantinos Koudounas³ , Valentina Passeri⁴ , Luciana Baldoni⁴ , Fiammetta Alagna⁵ , Ornella Calderini⁴  and Sarah E. O'Connor¹ 

¹Department of Natural Product Biosynthesis, Max Planck Institute for Chemical Ecology, Jena 07745, Germany; ²Research Group Biosynthesis/NMR, Max Planck Institute for Chemical Ecology, Jena 07745, Germany; ³Biomolécules et Biotechnologies Végétales, Université de Tours, Tours 37200, France; ⁴Institute of Biosciences and Bioresources, CNR, Perugia 06128, Italy; ⁵Trisaia Research Centre, ENEA, Rotondella 75026, Italy

Authors for correspondence:
Fiammetta Alagna
Email: fiammetta.alagna@enea.it

Ornella Calderini
Email: ornella.calderini@ibbr.cnr.it

Sarah E. O'Connor
Email: oconnor@ice.mpg.de

Received: 24 August 2020
Accepted: 18 September 2020

New Phytologist (2021) 229: 2288–2301
doi: 10.1111/nph.16975

Key words: C-C oxidative cleavage, cytochrome P450, iridoid biosynthesis, *Olea europaea*, oleoside methyl ester, oleuropein, olive, phenolic secoiridoids.

Summary

- Olive (*Olea europaea*) is an important crop in Europe, with high cultural, economic and nutritional significance. Olive oil flavor and quality depend on phenolic secoiridoids, but the biosynthetic pathway of these iridoids remains largely uncharacterized.
- We discovered two bifunctional cytochrome P450 enzymes, catalyzing the rare oxidative C-C bond cleavage of 7-*epi*-loganin to produce oleoside methyl ester (OeOMES) and secoxy-loganin (OeSXS), both through a ketologanin intermediary. Although these enzymes are homologous to the previously reported *Catharanthus roseus* secologanin synthase (CrSLS), the substrate and product profiles differ.
- Biochemical assays provided mechanistic insights into the two-step OeOMES and CrSLS reactions. Model-guided mutations of OeOMES changed the product profile in a predictable manner, revealing insights into the molecular basis for this change in product specificity.
- Our results suggest that, in contrast to published hypotheses, *in planta* production of secoxy-iridoids is secologanin-independent. Notably, sequence data of cultivated and wild olives point to a relation between domestication and OeOMES expression. Thus, the discovery of this key biosynthetic gene suggests a link between domestication and secondary metabolism, and could potentially be used as a genetic marker to guide next-generation breeding programs.

Introduction

Olive (*Olea europaea* L.) is an economically and culturally important crop in the Mediterranean Basin, comprising regions in Southern Europe, Northern Africa and Western Asia, which produce more than 95% of the olives in the world (FAO, 2018). The European Union alone produces two-thirds of the global supply of olive oil, and also consumes half of it (European Commission, 2020). Olives, along with wine and bread, are an important part of the Mediterranean diet, in the form of both its fermented, pickled fruit and its oil, which has been consumed for millennia. The specialized metabolites of olive fruit and oil have been associated with various health benefits (George *et al.*, 2019; Schwingshackl *et al.*, 2019; de la Torre *et al.*, 2020; Marrero *et al.*, 2020). The secoiridoids are the major class of metabolites present in olive fruit, with oleuropein, a phenolic secoiridoid,

making up 6–14% of the DW of the fruit (Amiot *et al.*, 1986; Ryan *et al.*, 1999). Oleuropein strongly influences the quality of olive products because of its bitterness, which is a desirable trait in olive oil, although table olives must be subjected to a lengthy curing process to degrade this compound (Garrido-Fernández *et al.*, 1997). *In planta*, the glucose of oleuropein can be hydrolyzed by a dedicated glucosidase (Koudounas *et al.*, 2015, 2017; Velázquez-Palmero *et al.*, 2017), as part of a defense response. The resulting aglycone has protein-crosslinking and lysine-alkylating activities, making this compound an effective deterrent of generalist herbivores (Konno *et al.*, 1999).

Obtaining olive varieties with ‘fine-tuned’ oleuropein concentrations is an important goal of olive breeding programs (Pérez *et al.*, 2018). The benefits of secoiridoid-related markers have recently been highlighted in the industrial processing of table olives, where oleuropein degradation via glucosidases was a major

biochemical marker for strain selection (Bavaro *et al.*, 2017). The importance of secoiridoids in systematic breeding of olive for quality virgin olive oil is increasingly being recognized, with lack of knowledge on their genetic determinants as the major hurdle for this cause (Pérez *et al.*, 2018). As genotype appears to be a stronger determinant of secoiridoid contents than environmental conditions (Pérez *et al.*, 2018), identification of the genes responsible for oleuropein biosynthesis could readily benefit the olive oil industry as a molecular marker.

Despite the agricultural, biomedical and ecological importance of iridoids, the biosynthetic pathway of only two iridoids has been solved to date: nepetalactone, an evolutionarily and structurally unrelated iridoid produced in catnip (Lichman *et al.*, 2019; Lichman *et al.*, 2020); and secologanin, a secoiridoid produced by *Catharanthus roseus* and several plants in the Gentianales order (Irmeler *et al.*, 2000; Salim *et al.*, 2013; Miettinen *et al.*, 2014). Secoiridoids are characterized by lack of the cyclopentane ring, which is subjected to an oxidative cleavage reaction. This reaction is catalyzed by the enzyme secologanin synthase (SLS), a cytochrome P450 (CYP72) first identified from *C. roseus*, which catalyzes the oxidative opening of loganin into secologanin (Irmeler *et al.*, 2000). The only additional homolog that has been characterized is secologanic acid synthase (SLAS) from *Camptotheca acuminata*, which catalyzes the same oxidative cleavage as SLS, but using des-methylated loganin (loganic acid) as the substrate (Yang *et al.*, 2019). Unlike secologanin-type iridoids, in which the oxidative cleavage produces a vinyl group, oleoside-type iridoids have a distinctive exocyclic olefin (Fig. 1), which increases the reactivity of the aglycone for defensive protein crosslinking (Konno *et al.*, 1999). Oleoside-type iridoids are a chemotaxonomic marker, being produced exclusively in the Oleaceae family, with olive producing both secologanin- and oleoside-type secoxy-iridoids.

The structure of oleuropein and secologanin suggests that the early biosynthetic steps of these two compounds are the same: namely biosynthesis of geraniol, cyclization to nepetalactol, followed by oxidation and glucosylation (Fig. 1). The identification of olive iridoid synthase (OeISY) supports this hypothesis (Alagna *et al.*, 2016). However, the mechanism of the chemotaxonomical divergence of secologanin- and oleoside-type iridoids remains unknown. Several biosynthetic precursors for oleuropein have been proposed: secologanin (Inouye *et al.*, 1971; Damtoft *et al.*, 1993; Alagna *et al.*, 2016), 8-*epi*-kinginside (Inouye *et al.*, 1971; Damtoft *et al.*, 1993) or ketologanin (Alagna *et al.*, 2016) (Supporting Information Fig. S1). Although several feeding experiments have been performed in olive plants (Inouye *et al.*, 1971; Damtoft *et al.*, 1993; Serrilli *et al.*, 2016), no conclusive evidence has been found for the oleoside pathway.

Here, we use olive transcriptomic data to identify the enzymes responsible for generating the secoiridoid scaffolds in olive. Using SLS from *C. roseus* as bait for a homology-based search, we identified three SLS homologs in olive, and showed that these catalyze formation of either oleoside methyl ester or secoxyloganin from 7-*epi*-loganin, through a ketologanin intermediary that can also be taken up directly by the enzymes. This work reports the key step for this important flavor agent and mutational analysis

reveals how enzyme substrate and product specificity are modulated to generate chemical diversity in plants.

Materials and Methods

Identification of enzyme candidates

To identify candidates for oleoside methyl ester (OME) biosynthesis, we blasted the protein sequence of *Catharanthus roseus* secologanin synthase 2 (CrSLS2) against a previously published, in house assembled olive transcriptome (Alagna *et al.*, 2009), from which expression correlations were also calculated. We designed primers for the three homologous sequences found (*OeSLS1-3*; GenBank MT909126, MT909123 and MT909125) and amplified them via PCR from a cDNA library derived from total RNA of olive (*O. europaea* ssp. *europaea*) young leaf tissue. As we were unable to obtain full-length clones of *OeSLS1* and *OeSLS3*, DNA fragments were ordered from IDT (Integrated DNA Technologies, BVBA; Belgium). The sequences were cloned into pUC19 and sequence verified, and later subcloned using an In-Fusion cloning system (Takara Bio Europe; Saint-Germain-en-Laye, France) into the Gal10 Multiple Cloning Site (MCS) of pESC-Leu2d containing an *Artemisia annua* cytochrome P450 reductase (AaCPR) under the Gal1 promoter (Ro *et al.*, 2008). Mutants were directly synthesized in the pESC-Leu2d::AaCPR by Twist (Twist Bioscience; San Francisco, CA, USA).

Heterologous expression and microsomal preparations

Protease-deficient *Saccharomyces cerevisiae* strain YPL 154C: Pep4KO was used for heterologous expression of p450s, as previously published (Ro *et al.*, 2002), with minor modifications. Namely, the strain, harboring the dual expression pESC-Leu2d::AaCPR-GOI plasmid, was plated, and single colonies inoculated in 5 ml Synthetic Defined medium without leucine (SD-Leu; Sigma-Aldrich) with 2% glucose, and cultivated overnight in an orbital shaker at 30°C and 250 rpm. This culture was used to inoculate 100 ml of SD-Leu with 2% glucose and left to grow for 24 h. This culture was diluted into 1 l SD-Leu with 2% galactose, for protein expression, to a final absorbance of OD₆₀₀ = 0.5 (*c.* 6–7 × 10⁶ cells ml⁻¹). After 24 h, cells were harvested, and microsomes prepared as previously mentioned (Dang *et al.*, 2017), using an LM20 microfluidizer (Microfluidics International Corp.; Westwood, MA, USA).

Microsome reactions were performed using 2 mg of total protein in 50 µl of 50 mM Tris buffer at pH 7.5, incubated at 30°C in the dark on a shaker (Vortemp56; Labnet International, Edison, NJ, USA) at 300 rpm. Exploratory reactions were incubated overnight with 25 µM substrate and 125 µM NADPH. Steady-state kinetics were determined by feeding with 10, 20, 40, 80, 160, 250, and 500 µM substrate, with three times as much NADPH, and stopping the reactions at 30, 60, 90, 120 and 180 min. All reactions were done in triplicate, started by adding the substrate, and stopped by adding 50 µl methanol. Samples were then centrifuged at 20 000 Relative Centrifugal Force (RCF) for 10 min and transferred to high-performance liquid chromatography (HPLC) vials or plates.

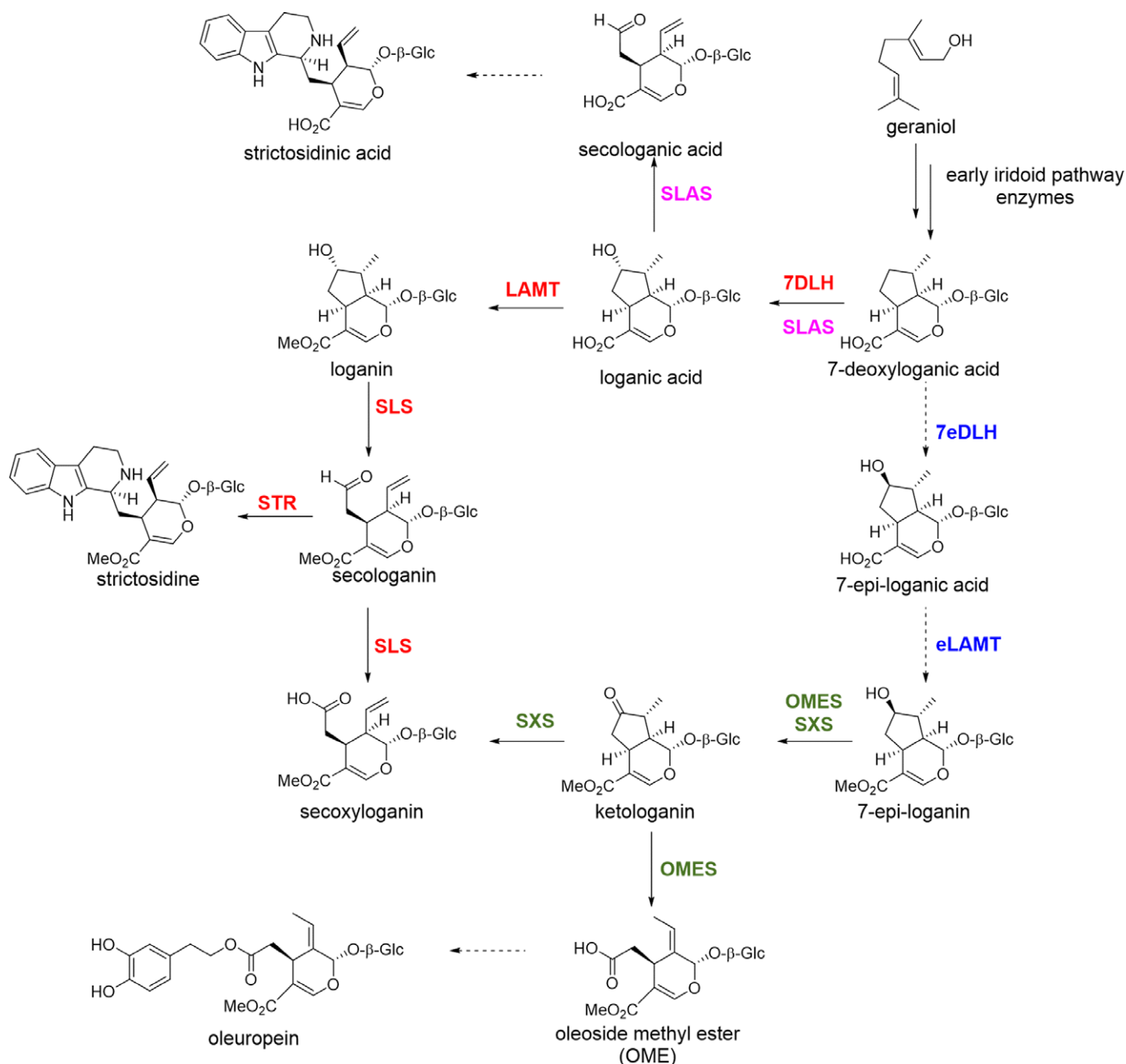


Fig. 1 Summary of existing knowledge of the secoiridoid pathways. Enzymes in red have been characterized in *Catharanthus roseus*, and enzymes in pink have been characterized in *Camptotheca acuminata*. Enzymes in blue are hypothesized to exist in *Olea europaea*, but not in *C. roseus*. Dashed arrows indicate steps for which a biosynthetic gene has not yet been discovered. Enzymes in olive green were characterized in this work in *O. europaea*. SLS, secologanin synthase; SLAS, secologanic acid synthase; STR, strictosidine synthase; 7DLH, 7-deoxy-loganic acid hydroxylase; LAMT, loganic acid methyl-transferase; 7eDLH, 7-deoxy-loganic acid 7-*epi*-hydroxylase; eLAMT, 7-*epi*-loganic acid methyl-transferase; OMES, oleoside methyl ester synthase; SXS, secoxyloganin synthase.

Iridoid profiling using HPLC-MS

Samples were analyzed using an Elute LC system (Bruker Daltonik, Bremen, Germany) coupled via electrospray ionization (ESI) to a Maxis II quadrupole-time-of-flight (qTOF) MS (Bruker Daltonik; Bremen, Germany). Iridoids were separated using an Acquity UPLC[®] C18 column (2.1 × 50 mm, 1.7 μm, 100 Å; Waters, Milford, MA, USA) at 40°C with a gradient

from water (A) to acetonitrile (B), both modified with 0.1% formic acid, using two different programs, both at a flow rate of 0.3 ml min⁻¹ and redirecting to waste for the first 90 s. For exploration of reaction products, an 11 min gradient was used, starting at 1% B for 1 min, then linearly increasing to 20% B for 5 min, then 40% B for 2 min, and finally 99% B for 1 min, and kept at 99% B for 2 min. The column was equilibrated in 1% B for another 2 min before the next injection. For the kinetic

experiments, the gradient was shortened to 4.5 min, starting at 1% B for 0.5 min, then linearly increasing to 20% B for 2.5 min, then 25% B for 0.5 min, and immediately changing to 99% B, which was kept for 1 min. The column was equilibrated in 1% B for 1 min.

Acquisition of MS data was performed using the same conditions for exploration and kinetics, but the latter with no fragmentation. Ionization was performed via pneumatic assisted ESI in negative mode (ESI[−]) with a capillary voltage of 3.5 kV, an end-plate offset of 500 V, and a nebulizer pressure of 3 bar. Nitrogen gas was used as drying gas at 350°C and a flow of 11 l min^{−1}. For the kinetics experiments, acquisition was performed at 2 Hz with no fragmentation, at a mass range from 250 to 1000 *m/z*. For exploration experiments, acquisition was done at 12 Hz following a mass range from 250 to 1000 *m/z*, with data-dependent MS/MS without active exclusion. Fragmentation was triggered on an absolute threshold of 400, and acquired on the most intense peaks using a target intensity of 2×10^4 counts, with a MS/MS spectra acquisition of 12 Hz, and limited to a total cycle time range of 0.5 s, and a dynamic collision energy of 20–50 eV. At the beginning of each run, while the LC input was redirected to waste, a sodium formate-isopropanol solution was injected at 18 ml h^{−1} for the first 90 s of each run, and the *m/z* values were recalibrated using the expected cluster ion *m/z* values.

Data analysis

Raw MS files were converted to mzXML using Bruker Data Analysis software (Bruker Daltonik) and, when needed, extracted ion chromatograms (XIC) were exported to EXCEL using MZMINE2 (v.2.40.1; Pluskal *et al.*, 2010). For the kinetics experiment, areas were obtained by an in-house built script, using the R programming language (R Core Team, 2019). The script accessed the raw data files, extracted the XIC within 0.05 *m/z* error, and integrated it in a window of 0.1 min within the retention time of each standard. Raw mzXML files were accessed with the aid of the MZR package (Chambers *et al.*, 2012) and parallel processing of data facilitated by the ‘parallel’ (R Core Team, 2019) and ‘snow’ (Tierney *et al.*, 2018) libraries. Figures were done using the R base package (R Core Team, 2019), swarmed boxplots using the ‘beeswarm’ library (Eklund, 2016) and chromatograms using EXCEL. Molecules were drawn using CHEMDRAW.

Steady-state kinetics

For quantification in the kinetics experiments, a 10-point calibration curve of 7-*epi*-loganin, ketologanin and OME, from 0.1 to 100 pmol, was run in between each replicate of all substrate concentrations for each enzyme. Concentrations were quantified with the enclosing calibration curves of each sample; secoxyloganin was quantified in OME equivalents. Linear velocities were calculated by running linear regressions, and steady-state constants were estimated by a nonlinear least squares fit to a Michaelis–Menten equation or Eqn 1, according to the case, using the R *stats* library (R Core Team, 2019). Derivation of Eqn 1 is shown in Notes S1 and S2.

Homology models

Protein sequences of biochemically tested CYP72 enzymes were obtained from Uniprot (The Uniprot Consortium, 2019), multiple alignments were performed using CLUSTALW (Madeira *et al.*, 2019), and phylogenetic trees generated using the IQ-TREE web server (Trifinopoulos *et al.*, 2016) and visualized using the interactive Tree of Life (Letunic & Bork, 2019). The homology model for OeOMES, OeSXS, CrSLS and CAC CYP72A612 was selected using PHYRE2 online software (Kelley *et al.*, 2015) with the default settings in the normal modeling mode; all models were aligned using the Partial Order Structure Alignment server (POSA; Li *et al.*, 2014). For visual exploration and selection of the regions closest to the substrate, we used UCSF CHIMERA v.1.14 (Pettersen *et al.*, 2004).

Analysis of RNA-seq public datasets

We selected publicly available RNA-seq data by searching in NCBI's Sequence Read Archive, and downloaded the raw fastq files of projects PRJNA525000 (Gros-Balthazard *et al.*, 2019), PRJNA256033 (Leyva-Pérez *et al.*, 2015), PRJNA378602 (Jiménez-Ruiz *et al.*, 2019), PRJNA401310 (Jiménez-Ruiz *et al.*, 2018), PRJNA514943 (Research Institute of Forestry) and PRJNA596876 (Research Institute of Forestry) using the European Nucleotide Archive FTP. Salmon (Patro *et al.*, 2017) was used for abundance estimation via quasi-mapping of reads against transcriptomes derived from the olive cv Farga genome (Cruz *et al.*, 2016), wild olive genome (Unver *et al.*, 2017), and the Reproline project (Carmona *et al.*, 2015), generated from vegetative and reproductive tissue. The cv Farga transcriptome was chosen for analysis, as it was the only one with mappings of > 70% (Fig. S2). Of the 112 samples, 103 had a mapping > 70%, having a total of 2×10^9 mapped reads, an 82% mapping total (Fig. S2).

Iridoid standards

Loganin and secologanin standards were purchased from Sigma-Aldrich, and secoxyloganin was obtained by incubating secologanin and CrSLS microsomes overnight, with five NADPH equivalents. 8-*epi*-loganin as well as OME were purchased from AnalytiCon Discovery GmbH (Potsdam, Germany). An inconsistency exists in the PubChem database and AnalytiCon Discovery catalog, both reporting OME with the exocyclic bond in the Z configuration, in contrast to the published literature (Inouye *et al.*, 1971; Damtoft *et al.*, 1993; Alagna *et al.*, 2016) which reports the E configuration. Thus, the determination of the absolute configuration of the OME standard was performed by nuclear magnetic resonance (NMR; Methods S1), which confirmed that the standard from AnalytiCon Discovery is in the E configuration (Figs S3, S4), despite it being reported in the Z configuration in their catalog.

The rest of the iridoid standards are not commercially available, and were synthesized. The detailed synthesis protocols, along with yields and spectroscopic data, are shown in the

Methods S2. The spectroscopic data are in accordance with the literature values reported for geniposide pentaacetate (Zhang *et al.*, 2013); 10-deoxy-geniposide tetraacetate, 10-deoxy-geniposidic acid, and 7-deoxy-loganic acid (Inoue *et al.*, 1992); 7-deoxy-8-*epi*-loganic acid (Nakamura *et al.*, 2000); loganin epoxide and ketologanin tetraacetate (Inoue *et al.*, 1970); 7-keto-loganin (Gross *et al.*, 1986); and 7-*epi*-loganin (Itoh *et al.*, 2005). Spectroscopic data for 7-*epi*-8-*epi*-loganin, which has not been previously reported, is shown here (see also Figs S5, S6, and Methods S3):

- 7-*epi*-8-*epi*-loganin: ^1H NMR (400 MHz, CD_3OD) δ 7.43 (s, 1H), 5.66 (d, J = 5.6 Hz, 1H), 4.70 (d, J = 7.9 Hz, 1H), 4.17 (dd, J = 12.7, 5.7 Hz, 1H), 3.89–3.84 (m, 1H), 3.69 (s, 3H), 3.68–3.61 (m, 1H), 3.39–3.34 (m, 1H), 3.24–3.17 (m, 3H), 2.84 (dd, J = 18.0, 8.4 Hz, 1H), 2.61 (ddd, J = 14.2, 9.8, 7.4 Hz, 1H), 2.22 (td, J = 12.1, 6.6 Hz, 2H), 1.42–1.32 (m, 1H), 1.10 (d, J = 7.0 Hz, 3H);
- ^{13}C NMR (100 MHz, CD_3OD) δ 168.2, 151.2, 112.2, 98.5, 95.7, 76.9, 76.6, 74.0, 73.5, 70.3, 61.5, 50.3, 42.2, 41.7, 40.5, 30.9, 9.4;

Results

Identification of enzyme candidates for secoiridoid formation in *Olea europaea*

We identified three SLS (*C. roseus*) homologs, identified as *OeSLS1*–3, in the olive transcriptome (Fig. 2). A coexpression analysis using transcriptome data from different olive tissues found that, while *OeSLS1* showed modest correlation (Pearson's r = 0.48), *OeSLS2* and *OeSLS3* showed high coexpression with the olive iridoid synthase (r = 0.76 and r = 0.95, respectively). *OeSLS1*–3 were cloned and heterologously expressed in yeast, and the subsequent yeast microsomes were assayed for biochemical activity. During amplification of these genes from olive cDNA, it became evident that *OeSLS2* had a variant, *OeSLS2A*, with 99.8% identity to *OeSLS2*. Subsequent enzyme assays indicated that *OeSLS2A* had identical biochemical activity to *OeSLS2*, and this was not investigated further.

Biochemical functions of SLS homologs

Oxidative cleavage of the iridoid scaffold is required to generate the secoiridoids. Feeding studies with olive plants (Inoue *et al.*, 1971) indicate that the final product of such an oxidative cleavage product would be OME, which is then esterified to form oleuropein. The most straightforward hypothesis is that an SLS homolog would catalyze such a reaction from an iridoid intermediate, but no experimental data suggested which specific substrate this putative enzyme would utilize as a starting material. The two characterized oxidative cleavage reactions in secoiridoid biosynthesis, namely SLS from *C. roseus* and SLAS from *C. acuminata*, utilize loganin and loganic acid, respectively (Fig. 1). Several substrates have been suggested for OME biosynthesis, such as secologanin (Inoue *et al.*, 1971; Damtoft *et al.*, 1993; Alagna *et al.*, 2016), 8-*epi*-kingiside (Inoue *et al.*, 1971; Damtoft *et al.*, 1993), and ketologanin (Alagna *et al.*, 2016) (Fig. S1).

We first assayed the olive SLS homologs with secologanin and loganin (Fig. S7). SLS from *C. roseus* (CrSLS) was used for comparison. While *OeSLS1* did not show activity, CrSLS, *OeSLS2* and *OeSLS3* converted secologanin to secoxyloganin, although the latter did so to a lesser extent (Fig. S7a). Notably, all three SLS homologs (*OeSLS1*, *OeSLS2* and *OeSLS3*) failed to react with loganin (Fig. S7b), the native substrate of CrSLS. Interestingly, although secoxyloganin production has been reported for CrSLS (de Bernonville *et al.*, 2015), *C. roseus* has no detectable concentrations of secoxyloganin; by contrast, secoxyloganin is abundant in olive (De Marino, *et al.*, 2014).

As these enzymes failed to produce OME from loganin and secologanin, we next assayed SLS homologs with ketologanin. Ketologanin was not implicated in any known secoiridoid biosynthetic pathway, but previous research (Damtoft *et al.*, 1993; Alagna *et al.*, 2016) has suggested that this compound may be a biosynthetic intermediate in oleoside biosynthesis. While CrSLS and *OeSLS1* failed to react with this substrate, both *OeSLS2* and *OeSLS3* turned over ketologanin; *OeSLS2* oxidatively cleaved ketologanin to OME and *OeSLS3* oxidatively cleaved it to secoxyloganin (Fig. 3a). We thus renamed *OeSLS2* as oleoside methyl ester synthase (*OeOMES*) and *OeSLS3* as secoxyloganin synthase (*OeSXS*).

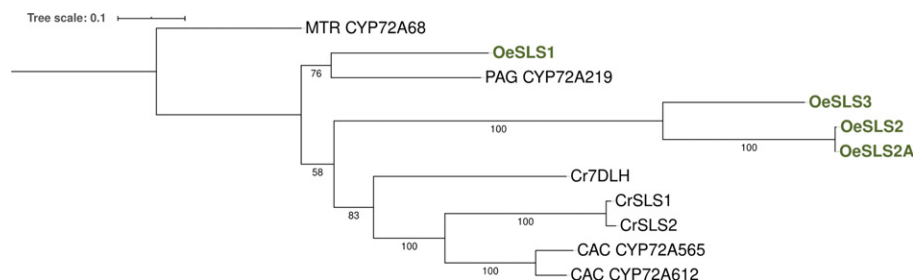


Fig. 2 Consensus alignment tree. Maximum-likelihood tree (LG + G4) of a CLUSTALW alignment of the protein sequences of CYP72 enzymes with proven biochemical activity. Nodal support was estimated using a 1000 bootstrap, and the tree rooted on *Medicago truncatula* (MTR) CYP72A68. *Catharanthus roseus* secologanin synthase (CrSLS) and *Camptotheca acuminata* (CAC) CYP72 enzymes catalyze oxidative cleavage of iridoid scaffolds to produce secologanin and secologanic acid, respectively, while *Panax ginseng* (PAG) and MTR CYP72 enzymes oxidize triterpenoid scaffolds, without cleavage. The enzymes covered in this work for *Olea europaea* (Oe) are shown in bold olive green.

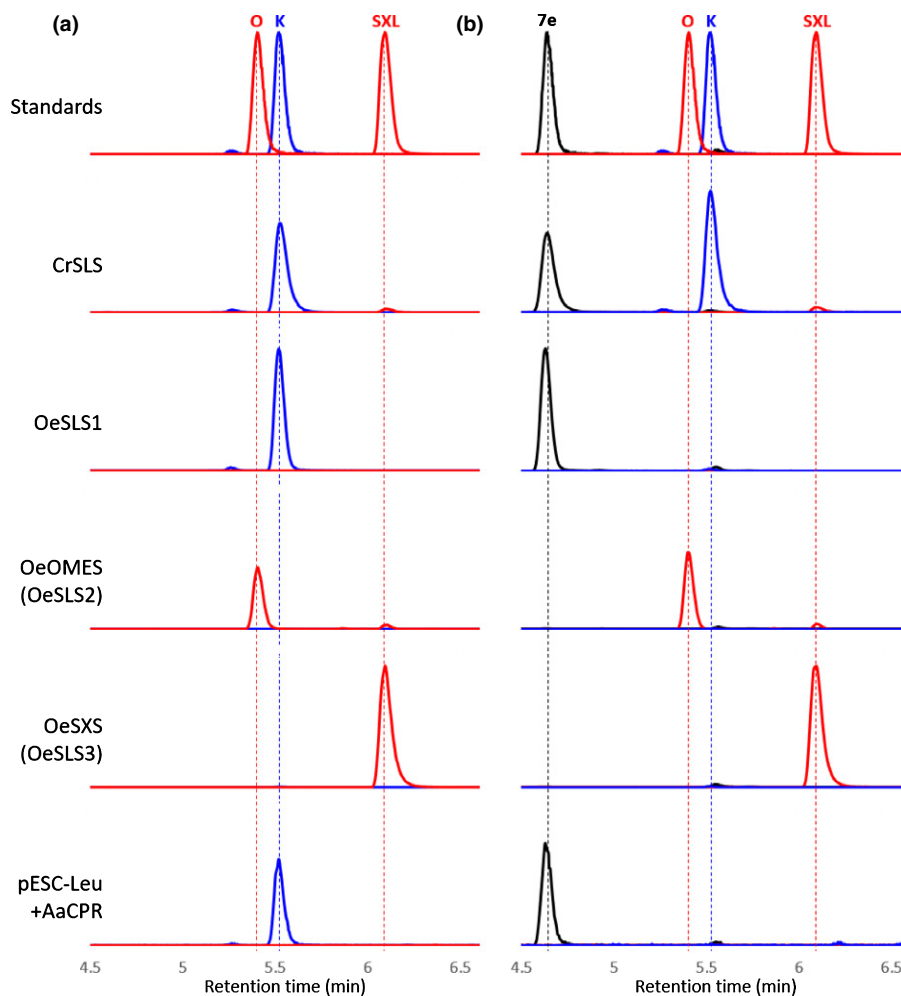


Fig. 3 Extracted ion chromatograms (XICs) of microsome substrate assays. (a, b) XICs of the microsomal incubations with ketologanin (a) and 7-*epi*-loganin (b). Three channels are depicted, corresponding to the most abundant adduct of oleoside methyl ester (OME) or secoxyloganin (red, $[M-H]^- = 403.1240 \pm 0.05$), ketologanin (blue, $[M + FA-H]^- = 433.1346 \pm 0.05$) and 7-*epi*-loganin (black, $[M + FA-H]^- = 435.1503 \pm 0.05$). Intensities are scaled to the highest intensity of the corresponding channel in all incubations. pESC-Leu_AaCPR refers to microsomes that express only the cytochrome reductase. O, OME; K, ketologanin; SXL, secoxyloganin.

Given that CrSLS can perform the sequential oxidation of loganin to secologanin to secoxyloganin, we next investigated whether the olive SLS homologs were also capable of performing two sequential oxidation steps. We synthesized a mixture of loganin diastereomers (7-*epi*-loganin, 8-*epi*-loganin, and 7,8-*epi*-loganin) to test if they could be oxidized by the SLS homologs. Of these epimers, only 7-*epi*-loganin was depleted in OeOMES and OeSXS, where peaks corresponding to OME and secoxyloganin, respectively, were detected (Fig. S8). Ketologanin was also detected in OeOMES and OeSXS reactions, and also, interestingly, when CrSLS was reacted with 7-*epi*-loganin. In these enzymatic reactions, an isomer of ketologanin was observed along with the ketologanin product (Fig. S8), although this product could not be isolated in sufficient quantity for characterization. OeSLS1, as with the other substrates, was largely inactive, but did produce trace amounts of ketologanin. These enzyme assays were then repeated with purified 7-*epi*-loganin, which confirmed the observation that OeOMES and OeSXS produce OME and secoxyloganin, respectively, while CrSLS produces ketologanin (Fig. 3b). Although only trace amounts of ketologanin were detected in long incubations with 7-*epi*-loganin and OeOMES and OeSXS, its presence in the early time points of the kinetic reactions supports its role as an intermediary.

We next tested 7-deoxy-loganin to assess whether these enzymes could catalyze the three-step oxidation: 7-deoxy-loganin to 7-*epi*-loganin to ketologanin, and then finally to either OME or SXS. When incubated with 7-deoxy-loganin, CrSLS produced loganin, subsequently oxidizing it to secologanin and secoxyloganin (Fig. S9). Notably, *C. roseus* has a dedicated hydroxylase (7DLH) that converts 7-deoxy-loganic acid to loganic acid that has been validated by gene silencing (Salim *et al.*, 2013), so it is not clear whether CrSLS turns over 7-deoxy-loganin *in planta*. OeOMES oxidized 7-deoxy-loganin to a product that could not be isolated in sufficient quantity for characterization. However, as only traces of ketologanin and OME were observed in this reaction, we assume that 7-deoxy-loganin is not an intermediate of the pathway (Fig. S9). OeSXS was also found to produce this molecule, and OeSLS1 showed no activity. No activity was observed when the microsomes were incubated with 7-deoxy-loganic acid, 7-deoxy-8-*epi*-loganic acid, and 7-deoxy-8-*epi*-loganin.

Steady-state kinetics of OeOMES and OeSXS

Kinetic parameters were calculated for OeOMES and OeSXS, incubated with either ketologanin or 7-*epi*-loganin (Fig. S10). If

enzyme regeneration is faster than the oxidation reaction, conditions that should be fulfilled by adding excess NADPH, the kinetic parameters for the ketologanin substrate should follow Michaelis–Menten kinetics (Michaelis & Menten, 1913). The exponential fit obtained (Fig. S10a,c) suggests that this assumption is reasonable, and can be used as a starting point to measure the kinetics of the consecutive reaction from 7-*epi*-loganin. The subject of the dissociation of intermediates in consecutive reactions catalyzed by cytochrome P450 enzymes remains contentious, but the kinetics of either case do not practically differ under a quasi-steady-state assumption (Notes S1 (Eqn S1.7), Notes S2 (Eqn S2.13)), and follow the equation:

$$v_{\text{comp}} = \frac{(v_K + v_{EK}) + v_P}{1 - \frac{v_P}{V_K}} = \frac{V_{\text{epi}}[S]}{K_{\text{Mepi}} + [S]} \quad \text{Eqn 1}$$

where $[S]$ is the concentration of 7-*epi*-loganin, V_{epi} and V_K are the maximum velocities of 7-*epi*-loganin and ketologanin, respectively; $(V_K + V_{EK})$ is the change in time of measured ketologanin (in solution plus complexed with the enzyme), and v_P the rate of production of OME (by OeOMES) or secoxyloganin (by OeSXS). For simplification, the scaled sum of velocities is called v_{comp} and follows Michaelis–Menten kinetics (Fig. S10b,d). The results of this model are shown in Fig. S10(b) for OeOMES and Fig. S5(d) for OeSXS. Individual velocities, used to calculate Fig. S10(b,d), can be seen in Fig. S11 for OeOMES (Fig. S11a) and OeSXS (Fig. S11b).

The relative turnover $\frac{(k_{\text{cat}})_{\text{epi}}}{(k_{\text{cat}})_{\text{keto}}}$ and specificity constants (V_{max}/K_M) of 7-*epi*-loganin vs ketologanin oxidation were compared (Table 1). These kinetic analyses indicate that OeOMES is more specific towards 7-*epi*-loganin and the V_{max} for oxidation of 7-*epi*-loganin to ketologanin is twice as fast as the oxidative cleavage to OME. This explains the nonMichaelis–Menten kinetics seen in the individual velocities (Fig. S11a), where, as there is more substrate presence ($> 250 \mu\text{M}$ 7-*epi*-loganin), the rate of ketologanin formation is higher than that of OME. By contrast, OeOSXS, although it has a similar relative specificity for both 7-*epi*-loganin and ketologanin, has a lower relative turnover for 7-*epi*-loganin, indicating that the oxidative cleavage of ketologanin to secoxyloganin ($V_{\text{max}} = 54.6 \pm 3.6 \mu\text{M s}^{-1}$) is faster than the formation of ketologanin from 7-*epi*-loganin ($V_{\text{max}} = 12.6 \pm 0.6 \mu\text{M s}^{-1}$). Again, this can be seen in the individual kinetic analyses where, unlike OeOMES, the rate of secoxyloganin is always higher than the rate of ketologanin formation (Fig. S11b).

Mutational analysis of OeOMES

Structural information for plant cytochrome P450s is limited, which hampers rational mutational analysis. Nevertheless, we decided to explore the effects of amino acid substitution, in an attempt to identify the molecular basis of substrate and/or product specificity. As OeOMES can produce both secoxyloganin (from secologanin) and OME (from ketologanin), we used OeOMES as a template to introduce amino acid mutations.

Using PHYRE2.0 (Kelley *et al.*, 2015), we modeled OeOMES, OeSXS, CrSLS and CAC CYP72A612 against the crystal structure of *Saccharomyces cerevisiae* lanosterol 14- α demethylase (PDB 4LXJ; Monk *et al.*, 2014). This was the only candidate in common, within the top five entries for OeOMES, OeSXS and CrSLS, that catalyzed a similar reaction: the oxidative cleavage of a carbon–carbon bond, resulting in formation of a vinyl group and a carboxylic acid (Monk *et al.*, 2014).

As simultaneous docking of the substrate and heme groups into the homology models would potentiate the error, we instead aligned the three-dimensional models to the original protein template using POSA (Li *et al.*, 2014). We explored the resulting superimposed homology models using CHIMERA (Pettersen *et al.*, 2004), and we selected five regions of interest within 10 Å of the substrate binding site or the iron reactive center (Fig. S12). As expected, the conserved region of the Cytochromes P450 (CYP) responsible of the coordination of iron aligned perfectly with the heme group in all the models. Within the five regions closest to the substrate and iron, named A–E, we chose the amino acids that were not conserved among OeOMES, OeSXS, CrSLS and CAC CYP72A612. We substituted all these amino acids in each region of OeOMES with the corresponding residues of OeSXS, yielding six mutant designs (Table S1). Each of these mutants was expressed in yeast, and the resulting microsomal fractions were incubated with ketologanin, 7-*epi*-loganin and secologanin (Figs 4, S13).

Wild-type OeOMES produces OME from 7-*epi*-loganin and ketologanin, with only trace amounts of secoxyloganin being formed. Production of secoxyloganin from ketologanin and 7-*epi*-loganin increased in three mutants (Fig. 4): mutants in region B (four amino acids switched), which is close to the Fe–ligand interaction; mutants in region C (four amino acids), a loop and sheet close to the substrate; and the fully switched mutant (22 amino acids). The latter had the highest production of secoxyloganin of the mutants consistently producing 75% as much secoxyloganin as OME from both 7-*epi*-loganin and ketologanin, along with a decrease in OME production (Fig. 4). Notably, a new product was also observed in the secoxyloganin-producing mutants (Fig. 4, yellow chromatogram). This new product has a m/z corresponding to the formic acid adduct of an oxidation product of ketologanin, $[\text{C}_{17}\text{H}_{24}\text{O}_{11} + \text{HCOO}^-]^-$.

Interestingly, the mutations in region A (five amino acids), which is close to the substrate, opposite the heme-coordinating region, had a strong negative effect in production of OME from 7-*epi*-loganin (Fig. 4b), which is ameliorated when fed ketologanin (Fig. 4a). Similarly, the mutant in region E (three amino acids), a loop physically close to region A and the substrate, when fed ketologanin produces as much OME as the wild-type, but not when incubated with 7-*epi*-loganin (Fig. S13). Mutations in region D (six amino acids), the heme-coordinating region, had no effect in any of the tested activities (Figs S13, S14). Secologanin was also tested as a substrate to monitor overall oxidative activity. All mutants, regardless of the observed product profile with ketologanin, efficiently produced secoxyloganin from secologanin (Fig. S14).

Table 1 Kinetic parameters of *Olea europaea* oleoside methyl ester synthase (OeOMES) and *O. europaea* secoxyloganin synthase (OeSXS).

	7-Epi-loganin			Ketologanin			
	K_M (μM)	$V_{\max}$ ($\mu\text{M s}^{-1}$)	$V_{\max}/K_M$ (s^{-1})	K_M (μM)	$V_{\max}$ ($\mu\text{M s}^{-1}$)	$V_{\max}/K_M$ (s^{-1})	$\frac{(k_{\text{cat}})_{\text{epi}}}{(k_{\text{cat}})_{\text{keto}}}$
OeOMES	170 $\pm$ 24	90.0 $\pm$ 5.4	0.53 $\pm$ 0.08	172 $\pm$ 37	39.0 $\pm$ 3.6	0.23 $\pm$ 0.05	2.31 $\pm$ 0.25
OeSXS	37.7 $\pm$ 9.3	12.6 $\pm$ 0.6	0.33 $\pm$ 0.08	254 $\pm$ 36	54.6 $\pm$ 3.6	0.21 $\pm$ 0.03	0.23 $\pm$ 0.022

Values are shown as the calculated regression coefficient $\pm$ SE. K_M , Michaelis-Menten constant; $V_{\max}$, maximum velocity; $(k_{\text{cat}})_{\text{epi}}$, catalytic rate constant of 7-*epi*-loganin transformation; $(k_{\text{cat}})_{\text{keto}}$, catalytic rate constant of ketologanin transformation.

Analysis of expression data of the olive iridoid pathway

To further explore the *in planta* properties of OeOMES and OeSXS, we analyzed the extensive publicly available expression data for olive. We separated the data in two subgroups: a study of 56 commercially cultivated varieties and wild olives (PRJNA525000, Gros-Balthazard *et al.*, 2019); and five studies, comprising 47 datasets, targeting different biotic and abiotic stresses on olive (Leyva-Pérez *et al.*, 2015; Jiménez-Ruiz *et al.*, 2018; Jiménez-Ruiz *et al.*, 2019). In the latter dataset, regardless of cultivar or stress, *OeISY* is always highly correlated with *OeOMES* ($r = 0.67$) and *OeSXS* ($r = 0.58$; Fig. S15). Interestingly, *OeISY* is negatively correlated ($r = -0.77$) with *OeSLS1*, providing further evidence that OeSLS1 does not have a role in iridoid-related metabolism (Fig. S15). The studies showed that the only stressors that increase *OeOMES* expression are in roots (cv Picual) 24 h after wounding (PRJNA256033), and in roots (cv Frantoio) at 2 d after a challenge with *Verticillium dahliae* (Fig. S16). Detailed experiments have measured expression in inflorescence tissue of cultivated and wild olives (two botanical varieties of *O. europaea* ssp. *europaea*, var. *europaea* and *sylvestris*) to obtain insights into the domestication of olive (study PRJNA525000, Gros-Balthazard *et al.*, 2019). Fascinatingly, *ISY*, iridoid oxidase and *OMES* (12-fold increase on average) were highly expressed in inflorescences of commercial cultivars, while lacking expression in most wild olives (Fig. 5). By contrast, almost no expression of *OeSXS* was found in either commercial or wild olives, and no pattern can be discerned (Fig. 5d). In all datasets, *OeOMES* had consistently higher expression than *OeISY*.

Discussion

Here we report the discovery of the enzyme OeOMES, which catalyzes the oxidative cleavage leading to the major secoiridoid scaffold in olive, OME, the penultimate step of oleuropein biosynthesis. Previously, reported feeding studies in olive failed to identify the biosynthetic intermediates that lead to OME. These biochemical studies now show that OeOMES converts 7-*epi*-loganin via a two-step oxidation to form OME. We also identified a variant of this enzyme that catalyzes secoxyloganin biosynthesis, a metabolite that is also present in olive. Unexpectedly, secoxyloganin is also derived from 7-*epi*-loganin, and in *O. europaea* is not synthesized via secologanin, as we anticipated. This discovery highlights that SLS homologs have evolved to

catalyze several oxidative cleavage reactions, as summarized in Fig. 6. Coexpression analysis shows that *OeISY*, *OeOMES* and *OeSXS* are coregulated, while *OeSLS1* is negatively correlated to *OeISY*. Along with the fact that we detected no activity for OeSLS1, this supports the notion that OeSLS1 is not involved in iridoid biosynthesis. The kinetic experiments have shown that the oxidative cleavage of ketologanin by OeOMES is slightly slower, compared with ketologanin formation from 7-*epi*-loganin by the same enzyme. Overall, however, the catalytic efficiencies for the two enzymes for both substrates are similar. As only Oleaceae and select members of the Loasaceae (Weigend *et al.*, 2000) produce oleoside-type secoiridoids, while also producing secologanin-type, we hypothesize that the emergence of the exocyclic olefin products is evolutionarily recent.

These biochemical findings are largely consistent with previously published feeding experiments, particularly the thorough studies of Damtoft *et al.* (1993), in which deuterated iridoids were fed to *Syringa josikaea* and *Fraxinus excelsior*, both members of the Oleaceae. These studies demonstrated that while 7-deoxyloganic acid was incorporated into secoiridoids in both *S. josikaea* (12.5%) and *F. excelsior* (11.5%), 7-deoxy-8-*epi*-loganic acid was not metabolized by either, consistent with the substrate selectivity observed for OeOMES and OeSXS. Notably, Damtoft *et al.* reported incorporation of both loganic acid and 7-*epi*-loganic acid into the oleoside ligstroside (18% and 15%, respectively) (ligstroside is an analog of oleuropein in which the OME is derivatized with a tyrosol ester instead of a dihydroxyl phenyl ester). Although loganic acid was not turned over by OeOMES and OeSXS, it is possible that loganic acid could be oxidized to ketologanin by another plant oxidoreductase, and ketologanin could then be incorporated into ligstroside via an OME synthase (Fig. 1).

Catharanthus roseus secologanin synthase, OeOMES and OeSXS all oxidize secologanin to secoxyloganin, and all also oxidize 7-*epi*-loganin to ketologanin. Notably, although the olive enzymes can turn over secologanin, we found no olive cytochrome P450 that generates secologanin, and there are no reports of secologanin in olive. Unexpectedly, OeOMES, which synthesizes the OME, is able to oxidize secologanin. Given the lack of secologanin in olive tissues, it is likely that the ability of OeOMES and OeSXS to turn over secologanin to secoxyloganin is not a physiologically relevant activity. Similarly, 7-deoxyloganic acid, 7-*epi*-loganin and ketologanin have not been detected in *C. roseus*, so the ability of CrSLS to turn over 7-*epi*-loganin and 7-deoxyloganic acid is also probably not relevant *in planta*.

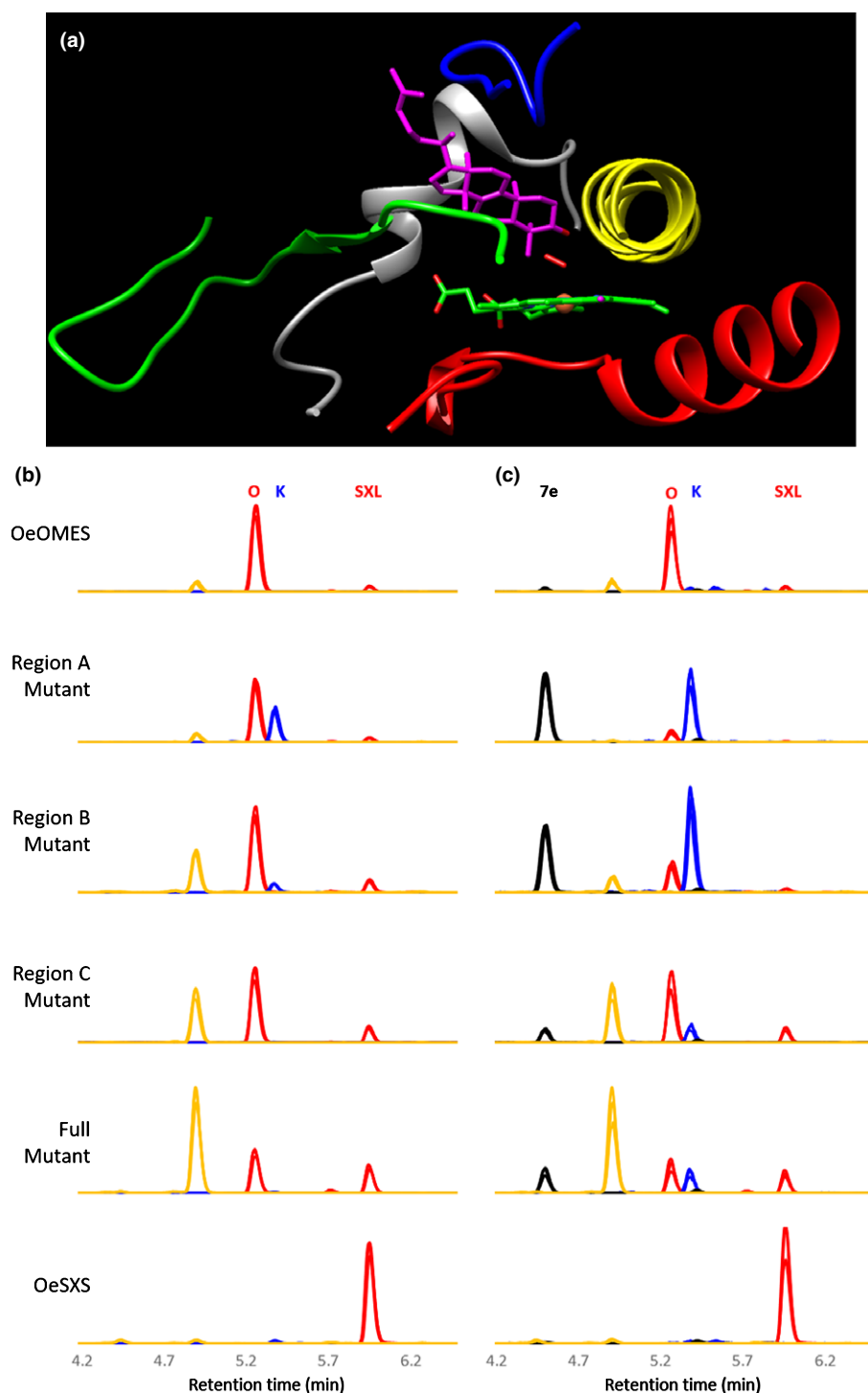


Fig. 4 Extracted ion chromatograms (XICs) of biochemical assays of oleoside methyl ester synthase (OeOMES) mutants. (a) Homology model of OeOMES, using the backbone of lanosterol 14- α demethylase (L14D; PDB 4LXJ; Monk *et al.*, 2014) as a guide, showing the original ligands. Only regions of interest are shown, in colors: gray (region A), yellow (region B), green (region C), red (region D) and blue (region E). (b, c) XICs of the microsomal incubations with ketologanin (b) and 7-*epi*-loganin (c) are shown for selected mutants. Four channels are depicted, corresponding to the most abundant adduct of oleoside methyl ester (OME) and secoxyloganin (red, $[M-H]^- = 403.1240 \pm 0.05$), ketologanin (blue, $[M + FA-H]^- = 433.1346 \pm 0.05$), 7-*epi*-loganin (black, $[M + FA-H]^- = 435.1503 \pm 0.05$), and what we hypothesize to be the formic acid adduct of $C_{17}H_{24}O_{11}$ (yellow, $[C_{17}H_{24}O_{11} + FA-H]^- = 449.1295 \pm 0.05$). Intensities are scaled to the highest intensity of the corresponding channel in all incubations. Three replicates are shown, overlaid, for each mutant. O, OME; K, ketologanin; SXL, secoxyloganin.

Although not all of these oxidation reactions may be physiologically relevant, they can shed light on the mechanism of these enzymes. We speculate that the stereochemistry of the hydroxyl group in position 7 of loganin determines the orientation of the substrate in the CrSLS binding site, such that the hydrogen of C10 can be abstracted by the iron cofactor, which ultimately leads to secologanin (Yamamoto *et al.*, 2000; Fig. 7a). Formation of secoxyloganin by OeSXS could follow a similar mechanism

(Fig. 7b), with the abstraction of the hydrogen of C10 ultimately leading to the ring opening. Conversely, when 7-deoxy-loganin binds to CrSLS, we hypothesize that the substrate binds such that the hydrogen of carbon 7 reacts with the iron cofactor, leading to the formation of loganin (7*S*-OH) (Fig. S17b). Similarly, it appears that 7-*epi*-loganin (7*R*-OH) also binds with the C7 H available to the iron cofactor, resulting in formation of ketologanin (Fig. S17c).

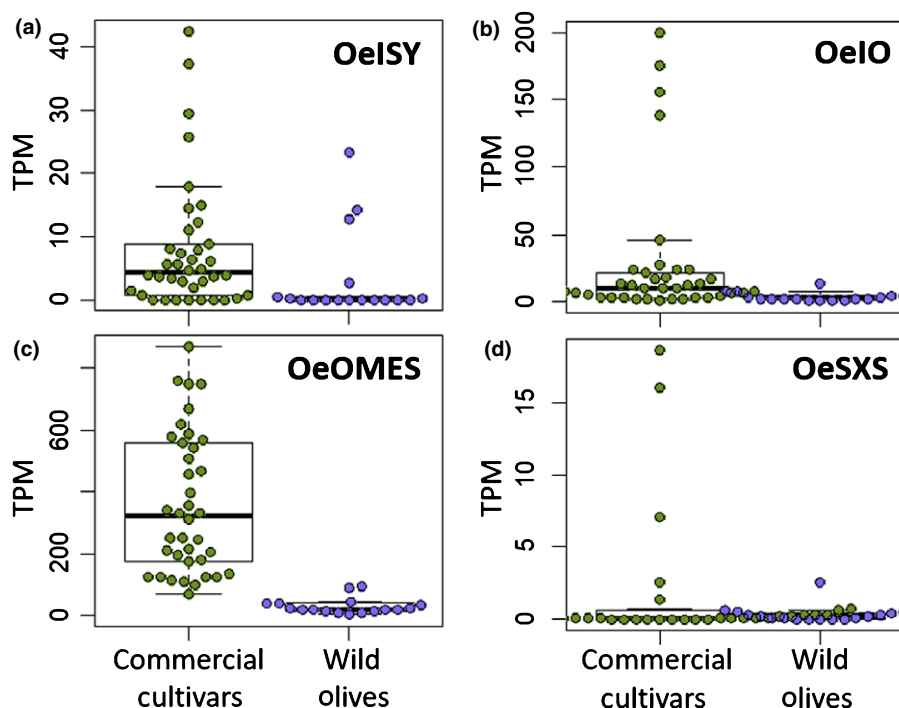
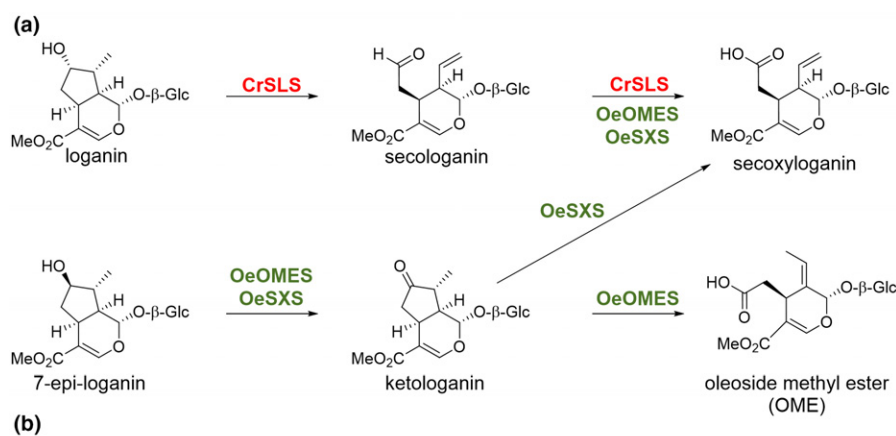


Fig. 5 Expression levels of relevant transcripts. (a–d) Swarmed boxplots are shown of the expression levels in inflorescences, in transcript per million (TPM), of 38 commercial cultivars (olive green) and 17 wild olives (light blue), of iridoid synthase (a), iridoid oxidase (b), oleoside methyl ester synthase (c) and secoxyloganin synthase (d). Raw data were obtained from Gros-Balthazard *et al.* (2019) via the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA). The box delimits the lower (25%) and upper (75%) quartiles, with the internal horizontal line depicting the median; the whiskers represent the 'reasonable extremes': $1.5 \times$ the interquartile range or, if no data-point exceeds this value, the minimum or maximum of the data points for the upper and lower whiskers. Each circle represents a cultivar, with the y-axis being the expression in TPM, and the distribution in the x-axis having no particular meaning, other than to avoid overlaps



(b)

Feeding →	Loganin	Secologanin	Ketologanin	7-Epiloganin
CrSLS	Secologanin	Secoxyloganin	X	Ketologanin
OeSLS1	X	X	X	Ketologanin (traces)
OeOMES (OeSLS2)	X	Secoxyloganin	OME	Ketologanin, OME
OeSXS (OeSLS3)	X	Secoxyloganin	Secoxyloganin	Ketologanin, Secoxyloganin

Fig. 6 Summary of reactions producing biologically relevant metabolites. A graphical summary of reactions leading to biologically relevant metabolites is shown (a), along with a table (b) summarizing the products of each enzyme, when incubated with different substrates. OeOMES, *Olea europaea* oleoside methyl ester synthase; CrSLS, *Catharanthus roseus* secoxyloganin synthase; OeSXS, *O. europaea* oleoside methyl ester secoxyloganin synthase.

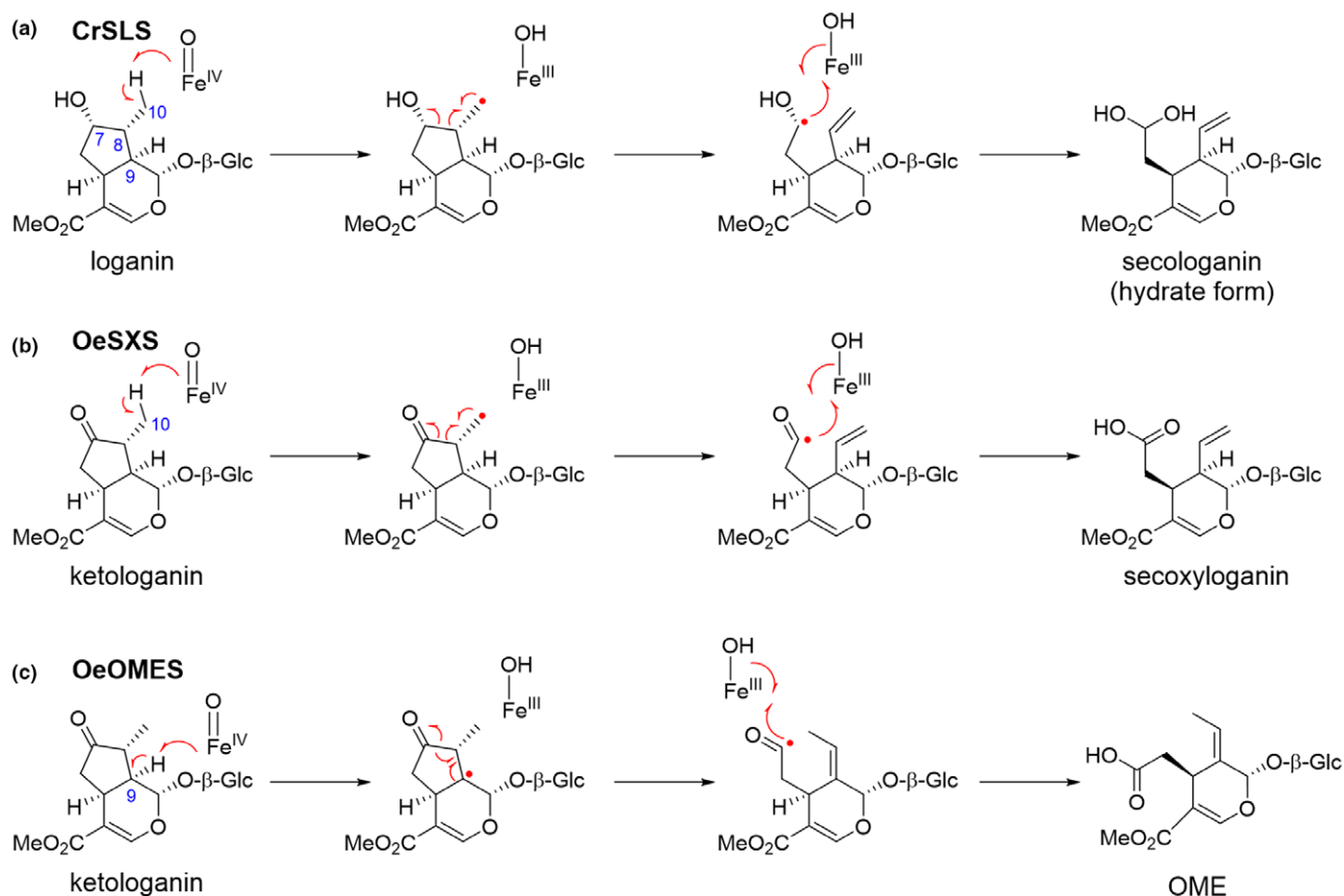


Fig. 7 Mechanistic scenarios of secoiridoid formation. (a–c) A summary of the proposed mechanistic scenarios of secologanin formation by *Catharanthus roseus* secologanin synthase (CrSLS) (a), secoxyloganin formation by *Olea europaea* secoxyloganin synthase (OeSXS) (b), and oleoside methyl ester (OME) formation by *O. europaea* OME synthase (OeOMES) (c).

It is likely that the oxidative cleavage of ketologanin by OeOMES follows a different mechanism from CrSLS. One mechanistic scenario would involve the abstraction of the hydrogen from C8 followed by rearrangement of the radical to C9, which could then open to form OME. Alternatively, the hydrogen of C9 could be abstracted directly (Fig. 7c). Notably, some of the mutants accumulated a side product, which, although they could not be isolated in sufficient quantity, displayed an *m/z* corresponding to a hydroxylated ketologanin derivative. We hypothesize that in these mutants (regions B and C; Fig. 4) one of these radical intermediates is quenched with a hydroxyl moiety to form this side product.

The mutational study of OeOMES provides some insights into which regions of the enzyme control substrate binding. Swaps of region A and region E led to a significant decrease in the oxidation, so these may be the major regions responsible for substrate coordination. Similarly, amino acid shifts in regions B and C caused a considerable increase in secoxyloganin production, revealing the role of these regions in product selectivity, possibly by modulating the position of the substrate relative to the activated iron species. The combination of all mutants, 22 amino acid shifts in total, led to a higher production of secoxyloganin

and a lower production of OME, probably a product of a synergy between the regions discussed.

Remarkably, inflorescences of commercial cultivars of olive show higher expression of genes in the oleuropein biosynthesis pathway, especially OeOMES, when compared with their wild counterparts. It has been hypothesized that domestication of olive was driven by gene expression, rather than genomic changes (Gros-Balthazard *et al.*, 2019). It has also been reported that secoiridoid contents are mainly controlled by the cultivar genotype (Miho *et al.*, 2018; Pérez *et al.*, 2018; Deiana *et al.*, 2019). Notably, published transcriptomic datasets for olive have shown that OeOMES expression is not substantially increased by environmental stresses, but is more highly expressed in domesticated vs wild olives.

For thousands of years, olive trees have been used to produce olive oil, and secoiridoids are a major effector in taste and quality. The expression profiles of OeOMES in the extensive transcriptomic data for olive points to a connection between domestication and iridoid metabolism. These newly identified genetic determinants of secoiridoid metabolism are of great value for breeding programs and can help to develop next-generation cultivars with ‘fine-tuned’ nutritional and taste characteristics.

Acknowledgements

Our work was supported by the Alexander von Humboldt Foundation (BH), the Max-Planck-Gesellschaft and the Mint Consortium (CER-L and SEO'C). Further support came from the European Union's Horizon 2020 Research and Innovation Program Marie Skłodowska-Curie – Before Project (grant agreement no. 645595; to OC, LB, VP and KK), the Italian National Research Council, Short-Term Mobility Program (to OC), the Rural Development Program of Umbria Region, 2014-2020 – Measure 16.2.1, Project INNO.V.O. (to OC, LB and VP), and MIUR- DM no. 856 del 10-10-2019 – Riparto Fondo Ordinario Enti di Ricerca – Anno 2019 – Assegnazione CNR per Progettualità di carattere Straordinario – Progetto 'Economia Circolare' (Green & Circular Economy – GECE) (OC and LB).

Author contributions

CER-L, OC and SEO'C designed experiments and wrote the manuscript; CER-L and OC identified, cloned and characterized enzymes; CER-L designed all mutants and performed all kinetic analyses; BH synthesized all standards and substrates; CP and YN determined the absolute configuration of the OME standard; KK, VP, LB and FA provided plant material and gene sequences and assisted with transcriptome analysis.

ORCID

Fiammetta Alagna  <https://orcid.org/0000-0003-1878-2023>
Luciana Baldoni  <https://orcid.org/0000-0002-6636-0055>
Ornella Calderini  <https://orcid.org/0000-0001-9516-8997>
Konstantinos Koudounas  <https://orcid.org/0000-0002-6000-6565>
Sarah E. O'Connor  <https://orcid.org/0000-0003-0356-6213>
Christian Paetz  <https://orcid.org/0000-0002-5776-7574>
Valentina Passeri  <https://orcid.org/0000-0002-3744-2246>
Carlos E. Rodríguez-López  <https://orcid.org/0000-0001-9224-0455>

Data availability

The data that support the findings of this study are publicly available or available from the corresponding author upon reasonable request.

References

- FAO – Food and Agriculture Organization of the United Nations. 2018. Olive – Crop Production. [WWW document] URL <http://www.fao.org/faostat/en/#search/olive%20production> [accessed 12 September 2020].
- European Commission. 2020. Olive oil: An overview of the production and marketing of olive oil in the EU. [WWW document] URL https://ec.europa.eu/info/food-farming-fisheries/plants-and-plant-products/plant-products/olive-oil_en [accessed 22 July 2020].
- Alagna F, D'Agostino N, Torchia L, Servili M, Rao R, Pietrella M, Giuliano G, Chiusano ML, Baldoni L, Perrotta G. 2009. Comparative 454 pyrosequencing of transcripts from two olive genotypes during fruit development. *BMC Genomics* 10: 399.
- Alagna F, Geu-Flores F, Kries H, Panara F, Baldoni L, O'Connor SE, Osbourn A. 2016. Identification and characterization of the iridoid synthase involved in oleuropein biosynthesis in olive (*Olea europaea*) fruits. *Journal of Biological Chemistry* 291: 5542–5554.
- Amiot MJ, Fleuriot A, Macheix JJ. 1986. Importance and evolution of phenolic compounds in olive during growth and maturation. *Journal of Agriculture and Food Chemistry* 34: 823–826.
- Bavaro SL, Susca A, Frisvad JC, Tufariello M, Chytiri A, Perrone G, Mita G, Logrieco AF, Bleve G. 2017. Isolation, characterization, and selection of molds associated to fermented black table olives. *Frontiers in Microbiology* 8: 1356.
- Carmona R, Zafra A, Seoane P, Castro AJ, Guerrero-Fernández D, Castillo-Castillo T, Medina-García A, Cánovas FM, Aldana-Montes JF, Navas-Delgado I *et al.* 2015. ReprOlive: a database with linked data for the olive tree (*Olea europaea* L.) reproductive transcriptome. *Frontiers in Plant Science* 6: 625.
- Chambers CM, Maclean B, Burke R, Amodi D, Ruderman LD, Neumann S, Gatto L, Fischer B, Pratt B, Egerton J *et al.* 2012. A cross-platform toolkit for mass spectrometry and proteomics. *Nature Biotechnology* 30: 918–920.
- Cruz F, Julca I, Gómez-Garrido J, Loska D, Marcet-Houben M, Cano E, Galán B, Frias L, Ribeca P, Derdak S *et al.* 2016. Genome sequence of the olive tree, *Olea europaea*. *GigaSci* 5: 29.
- Damtoft S, Franzzyk H, Jensen SR. 1993. Biosynthesis of secoiridoid glucosides in Oleaceae. *Phytochemistry* 34: 1291–1299.
- Dang TTT, Franke J, Tatsis E, O'Connor SE. 2017. Dual catalytic activity of a cytochrome P450 controls bifurcation at a metabolic branch point of alkaloid biosynthesis in *Rauwolfia serpentina*. *Angewandte Chemie (International ed. in English)* 56: 9440–9444.
- de Bernonville TD, Foureau E, Parage C, Lanoue A, Clastre M, Arias Londono M, Oudin A, Houllé B, Papon N, Besseau S *et al.* 2015. Characterization of a second secologanin synthase isoform producing both secologanin and secoxyloganin allows enhanced *de novo* assembly of a *Catharanthus roseus* transcriptome. *BMC Genomics* 16: 619.
- de la Torre R, Carbó M, Pujadas M, Biel S, Mesa MD, Covas MI, Expósito M, Espejo JA, Sanchez-Rodriguez E, Díaz-Pellicer P *et al.* 2020. Pharmacokinetics of maslinic and oleonic acids from olive oil - effects on endothelial function in healthy adults. A randomized, controlled, dose-response study. *Food Chemistry* 322: 126676.
- De Marino S, Festa C, Zollo F, Nini A, Antenucci L, Raimo G, Iorizzi M. 2014. Antioxidant activity and chemical components as potential anticancer agents in the olive leaf (*Olea europaea* L. cv Leccino) decoction. *Anti-Cancer Agents in Medicinal Chemistry* 14: 1376–1385.
- Deiana P, Santona M, Dettori S, Culeddu N, Dore A, Molinu MG. 2019. Multivariate approach to assess the chemical composition of Italian virgin olive oils as a function of variety and harvest period. *Food Chemistry* 300: 125243.
- Eklund A. 2016. beeswarm: The Bee Swarm Plot, an Alternative to Stripchart. [R package] <https://CRAN.R-project.org/package=beeswarm> [accessed 20 July 2020 – v0.2.3].
- Garrido-Fernández A, Adams MR, Fernández-Díez MJ. 1997. *Table olives: production and processing*. London, UK: Chapman & Hall.
- George ES, Marshall S, Mayr HL, Trakman GL, Tatucu-Babet OA, Lasemillante ACM, Bramley A, Reddy AJ, Forsyth A, Tierney AC *et al.* 2019. The effect of high-polyphenol extra virgin olive oil on cardiovascular risk factors: a systematic review and meta-analysis. *Critical Reviews in Food Science and Nutrition* 59: 2772–2795.
- Gros-Balthazard M, Besnard G, Sarah G, Holtz Y, Leclercq J, Santoni S, Wegmann D, Glémin S, Khadari B. 2019. Evolutionary transcriptomics reveals the origins of olives and the genomic changes associated with their domestication. *The Plant Journal* 100: 143–157.
- Gross G, Sticher O, Anklin C. 1986. Ein neues Esteriridoidglycosid aus *Sambucus ebulus* L. (Caprifoliaceae). 1. Mitteilung über die inhaltsstoffe der Zwergholunderwurzel. *Helvetica Chimica Acta* 69: 156–162.
- Inoue K, Ono M, Nakajima H, Fijie I, Inouye H, Fujita T. 1992. Radioimmunoassay of iridoid glucosides: part 1. General method for preparation of the haptens and the conjugates with a protein of this series of glucosides. *Heterocycles* 33: 673–695.

- Inouye H, Ueda S, Inoue K, Takeda Y. 1971. Über die biosynthese der oleuropein-typ-secoiridoidglucoside der Oleaceae. *Tetrahedron Letters* 12: 4073–4076.
- Inouye H, Yoshida T, Tobita S, Okigawa M. 1970. Studies on monoterpene glucosides—IX: Chemical correlation between asperuloside and loganin. *Tetrahedron* 26: 3905–3915.
- Irmeler S, Schröder G, St-Pierre B, Crouch NP, Hotze M, Schmidt J, Strack D, Matern U, Schröder J. 2000. Indole alkaloid biosynthesis in *Catharanthus roseus*: new enzyme activities and identification of cytochrome P450 CYP72A1 as secologanin synthase. *The Plant Journal* 24: 797–804.
- Itoh A, Kumashiro T, Yamaguchi M, Nagakura N, Mizushima Y, Nishi T, Tanahashi T. 2005. Indole alkaloids and other constituents of *Rauwolfia serpentina*. *Journal of Natural Products* 68: 848–852.
- Jiménez-Ruiz J, Leyva-Pérez MO, Gómez-Lama Cabanás C, Barroso JB, Luque F, Mercado-Blanco J. 2019. The transcriptome of *Verticillium dahliae* responds differentially depending on the disease susceptibility level of the olive (*Olea europaea* L.) cultivar. *Genes* 10: 251.
- Jiménez-Ruiz J, Leyva-Pérez MO, Vidoy-Mercado I, Barceló A, Luque F. 2018. Transcriptomic time-series analysis of early development in olive from germinated embryos to juvenile tree. *BMC Genomics* 19: 824.
- Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. 2015. The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols* 10: 845–858.
- Konno K, Hirayama C, Yasui H, Nakamura M. 1999. Enzymatic activation of oleuropein: A protein crosslinker used as a chemical defense in the privet tree. *Proceedings of the National Academy of Sciences, USA* 96: 9159–9164.
- Koudounas K, Banilas G, Michaelidis C, Demoliou C, Rigas S, Hatzopoulos P. 2015. A defence-related *Olea europaea* β -glucosidase hydrolyses and activates oleuropein into a potent protein cross-linking agent. *Journal of Experimental Botany* 66: 2093–106.
- Koudounas K, Thomopoulou M, Michaelidis C, Zevgiti E, Papakostas G, Tserou P, Daras G, Hatzopoulos P. 2017. The C-domain of oleuropein β -glucosidase assists in protein folding and sequesters the enzyme in nucleus. *Plant Physiology* 174: 1371–1383.
- Letunic I, Bork P. 2019. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Research* 47: W256–W259.
- Leyva-Pérez MO, Valverde-Corredor A, Valderrama R, Jiménez-Ruiz J, Muñoz-Merida A, Trelles O, Barroso JB, Mercado-Blanco J, Luque F. 2015. Early and delayed long-term transcriptional changes and short-term transient responses during cold acclimation in olive leaves. *DNA Research* 22: 1–11.
- Li Z, Natarajan P, Ye Y, Hrabec T, Godzik A. 2014. POSA: a user-driven, interactive multiple protein structure alignment server. *Nucleic Acids Research* 42: W240–W245.
- Lichman BR, Godden GT, Hamilton JP, Palmer L, Kamileen MO, Zhao D, Vaillancourt B, Wood JC, Sun M, Kinser TJ *et al.* 2020. The evolutionary origins of the cat attractant nepetalactone in catnip. *Science Advances* 6: eaba0721.
- Lichman BR, Kamileen MO, Titchiner GR, Saalbach G, Stevenson CEM, Lawson DM, O'Connor SE. 2019. Uncoupled activation and cyclization in catmint reductive terpenoid biosynthesis. *Nature Chemical Biology* 15: 71–79.
- Madeira F, Park YM, Lee J, Buso N, Gur T, Madhusoodanan N, Basutkar P, Tivey ARN, Potter SC, Finn RD *et al.* 2019. The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Research* 47: W636–W641.
- Marrero AD, Castilla L, Espartero JL, Madrona A, Quesada AR, Medina MA, Martínez-Poveda B. 2020. A comparative study of the antiangiogenic activity of hydroxytyrosyl alkyl ethers. *Food Chemistry* 333: 127476.
- Michaelis L, Menten ML. 1913. Die Kinetik der Invertinwirkung. *Biochem Z* 49: 333–369.
- Miettinen K, Dong L, Navrot N, Schneider T, Burlat V, Pollier J, Woittiez L, van der Krol S, Lugin R, Ilc T *et al.* 2014. The seco-iridoid pathway from *Catharanthus roseus*. *Nature Communications* 5: 3606.
- Miho H, Díez CM, Mena-Bravo A, de Medina VS, Moral J, Melliou E *et al.* 2018. Cultivar influence on variability in olive oil phenolic profiles determined through an extensive germplasm survey. *Food Chemistry* 266: 192–199.
- Monk BC, Tomasiak TM, Keniya MV, Huschmann FU, Tyndall JD, O'Connell JD, Cannon RD, McDonald JG, Rodriguez A, Finer-Moore JS *et al.* 2014. Architecture of a single membrane spanning cytochrome P450 suggests constraints that orient the catalytic domain relative to a bilayer. *Proceedings of the National Academy of Sciences, USA* 111: 3865–3870.
- Nakamura M, Kido K, Kinjo J, Nohara T. 2000. Antinociceptive substances from *Incarvillea delavayi*. *Phytochemistry* 53: 253–256.
- Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. 2017. Salmon provides fast and bias-aware quantification of transcript expression. *Nature Methods* 14: 417–419.
- Pérez AG, León L, Sanz C, de la Rosa R. 2018. Fruit phenolic profiling: a new selection criterion in olive breeding programs. *Frontiers in Plant Science* 9: 241.
- Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF Chimera – a visualization system for exploratory research and analysis. *Journal of Computational Chemistry* 25: 1605–1612.
- Pluskal T, Castillo S, Villar-Briones A, Orešič M. 2010. MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics* 11: 395.
- R Core Team. 2019. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Ro D-K, Ehrling J, Douglas CJ. 2002. Cloning, functional expression, and subcellular localization of multiple NADPH-cytochrome P450 reductases from hybrid poplar. *Plant Physiology* 130: 1837–1851.
- Ro D-K, Ouellet M, Paradise EM, Burd H, Eng D, Paddon CJ, Newman JD, Keasling JD. 2008. Induction of multiple pleiotropic drug resistance genes in yeast engineered to produce an increased level of anti-malarial drug precursor, artemisinic acid. *BMC Biotechnology* 8: 83.
- Ryan D, Robards K, Lavee S. 1999. Changes in phenolic content of olive during maturation. *International Journal of Food Science & Technology* 34: 265–274.
- Salim V, Yu F, Altarejos J, De Luca V. 2013. Virus-induced gene silencing identifies *Catharanthus roseus* 7-deoxyloganic acid-7-hydroxylase, a step in iridoid and monoterpene indole alkaloid biosynthesis. *The Plant Journal* 76: 754–65.
- Schwingshackl L, Krause M, Schmucker C, Hoffmann G, Rücker G, Meerpohl JJ. 2019. Impact of different types of olive oil on cardiovascular risk factors: a systematic review and network meta-analysis. *Nutrition, Metabolism and Cardiovascular Diseases* 29: 1030–1039.
- Serrilli AM, Maggi A, Casagrande V, Bianco A. 2016. Synthesis of deuterium-labelled substrates for the study of oleuropein biosynthesis in *Olea europaea* callus cultures. *Natural Product Research* 30: 926–934.
- The UniProt Consortium. 2019. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Research* 47: D506–D515.
- Tierney L, Rossini AJ, Li N, Sevcikova H. 2018. snow: Simple Network of Workstations. [R package] <https://CRAN.R-project.org/package=snow> [accessed 09 July 2019 – v0.4.3].
- Trifinopoulos J, Nguyen LT, von Haeseler A, Minh BQ. 2016. W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Research* 44: W232–W235.
- Unver T, Wu Z, Sterck L, Turktaş M, Lohaus R, Li Z, Yang M, He L, Deng T, Escalante FJ *et al.* 2017. Genome of wild olive and the evolution of oil biosynthesis. *Proceedings of the National Academy of Sciences, USA* 114: E9413–E9422.
- Velázquez-Palmero D, Romero-Segura C, García-Rodríguez R, Hernández ML, Vaistij FE, Graham IA, Pérez AG, Martínez-Rivas JM. 2017. An oleuropein β -glucosidase from olive fruit is involved in determining the phenolic composition of virgin olive oil. *Frontiers in Plant Science* 8: 1902.
- Weigend M, Kufer J, Müller AA. 2000. Phytochemistry and the systematics and ecology of Loasaceae and Gronoviaceae (Loasales). *American Journal of Botany* 87: 1202–1210.
- Yamamoto H, Katano N, Ooi A, Inoue K. 2000. Secologanin synthase which catalyzes the oxidative cleavage of loganin into secologanin is a cytochrome P-450. *Phytochemistry* 53: 7–12.
- Yang Y, Li W, Pang J, Jiang L, Qu X, Pu X, Zhang G, Luo Y. 2019. Bifunctional cytochrome P450 enzymes involved in camptothecin biosynthesis. *ACS Chemical Biology* 14: 1091–1096.

Zhang H, Shi T, Wang J, Li R, Tang W. 2013. Protective effect of penta-acetyl geniposide on acute liver injury induced by D-galactosamine in mice. *British Journal of Pharmacology* 4: 256–261.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

Fig. S1 Previous hypotheses on the biosynthesis of secoiridoids in olive.

Fig. S2 Mapping of publicly available datasets against available transcriptomes.

Fig. S3 NMR spectra of the oleoside methyl ester standard.

Fig. S4 Confirmation of the oleoside methyl ester configuration.

Fig. S5 NMR spectra of 7-*epi*-8-*epi*-loganin.

Fig. S6 COSY and NOESY spectra of 7-*epi*-8-*epi*-loganin.

Fig. S7 Extracted ion chromatograms (XIC) of microsome incubations with secologanin and loganin.

Fig. S8 Extracted ion chromatograms (XIC) of microsome incubations with a mix of loganin epimers.

Fig. S9 Extracted ion chromatograms of microsome incubations with 7-deoxy-loganin.

Fig. S10 Kinetic modeling of enzyme activity.

Fig. S11 Individual enzyme activities.

Fig. S12 Regions of interest in the homology model.

Fig. S13 Extracted ion chromatograms (XIC) of microsome incubations of mutant proteins with ketologanin and 7-*epi*-loganin.

Fig. S14 Extracted ion chromatograms (XIC) of microsome incubations of mutant proteins with secologanin.

Fig. S15 Coexpression analysis of olive secologanin synthase homologs with iridoid synthase.

Fig. S16 Effect of different stresses on OeISY and OeOMES expression.

Fig. S17 Proposed mechanisms of CrSLS catalyzed reactions.

Methods S1 Standard compound characterization.

Methods S2 Chemical synthesis of iridoid standards.

Methods S3 NMR spectroscopical data for 7-*epi*-8-*epi*-loganin.

Notes S1 Steady-state kinetics of a coupled reaction.

Notes S2 Steady-state kinetics of a dissociative, sequential reaction.

Table S1 Amino acid shifts in each of the selected regions.

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.