





Article

# Tuscan Varieties of Sweet Cherry Are Rich Sources of Ursolic and Oleanolic Acid: Protein Modeling Coupled to Targeted Gene Expression and Metabolite Analyses

Roberto Berni <sup>1,2</sup>, Mubasher Zahir Hoque <sup>3,4</sup>, Sylvain Legay <sup>5</sup>, Giampiero Cai <sup>1</sup>,  
Khawar Sohail Siddiqui <sup>4,\*</sup>, Jean-Francois Hausman <sup>5</sup>, Christelle M. Andre <sup>5</sup>  
and Gea Guerriero <sup>5,\*</sup>

<sup>1</sup> Department of Life Sciences, University of Siena, via P.A. Mattioli 4, 53100 Siena, Italy; berni10@student.unisi.it (R.B.); giampiero.cai@unisi.it (G.C.)

<sup>2</sup> Trees and Timber Institute, National Research Council of Italy (CNR-IVALSA), via Aurelia 49, 58022 Follonica (GR), Italy

<sup>3</sup> Bio-Bio-1 Research Foundation, Sangskriti Bikash Kendra Bhaban, 1/E/1 Poribagh, Dhaka 1000, Bangladesh; mubasher\_zahir856@hotmail.com

<sup>4</sup> Life Sciences Department, King Fahd University of Petroleum and Minerals (KFUPM), Dhahran 31261, Saudi Arabia

<sup>5</sup> Environmental Research and Innovation Department, Luxembourg Institute of Science and Technology, 5 avenue des Hauts-Fourneaux, L-4362 Esch/Alzette, Luxembourg; sylvain.legay@list.lu (S.L.); jean-francois.hausman@list.lu (J.-F.H.); andre.christellem@gmail.com (C.M.A.)

\* Correspondence: sohailsiddiqui1995@yahoo.com (K.S.S.); gea.guerriero@list.lu (G.G.)

Received: 11 March 2019; Accepted: 20 April 2019; Published: 22 April 2019



**Abstract:** The potential of six ancient Tuscan sweet cherry (*Prunus avium* L.) varieties as a source of health-promoting pentacyclic triterpenes is here evaluated by means of a targeted gene expression and metabolite analysis. By using a sequence homology criterion, we identify five oxidosqualene cyclase genes (OSCs) and three cytochrome P450s (CYP85s) that are putatively involved in the triterpene production pathway in sweet cherries. We performed 3D structure prediction and induced-fit docking using cation intermediates and reaction products for some OSCs to predict their function. We show that the Tuscan varieties have different amounts of ursolic and oleanolic acids and that these variations are related to different gene expression profiles. This study stresses the interest of valorizing ancient fruits as alternative sources of functional molecules with nutraceutical value. It also provides information on sweet cherry triterpene biosynthetic genes, which could be the object of follow-up functional studies.

**Keywords:** *Prunus avium* L.; Tuscan varieties; pentacyclic triterpenes; oxidosqualene cyclase; cytochrome P450; gene expression; bioinformatics

## 1. Introduction

The fruits of *Prunus avium* L. are rich sources of health-promoting compounds [1–3]. Among the beneficial phytochemicals are polyphenols [4,5] and triterpenes [6]. The latter compose the cuticle of sweet cherry fruits and, more specifically, they are found almost exclusively associated with the intracuticular waxes [7,8]. Triterpenes attract a lot of interest in the nutraceutical and pharmaceutical sectors, in virtue of their beneficial impact on human health [9].

The consumption of triterpene-rich fruits and vegetables has indeed a positive correlation with the prevention/decreased susceptibility to develop chronic diseases [7]. Among triterpenes, pentacyclic

triterpenes (and their derivatives) are produced and commercialized as dietary supplements or therapeutic agents [9].

The enzymes involved in the synthesis of the three major classes of plant pentacyclic triterpene acids (ursolic, oleanolic and betulinic acid deriving from  $\alpha$ -amyrin,  $\beta$ -amyrin and lupeol, respectively) and their structural modifications are the oxidosqualene cyclases (OSCs, or triterpene synthases) and cytochrome P450 (CYP85s) [10,11]. Numerous OSCs have been characterized in plants but, to the best of our knowledge, no information is available on those from sweet cherries. Indeed, only limited knowledge is available regarding the triterpene profile of sweet cherry fruits. Two studies reported that the major pentacyclic triterpene acids found in the cuticle of sweet cherry fruits at 85 DAFB (days after full bloom) are ursolic (60%) and oleanolic acid (7.5%) [6], with concentrations varying according to the cultivars. Peschel et al. (2007) compared four cultivars, i.e., ‘Hedelfinger’, ‘Kordia’, ‘Sam’ and ‘Van’, and reported no qualitative differences in terms of wax-associated components, but highlighted a lower amount of triterpenes in the waxes of the cv. ‘Hedelfinger’ [6]. Additionally, ursolic acid was reported to be the most abundant triterpene in cold-stored sweet cherry fruits of the cv. ‘Celeste’ and ‘Somerset’ and to increase upon cold storage in the former [12].

Italy is an important producer of sweet cherry fruits, with a stable production of around 110,000–120,000 tons over ca. 30,000 ha of orchards [13]. At the regional level, Italy has started conservation programs aimed at preserving the local biodiversity via *ex situ* collections; in this respect Tuscany is at the forefront, with a regional law (law 64/04; [14]) created for the recovery and preservation of local varieties of woody and herbaceous plant species [15]. These varieties are reported in the Regional Bank of Germplasm, which provides, when possible, both phenotypic and genotypic information [16]. All of the reported plant biodiversity comprises ancient varieties, which designate plants that were grown in the past, but that have not been subject to a market-driven pressure, thus falling out of agricultural interest. These varieties can thrive in soils where the human input is minimal and are sources of interesting agronomic characters, notably higher resilience to adverse environmental conditions.

Sustainable agricultural practices based on the valorization of the local biodiversity help prevent habitat loss due to urbanization. Additionally, they provide a boost to the local economy, by promoting the commercialization of products deriving from regional varieties, thereby using a “0 km” concept that reduces the C footprint. The local agrobiodiversity comprises crops (woody and herbaceous), as well as indigenous soil microbiota and its valorization contributes, on a wider perspective, to ecological restoration, i.e., the recovery of the local ecosystem. Such an awareness of the importance that local bioresources have is already reflected in research projects involving local farmers, enterprises and universities and aimed at promoting fully traceable products, from the raw material to the final product [17].

With a view to valorize non-commercial regional plant varieties of Tuscany, we have analyzed the content of ursolic and oleanolic acid in six ancient *P. avium* varieties. We show different levels of pentacyclic triterpenes in the fruits of Tuscan varieties. Furthermore, we identify five OSCs and two CYP85s that are putatively implicated in the sweet cherry triterpene biosynthetic pathway and analyze their expression in the six varieties at the stage commercial harvest (ca. 60 dpa, days post anthesis). Bioinformatics carried out to perform 3D structure prediction and induced-fit docking of cation intermediates, as well as of products from three sweet cherry OSCs complement the analyses.

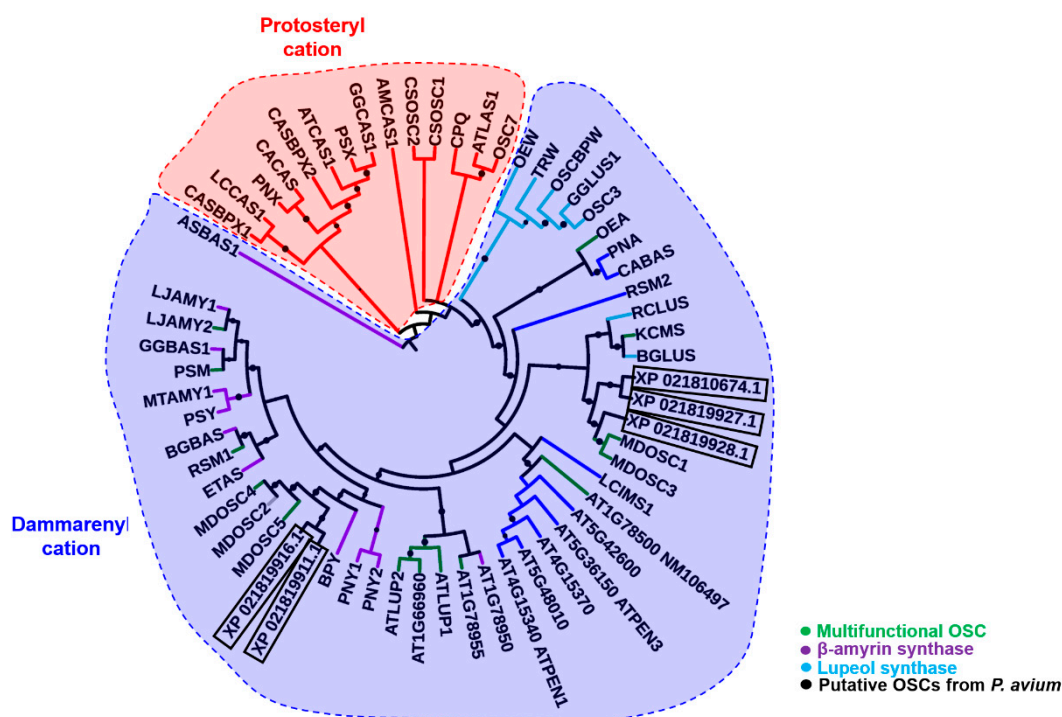
Our results highlight the significant nutraceutical value of ancient Tuscan sweet cherries. On a longer-term perspective, this study sensitizes to the preservation and study of ancient local plant varieties.

## 2. Results and Discussion

### 2.1. Bioinformatics

NCBI BLAST searches of the apple OSC orthologs [10,11] (nucleotide sequences) in the sweet cherry genome [18] led to the identification of five transcripts coding for XP\_021819927.1, XP\_021819916.1, XP\_021819911.1, XP\_021810674.1, XP\_021819928.1 (Supplementary Information). Of the corresponding protein sequences, three are full length putative OSCs, XP\_021819911.1, XP\_021819928.1 and XP\_021819927.1. NCBI BLAST of the apple *CYP716A175*, a known triterpene C-28 oxidase [11], resulted in three transcripts coding for the full-length proteins, XP\_021816761.1, XP\_021823674.1 and XP\_021815663.1 (Supplementary Information).

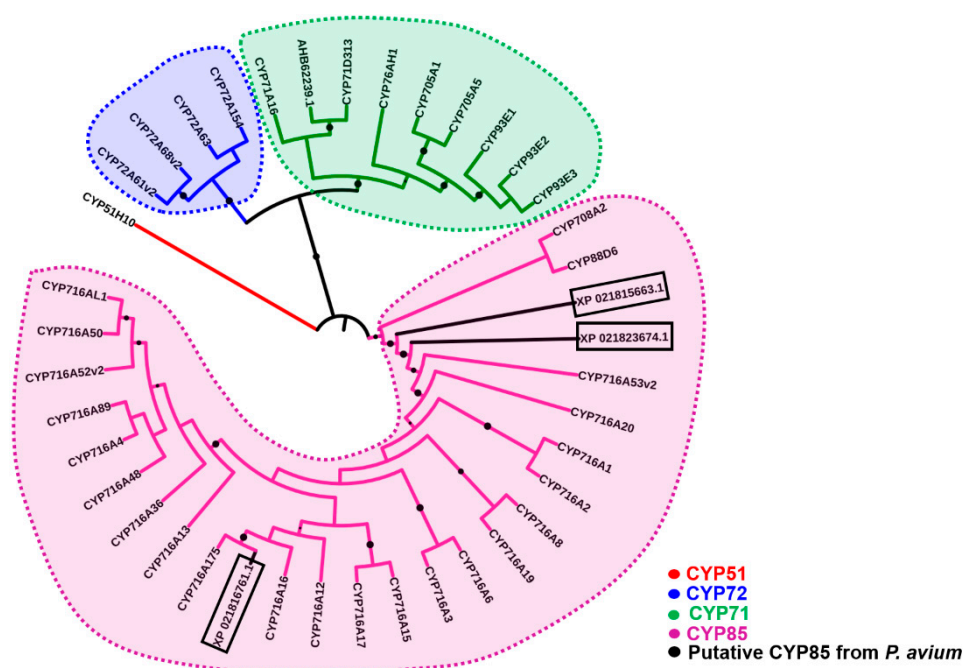
Alignment of the *P. avium* putative OSCs with those from apple revealed the presence of the conserved sequence QX<sub>3</sub>GXW, together with the substrate-binding motif DCTAE [11,19–21] (Figure S1). The maximum likelihood phylogenetic analysis shows that the *P. avium* putative OSCs XP\_021819911.1 and XP\_021819916.1 branch together with the cluster formed by monofunctional  $\beta$ -amyrin synthases (indicated in violet) and the multifunctional apple MdOSC4 (*Malus domestica* oxidosqualene cyclase) and 5 (indicated in green) (Figure 1). The putative OSCs XP\_021810674.1, XP\_021819927.1 and XP\_021819928.1 cluster instead in the clade formed by multifunctional OSCs and lupeol synthases (Figure 1). From the phylogenetic analysis, we can speculate that the identified OSCs belong to the multifunctional type, given their stricter clustering with apple multifunctional orthologs.



**Figure 1.** Maximum likelihood phylogenetic tree (number of bootstrap alignments in ultrafast mode: 1000) of different OSC protein sequences (conserved regions only) from the plant species used by Brendolise and colleagues [10] and from sweet cherry (boxed). The branch with MdOSC2 is in gray, as it may represent a non-functional pseudogene. The FASTA protein sequences used to build the tree are provided in Supplementary Information. Bootstrap values ranging from 0.8 to 1 are displayed as black circles; the bigger the circle, the higher the bootstrap value.

The phylogenetic analysis of the putative sweet cherry CYPs shows their clustering in the clade composed by CYP85s (Figure 2). XP\_021816761.1 clusters together with the recently characterized

apple CYP716A175 [11] and is therefore the closest, in terms of sequence, among the three CYP85s found in sweet cherry.

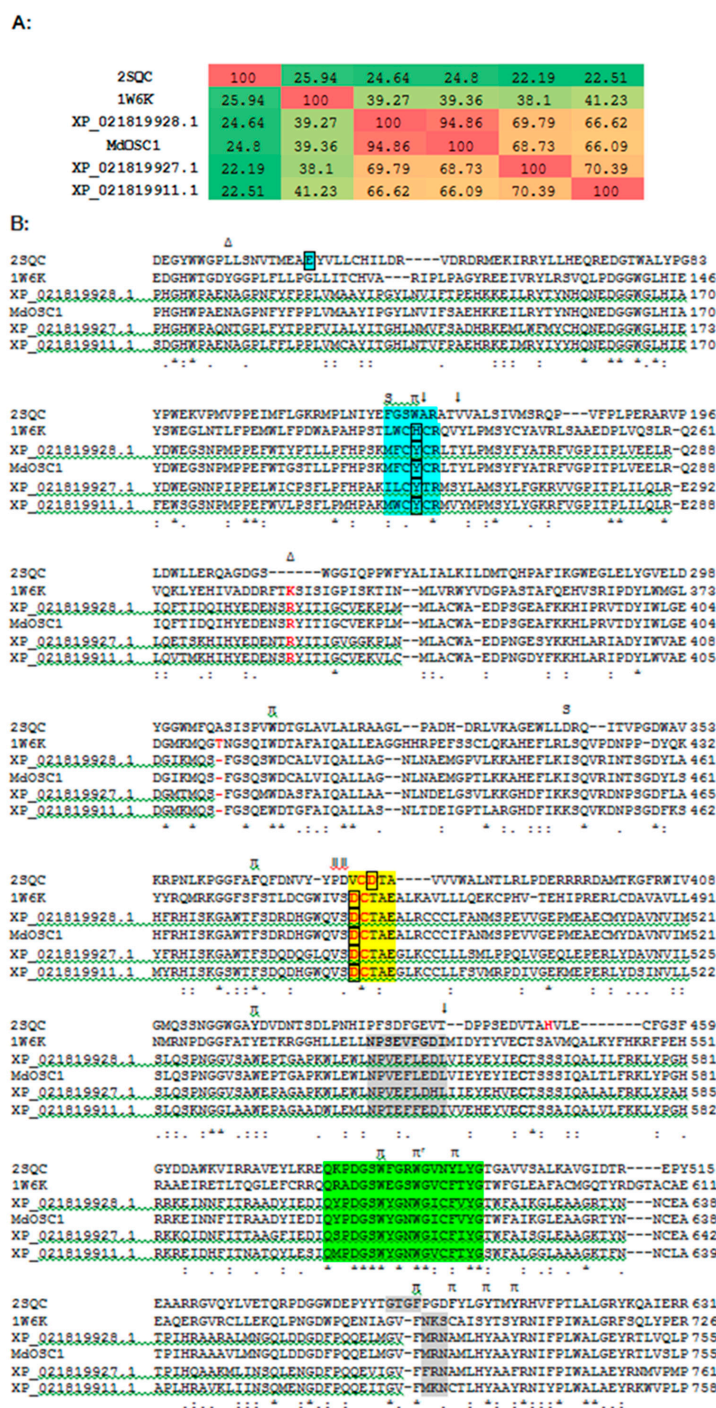


**Figure 2.** Maximum likelihood phylogenetic tree (number of bootstrap alignments in ultrafast mode: 1000) of different CYP protein sequences (conserved regions only) from the plant species used by Andre and colleagues [11] and from sweet cherry (boxed). The FASTA protein sequences used to build the tree are provided in Supplementary Information. Bootstrap values ranging from 0.8 to 1 are displayed as black circles; the bigger the circle, the higher the bootstrap value.

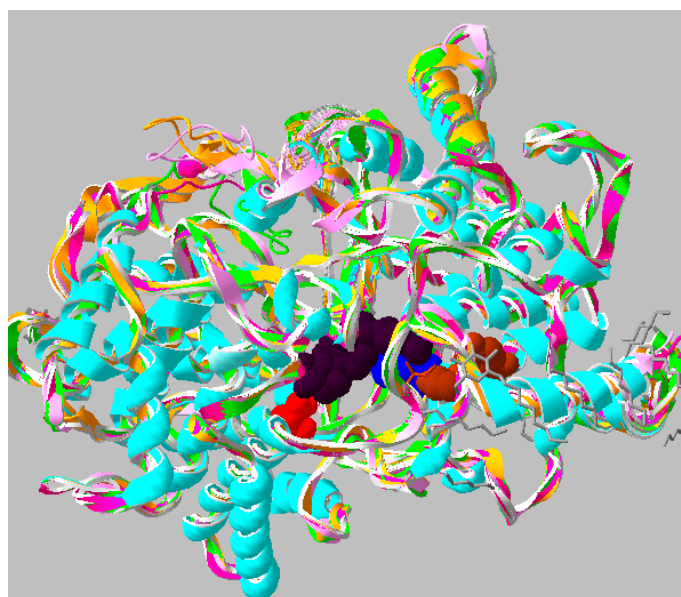
## 2.2. Protein Modeling

In the absence of a plant OSC X-ray structure, fruit (cherry and apple) enzymes were modeled on the human template (PDB 1w6k). Our results indicate that the sequence alignment (Figure 3), overall protein fold (Figure 4) and ligand interacting residues (Figure 5A,B) of fruit enzymes are very similar to the human OSC that produces tetracyclic products, rather than to the bacterial enzyme (1ump) that produces pentacyclic products.

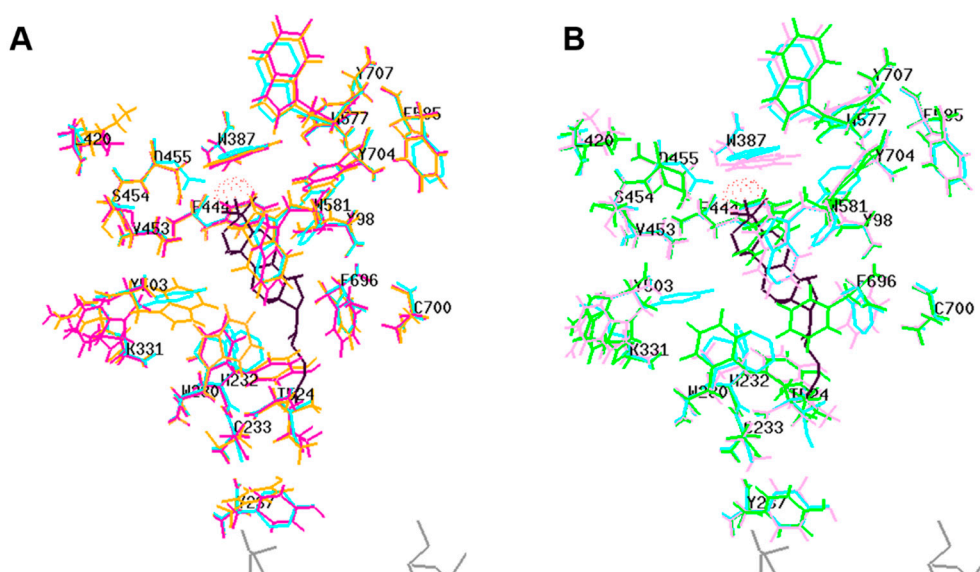
Interestingly, the percentage identity within all three cherry isoforms ranges from 66–70%, whereas the identity between cherry (XP\_021819928.1) and apple (MdOSC1) was 95%. The percentage identity of fruit OSCs with the bacterial and human ones ranges from 22–25% and 38–41%, respectively (Figure 3A).



**Figure 3.** Percentage identity (A) and multiple alignment (B) of OSCs from fruits, bacteria (2sqc) and human (1w6k). Apple, Cherry9911, 9927 and 9928 indicate MdOSC1, XP\_021819911.1, XP\_021819927.1 and XP\_021819928.1, respectively. Cyan highlighted, conserved M(W/Y)CY(C/S)R motif; yellow highlighted, conserved DCTAE motif; green highlighted, conserved Q-X3-G-X-W-X3-W-G motif; grey highlighted, loops surrounding substrate channel; π, FYWH may be involved in carbocation-pi interactions, whereas π<sup>r</sup> is used as binding reference in docking experiments; red boxed, Asp involved in the initial protonation of linear substrate; red Cys/His, H-bonding partners of Asp involved in increasing their acidity; cyan boxed, involved in deprotonation of the cyclic substrate at the end; ↓, substrate entry; Δ, may be involved in steric deformation; ||, may be involved in correct substrate orientation (Asp in 2SQC; Val in all others); S, product specificity. Unimportant sequence blocks have been removed.



**Figure 4.** Homology models of fruit OSCs superimposed on the human crystal structure (1w6k). Cyan, human; yellow, orange and pink, XP\_021819911.1, XP\_021819927.1 and XP\_021819928.1; green, MdOSC1. The product, along with important residues, are shown only for human OSC. Black, lanosterol; blue, H232; brown, substrate entry point (C233, Y237, I524); red, active site D455 involved in protonation of oxidosqualene substrate. Grey,  $\beta$ -octylglucoside showing enzyme regions interacting with hydrophobic membrane-inserting surface.



**Figure 5.** Critical residues of human OSC crystal structure (PDB, 1w6k) superimposed on fruit models. **A:** human (cyan); XP\_021819911.1 (orange) and XP\_021819928.1 (dark pink). **B:** XP\_021819927.1 (light pink) and apple MdOSC1 (green). The ligand (black) is lanosterol. The red dots correspond to hydroxyl group in the substrate that makes H-bond with the catalytic Asp residues (D455 in human). Readers are referred to Figure 4 for a close-up of the interaction between D455 and the ligand. The residue numbering is as in human OSC. Grey,  $\beta$ -octylglucoside showing enzyme region interacting with hydrophobic membrane-inserting surface. The reader should refer to legend of Figure 1B (multiple alignment) for the significance of each residue.

A Tm-score > 0.5 indicates a model of correct topology. The Tm-scores and root-mean-square deviation of atomic positions (RMSD) of fruit models using the human template (1w6k) range from

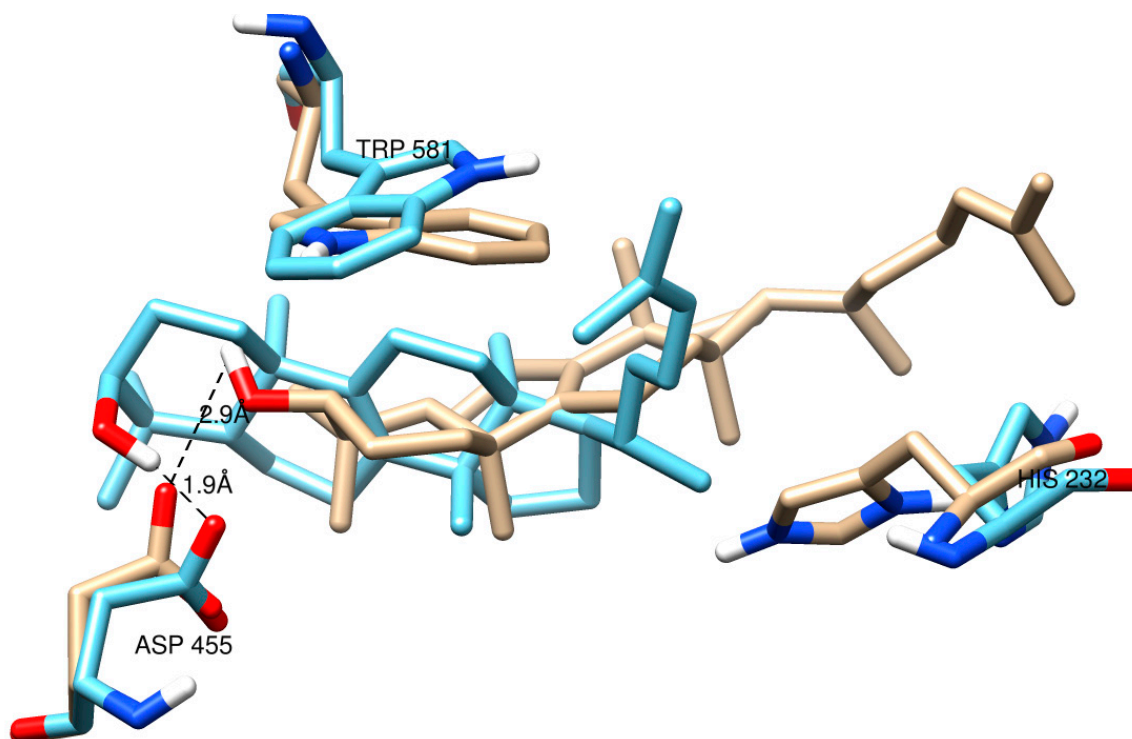
0.93–0.97 and 0.64–1.08 Å, respectively, whereas Tm-scores and RMSD of fruit models using bacterial squalene-hopene cyclase template (1ump, 2sqc) are around 0.75 and 2.9 Å, respectively. Based on the higher sequence identity (Figure 3A) and modeling scores (Tm-score and RMSD) of fruit OSCs with the human template (1w6k), rather than with the bacterial squalene-hopene cyclase (1ump, 2sqc), all the modeling was carried out using the human OSC template. The quality of fruit models based on the Ramachandran plots (not shown) indicate that for all cherry and apple models, 93.1–94.6 and 97% respectively of the residues are found in favored + allowed regions. More importantly, none of the ligand-interacting and catalytically critical residues shown in Figures 3B and 5 are located in the outlier region for any of the fruit models.

The multiple alignment shows various conserved motifs and residues involved in the ligand interaction in fruit, bacterial and human OSCs (Figure 3B). The catalytic Asp (D455 in the human enzyme and equivalent residues in the fruit ones) involved in the protonation of oxidosqualene to form a hydroxyl group is structurally conserved in all fruit and human OSCs (Figure 3B, red boxed; Figure 4, red residue; Figure 5). In the bacterial OSC, it is shifted two residues towards the C-terminus (Figure 3B). Another important catalytic residue is His232 in human OSC (Figure 3B, cyan boxed; Figure 4, blue residue; Figure 5, H232 and equivalent Tyr residues in fruits) that is involved in the deprotonation of the cyclic substrate at the end. While this role has been proposed to be performed by a Glu and/or water molecule in the bacterial OSC [22] (Figure 3B, cyan boxed), it is interesting that in fruit OSCs H232 has been replaced by Tyr residues (Figure 3B, cyan boxed; Figure 5). The pKa of the side chain of Tyr residues is ~10.5, which is expected to further increase inside the hydrophobic active site. This results in the inability of Tyr to be ionized at physiological pH and deprotonate the cation intermediate, thus leading to pentacyclic (lupeol,  $\alpha$ - and/or  $\beta$ -amyrin), rather than tetracyclic products, such as cycloartenol and lanosterol [22–24]. Conserved aromatic residues involved in stabilizing intermediate cations by cation- $\pi$  interactions are also shown, with the noteworthy exception that F605 in bacterial OSC is replaced with Cys or Ala residues in plant and fruit enzymes (Figure 3B). W169 and F605 stabilize the cation intermediate required for the cyclization of the 5th ring in bacterial enzyme [24]. Although W169 is replaced by Tyr residues in fruits, there is no aromatic residue in lieu of F605 in fruit OSCs (Figure 3B).

Minor changes in amino acids have been implicated in major changes in the type of products formed by OSCs [25]. For example, Trp within MWCYCR motif (as in XP\_021819911.1) has been linked to  $\beta$ -amyrin specificity, whereas Leu (as in XP\_021819927.1) and Phe (in XP\_021819928.1 and MdOSC1) at this position are implicated in lupeol synthesis [10] (Figure 3B). Similarly, Lys residue (K449 in XP\_021819911.1, K452 in XP\_021819927.1) have been implicated in  $\alpha/\beta$ -amyrin specificity, whereas Ala, Asn or Ile (I448 in XP\_021819928.1 and MdOSC1) have been proposed to be involved in lupeol activity [10]. The mutation of Ser or Thr (equivalent to S699 in human) to Phe in plants has been shown to convert OSC to make tetracyclic rather than pentacyclic products [20]. However, these rules are not unambiguous, as all fruit OSCs are supposed to be amyirin producers based on Lys and Ile residues, but should be lupeol synthases based on MFCYCR/ILCYCR motifs in XP\_021819927.1, XP\_021819928.1 and MdOSC1. This implies that these fruit enzymes are either multifunctional OSCs (as the phylogenetic analysis shown before would suggest, Figure 1), or other residues may be involved in the product specificity [10,25]. It is noteworthy that MdOSC1, which shows 95% identity with XP\_021819928.1, has been proposed to be a mixed amyirin synthase [10], able to produce  $\alpha$ -amyrin,  $\beta$ -amyrin and lupeol with a ratio 85:13:2 [11].

Although sequence comparisons give useful insight into the functions of OSCs, it cannot accurately predict the product specificity. In order to identify which products could possibly be formed by sweet cherry OSCs, we subjected human and fruit structures to docking analysis using intermediate cations [10,23]. Enzymes bind high-energy intermediate states much better than ground-state substrates or products and can help predict the structures of products [26]. In order to validate docking, we compared the lanosterol-complexed human OSC (1w6k) with the OSC structure docked with the same ligand. The result shows that the docking of lanosterol (ligand-OH ... OOC-Asp distance, 1.9 Å)

mimics the pose of the ligand, as obtained by X-ray crystallography (ligand-OH . . . OOC-Asp distance, 2.9 Å). Both are within the allowed H-bond distance (2–4 Å), except the linear chain with rotatable bonds where the docked ligand shows different orientation (Figure 6).



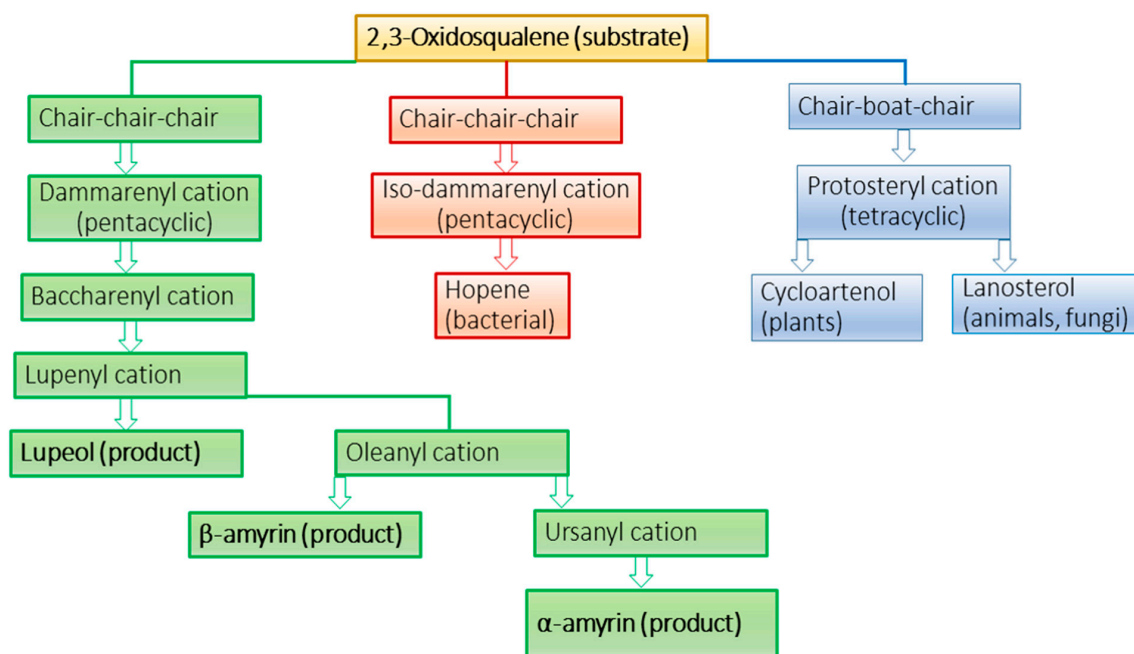
**Figure 6.** Validity of docking using ROSIE (Rosetta Online Server that Includes Everyone) Ligand Docking. Lanosterol was docked using the human OSC crystal structure (1w6k) without the ligand and then superimposed on 1w6k-lanosterol complex using CZ3 atom of W581 as the docking center. Brown, ligand pose in the crystal structure showing oxygen atom (red) of hydroxyl group; cyan, ligand docked using ROSIE. Hydrogen bonds between lanosterol and D455 (involved in initial protonation of the substrate) are shown for both structures. The deprotonating H232 (lower right) is also shown.

Product specificity in OSCs has been proposed to be achieved by enforcing 2,3-oxidosqualene to bind the hydrophobic active site in either chair-chair-chair (CCC) or chair-boat-chair (CBC) conformations leading to partially cyclized intermediate cations that are stabilized by aromatic residues via cation- $\pi$  interactions [27]. The CCC and CBC forms lead to various products via cation intermediates (Figure 7) that prevent early termination of the cyclization process. The termination of cyclization in humans occurs by deprotonation by His232 and in bacteria by Glu45 [27] (Figure 3B), or water molecule. However, corresponding His or Glu residues are missing in cherry and MdOSC1 enzymes (Figure 3B). The docking analysis (distance between ligand-OH and O atom of catalytic Asp and protein-ligand affinities) of OSCs with various cation intermediates, such as protosteryl, iso-dammarenyl and dammarenyl is expected to discriminate between various pathways leading to different products in human, bacterial and fruit enzymes (Figure 7).

Previously, we successfully used the Mcule tool to analyze the docking of various substrates in laccases from Tuscan sweet cherry varieties [29]. However, compared to the accessible active site of laccases, OSCs have a more hydrophobic and closed catalytic pocket (Figure 4). Docking of the cation intermediates using the advanced Mcule version showed incorrect poses and low affinities for all, but a couple of OSC–ligand complexes, where the -OH group of the ligand either does not face the O atom of catalytic Asp, or displayed very large ligand-OH . . . OOC-Asp-protein distances (results not shown). The inability to dock in the correct pose with Mcule (-OH of the ligand able to form a hydrogen bond within a distance of 1.5–4 Å with the oxygen atom of catalytic Asp residue) [24,28]



and the low ligand-OSC affinities are probably either due to the high number of rotatable bonds giving alternative catalytically non-competent orientations [26], and/or the inability of the protein backbone to undergo large movements (free-energy of orientation) upon binding of cation intermediates [30]. It is noteworthy that the cyclization reaction (where double bonds are replaced by single bonds) is highly exothermic [27] and this heat can be exploited for the conformational changes in both the ligand and the active site [22].



**Figure 7.** Conversion of substrate to various products via cation intermediates in fruits compared to that in other organisms, according to the previous literature [10,20,28]. The three main routes are colored green, pink and blue based on various triterpene skeletons. Products of OSCs (amyrins) can subsequently be converted to oleanolic and ursolic acids by C-28 oxidases.

Recent work in *Ononis spinosa* OSC found that by using 2,3-oxidosqualene dioxide with epoxide groups on either end, the hydroxylated intermediate is able to re-enter the active site for hydroxylation of the second epoxide ring to give  $\alpha$ -onocerin, implying conformational changes in the active site to accommodate the intermediate [31]. Protein flexibility is very crucial, as it can compensate for errors in the homology models, as well as allow ligands to bind to the active site [28]. Some workers have got a higher level of correct poses due to the application of strict restraints, such as H-bond between the ligand and catalytic Asp as in OSCs [28], or molecular dynamics simulation to allow conformational changes in the protein molecule [30]. However, putting H-bond restraint will preclude from getting realistic poses, variable ligand to OSC distances and binding affinities.

In order to assess if induced-fit docking of cation intermediates can improve the ligand-OH ... ..OOC-Asp distances and protein-ligand affinities, all OSCs were subjected to docking analysis using ROSIE Ligand Docking (RLD) tool (Figure 8). In contrast to Mcule that considers the ligand and protein side-chains as flexible and the protein backbone as rigid, RLD additionally takes into account the backbone flexibility of the active site, thus improving the ligand-protein fit via induced-fit mechanism [32].

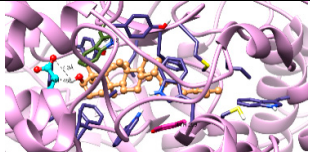
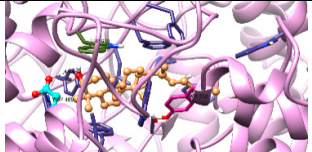
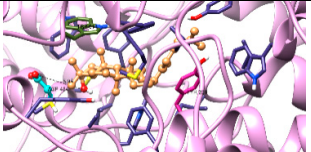
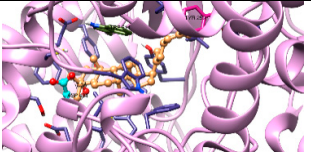
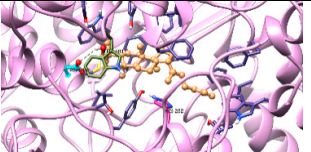
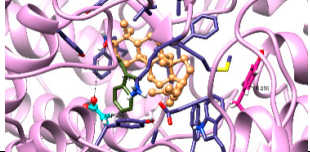
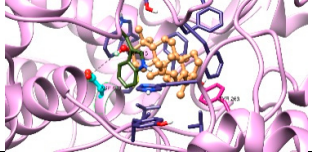
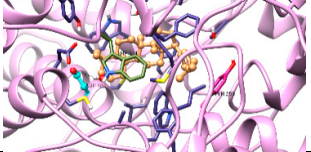
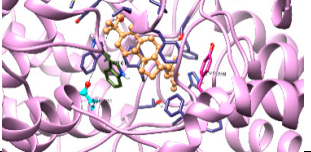
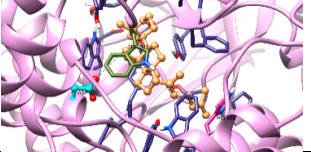
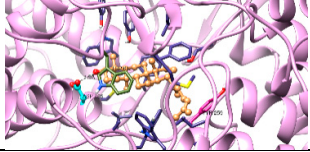
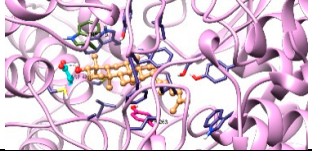
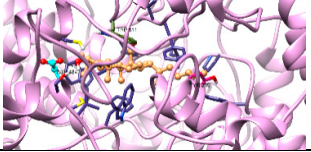
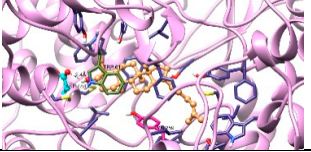
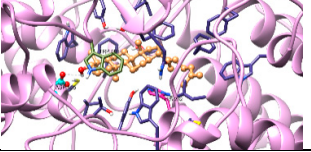
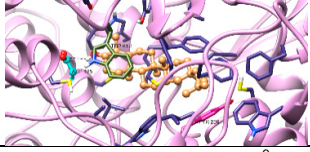
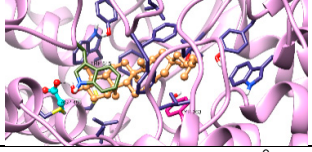
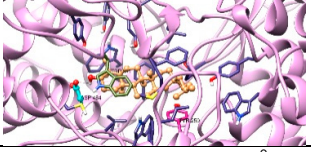
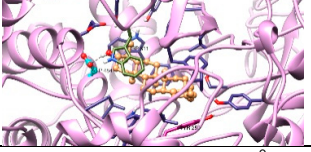
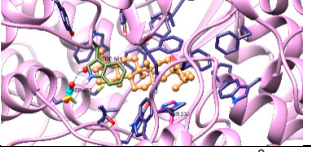
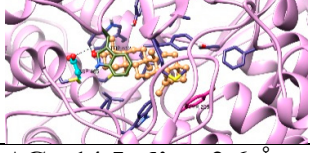
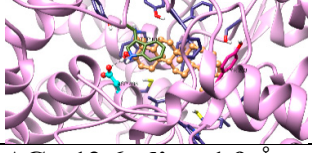
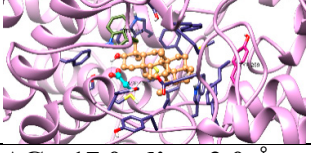
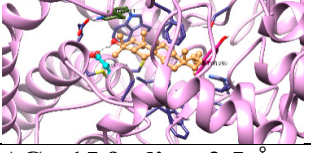
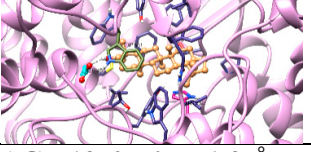

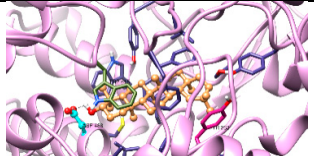
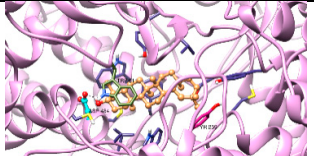
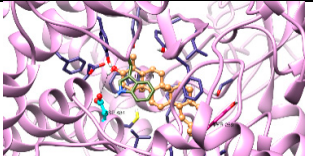
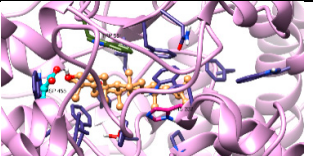
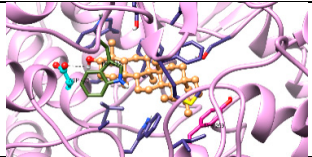
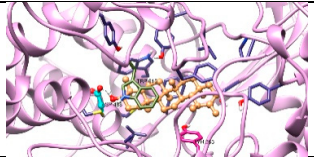
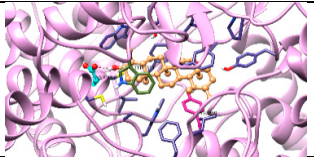
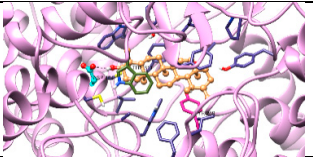
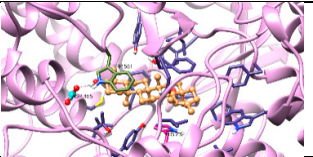
Ligand	Cherry			MdOSC1	Human 1w6k
	XP_021819911.1	XP_021819927.1	XP_021819928.1		
Protosteryl					
	$\Delta G, -12.2$ ; dist., 4.3 Å	$\Delta G, -15.1$ ; dist., 4.3 Å	$\Delta G, -13.8$ ; dist., 5.9 Å	$\Delta G, -16.2$ ; dist., 7.0 Å	$\Delta G, -10.1$ ; dist., 3.5 Å
Iso-dammarenyl					
	$\Delta G, -13.7$ ; dist., 5.7 Å	$\Delta G, -11.6$ ; dist., 3.9 Å	$\Delta G, -11.6$ ; dist., 4.1 Å	$\Delta G, -15.4$ ; dist., 9.4 Å	$\Delta G, -13.3$ ; dist., 10.3 Å
Dammarenyl					
	$\Delta G, -12.9$ ; dist., 1.9 Å	$\Delta G, -12.3$ ; dist., 2.3 Å	$\Delta G, -12.0$ ; dist., 2.7 Å	$\Delta G, -16.2$ ; dist., 2.4 Å	$\Delta G, -15.2$ ; dist., 2.8 Å
Baccharenyl					
	$\Delta G, -17.3$ ; dist., 2.8 Å	$\Delta G, -12.5$ ; dist., 2.1 Å	$\Delta G, -14.7$ ; dist., 2.0 Å	$\Delta G, -15.6$ ; dist., 2.0 Å	$\Delta G, -18.4$ ; dist., 1.9 Å
Lupenyl					
	$\Delta G, -14.5$ ; dist., 2.6 Å	$\Delta G, -12.6$ ; dist., 1.9 Å	$\Delta G, -17.0$ ; dist., 2.0 Å	$\Delta G, -15.9$ ; dist., 2.5 Å	$\Delta G, -12.6$ ; dist., 1.9 Å

Figure 8. Cont.

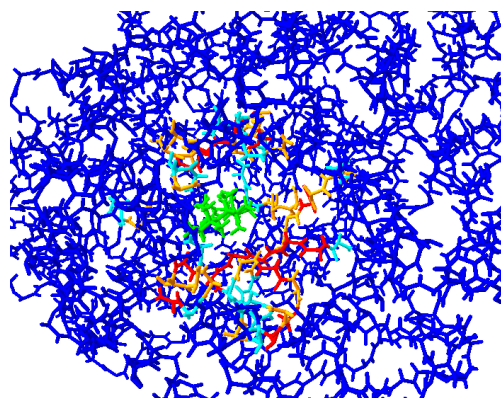
Oleanyl					
	$\Delta G$ , 13.9; dist., 2.6 Å	$\Delta G$ , -13.2; dist., 2.6 Å	$\Delta G$ , -13.5; dist., 1.9 Å	$\Delta G$ , -14.6; dist., 3.2 Å	$\Delta G$ , -15.3; dist., 2.0 Å
Ursanyl					
	$\Delta G$ , -12.0; dist., 2.6 Å	$\Delta G$ , -13.5; dist., 2.0 Å	$\Delta G$ , -17.5; dist., 2.0 Å	$\Delta G$ , -15.4; dist., 1.8 Å	$\Delta G$ , -18.7; dist., 1.8 Å

**Figure 8.** Poses, ligand to protein distances and affinities of oxidosqualene cyclases (OSCs) docked with the intermediate cations from fruit models and human X-ray structure.  $\Delta G$ , Gibbs free-energy of binding ( $\text{kcal mol}^{-1}$ ); dist, ligand-OH. . . . OOC-Asp-protein distances, ligand, brown; catalytic Asp, cyan; Trp docking reference, green; Tyr/His: magenta. Residues surrounding the active site, dark blue; oxygen atoms, red; nitrogen atoms, light blue and hydrogens, white.

In order to validate that RLD is taking into account the flexibility of the active site during ligand docking, we determined the RMSD of the overall structures and key residues surrounding the catalytic-site. The overall RMSD (spread over the entire protein molecule) of all OSCs docked with cation intermediates and products varied from 0.04–0.076 Å, whereas RMSD of residues away from the active site were close to 0 Å and RMSD of residues surrounding the active site were > 0 Å (results not shown). For example, the RMSD values of the residues surrounding the catalytic site of a representative docked cation-intermediate and a product are given in Table 1, whereas a color representation of cherry XP\_021819928.1 docked with dammarenyl cation is shown in Figure 9. The RMSD of residues away from the active site were close to 0 Å (Figure 9), whereas the RMSD values of residues surrounding the catalytic-site varied from 0.102–1.245 Å (Table 1), thereby confirming that backbone conformational changes within the active site took place upon docking.

**Table 1.** RMSD values of some key residues surrounding the active site after the induced-fit docking of a cation intermediate and a product on fruit OSCs.

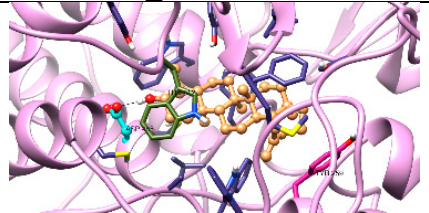
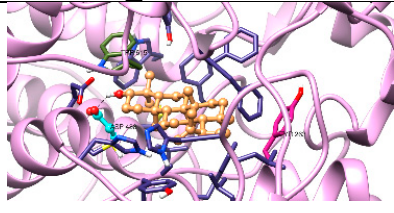
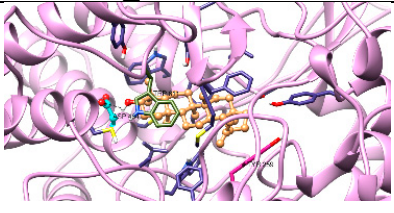
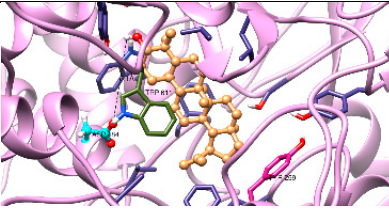
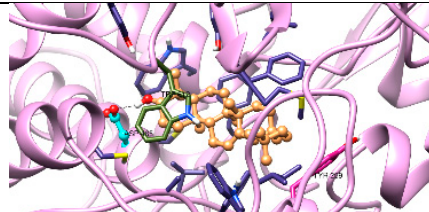
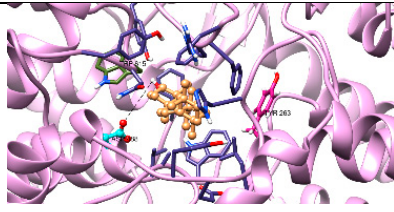
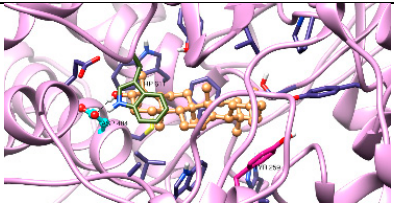
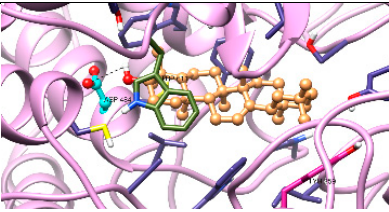
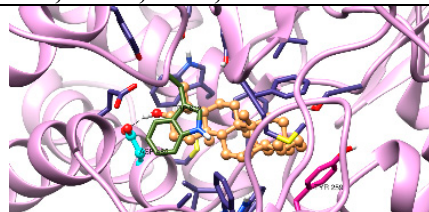
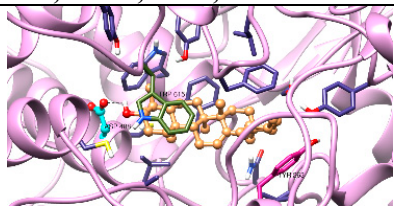
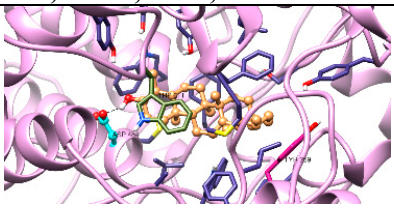
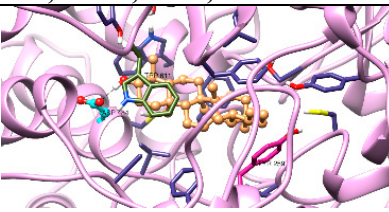
OSC	Residue	RMSD (Å)	
		Dammarenyl Cation	β-Amyrin Product
XP_021819911.1	Y 259	0.368	0.254
	G 414	0.109	0.350
	D 485 (protonating Asp)	0.089	0.105
	D 554	0.242	0.690
	W 612 (binding center)	0.395	0.288
	G 613	0.567	0.361
	W 621	0.384	0.001
	K 730	0.151	0.251
XP_021819927.1	Y 263	0.266	0.105
	V 373	0.321	0.001
	G 374	0.318	0.001
	D 488 (protonating Asp)	0.123	0.238
	A 536	0.216	0.409
	W 537	0.041	0.483
	W 615 (binding center)	0.161	0.245
F 732	0.124	0.336	
XP_021819928.1	Y 259	0.187	0.132
	C 369	1.245	0.793
	F 473	0.354	0.237
	D 484 (protonating Asp)	0.578	0.144
	A 532	0.301	0.246
	W 611 (binding center)	0.102	0.160
	M 726	0.302	0.305
H 732	0.554	0.044	
MdOSC1	Y 259	0.198	0.161
	C 369	0.332	0.503
	E 371	0.388	0.432
	K 372	0.876	0.786
	P 373	0.572	0.020
	F 412	0.250	0.417
	D 484 (protonating Asp)	0.151	0.163
W 611 (binding center)	0.115	0.214	



**Figure 9.** Cherry XP\_021819928.1 model was superimposed on its docked structure with dammarenyl cation for the determination of backbone RMSD of all residues due to induced-fit docking. Blue, 0 RMSD; cyan, RMSD between 0.05–0.1 Å; orange, RMSD between 0.1–0.2 Å; red, RMSD > 0.2 Å. Dammarenyl cation ligand is shown in green.

The docking poses, ligand to Asp distances and OSC–ligand binding affinities are given in Figure 8. The results indicate that for protosteryl cation, only human OSC gave the distance between the ligand and catalytic Asp within the hydrogen bond distance (3.5 Å, Figure 8), whereas none of the distances in fruit OSCs were within the hydrogen bond distances. This indicates that fruit OSC do not synthesize tetracyclic products via protosteryl cation route (Figure 6). For iso-dammarenyl cation, the ligand–Asp distances were large for all OSCs, except for cherry XP\_021819927.1 and XP\_021819928.1, which were borderline (~4 Å), with relatively low ligand–OSC affinities (Figure 8) implying that fruit and human enzymes do not likely form hopene-type products (Figure 6). For all other intermediates, leading from dammarenyl to ursanyl cations, docking showed very short hydrogen bond distances and high ligand to protein affinities (Figure 8) implying that fruit OSCs likely forms lupeol,  $\alpha$ -amyrin and  $\beta$ -amyrin products (Figure 6). The likely reason that human OSC binds these intermediates with short hydrogen bonds and high affinities (Figure 8) implies a flexible active site that can undergo conformational change to accommodate these cations. In spite of very good binding between these cations and human enzyme, the formation of four-ringed lanosterol product in human OSC is due to the active site His232 that deprotonates the cation intermediate thus terminating the reaction at the tetracyclic stage [24].

In order to discriminate between three possible types of products (lupeol,  $\alpha$ -amyrin and  $\beta$ -amyrin) formed by the Tuscan sweet cherry OSCs, we subjected fruit OSCs to docking analysis using products (Figure 10) along with lanosterol and human OSC as a control (not shown). While tetracyclic lanosterol binds well to sweet cherry XP\_021819911.1 and MdOSC1, the minimum ligand to Asp distance (1.9 Å) was obtained with human OSC (results not shown) implying that lanosterol is the preferred substrate for human enzyme. Lupeol binds to all cherry enzymes with low ligand to Asp distances (1.9–2.6 Å) and high ligand–protein affinities (−13.4 to −14.2 kcal mol<sup>−1</sup>), compared to apple MdOSC1 that showed larger distance (Figure 10), implying that all cherry OSCs have correct active-site topology to synthesize lupeol. Similarly, Figure 10 shows that all fruit OSCs also have potential to form  $\alpha$ -amyrin, as well as  $\beta$ -amyrin, except cherry XP\_021819927.1.

Ligand	Cherry			MdOSC1
	XP 021819911.1	XP 021819927.1	XP 021819928.1	
Lupeol				
	$\Delta G$ , -14.4; dist., 1.9 Å	$\Delta G$ , -13.4; dist., 1.9 Å	$\Delta G$ , -14.2; dist., 2.6 Å	$\Delta G$ , -15.8; dist., 7.1 Å
$\beta$ -Amyrin				
	$\Delta G$ , -14.9; dist., 1.8 Å	$\Delta G$ , -12.3; dist., 5.8 Å	$\Delta G$ , -14.0; dist., 2.1 Å	$\Delta G$ , -13.9; dist., 3.0 Å
$\alpha$ -Amyrin				
	$\Delta G$ , -13.7; dist., 1.9 Å	$\Delta G$ , -13.6; dist., 2.2 Å	$\Delta G$ , -15.0; dist., 1.8 Å	$\Delta G$ , -14.3; dist., 1.8 Å

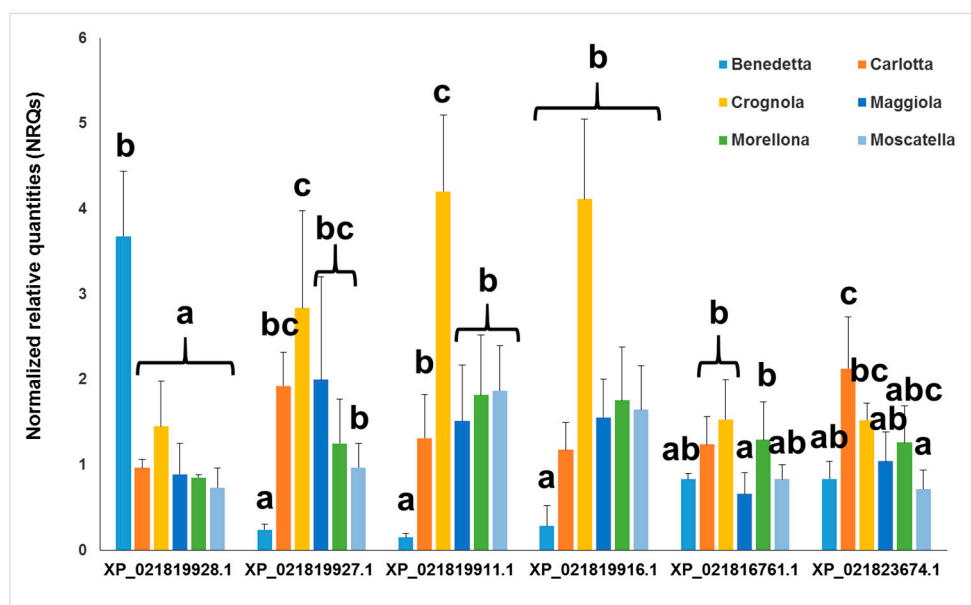
**Figure 10.** Poses, ligand-OH—OOC-Asp-protein distances and affinities of oxidosqualene cyclases (OSCs) docked with the reaction products from fruit models.  $\Delta G$ , Gibbs free-energy of binding ( $\text{kcal mol}^{-1}$ ); dist, ligand-OH. . . . OOC-Asp-protein distances. Ligand, brown; catalytic Asp, cyan; Trp docking reference, green; Tyr/His: magenta. Other active site residues, dark blue; oxygen atoms, red; nitrogen atom, light blue and hydrogens, white.

Based on the present docking results employing cation intermediates and products, all cherry OSCs are competent to produce all three products (lupeol,  $\alpha$ -amyrin,  $\beta$ -amyrin), except XP\_021819927.1 that is unable to make  $\beta$ -amyrin (Figures 8 and 10). Based on the above-described role of amino acids in product specificity (Figure 3B), it can be predicted that sweet cherry XP\_021819928.1 is likely to be a lupeol synthase, XP\_021819927.1 a mixed amyrin/lupeol synthase and XP\_021819911.1 an amyrin-synthesizing OSC. Collectively, the results obtained via bioinformatics imply that the OSCs from the Tuscan sweet cherries are able to synthesize a mixture of all three products; this awaits experimental confirmation, as these enzymes may each show capability to produce all the products in different ratios, as shown experimentally for MdOSC1 [11].

### 2.3. Gene Expression Analysis

The primers designed on the genes encoding the OSC XP\_021810674.1 and the CYP85 XP\_021815663.1 did not amplify with suitable efficiencies (comprised between 90–110%); they were therefore discarded and not included in the gene expression study.

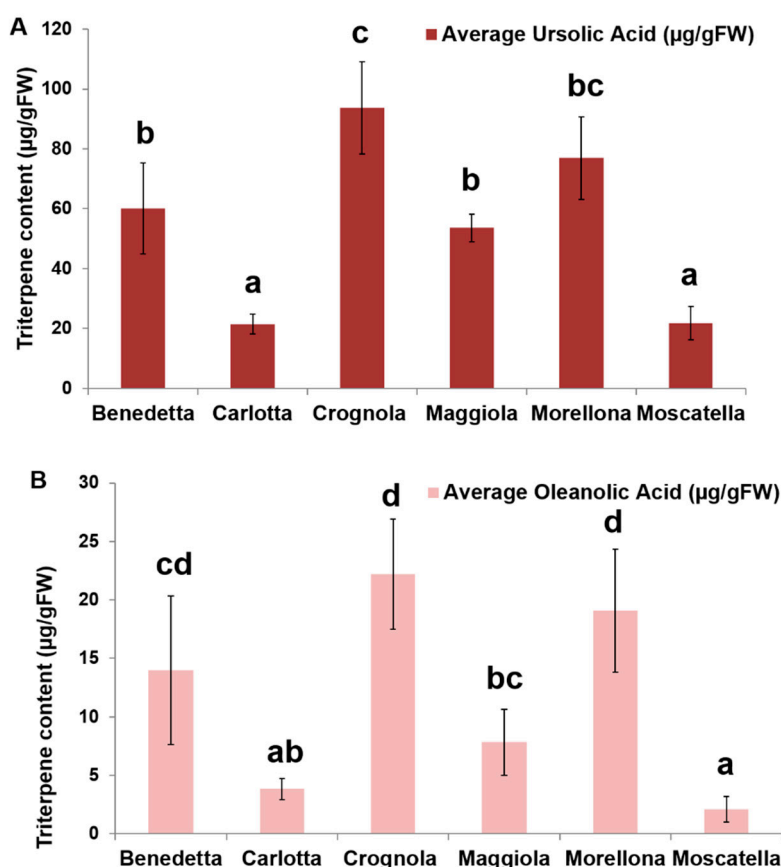
The expression analysis highlights different values among the Tuscan varieties (Figure 11). ‘Crognola’ and ‘Carlotta’ are among the varieties showing the highest expression levels of OSCs and CYP85s. This finding is very interesting if one considers the recently reported higher content of phenolic compounds in these ancient Italian sweet cherries [33] and the higher expression levels of genes encoding laccase-like multicopper oxidases [29]. ‘Benedetta’ shows higher levels than all the other varieties only for XP\_021819928.1. As discussed in the next paragraph, for the sake of comparison we also included a commercial variety (‘Durone’, purchased at a local grocery shop in Siena, Tuscany), which displayed lower expression levels (not shown). The data for the commercial variety are not given because of the different storage conditions of the fruits: the post-harvest storage may indeed have had an effect on the endogenous mRNA concentrations, thereby affecting the expression levels. The results here shown on the expression levels of OSCs and CYPs suggest higher levels of pentacyclic triterpenes. It was therefore necessary to quantify their amounts in the ancient fruits from Tuscany.



**Figure 11.** Gene expression analysis of the putative OSCs and CYP85s in the ancient sweet cherry varieties. For linearity with the previous phylogenetic data, the accession numbers of the corresponding proteins are used to identify the OSCs and CYP85s. Error bars indicate standard deviations ( $n = 4$ ). Different letters indicate statistically significant differences ( $p < 0.05$ ) calculated with the Tukey's post-hoc test.

#### 2.4. Pentacyclic Triterpene Content

The quantification of the pentacyclic triterpenes from the Tuscan sweet cherries indicates a higher content of ursolic acid, as compared to oleanolic acid (Figure 12). No betulinic acid is found in the fruits. This is in agreement with what previously reported in the literature, where a higher % w/w was quantified in *P. avium* fruits [6]. The average content of ursolic acid is significantly higher in some of the Tuscan cherries; in particular, the varieties ‘Crognola’ and ‘Morellona’ show the highest contents, followed by ‘Maggiola’, ‘Benedetta’, ‘Carlotta’ and ‘Moscatella’. The same ranking in terms of average oleanolic acid content is observed among the ancient varieties: ‘Crognola’ and ‘Morellona’ show the highest contents, followed by ‘Benedetta’, ‘Maggiola’, ‘Carlotta’ and ‘Moscatella’. We also determined the concentrations of ursolic and oleanolic acid in the commercial variety. The commercial fruits displayed significantly lower contents (7-fold and 31-fold lower than ‘Crognola’ for ursolic and oleanolic acid, respectively). We aimed at making a first comparison to confirm the nutraceutical value of the ancient fruits. Future studies on the commercial fruits performed under the same sampling conditions here described will confirm these preliminary data. It is interesting to note that post-harvest studies on two commercial varieties showed very little variation (1–10%) in both ursolic and oleanolic content after separate (20 °C, 3 days and 0 °C, 14 days), as well as combined (20 °C + 0 °C, 17 days) storage [12].



**Figure 12.** Triterpene composition of the Tuscan sweet cherry fruits. **A:** ursolic acid content; **B:** oleanolic acid content. Data are expressed in  $\mu\text{g g}^{-1}$  FW ( $n = 5$ , with the exception of ‘Crognola’, where  $n = 4$ ). Different letters indicate statistically significant differences ( $p < 0.05$ ) calculated with the Tukey’s post-hoc test.

The gene expression analysis and targeted metabolomics here performed show varying levels of the pentacyclic triterpenes ursolic and oleanolic acids in the ancient Tuscan fruits. We identify high producing varieties, such as ‘Crognola’ and ‘Morellona’ and our preliminary comparison with



a widespread commercial counterpart indicates higher levels of triterpenes in the ancient fruits. This result should however be confirmed by studies carried out on commercial fruits sampled under the same conditions here reported for the Tuscan varieties. Our data confirm the value of ancient local varieties of fruit-trees as alternative sources of secondary metabolites of pharmacological and nutraceutical interest.

### 3. Materials and Methods

#### 3.1. Fruit Harvesting

The sweet cherry fruits used in this study were sampled in 2017, as previously described [29], from eighteen-year-old cherry trees (ancient local varieties of *P. avium* on P-HL-B rootstocks) grown in the experimental field of the CNR in Ivalsa Follonica (GR, Italy, coordinates 42° 55' 59" N, 10° 45' 57" E). Sampling took place in May 2017. However, the variety 'Benedetta' was sampled in May 2016, as the trees gave no fruits in 2017. At the sampling time (between 9:00–10.00 a.m.), the average temperature was of 17–20 °C, 71–74% of humidity in 2016 and 2017, respectively. Twenty sweet cherry fruits (growth stage: 60 dpa) were sampled from each tree to have enough biological replicates (four in this study), each represented by a pool of five fruits. Fruits were sampled from different places on the tree canopy (at around 1.70–1.90 m from the soil), to minimize eventual bias due to variations in solar exposure. Fruits were sampled from the trees with the stems, which was rapidly removed after picking. Then the samples were immediately plunged in liquid nitrogen in tubes, brought to the laboratory and stored at –80 °C in Ziploc bags until processing.

#### 3.2. RNA Extraction, Primer Design and Real-Time PCR Data Analysis

The RNA extraction procedure, quality control and quantification, as well as reverse transcription and real-time PCR conditions have been previously detailed [29]. The primers used for the reference genes have been previously published [29], those for *eTIF4E*, *GAPDH*, *OSCs* and *CYP85s* are indicated in Table 2. They were designed with Primer3Plus (<http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi>) and checked with OligoAnalyzer 3.1 (<http://eu.idtdna.com/calc/analyzer>).

**Table 2.** List of primers with details of the sequences, amplicon sizes and amplification efficiencies.

Name	Sequence (5'→3')	Amplicon Size	Amplification Efficiency
PaveTIF4E Fwd PaveTIF4E Rev	GGCAAAGCCTCGATAACAATG TTGGTTATGGAGAGCGAAGAC	71	1.90
PavGAPDH Fwd PavGAPDH Rev	TATCAAAGCCACAGCCACTG TGCTATTCGGAGAACCAACC	118	1.88
XP_021819928.1 Fwd XP_021819928.1 Rev	AAGGCAGACATGGGAGTTTG ATCTGAAAACGCCAGAGGAG	135	1.94
XP_021819916.1 Fwd XP_021819916.1 Rev	ATGAGGGTTCAGCTTGATGC CCACCGTGATTGTGTGAATG	104	2.09
XP_021819927.1 Fwd XP_021819927.1 Rev	CGAAGAGTGTTGTCTGCTTAACG AAAGAGAATAGTACCCCTTG	109	2.04
XP_021819911.1 Fwd XP_021819911.1 Rev	TGTATTCCCAGCAGAGCATC CCCAACCACCATCTTCATTC	74	1.90
XP_021816761.1 Fwd XP_021816761.1 Rev	TGAGAAAGATGCTCCCAAC CCTGTTTCCCAACCATCAG	107	1.90
XP_021823674.1 Fwd XP_021823674.1 Rev	GAGGTTCAAATGGGAAGTGC TGACGATGAAGACGAACTGG	96	1.90

The expression values of the *OSCs* and *CYP85s* were calculated with qBase<sup>PLUS</sup> (version 2.5, Biogazelle, Ghent, Belgium) by using the reference genes indicated by geNorm<sup>PLUS</sup>. Eight reference

genes were tested for stability; *PavAct7* and *PaveTIF4E* were identified as sufficient for data normalization when tested together with *PavPP2A*, *PavPolyUbg*, *PavSerThr*, *PavAP4*, *PavGAPDH* and *PavTIP41*. A one-way ANOVA with a Tukey's post-hoc test was performed on log<sub>2</sub> transformed NRQs (Normalized Relative Quantities) by using IBM SPSS Statistics v19.

### 3.3. Bioinformatics

The pair-wise multiple alignment of OSCs and CYP85s to identify conserved residues was carried out with CLUSTAL-Ω (<http://www.ebi.ac.uk/Tools/msa/clustalo/>) [34]. The alignment was then used to build the maximum likelihood phylogenetic tree using W-IQ-TREE [35] (number of bootstrap alignments in ultrafast mode: 1000), available at <http://iqtree.cibiv.univie.ac.at>. The tree in Newick format was visualized with iTOL (available at <https://itol.embl.de/>).

The OSC sequences from human, bacteria and fruits to identify conserved residues and motifs were aligned using CLUSTAL-Ω (<http://www.ebi.ac.uk/Tools/msa/clustalo/>) [34]. The 3D homology models were generated with the I-TASSER Suite (<http://zhanglab.ccmb.med.umich.edu/I-TASSER/>) [36] utilizing LOMETS, SPICKER, and TM-align. The models based on human OSC (PDB 1w6k) were then refined using REMO by optimizing the backbone hydrogen-bonding networks and FG-MD by removing the steric clashes and improving the torsion angles. The quality of the models was checked by RAMPAGE [37].

Molecular docking of intermediate cations (protosteryl, dammarenyl, iso-dammarenyl, baccharenyl, lupenyl, oleanyl and ursanyl) and the product (lanosterol, lupeol, α-amyrin and β-amyrin) with OSCs was carried out with the advanced version of online Mcule tool (<http://doc.mcule.com/doku.php?id=1clickdocking>), based on AutoDock Vina [38] and Ligand Docking Tool [32] based on the Rosetta Online Server that Includes Everyone, ROSIE ([https://rosie.graylab.jhu.edu/ligand\\_docking](https://rosie.graylab.jhu.edu/ligand_docking)) [39], using the coordinates of atom CZ3 of Trp581 in 1w6k as the binding center and equivalent CZ3 atoms of conserved Trp residues in all other structures [40]. The SMILES codes of products were imported from PubChem (<https://pubchem.ncbi.nlm.nih.gov/>). The structures of the cation intermediates were drawn within Mcule tool [10,23]. For ROSIE Ligand Docking (RLD), the structures of the cation intermediates were either drawn in PubChem Sketcher (<https://pubchem.ncbi.nlm.nih.gov/edit2/index.html>) and the SMILES codes imported into NIH Cactus Online SMILES translator (<https://cactus.nci.nih.gov/translate/>) developed by National Cancer Institute or the structures were directly drawn in NIH Cactus Online Translator and converted to 3D SDF files with hydrogen atoms added according to standard valences. The overall structures showing critical residues were visualized by superimposing fruit models on the human OSC (PDB 1w6k) with DeepView Swiss-PdbViewer v4.1 (<http://www.expasy.org/spdbv/>) [41]. Molecular graphics and analyses of OSC structures docked with cation intermediates were performed with UCSF Chimera, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco [42] available at <http://www.rbvi.ucsf.edu/chimera>. RMSD of the overall structures and key residues surrounding the active site was determined by PDB viewer after superimposing OSC structures before and after induced-fit docking employing cation intermediates and products.

### 3.4. Extraction and Analysis of Pentacyclic Triterpenes

Sample preparation and HPLC analysis were performed based on the previously described method by Andre and colleagues [11]. Five hundred mg of powdered freeze-dried material (exocarp and ca. 5 mm of the mesocarp tissue) were mixed with 10 mL of ethyl acetate/hexane (50:50 v/v). The mixture was homogenized using a vortex for 30 s and then shaken for 1 h at room temperature. After centrifugation at 10,000 g for 15 min, the supernatant was collected and evaporated to dryness using a centrifugal vacuum evaporator. The pellet was re-extracted using 10 mL of ethanol/H<sub>2</sub>O (80:20 v/v) solution, homogenized and shaken for 2 h at room temperature and centrifuged at 10,000g for 15 min. The supernatant was collected, combined with the lipophilic dried extract and evaporated to dryness. Triterpenes were re-suspended in 1 mL of EtOH and filtered (0.45 μm) before HPLC

analysis. Each cherry sample was analyzed in triplicate ( $n = 3$ ). Extracts were analyzed with a Waters Acquity UPLC system (Waters, Milford, MA, USA) hyphenated to a Diode Array Detector (UPLC-DAD). The separation of the 5  $\mu\text{L}$  aliquot was performed on a reverse-phase Acquity UPLC BEH C18 column (2.1  $\times$  100 mm, 1.7  $\mu\text{m}$  particle size, Waters, Milford, MA, USA). The eluents were 0.05% *o*-phosphoric acid in water (A) and 0.05% *o*-phosphoric acid in methanol (B). The gradient was as follows: 0 min, 75% B; 2 min, 75% B; 16 min, 82% B; 25 min, 100% B; 26.5 min, 100% B; 27 min, 75% B; 30 min, 75% B. The flow rate was of 0.3  $\text{mL min}^{-1}$  and the column temperature was 40  $^{\circ}\text{C}$ . Ursolic acid and oleanolic acid were identified by their retention times and spectral data by comparison with standards and they were quantified at 210 nm using five-point calibration curves. Excellent linearity ( $R^2 > 0.99$ ) was obtained in the concentration range 100–6.25  $\mu\text{g mL}^{-1}$  for both compounds.

#### 4. Conclusions

We identified OSCs and CYP85s involved in pentacyclic triterpene biosynthesis in sweet cherry and measured their expression in ancient Tuscan varieties. The *in silico* analysis performed on three OSCs suggests that these isozymes can synthesize all the three classes ( $\alpha$ - and  $\beta$ -amyrins, lupeol) of pentacyclic triterpenes. The high gene expression levels and the increased contents of oleanolic and ursolic acids indicate that the ancient sweet cherries are rich sources of secondary metabolites.

**Supplementary Materials:** The following are available online, Figure S1: Alignment of the sweet cherry and apple OSCs showing the conserved DCTAE (green highlight), M(W/Y)CY(C/S)R (cyan highlight), as well as the Q-X3-G-X-W motifs (yellow highlight) [11,18–20].

**Author Contributions:** Conceptualization, R.B., C.M.A., G.G.; Methodology, R.B., M.Z.H., K.S.S., C.M.A., G.G.; formal analysis, R.B., M.Z.H., K.S.S., C.M.A., G.G.; investigation, R.B., G.C., K.S.S., J.-F.H., C.M.A., G.G.; data curation, R.B., M.Z.H., K.S.S., C.M.A., G.G.; writing—original draft preparation, R.B., K.S.S., C.M.A., G.G.; writing—review and editing, R.B., S.L., G.C., K.S.S., J.-F.H., C.M.A., G.G.

**Funding:** This research received no external funding.

**Acknowledgments:** The authors are grateful to the Tuscany Region and the National Research Council (CNR-Italy) for support. RB acknowledges the Tuscany Region for financial support through the PhD fellowship “Pegaso”. Personal assistance of KSS by KFUPM is acknowledged. The authors are grateful to the Mcule team for providing access to the advanced version of the docking tool.

**Conflicts of Interest:** The authors declare no conflict of interest.

#### References

1. Ballistreri, G.; Continella, A.; Gentile, A.; Amenta, M.; Fabroni, S.; Rapisarda, P. Fruit quality and bioactive compounds relevant to human health of sweet cherry (*Prunus avium* L.) cultivars grown in Italy. *Food Chem.* **2013**, *140*, 630–638. [[CrossRef](#)] [[PubMed](#)]
2. Gonçalves, A.C.; Rodrigues, M.; Santos, A.O.; Alves, G.; Silva, L.R. Antioxidant Status, Antidiabetic Properties and Effects on Caco-2 Cells of Colored and Non-Colored Enriched Extracts of Sweet Cherry Fruits. *Nutrients* **2018**, *10*, 1688. [[CrossRef](#)]
3. Correia, S.; Schouten, R.; Silva, A.P.; Gonçalves, B. Factors Affecting Quality and Health Promoting Compounds during Growth and Postharvest Life of Sweet Cherry (*Prunus avium* L.). *Front. Plant Sci.* **2017**, *8*, 2166. [[CrossRef](#)] [[PubMed](#)]
4. Kim, D.-O.; Heo, H.J.; Kim, Y.J.; Yang, H.S.; Lee, C.Y. Sweet and Sour Cherry Phenolics and Their Protective Effects on Neuronal Cells. *J. Agric. Food Chem.* **2005**, *53*, 9921–9927. [[CrossRef](#)] [[PubMed](#)]
5. Commisso, M.; Bianconi, M.; Di Carlo, F.; Poletti, S.; Bulgarini, A.; Munari, F.; Negri, S.; Stocchero, M.; Ceoldo, S.; Avesani, L.; et al. Multi-approach metabolomics analysis and artificial simplified phytocomplexes reveal cultivar-dependent synergy between polyphenols and ascorbic acid in fruits of the sweet cherry (*Prunus avium* L.). *PLoS ONE* **2017**, *12*, e0180889. [[CrossRef](#)]
6. Peschel, S.; Franke, R.; Schreiber, L.; Knoche, M. Composition of the cuticle of developing sweet cherry fruit. *Phytochemistry* **2007**, *68*, 1017–1025. [[CrossRef](#)] [[PubMed](#)]
7. Szakiel, A.; Pączkowski, C.; Pensec, F.; Bertsch, C. Fruit cuticular waxes as a source of biologically active triterpenoids. *Phytochem. Rev.* **2012**, *11*, 263–284. [[CrossRef](#)]

8. Buschhaus, C.; Jetter, R. Composition differences between epicuticular and intracuticular wax substructures: how do plants seal their epidermal surfaces? *J. Exp. Bot.* **2011**, *62*, 841–853. [CrossRef]
9. Furtado, N.A.J.C.; Pirson, L.; Edelberg, H.; M Miranda, L.; Loira-Pastoriza, C.; Preat, V.; Larondelle, Y.; André, C.M. Pentacyclic Triterpene Bioavailability: An Overview of In Vitro and In Vivo Studies. *Mol. Basel Switz.* **2017**, *22*, 400. [CrossRef] [PubMed]
10. Brendolise, C.; Yauk, Y.-K.; Eberhard, E.D.; Wang, M.; Chagne, D.; Andre, C.; Greenwood, D.R.; Beuning, L.L. An unusual plant triterpene synthase with predominant  $\alpha$ -amyrin-producing activity identified by characterizing oxidosqualene cyclases from *Malus  $\times$  domestica*. *FEBS J.* **2011**, *278*, 2485–2499. [CrossRef]
11. Andre, C.M.; Legay, S.; Deleruelle, A.; Nieuwenhuizen, N.; Punter, M.; Brendolise, C.; Cooney, J.M.; Lateur, M.; Hausman, J.-F.; Larondelle, Y.; et al. Multifunctional oxidosqualene cyclases and cytochrome P450 involved in the biosynthesis of apple fruit triterpenic acids. *New Phytol.* **2016**, *211*, 1279–1294. [CrossRef]
12. Belge, B.; Llovera, M.; Comabella, E.; Gatius, F.; Guillén, P.; Graell, J.; Lara, I. Characterization of Cuticle Composition after Cold Storage of “Celeste” and “Somerset” Sweet Cherry Fruit. *J. Agric. Food Chem.* **2014**, *62*, 8722–8729. [CrossRef] [PubMed]
13. Quero-Garcia, J.; Iezzoni, A.; Pulawska, J.; Lang, G.A. *Cherries: Botany, Production and Uses*; CABI Editions: Oxford, UK, 2017; ISBN 978-1-78064-837-8.
14. Ghiselli, L.; Rossi, E.; Whittaker, A.; Dinelli, G.; Baglio, A.P.; Andrenelli, L.; Benedettelli, S. Nutritional characteristics of ancient Tuscan varieties of *Triticum aestivum* L. *Ital. J. Agron.* **2016**, *11*, 237–245. [CrossRef]
15. Agnoletti, M.; Emanuelli, F. *Biocultural Diversity in Europe*; Springer International Publishing: Cham, Switzerland, 2016; ISBN 978-3-319-26315-1.
16. Berni, R.; Cantini, C.; Romi, M.; Hausman, J.-F.; Guerriero, G.; Cai, G. Agrobiotechnology Goes Wild: Ancient Local Varieties as Sources of Bioactives. *Int. J. Mol. Sci.* **2018**, *19*, 2248. [CrossRef]
17. BASIQ – Bottega Alimentare (grocery) for Sustainability, Identity and Quality | ecodynamicsgroup. Available online: <http://www.valdimersegreen.com/en/basiq/> (accessed on 20 April 2019).
18. Shirasawa, K.; Isuzugawa, K.; Ikenaga, M.; Saito, Y.; Yamamoto, T.; Hirakawa, H.; Isobe, S. The genome sequence of sweet cherry (*Prunus avium*) for use in genomics-assisted breeding. *DNA Res. Int. J. Rapid Publ. Rep. Genes Genomes* **2017**, *24*, 499–508. [CrossRef]
19. Siedenburg, G.; Jendrossek, D. Squalene-Hopene Cyclases. *Appl. Environ. Microbiol.* **2011**, *77*, 3905–3915. [CrossRef] [PubMed]
20. Salmon, M.; Thimmappa, R.B.; Minto, R.E.; Melton, R.E.; Hughes, R.K.; O’Maille, P.E.; Hemmings, A.M.; Osbourn, A. A conserved amino acid residue critical for product and substrate specificity in plant triterpene synthases. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, E4407–E4414. [CrossRef] [PubMed]
21. Ito, R.; Masukawa, Y.; Hoshino, T. Purification, kinetics, inhibitors and CD for recombinant  $\beta$ -amyrin synthase from *Euphorbia tirucalli* L and functional analysis of the DCTA motif, which is highly conserved among oxidosqualene cyclases. *FEBS J.* **2013**, *280*, 1267–1280. [CrossRef]
22. Reinert, D.J.; Balliano, G.; Schulz, G.E. Conversion of squalene to the pentacarboxylic hopene. *Chem. Biol.* **2004**, *11*, 121–126. [CrossRef]
23. Iturbe-Ormaetxe, I.; Haralampidis, K.; Papadopoulou, K.; Osbourn, A.E. Molecular cloning and characterization of triterpene synthases from *Medicago truncatula* and *Lotus japonicus*. *Plant Mol. Biol.* **2003**, *51*, 731–743. [CrossRef] [PubMed]
24. Thoma, R.; Schulz-Gasch, T.; D’Arcy, B.; Benz, J.; Aebi, J.; Dehmlow, H.; Hennig, M.; Stihle, M.; Ruf, A. Insight into steroid scaffold formation from the structure of human oxidosqualene cyclase. *Nature* **2004**, *432*, 118–122. [CrossRef]
25. Segura, M.J.R.; Jackson, B.E.; Matsuda, S.P.T. Mutagenesis approaches to deduce structure–function relationships in terpene synthases. *Nat. Prod. Rep.* **2003**, *20*, 304–317. [CrossRef] [PubMed]
26. Hermann, J.C.; Ghanem, E.; Li, Y.; Raushel, F.M.; Irwin, J.J.; Shoichet, B.K. Predicting substrates by docking high-energy intermediates to enzyme structures. *J. Am. Chem. Soc.* **2006**, *128*, 15882–15891. [CrossRef]
27. Schulz-Gasch, T.; Stahl, M. Mechanistic insights into oxidosqualene cyclizations through homology modeling. *J. Comput. Chem.* **2003**, *24*, 741–753. [CrossRef]
28. Tian, B.-X.; Wallrapp, F.H.; Holiday, G.L.; Chow, J.-Y.; Babbitt, P.C.; Poulter, C.D.; Jacobson, M.P. Predicting the functions and specificity of triterpenoid synthases: a mechanism-based multi-intermediate docking approach. *PLoS Comput. Biol.* **2014**, *10*, e1003874. [CrossRef]

29. Berni, R.; Piasecki, E.; Legay, S.; Hausman, J.-F.; Siddiqui, K.S.; Cai, G.; Guerriero, G. Identification of the laccase-like multicopper oxidase gene family of sweet cherry (*Prunus avium* L.) and expression analysis in six ancient Tuscan varieties. *Sci. Rep.* **2019**, *9*, 3557. [[CrossRef](#)] [[PubMed](#)]
30. Feenstra, K.A.; Starikov, E.B.; Urlacher, V.B.; Commandeur, J.N.M.; Vermeulen, N.P.E. Combining substrate dynamics, binding statistics, and energy barriers to rationalize regioselective hydroxylation of octane and lauric acid by CYP102A1 and mutants. *Protein Sci. Publ. Protein Soc.* **2007**, *16*, 420–431. [[CrossRef](#)]
31. Almeida, A.; Dong, L.; Khakimov, B.; Bassard, J.-E.; Moses, T.; Lota, F.; Goossens, A.; Appendino, G.; Bak, S. A Single Oxidosqualene Cyclase Produces the Seco-Triterpenoid  $\alpha$ -Onocerin. *Plant Physiol.* **2018**, *176*, 1469–1484. [[CrossRef](#)]
32. DeLuca, S.; Khar, K.; Meiler, J. Fully Flexible Docking of Medium Sized Ligand Libraries with RosettaLigand. *PLoS ONE* **2015**, *10*, e0132508. [[CrossRef](#)]
33. Berni, R.; Romi, M.; Cantini, C.; Hausman, J.-F.; Guerriero, G.; Cai, G. Functional Molecules in Locally-Adapted Crops: The Case Study of Tomatoes, Onions, and Sweet Cherry Fruits from Tuscany in Italy. *Front. Plant Sci.* **2019**, *9*, 1983. [[CrossRef](#)] [[PubMed](#)]
34. McWilliam, H.; Li, W.; Uludag, M.; Squizzato, S.; Park, Y.M.; Buso, N.; Cowley, A.P.; Lopez, R. Analysis Tool Web Services from the EMBL-EBI. *Nucleic Acids Res.* **2013**, *41*, W597–W600. [[CrossRef](#)]
35. Trifinopoulos, J.; Nguyen, L.-T.; von Haeseler, A.; Minh, B.Q. W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Res.* **2016**, *44*, W232–W235. [[CrossRef](#)] [[PubMed](#)]
36. Yang, J.; Yan, R.; Roy, A.; Xu, D.; Poisson, J.; Zhang, Y. The I-TASSER Suite: protein structure and function prediction. *Nat. Methods* **2015**, *12*, 7–8. [[CrossRef](#)] [[PubMed](#)]
37. Lovell, S.C.; Davis, I.W.; Arendall, W.B.; de Bakker, P.I.W.; Word, J.M.; Prisant, M.G.; Richardson, J.S.; Richardson, D.C. Structure validation by C $\alpha$  geometry: phi,psi and Cbeta deviation. *Proteins* **2003**, *50*, 437–450. [[CrossRef](#)] [[PubMed](#)]
38. Trott, O.; Olson, A.J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. *J. Comput. Chem.* **2010**, *31*, 455–461. [[CrossRef](#)] [[PubMed](#)]
39. Lyskov, S.; Chou, F.-C.; Conchúir, S.Ó.; Der, B.S.; Drew, K.; Kuroda, D.; Xu, J.; Weitzner, B.D.; Renfrew, P.D.; Sripakdeevong, P.; et al. Serverification of molecular modeling applications: the Rosetta Online Server that Includes Everyone (ROSIE). *PLoS ONE* **2013**, *8*, e63906. [[CrossRef](#)] [[PubMed](#)]
40. Kiss, R.; Sandor, M.; Szalai, F.A. <http://Mcule.com>: a public web service for drug discovery. *J. Cheminform.* **2012**, *4*, P17. [[CrossRef](#)]
41. Guex, N.; Peitsch, M.C. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* **1997**, *18*, 2714–2723. [[CrossRef](#)] [[PubMed](#)]
42. Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Couch, G.S.; Greenblatt, D.M.; Meng, E.C.; Ferrin, T.E. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **2004**, *25*, 1605–1612. [[CrossRef](#)]

