

## Supplementary information

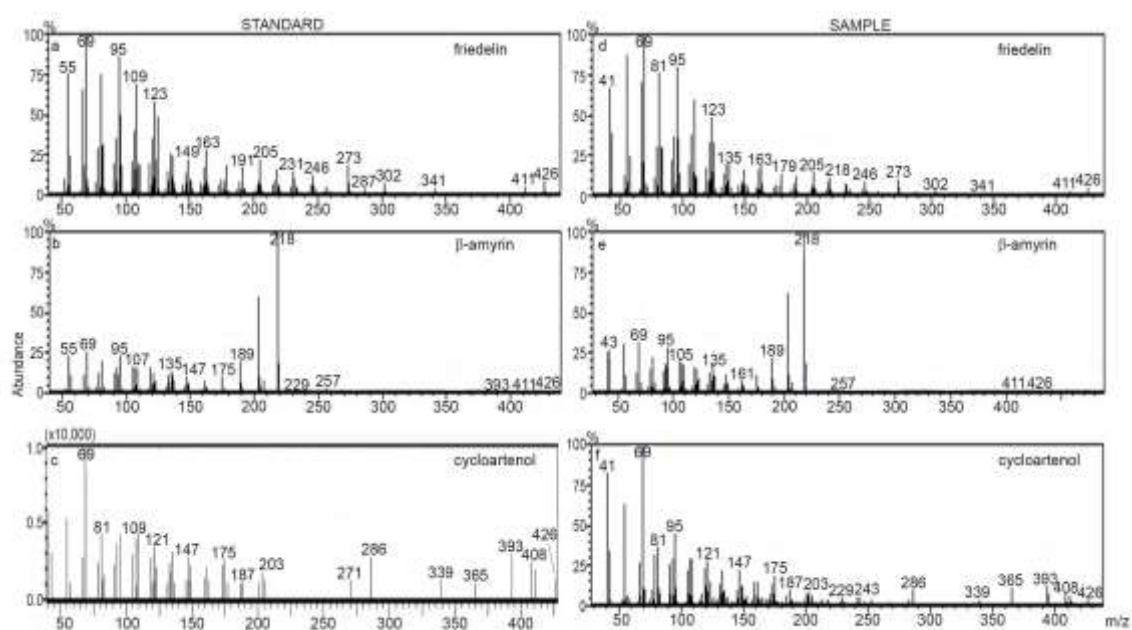
**Friedelin Synthase from *Maytenus ilicifolia*:**

**Leucine 482 Plays an Essential Role in the Production of the Most Rearranged Pentacyclic  
Triterpene**

Tatiana M. Souza-Moreira, Thaís B. Alves, Karina A. Pinheiro, Lidiane G. Felipe, Gustavo M.

A. De Lima, Tatiana F. Watanabe, Cristina C. Barbosa, Vânia A. F. F. M. Santos, Norberto P.

Lopes, Sandro R. Valentini, Rafael V. C. Guido, Maysa Furlan, Cleslei F. Zanelli

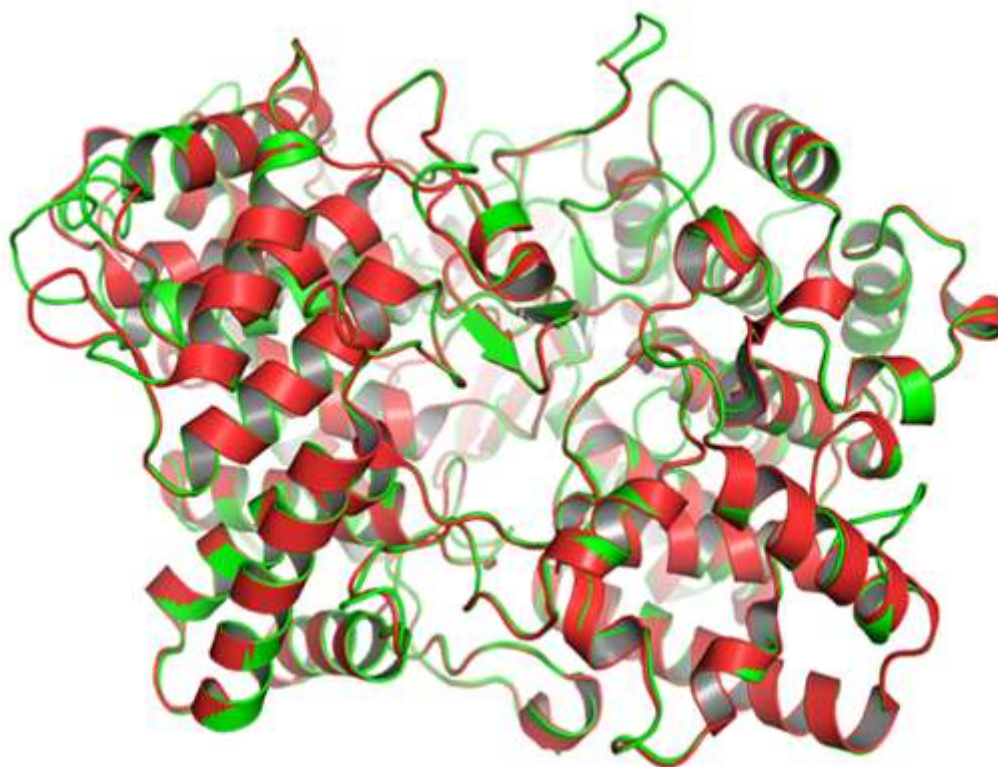


**Supplementary Figure S1.** The mass spectra for the triterpenoids analysed in this work. The fragmentation patterns of the standards (a) friedelin, (b) cycloartenol and (c)  $\beta$ -amyryn.

Triterpene compounds produced by *Saccharomyces cerevisiae* heterologously expressing the plant oxidosqualene cyclases ((d) friedelin, (e) cycloartenol and (d)  $\beta$ -amyryn) show similar fragmentation patterns as standards. Mass spectrum of standard cycloartenol was obtained from the NIST08 library.

PEN5	481	GLPISDGTAESECC
PEN6	486	GWPEVS DCTAESECC
PEN1	485	GWPEVS DCTAEGLKCC
PEN4	485	GWPEVS DCTAESECC
LcOSC2	477	GWPEVS DCTSEGLRVA
MiCAS2	441	GWPEVS DCTAEGLRAV
BPX1	487	GWPEVS DCTAEGLRAV
MiCAS1	482	GWPEVS DCTAEGLRAA
KdCAS	478	GWPEVS DCTAEGLRVV
CAS1	477	GWPEVS DCTAEGLRAA
PNX	477	GWPEVS DCTAEGLRAV
PSX	477	GWPEVS DCTAEGLRAV
OSC5	477	GWPEVS DCTAEGLRAI
GgCAS1	477	GWPEVS DCTAEGLRAV
BPW	476	GWQVS DCTAEGLRVA
OSC3	476	GWQVS DCTAEGLRVA
GgLUS1	476	GWQVS DCTAEGLRAA
OEW	477	GWQVS DCTAEGLRAA
TRW	479	GWQVS DCTAEGLRVS
RsM2	481	GVIAIDCTAESECC
LcIMS1	476	GWQVS DCTAENLKCC
MiFRS	478	GWQVS DCTAENLKCC
LUP1	477	GWQVS DCTAENLKCC
LUP2	480	GWQVS DCTAENLKCC
RcLUS	478	GWQVS DCTAENLKCC
BgLUS	478	GWQVS DCTAENLKCC
KcMS	478	GWQVS DCTAENLKCC
KdFRS	479	GWQVS DCTAENLKCC
KdGLS	479	GWQVS DCTAENLKCC
KdLUS	479	GWQVS DCTAENLKCC
PNY1	480	GWQVS DCTAENLKCC
PNY2	478	GWQVS DCTAENLKCC
EIAS	479	GWQVS DCTAENLKCC
RsM1	479	GWQVS DCTAENLKCC
BPY	479	GWQVS DCTAENLKCC
KdTAS	498	GWQVS DCTAENLKCC
PSM	479	GWQVS DCTAENLKCC
AMY2	479	GWQVS DCTAENLKCC
PSY	479	GWQVS DCTAENLKCC
AMY1	479	GWQVS DCTAENLKCC
GgbAS1	479	GWQVS DCTAENLKCC
OSC1	479	GWQVS DCTAENLKCC

**Supplementary Figure S2.** The alignment of the different oxidosqualene cyclases (OSCs) around the Asp-Cys-Thr-Ala-Glu (DCTAE) conserved motif in the catalytic site. The leucine 2 positions before the DCTAE motif is present only in the friedelin-synthesizing enzymes. The alignment was performed using CLUSTALW (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>). The oxidosqualene cyclase (OSC) identification and accession numbers are shown in Supplemental Table 3.



**Supplementary Figure S3.** The structural overlay of the friedelin synthase from *Maytenus ilicifolia* (MiFRS, green, homology model) and human OSCs (red, template).

**Supplementary Table S1.** The features of the cloned OSC sequences and their translated protein sequences.

<b>ORF name</b>	<b><i>MiFRS</i></b>	<b><i>MiCAS1</i></b>
<b>ORF length</b>	2316 bp	2289 bp
<b>AA length</b>	771	762
<b>Predicted MW</b>	89.2 kDa	86.8 kDa
<b>Predicted pI</b>	6.04	5.88
<b>Identity to each other</b>	52% <i>MiCAS1</i>	-
<b>Highest protein Blast identity</b>	65-71% $\beta$ -amyrin synthases	72-81% cycloartenol synthases

*MiFRS* = friedelin synthase from *Maytenus ilicifolia*; *MiCAS1* cycloartenol synthase 1 from

*Maytenus ilicifolia*; bp = base pair; kDa = kilo daltons.

**Supplementary Table S2.** 1D NMR spectral data of friedelin (CDCl<sub>3</sub>, 14.1T).

Position	<sup>13</sup> C*	<sup>1</sup> H
1	22.3	1.97 (m)
2	41.5	2.38 (m)
3	213.3	---
4	58.2	2.25 (m)
5	42.1	---
6	41.2	1.28 / 1.77 (m)
7	18.2	1.40 (m) / 1.50 (m)
8	53.1	1.40 (m)
9	37.4	---
10	59.4	1.53 (dd, <i>J</i> = 2.5, 12.5)
11	35.6	1.26 (m) / 1.48 (m)
12	30.5	1.19 (m) / 1.35 (m)
13	39.7	---
14	38.3	---
15	32.4	1.28 (m) / 1.50 (m)
16	36.0	1.36 (m) / 1.58 (m)
17	30.0	---
18	42.8	1.57 (d, <i>J</i> = 5.0)
19	35.3	1.21(m) / 1.38 (m)
20	28.1	---
21	32.7	1.53 (m) / 1.30 (m)
22	39.2	0.94 (m) / 1.51(m)
23	6.8	1.00 (d, <i>J</i> = 6.8)
24	14.6	0.73 (s)

**Supplementary Table S2. Continuation**

---

<b>25</b>	17.9	0.87 (s)
<b>26</b>	20.2	0.88 (s)
<b>27</b>	18.6	1.05 (s)
<b>28</b>	32.1	1.18 (s)
<b>29</b>	35.0	0.95 (s)
<b>30</b>	31.8	1.01 (s)

---

\* Spectral data is in accordance to the literature<sup>1</sup>.

**Supplementary Table S3.** The primers used in this study.

Name	Sequence 5' - 3'	Target	Notes
<b>Primers for cloning of the OSCs</b>			
Core amplification			
TS1F <sup>2</sup>	TTYGGNWDGYCARNHRTGGGAT <sup>1</sup>	PTS	Degenerated
CS1F <sup>2</sup>	TAYAATGGMAGYCARYTVTGGGA <sup>1</sup>	CAS	Degenerated
CS1R <sup>2</sup>	TRRCTYTCHCYCCAHCACC <sup>1</sup>	CAS/PTS	Degenerated
Core sequence extension			
BWN1F <sup>2</sup>	CARWAYRDDGAYGGNGGNTGGG <sup>1</sup>	all	Degenerated
GSP1R	GAGTAGGAGCTACTTCATCTGTGAG	MiFRS	
GSP3R	GACACAGACCATACTCTTCAAC	MiCAS1	
3'RACE			
GSP1F	GGAGGGCTAGCAGCAGCTGG	MiFRS	
GSP3F	AGGGGCTTGTAGCAGCTGGG	MiCAS1	
NGSP1F	TGTCAGGCTGTGCGCAGAG	MiFRS	Nested PCR
NGSP3F	AGCCTTCGTAAGGCTTGCGC	MiCAS1	Nested PCR
AUAP	GGCCACGCGTCGACTAGTAC	all	
5'-end RACE			
GSP5R	GCATCTTCGTCTGGTCCCCTC	MiFRS	cDNA
GSP7R	GTCCAAGATCCATTTCTTCC	MiCAS1	cDNA
GSP9R	GCAAACATAAGTGAATGCTGTGGC	MiFRS	
GSP11R	CACCCTCTCCACCCTCAG	MiCAS1	
NGSP5R	GCAAACACTGTGCTGTGTCCC	MiFRS	Nested PCR
NGSP7R	AGCAACCTCAAAGTGACATAGTTC	MiCAS1	Nested PCR
AUAP	GGCCACGCGTCGACTAGTAC	all	
Full length cDNAs			
MiFRS1F	CGCGGATCCATCATGTGGAAGATAAAGATTGC	MiFRS	<i>Bam</i> HI site
MiFRS1R	CGGAATTCCTGTGTAAGTGGACATATCATATGA	MiFRS	<i>Eco</i> RI site
MiCAS1F	CGCGGATCCGCGACAATGTGGAGGCTCAAATT	MiCAS1	<i>Bam</i> HI site
MiCAS1R	TGCTCTAGACAAATTCAAAGAGACTGTAATAC	MiCAS1	<i>Xba</i> I site
<b>q-RT-PCR</b>			
MI40SF	TCCTCTTGACCTGATTTGGG	40S	
MI40SR	ACTTTTGGGGGATGGACTTC	40S	
RTMiFRS1F	CCTCCAAGAGCCAAAACTG	MiFRS	
RTMiFRS1R	CCATAGGGTTGAGCATCTCC	MiFRS	
RTMiCAS1F	GTTGGGGAACCTTTGGATG	MiCAS1	
RTMiCAS1R	GTTGCAAACCCACCATCAG	MiCAS1	
<b>Site-directed mutagenesis</b>			
L482IF	GATCATGGATGGCAAATTTCTGATTGCACTG	MiFRS	
L482IR	CAGTGCAATCAGAAATTTGCCATCCATGATC	MiFRS	
L482TF	GATCATGGATGGCAAACCTTCTGATTGCACTGC	MiFRS	
L482TR	GCAGTGCAATCAGAAAGTTTGCCATCCATGATC	MiFRS	
L482VF	GATCATGGATGGCAAAGTTTCTGATTGCACTG	MiFRS	
L482VR	CAGTGCAATCAGAAACTTGGCCATCCATGATC	MiFRS	

The letter F or R at the end of the name of each primer indicates if the primer is forward or reverse, respectively. PTS = pentacyclic triterpene synthase. CAS = cycloartenol synthase. GSP



= gene specific primer; NGSP = nested gene specific primer. Endonuclease restriction sites are underlined. Changed codons in the primers for site-directed mutagenesis are in bold.

**Supplementary Table S4.** The sequences used in the phylogenetic analysis with their GenBank accession number, species and function.

<b>Accession number</b>	<b>Species</b>	<b>Function</b>
AB257562.1	<i>Arabidopsis thaliana</i>	Arabidiol synthase (PEN1)
AY327541.1	<i>Arabidopsis thaliana</i>	Thalianol synthase (PEN4)
BT020456.1	<i>Arabidopsis thaliana</i>	Marneral synthase (PEN5)
AB274959.1	<i>Arabidopsis thaliana</i>	Multifunctional triterpene synthase (PEN6)
NM_179572.1	<i>Arabidopsis thaliana</i>	Lupeol synthase 1 (LUP1)
NM_106545.3	<i>Arabidopsis thaliana</i>	Lupeol synthase 2 (LUP2)
NM_126681.2	<i>Arabidopsis thaliana</i>	Cycloartenol synthase (CAS1)
AB263204.1	<i>Rhizophora stylosa</i>	Multifunctional triterpene synthase (RsM2)
AB289586.1	<i>Bruguiera gymnorhiza</i>	Lupeol synthase (BgLUS)
AB257507.1	<i>Kandelia candel</i>	Multifunctional triterpene synthase (KcMS)
DQ268869.1	<i>Ricinus communis</i>	Lupeol synthase (RcLUS)
AB058643.1	<i>Luffa cylindrica</i>	Isomultiflorenol synthase (LcIMS1)
AB037203.1	<i>Glycyrrhiza glabra</i>	$\beta$ -amyrin synthase (GgbAS1)
AB181244.1	<i>Lotus japonicus</i>	$\beta$ -amyrin synthase (OSC1)
AB034802.1	<i>Pisum sativum</i>	$\beta$ -amyrin synthase (PSY)
AF478453.1	<i>Medicago truncatula</i>	$\beta$ -amyrin synthase (AMY1)
AF478455.1	<i>Lotus japonicus</i>	Multifunctional $\beta$ -amyrin synthase (AMY2)
AB034803.2	<i>Pisum sativum</i>	Mixed-amyrin synthase (PSM)

**Supplementary Table S4. Continuation**

---

AB009030.1	<i>Panax ginseng</i>	$\beta$ -amyrin synthase (PNY1)
AB014057.1	<i>Panax ginseng</i>	$\beta$ -amyrin synthase (PNY2)
AB055512.1	<i>Betula platyphylla</i>	$\beta$ -amyrin synthase (BPY)
HM623868.1	<i>Kalanchoe daigremontiana</i>	Taraxerol synthase (KdTAS)
AB263203.1	<i>Rhizophora stylosa</i>	Multifunctional triterpene synthase (RsM1)
AB206469.1	<i>Medicago tirucalli</i>	$\beta$ -amyrin synthase (EtAS)
HM623870.1	<i>Kalanchoe daigremontiana</i>	Friedelin synthase (KdFRS)
HM623869.1	<i>Kalanchoe daigremontiana</i>	Glutinol synthase (KdGLS)
HM623871.1	<i>Kalanchoe daigremontiana</i>	Lupeol synthase (KdLUS)
AB025343.1	<i>Olea europaea</i>	Lupeol synthase (OEW)
AB025345.1	<i>Taraxacum officinale</i>	Lupeol synthase (TRW)
AB181245.1	<i>Lotus japonicus</i>	Lupeol synthase (OSC3)
AB116228.1	<i>Glycyrrhiza glabra</i>	Lupeol synthase (GgLUS1)
AB055511.1	<i>Betula platyphylla</i>	Lupeol synthase (BPW)
AB181246.1	<i>Lotus japonicus</i>	Cycloartenol synthase (OSC5)
AB025968.1	<i>Glycyrrhiza glabra</i>	Cycloartenol synthase (GgCAS1)
D89619.1	<i>Pisum sativum</i>	Cycloartenol synthase (PSX)
AB009029.1	<i>Panax ginseng</i>	Cycloartenol synthase (PNX)
HM623872.1	<i>Kalanchoe daigremontiana</i>	Cycloartenol synthase (KdCAS)
AB033335.1	<i>Luffa cylindrica</i>	Cycloartenol synthase (LcCAS1)
AB055509.1	<i>Betula platyphylla</i>	Cycloartenol synthase (BPX1)
	<i>Maytenus ilicifolia</i>	Friedelin synthase (MiFRS)
	<i>Maytenus ilicifolia</i>	Cycloartenol synthase (MiCAS1)
	<i>Maytenus ilicifolia</i>	Cycloartenol synthase (MiCAS2)

---

## REFERENCES

- 1 Oliveira, D. M. *et al.* Combined experimental powder X-ray diffraction and DFT data to obtain the lowest energy molecular conformation of friedelin. *Quím. Nova* **35**, 1916-1921, (2012).
- 2 Guhling, O., Hobl, B., Yeats, T. & Jetter, R. Cloning and characterization of a lupeol synthase involved in the synthesis of epicuticular wax crystals on stem and hypocotyl surfaces of *Ricinus communis*. *Arch Biochem Biophys* **448**, 60-72 (2006).