ORGANIC LETTERS

2008
Vol. 10, No. 10
1897−1900

Trinorlupeol: A Major Nonsterol Triterpenoid in *Arabidopsis*

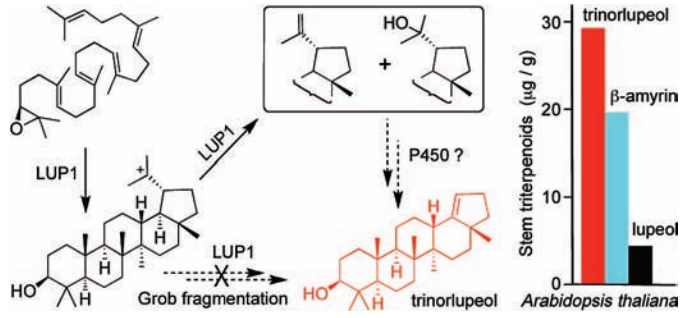
Hui Shan,[†] William K. Wilson,[†] Dereth R. Phillips,[†] Bonnie Bartel,[†] and Seiichi P. T. Matsuda*[,†,‡]

Department of Chemistry and Department of Biochemistry and Cell Biology, Rice University, Houston, Texas 77005

matsuda@rice.edu

Received February 19, 2008

ABSTRACT



We report the structure determination of 20,29,30-trinorlup-18-en-3β-ol (trinorlupeol) and establish this novel C_{27} metabolite as a major nonsterol triterpenoid in *Arabidopsis thaliana*. Trinorlupeol was concentrated in cuticular waxes, notably in the plant stem, floral buds, and seedpods, but not in leaves. Based on expression data and functional characterization of *A. thaliana* oxidosqualene cyclases, we propose that LUP1 is the cyclase responsible for trinorlupeol biosynthesis. Also described are two oxidized trinorlupeols and additional biosynthetic insights.

A global understanding of how and why plants make natural products is most likely to be achieved for the reference flowering plant *Arabidopsis thaliana*. Among thousands of known triterpenoid natural products, the only nonsterol triterpenoids reported in *A. thaliana* are β-amyrin (**1**), lupeol (**2**), and α-amyrin (**3**) (Figure 1).[1,2] We confirmed the presence of **1**−**3** in *A. thaliana* var. Columbia and found them to be concentrated in the cuticle, the protective waxy layer that covers the aerial surfaces of plants.[3] Our analyses of cuticular lipids showed very little sterol but revealed a hydrophobic C_{27} nonsterol (**4**) as a major triterpenoid in *A. thaliana* stem wax. Herein, we describe the isolation of **4**, its structure determination, its presence in *A. thaliana* tissues, and insights into its biosynthesis.

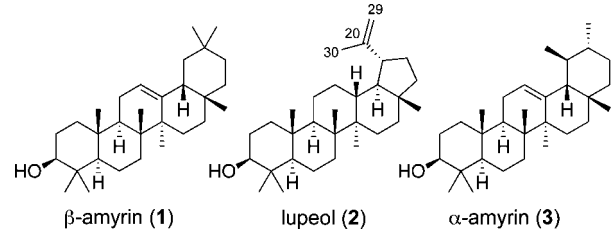


Figure 1. Nonsterol triterpenoids reported in *A. thaliana*.

A sample of **4** was isolated by extracting aerial *A. thaliana* tissue with hexane and purifying the extract by solid-phase extraction (SPE), preparative TLC, and reversed-phase

[†] Department of Biochemistry and Cell Biology.
[‡] Department of Chemistry.

(1) Reviews: (a) D'Auria, J. C.; Gershenzon, J. *Curr. Opin. Plant Biol.* **2005**, *8*, 308–316. (b) Phillips, D. R.; Rasbery, J. M.; Bartel, B.; Matsuda, S. P. T. *Curr. Opin. Plant Biol.* **2006**, *9*, 305–314.

(2) (a) Husselstein-Muller, T.; Schaller, H.; Benveniste, P. *Plant Mol. Biol.* **2001**, *45*, 75–92. (b) Observation of α- and β-amyrin in stem wax: Jenks, M. A.; Tuttle, H. A.; Eigenbrode, S. D.; Feldmann, K. A. *Plant Phys.* **1995**, *108*, 359–377. (c) Ohyama, K.; Suzuki, M.; Masuda, K.; Yoshida, S.; Muranaka, T. *Chem. Pharm. Bull.* **2007**, *55*, 1518–1521.

(3) Kunst, L.; Samuels, A. L. *Prog. Lipid Res.* **2003**, *42*, 51–80.

10.1021/ol800389q CCC: $40.75
Published on Web 04/15/2008

HPLC.[4] An analytical HPLC separation provided **4** in >99% purity. Superficially, **4** resembled a $C_{27}H_{44}O$ sterol[5] as judged by its GC retention time, molecular ion (*m/z* 384), number of ^{13}C NMR signals, and NMR data indicating a single oxygenated carbon atom (C3). However, six upfield methyl singlets in the ^1H NMR spectrum of **4** suggested a nonsterol triterpenoid, and a literature search indicated that **4** was novel.[6]

The structure of **4** was elucidated by 1D and 2D NMR (HSQC, HMBC, COSYDEC, and NOESY). The ^{13}C signals for **4** in rings A, B, and C could be matched within 1 ppm to assigned signals for lupeol,[7] and these assignments were consistent with 2D NMR correlations (Figure 2). Together

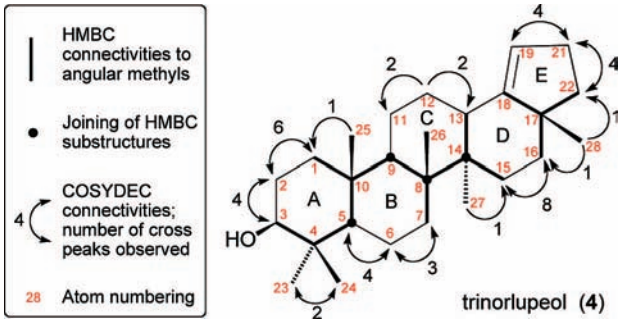


Figure 2. NMR evidence for the structure of **4**.

with $^1H-^1H$ coupling patterns and NOESY results, this reasoning secured the carbon connectivity and stereochemistry of **4** in rings A-C.

The molecular formula for **4** indicated an unsaturation number of 6, corresponding to five rings and one C=C bond. The carbon connectivity of ring E and its link to ring C via C15 and C16 were determined from HMBC and COSYDEC results (Figure 2). The final connection (C13 to C18) was deduced from the unfilled valences for these atoms, NOESY correlations between H19 and the C12 protons, and the deshielded position of H13β. The biosynthetically anticipated configuration of the C17 methyl was confirmed by its NOESY correlations to H13β, H15β, and H22β. The structure of **4** was thus established as 20,29,30-trinorlup-18-en-3β-ol (trinorlupeol) and confirmed by quantum mechanical modeling, in which observed and calculated NMR chemical shifts had root-mean-square deviations of <0.6 ppm for ^{13}C and <0.05 ppm for ^1H (see the Supporting Information). ^1H and ^{13}C NMR spectral data are given in Table 1.

To assay trinorlupeol in aerial *A. thaliana* tissues, we developed a simple method that provides separate extracts

Table 1. ^1H and ^{13}C NMR Data for Trinorlupeol (**4**)[a]

^{13}C NMR		^1H NMR		
atom	δ_C	atom	δ_H	$^1H-^1H$ couplings (Hz)
1	38.90	1α	0.961	tdd, 13.2, 4.0, 1.0
2	27.43	1β	1.739	dt, 13.1, 3.6
3	78.97	2α	1.643	dddd, 13.3, 4.8, 4.1, 3.4
4	38.96	2β	1.587	tdd, 13.3, 11.8, 4.0
5	55.54	3α	3.206	ddd, *11, 6, 5*
6	18.28	5α	0.710	dd, 11.8, 2.2
7	34.51	6α	1.525†	m
8	40.47	6β	1.389	tdd, 12.8, 11.8, 2.8
9	51.20	7α	1.337	td, *12.5, 3.4, 0.9*
10	37.31	7β	1.483	dt, 12.3, 3.0
11	20.86	9α	1.328†	dd, *12.6, 2.6*
12	25.85	11α	1.542†	m
13	37.42	11β	1.294	qd, 12.1, 3.6
14	42.75	12α	1.248†	qd, ~*12.5, 4*
15	28.32	12β	1.623	dq, 12.5, 3.3
16	37.09	13β	2.233	ddq, 11.9, 3.6, 1.8
17	45.62	15α	1.146	ddd, 13.2, 4.1, 2.8
18	153.22	15β	1.719	dddd, 13.9, 13.3, 4.3, 0.8
19	118.95	16α	1.437	td, 13.4, 4.0
21	29.82	16β	1.534	m
22	41.51	19	5.052	br q, 2.1
23	27.95	21α	2.198	dddt, 15.5, 9.5, 2.6, 1.7
24	15.37	21β	2.274	ddddd, 15.5, 9.5, 7.9, 3.7, 1.7
25	16.61	22α	1.568	dt, 12.4, 9.5
26	15.74	22β	1.765	ddd, 12.4, 7.9, 1.7
27	14.77	4α-Me	0.970	s
28	23.71	4β-Me	0.771	s
		10-Me	0.877	d, 0.9
		8-Me	1.069	d, 0.9
		14-Me	0.775	d, 0.8
		17-Me	0.995	t, 0.7

[a] Spectra were acquired at 800 MHz (^1H) or 125 MHz (^{13}C) at 25 °C in $CDCl_3$ solution containing 0.5−5 mM **4**. Chemical shifts were referenced to TMS (^1H) or $CDCl_3$ at 77.0 ppm (^{13}C) and corrected for strong coupling effects; accuracy is about ±0.03 ppm (^{13}C) or ±0.001 ppm (^1H) except for ^1H values marked by † (±0.003 ppm). Coupling constants are accurate to ca. ±0.2 Hz except for values in italics (±1 Hz).

of cuticular and internal lipids.[8,9] Plant tissue was first soaked in hexane, which furnished a cuticular lipid extract containing triterpenes. The near absence of sterols in this extract indicated that the hexane extraction did not rupture the plasma membrane. Further extraction of the residual green plant material with 1:1 dichloromethane-methanol gave the "internal" lipid extract, which comprised mainly sterols, chlorophyll, and acylglycerols. This second extract presumably included all cell membranes and cellular contents. The selectivity of the differential extraction indicated that minimal tissue damage occurred during the harvest of plant organs.

(4) The chromatographic behavior of **4** relative to sterols and triterpene alcohols is described in the Supporting Information.

(5) Perhaps **4** was not discovered previously because of its similarity to sterols.

(6) Among 731 compounds in SciFinder Scholar with a formula of $C_{27}H_{44}O$, the only 3β-hydroxy-4,4-dimethyl structure (25,26,27-trinordihydrolanosterol) would not have six upfield ^1H NMR singlets.

(7) Goad, L. J.; Akihisa, T. *Analysis of Sterols*; Blackie (Chapman & Hall): London, 1997; p 419.

(8) Cuticular lipids have commonly been obtained by brief contact of plant material with chloroform, hexane,[2b] or other solvents: (a) Walton, T. J. In *Methods in Plant Biochemistry*; Dey, P. M., Harborne, J. B., Eds.; Academic Press: London, 1990; Vol. 4, pp 105–158. (b) Dragota, S.; Riederer, M. *Ann. Bot.* **2007**, *100*, 225–231.

(9) (a) Our hexane extraction resembles traditional defatting methods but is uncommon for lipid analyses: ref, 7 Chapter 2, pp 43−68. (b) We did not attempt to separate intracuticular and epicuticular lipids, as described in: Buschhaus, C.; Herz, H.; Jetter, R. *Ann. Bot.* **2007**, *100*, 1557–1564, and references therein.

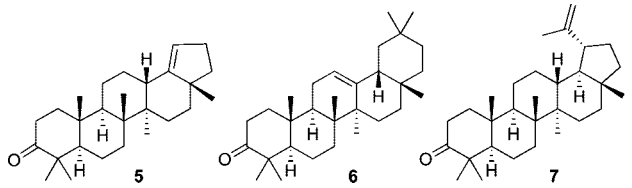


Figure 3. 3-Keto derivatives of trinorlupeol, β-amyrin, and lupeol detected in cuticular and internal lipids of *A. thaliana*.

Nonpolar fractions of the cuticular lipids contained small amounts of 3-keto derivatives of trinorlupeol, β-amyrin, and lupeol (Figure 3), but no triterpenoid esters. Thus, the cuticular lipids were not routinely saponified. In contrast, the internal lipids contained a variety of esters and were necessarily saponified. Subsequent chromatography on silica gel SPE cartridges was used to separate triterpenes from other lipids, followed by GC−MS analysis to quantify the individual triterpenes.

Trinorlupeol and β-amyrin were easily detected in stems, siliques (seedpods), and floral buds (Figure 4).[10] In each

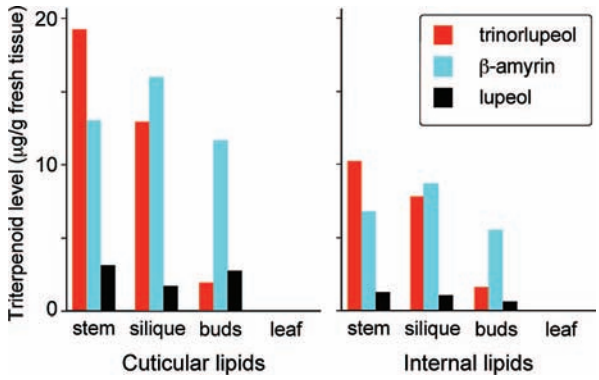


Figure 4. Levels of trinorlupeol and triterpene alcohols in aerial *A. thaliana* organs. Triterpene levels were negligible for both cauline and rosette leaves (leaf).

organ, these triterpenoids were found mainly (2/3) in the cuticle. Lupeol also was detected in these organs but at generally lower levels than trinorlupeol and β-amyrin. Cauline and rosette leaves contained negligible amounts of triterpenes, consistent with the low triterpene levels reported in leaf wax.[2b]

The data in Figure 4 represent a developmental time point that corresponds to plants grown in a short-day photoperiod and containing flowers and many siliques. Although the ratio of trinorlupeol to β-amyrin may vary during growth, trinorlupeol is clearly a major nonsterol triterpenoid in mature *A. thaliana* plants. Trinorlupeol and β-amyrin together comprised >90% of all nonsterol triterpenoids in aerial parts of mature *A. thaliana* plants.[11]

The biosynthetic pathway to **4** may involve the oxidosqualene cyclase LUP1 (At1g78970). Eleven of the 13 *A. thaliana* oxidosqualene cyclases have been at least partially characterized,[12] and only LUP1 makes primarily 6/6/6/6/5 triterpenes (i.e., lupane-3β,20-diol (**9**) and lupeol).[13] Consistent with this proposal, LUP1 is the most highly expressed nonsterol cyclase in stem,[14] where **4** predominates (Figure 4). Trinorlupeol may be the unknown triterpene metabolite hypothesized[2a] to explain the high *LUP1* expression in *A. thaliana* relative to modest levels of lupeol in the plant.

The recent discovery of oxidosqualene-cyclase mediated Grob fragmentations[15] opens the possibility that trinorlupeol could be an oxidosqualene cyclization product (Figure 5).

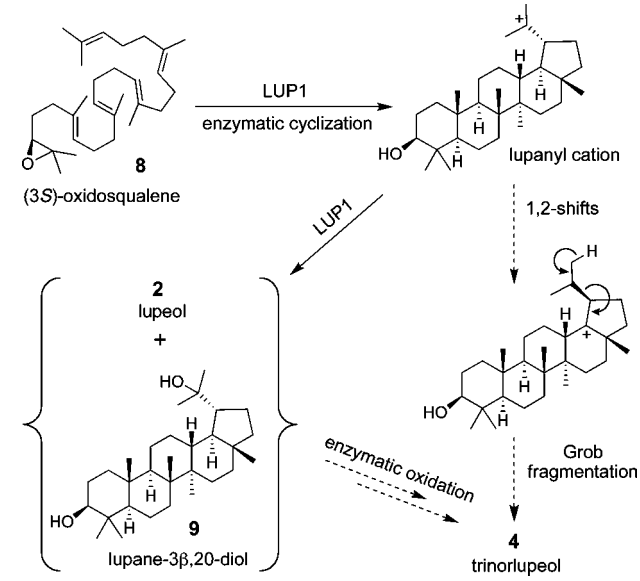


Figure 5. Possible routes to **4** by Grob fragmentation within LUP1 or enzymatic oxidation of LUP1 products **2** and **9**.

However, detailed examination of the products of LUP1 and several other heterologously expressed *A. thaliana* cyclases that access the lupanyl cation failed to reveal even minor amounts of **4**.[16] The inability of LUP1 to generate **4** may reflect that 1,2-hydride shifts for the lupanyl cation are

(10) These patterns of triterpenoid distribution among organs were similar in related *A. thaliana* experiments.

(11) The predominance of **1** and **4** was supported by triterpene analyses for seven sets of aerial *A. thaliana* tissue from flowering plants.

(12) Kolesnikova, M. D.; Wilson, W. K.; Lynch, D. A.; Obermeyer, A. C.; Matsuda, S. P. T. *Org. Lett.* **2007**, *9*, 5223–5226, and references therein.

(13) (a) Segura, M. J. R.; Meyer, M. M.; Matsuda, S. P. T. *Org. Lett.* **2000**, *2*, 2257–2259. (b) Kushiro, T.; Shibuya, M.; Masuda, K.; Ebizuka, Y. *J. Am. Chem. Soc.* **2000**, *122*, 6816–6824.

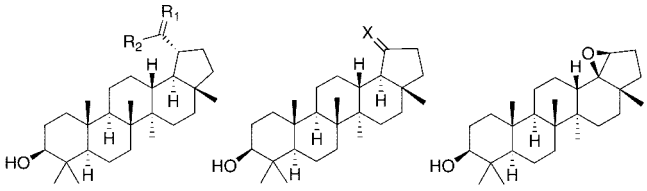
(14) (a) See Table S2 of: Suh, M. C.; Samuels, A. L.; Jetter, R.; Kunst, L.; Pollard, M.; Ohlrogge, J.; Beisson, F. *Plant Physiol.* **2005**, *139*, 1649–1665. (b) Genevestigator: Zimmermann, P.; Hennig, L.; Gruissem, W. *Trends Plant Sci.* **2005**, *10*, 407–409.

(15) (a) Xiong, Q.; Wilson, W. K.; Matsuda, S. P. T. *Angew. Chem., Int. Ed.* **2006**, *45*, 1285–1288. (b) Shibuya, M.; Xiang, T.; Katsube, Y.; Otsuka, M.; Zhang, H.; Ebizuka, Y. *J. Am. Chem. Soc.* **2007**, *129*, 1450–1455.

(16) Morlacchi, P.; Matsuda, S. P. T. Unpublished results.

unfavorable or that no basic residue is suitably positioned to promote Grob fragmentation. The hopene analogue of trinorlupeol (22,29,30-trinorhop-17(21)-ene), which has been identified only from GC−MS spectra of geological sediments,[17] also is evidently not generated directly by cyclases.[18]

Trinorlupeol could alternatively arise through oxidative degradation of the LUP1 products lupeol or lupane-3β,20-diol. A number of side-chain oxidized lupeol derivatives have been reported, often together with C_{27} and C_{29} analogues (Figure 6).[19] Similar oxidative degradations of isopropyl



10 $R_1 = CH_2$ $R_2 = CH_2OH$
11 $R_1 = CH_2$ $R_2 = CHO$
12 $R_1 = O$ $R_2 = CH_3$

13 X = H, OH
14 X = O

15

Figure 6. Structures of side-chain oxidized or degraded lupeols found in lichen (**10**−**13**),[19a] diverse eurosid I species (**10**–**14**),[19b–e] and *A. thaliana* (**15**).

substituents are known in organic synthesis.[20] The *A. thaliana* genome encodes many cytochrome P450 enzymes[21] that might perform the oxidation steps.

Internal extracts of *A. thaliana* contained small amounts of a trinorlupeol isomer and related C_{27} companions that coelute with trinorlupeol on silica gel. We determined the structure of the most abundant such trinorlupeol companion, the 18β,19β-epoxide **15**.[22] Because traces of **15** were also detected in an SPE fraction of cuticular lipids, **15** is unlikely to be a workup artifact. The trinorlupeol companions, which were found mainly in the internal extracts, may be biosynthetic intermediates or minor enzymatic byproducts. Their low levels suggest coordinated cyclization and oxidation to **4**, perhaps through a complex of LUP1 and oxidases.

Is trinorlupeol found in other plants? A promising candidate is birch bark, which is rich in lupeol and its 28-hydroxy derivative betulin and may contain norlupanes, such as **12**.[19d] However, our analysis of birch bark showed only lupeol and betulin, with no trace of trinorlupeol (<0.2% of the amount of betulin). Cruciferous vegetables, a significant component of the human diet, represent additional candidates for trinorlupeol content. Although these food crops are in the same family (Brassicaceae) as *A. thaliana*, our analyses of several vegetables representing a variety of species showed only common triterpenes (e.g., **1**−**3**), with no trace of trinorlupeol in any sample (see the Supporting Information). This finding is consistent with the apparent rarity of LUP1 orthologs in the Brassicaceae.[23] Oxidosqualene cyclases readily evolve to give new products,[1b,24] and trinorlupeol synthesis may have evolved narrowly, conceivably within *Arabidopsis*.[25,26]

In summary, we have isolated and identified trinorlupeol as a major nonsterol triterpenoid in *A. thaliana*. Many intriguing questions remain about the metabolism of trinorlupeol, its distribution in nature, and its potential biological function in defense against pathogens and herbivores.

Acknowledgment. The National Science Foundation (MCB-0209769), The Robert A. Welch Foundation (C-1323 and C-1309), and the Herman Frasch Foundation supported this work. The 800 MHz NMR spectra were acquired with instrumentation purchased by the John S. Dunn, Sr. Gulf Coast Consortium for Magnetic Resonance. We thank J. R. Rasbery and L. C. Strader (Rice University) for help with tissue collection.

Supporting Information Available: Details of experimental procedures; GC−MS and NMR spectra; quantum mechanical NMR predictions; characterization of 3-keto-triterpenoids, epoxide **15**, and other C_{27} triterpenoids; and analysis of triterpenes in cruciferous vegetables and birch bark. This material is available free of charge via the Internet at http://pubs.acs.org.

OL800389Q

(17) Simoneit, B. R. T. *Deep-Sea Res.* **1977**, *24*, 813–820.

(18) Extensive study of squalene-hopene cyclases and their mutants has not revealed any C_{27} products: (a) Hoshino, T.; Sato, T. *Chem. Commun.* **2002**, 291–301. (b) Pale-Grosdemange, C.; Feil, C.; Rohmer, M.; Poralla, K. *Angew. Chem., Int. Ed.* **1998**, *37*, 2237–2240.

(19) (a) Corbett, R. E.; Cong, A. N.; T; Holland, P. T.; Wilkins, A. L. *Aust. J. Chem.* **1987**, *40*, 461–468. (b) Chung, M.-I.; Su, H.-J.; Lin, C.-N. *J. Nat. Prod.* **1998**, *61*, 1015–1016. (c) Silva, S. R. S.; Silva, G. D. F.; Barbosa, L. C. A.; Duarte, L. P.; Vieira Filho, S. A. *Helv. Chim. Acta* **2005**, *88*, 1102–1109. (d) Habiyaremye, I.; Stevanovic-Janezic, T.; Riedl, B.; Garneau, F.-X.; Jean, F.-I. *J. Wood Chem. Technol.* **2002**, *22*, 83–91. (e) Wijeratne, D. B. T.; Kumar, V.; Sultanbawa, M. U. S. *J. Chem. Soc., Perkin Trans. 1* **1981**, 2724–2726. Analogous oxidized norhopanes are also known: (f) Shinozaki, J.; Shibuya, M.; Masuda, K.; Ebizuka, Y. *FEBS Lett.* **2008**, *582*, 310–318, and references therein.

(20) Herz, J. E.; Swaminathan, S.; Wilson, W. K.; Schroepfer, G. J., Jr. *Tetrahedron Lett.* **1991**, *32*, 3923–3926, and references therein.

(21) (a) Schuler, M. A.; Duan, H.; Bilgin, M.; Ali, S. *Phytochem. Rev.* **2006**, *5*, 205–237. (b) Ehlting, J.; Provart, N. J.; Werck-Reichhart, D. *Biochem. Soc. Trans.* **2006**, *34*, 1192–1198.

(22) The Supporting Information contains the structure elucidation of **15** and spectral data for other trinorlupeol companions.

(23) Although the only fully sequenced genome in Brassicaceae is currently *A. thaliana*, expressed sequence tag (EST) data are available for other species. BLAST searches of ESTs in Brassicaceae revealed many genes from the LUP clade but did not uncover obvious LUP1 orthologs.

(24) Xiong, Q.; Rocco, F.; Wilson, W. K.; Xu, R.; Ceruti, M.; Matsuda, S. P. T. *J. Org. Chem.* **2005**, *70*, 5362–5375.

(25) The almost identical lengths of nonterminal exons among genes in the LUP clade suggest that the cyclase portion of trinorlupeol synthesis may have arisen largely through point mutations. In contrast, *A. thaliana* evolved mainly by gene deletion, karyotype reduction, and chromosomal rearrangements: Schranz, M. E.; Song, B.-H.; Windsor, A. J.; Mitchell-Olds, T. *Curr. Opin. Plant Biol* **2007**, *10*, 168–175.

(26) Many secondary metabolites have a narrow biological distribution. For example, avenacin triterpenoids are found only in a single genus: Haralampidis, K.; Bryan, G.; Qi, X.; Papadopoulou, K.; Bakht, S.; Melton, R.; Osbourn, A. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 13431–13436.