This document is confidential and is proprietary to the American Chemical Society and its authors. Do not copy or disclose without written permission. If you have received this item in error, notify the sender and delete all copies.

**Investigation of Premyrsinane and Myrsinane Esters in
Euphorbia cupanii and Euphorbia pithyusa with MS2LDA and
Combinatorial Molecular Network Annotation Propagation**

Journal:	<i>Journal of Natural Products</i>
Manuscript ID	np-2018-00916g
Manuscript Type:	Full Paper
Date Submitted by the Author:	02-Nov-2018
Complete List of Authors:	<p>Nothias-Esposito, Mélissa; ICSN-CNRS, Nothias, Louis-Félix; University of California San Diego, Da Silva, Ricardo ; University of California San Diego, Skaggs School of Pharmacy Retailleau, Pascal; Institut de Chimie des Substances Naturelles, UPR 2301, CNRS Zhang, Zheng; University of California San Diego, Skaggs School of Pharmacy LEYSEN, Pieter; Katholieke Universiteit Leuven, Microbiology and Immunology Touboul, David; CNRS, ICSN Roussi, Fanny; CNRS, Institut de Chimie des Substances Naturelles Paolini, Julien; Université de Corse Pascal Paoli, UMR CNRS 6134 Dorrestein, Pieter; University of California San Diego, Skaggs School of Pharmacy Litaudon, Marc; CNRS, Essonne</p>

SCHOLARONE[®]
Manuscripts

Investigation of Premyrsinane and Myrsinane Esters in
Euphorbia cupanii and *Euphorbia pithyusa* with MS2LDA and
Combinatorial Molecular Network Annotation Propagation

Mélissa Nothias-Esposito,^{†,‡} Louis-Félix Nothias,[§] Ricardo R. da Silva,[§] Pascal

Retailleau,[‡] Zheng Zhang,[§] Pieter Leyssen,^{||} Fanny Roussi,[‡] David Touboul,[‡] Julien

Paolini,[†] Pieter C. Dorrestein,^{*,§} and Marc Litaudon^{*,‡}

[†]Laboratory of Natural Products Chemistry, UMR CNRS SPE 6134, University of
Corsica, 20250, Corte, France

[‡]Institute of Natural Substances Chemistry, CNRS UPR 2301, University of Paris-
Saclay, 91198, Gif-sur-Yvette, France

[§]Collaborative Mass Spectrometry Innovation Center, and Skaggs School of
Pharmacy and Pharmaceutical Sciences, University of California San Diego, La Jolla,
92093, California, USA *United States*

MANUSCRIPT
MARKED BY
EDM

1
2
3 "Laboratory for Virology and Experimental Chemotherapy, Rega Institute for Medical
4
5
6 Research, KU Leuven, 3000 Leuven, Belgium
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

ABSTRACT: The species *Euphorbia pithyusa* and *Euphorbia cupanii* are two closely

related Mediterranean spurge for which ~~taxonomical~~ ^{their} relationships are still ^{being} debated.

Herein, ~~we investigated~~ ^{was investigated} the diterpene esters ~~content~~ ^{content} of *E. cupanii* using liquid

chromatography coupled to tandem mass spectrometry. The use of molecular

networking coupled to unsupervised substructure annotation (*MS2LDA*) ^{demonstrated} ~~exposed~~

^{readily} ^{some} rapidly new premyrsinane/myrsinane type esters ^{the} in *E. cupanii* samples, and a

structure-guided isolation procedure yielded sixteen myrsinane (**11a-h**, **12** and **13**)

and premyrsinane esters (**14a-c** and **15a-c**). In addition, four 4β -phorbol esters

(**16a-c**, **17**) were also isolated and showed inhibitory activity against ~~the~~ chikungunya

virus replication. The structures of the sixteen new compounds (**11a-c**, **11h**, **12**, **13**,

14a-c, **15a-c**, **16a-c**, and **17**) were characterized by NMR ^{spectroscopy} and X-ray ~~diffraction~~

crystallography. To further uncover the diterpene esters content of these two species,

we introduced the concept of Combinatorial Network Annotation Propagation (~~C-~~

NAP). By leveraging the fact that ^{the} ~~Euphorbia's~~ ^{of Euphorbia species} diterpene esters are made of limited

building blocks, ~~we created~~ ^{was created} a combinatorial database of theoretical structures that

1
2
3 were used for C-NAP that made possible the annotation of 123 premyrsinane or
4
5
6
7 myrsinane esters, from which 74% are not found in any structure^{al} database.
8
9

10
11
12 Keywords

13
14
15 *Euphorbia cupanii*; LC-MS/MS; Molecular Networking; Diterpenoids; Combinatorial

16
17
18 Structure Generation; Combinatorial Network Annotation Propagation;
19
20
21

22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
omit - not part of
journal style

The taxonomic classification of *Euphorbia* species (Euphorbiaceae) has been subject to intense debate in scientific literature due to morphological diversification

within the genus and ~~of~~ frequent intra-specific polymorphism. Indeed, spurge

morphotypes can range from prostrate to annuals, to trees of more than twenty

meters, and even cactiform succulent.¹ *Euphorbia* is one of the largest angiosperms

genus with more than 2000 taxa widespread throughout the world, but particularly

diverse in the dry areas of the African and Asian continents, and in Mediterranean

basin.^{2,3} The systematic classification of *Euphorbia* species has been achieved by

morphological comparison, in particular of seeds and glands.⁴ Recent phylogenetic

studies⁵⁻¹¹ showed that the *Euphorbia* genus is composed of four subgenus clades

(*Rhizanthium*, *Esula*, *Euphorbia*, and *Chamaesyce*) and established some of the key

intra-clades relationships. These studies relied on the partial sequencing of the

nuclear and/or chloroplast genes to establish the relatedness between species, but

do not inform on the specialized metabolites production, which are encoded by other

genomic sequences. In addition, as the speciation and phenotype formation

processes are driven by environmental adaptation, they are also shaped by the

modulation of the genome transcriptional activity.¹² For those reasons, transcriptomics and/or metabolomics-based approaches are well suited to functionally characterize variation in specialized metabolite production which dictates species adaptation to its ecosystem.

Mass spectrometry-based metabolomics, in particular molecular networking, has been proposed as an efficient tool to characterize the production of specialized metabolites in plants.¹³ This approach has been used to investigate the molecular profile of species from the *Euphorbia* genus^{14,15} and New-Caledonian Euphorbiaceae,¹⁶ and has ~~even~~ been employed to ~~uncover~~ the production of specialized metabolites in 43 species across the four *Euphorbia* subgenera.¹⁷ In a recent work, we have studied the aerial parts extract of *Euphorbia pithyusa* L. using molecular networking, and isolated seven premyrsinane, one myrsinane, and two tiglane diterpene esters.¹⁴ This species, found on the coastal strip of Western

Mediterranea, is closely related to *Euphorbia cupanii* Guss. ex Bertol. [synonym: *Euphorbia pithyusa* ssp. *cupanii* (Guss. ex Bertol.) (Radcl.-Sm)], an endemic taxon found in the Corsica, Sardinia, and Sicily islands.¹⁸ Appendino and colleagues

studied the aerial parts of *E. cupanii* collected in Sardinia,¹⁹ and discovered eleven new diterpenoids belonging to the lathyrane, premyrsinane and tiglane types, which suggest some similarities with *E. pithyusa* metabolites.

The taxonomical relationship between endemic *E. cupanii* and *E. pithyusa* has been extensively debated over the last two centuries. Initially mentioned by Gussone in 1827 (*E. cupanii* Guss.) and described by Bertoloni in 1842 (*E. cupanii* Guss. ex Berthol.), *E. cupanii* was later proposed as a variety in 1901 [*E. pithyusa* var. *cupanii*

(Guss. ex Berthol.) Fiori],²⁰ and subsequently as a subspecies in 1968 [*E. pithyusa*

subsp. *cupanii* (Guss. ex Berthol.) Radcl.-Sm]. Nowadays, the sub-genic ranking

appears to be accepted by numerous taxonomic resources,^{21,22} while others accept

the species rank.^{23,24} Contemporary botanical studies have shown that these two taxa

have distinctive morphological characters (height, leaves, cyathium, glands,

capsules, and seeds), along with both different biotopes (Table S1 and Figure S2,

Supporting Information), and a different number of chromosomes,²⁵ which taken

together, led the authors to consider *E. cupanii* Guss. ex Berthol. as a paleoendemic

taxon, distinct of *E. pithyusa* L.²⁰

In the present study, we ~~conducted~~ ^{detailed} a fine chemotaxonomic comparison of *E. pithyusa* and *E. cupanii* extracts ^{was conducted} from specimen collected in Corsica using liquid chromatography and tandem mass spectrometry (LC-MS/MS) to compare and characterize the molecular profiles of these two closely related species. By leveraging ~~our~~ previous work on *E. pithyusa*, their diterpene esters ~~content~~ was investigated using computational mass spectrometry annotation tools (molecular networking and unsupervised substructure annotation), which enabled the targeted-isolation of putatively new diterpene esters in *E. cupanii*. Moreover, ~~we~~ ^{is} introduced here the concept of ~~Combinatorial Network Annotation Propagation~~ ^{process of determination} (C-NAP) that made possible to guide the ~~annotation~~ ^{it} of the diterpene esters ~~content~~ ^{previously} for these two species, including molecules ~~never~~ ^{not} described before.

RESULTS AND DISCUSSION

^{Comparison of} ~~Comparing the~~ ^{D E P 3} diterpene ester profile of *Euphorbia pithyusa* and *Euphorbia cupanii* using MS/MS ^{M N} molecular networking and ^{U S} unsupervised substructure ^{C NAP} annotation. ^{START TEXT IN THIS LINE}

MOVE TO LDT LINE OF PAGE 8

The aerial parts of *E. cupanii* were fractionated using the experimental protocol previously employed for *E. pithyusa*.¹⁴ Both plants were collected at similar vegetative states (end of the flowering cycle). The ethyl acetate extract was partitioned by successive liquid-liquid extractions, allowing the obtention of a diterpenoid-rich extract. This latter was subjected to normal-phase flash chromatography resulting in 14 fractions. To compare the diterpenoid profiles of *E. pithyusa* and *E. cupanii* collected in Corsica, the chromatographic fractions of the diterpenoid-rich extracts from both species were analyzed by untargeted LC-ESI-MS/MS analysis at the same concentration. The LC-MS/MS data were analyzed using the *feature-based molecular networking workflow*^{26,27} using MZmine2,²⁸ on the GNPS web-platform.²⁹ The generated molecular networks (MWS) were visualized in Cytoscape,³⁰ with the node pie chart diagram represent the spectral feature intensity (relative quantification using LC-MS peak area) between the two species (Figure S3, Supporting Information). The molecular networks showed that many of the spectral features were species-specific, with 81% of the spectral ion features being specific of *E. pithyusa* or *E. cupanii*, while only 19% were shared (Figure S4A, Supporting

Information). By performing ^(a) spectral library search against ^(b) public MS/MS reference spectra available on GNPS, including reference spectra of the molecules previously isolated from *E. pithyusa* (1–10),¹⁴ the main molecular networks (A, B, and C) were annotated as triterpenoids, *Euphorbia* diterpene esters, and tiglane esters, respectively (Figure S3, Supporting Information). Interestingly, it was observed that the molecular profile of *E. pithyusa* was characterized by a large diversity of triterpenoid derivatives, while *E. cupanii*'s profile appeared to be dominated by diterpene esters. Regarding the diterpene esters ~~the~~ molecular networks (Figure 1A), ^(c) ~~we~~ it was observed that 93% were found exclusively in only one of the two species (Figure S4B, Supporting Information). These results showed that the speciation process between these two species was associated with differential production of specialized metabolites.

The annotation of the diterpene esters was further detailed using the reference MS/MS spectra of premysinol, myrsinol, and dideoxyphorbol esters (1–10) obtained from *E. pithyusa* (Figure 1B).¹⁴ The spectral library search in analogue mode was used to propagate the annotation to related nodes. As previously described,¹⁴ the

spectra of the 14-oxoprenylsinane esters (1–7) were observed in cluster I (characterized by fragment ions at m/z 295, 277, 267, and 149), while the cluster II contained spectra of myrsinane derivatives (fragment ions at m/z 295, 277, 265, and 247). The MS/MS spectra for the previously isolated dideoxyphorbol esters (9 and 10) were found in a separate cluster (III), and presented the fragment ions at m/z 335, 307, 295, and 277.¹⁴ Overall, results showed that the 14-oxoprenylsinol and dideoxyphorbol type esters (cluster I and III, respectively) were predominantly found in *E. pithyusa*, while the myrsinane esters (cluster II), along with other unannotated diterpene esters, were found essentially only in *E. cupanii*, as depicted in the Figure 1.

Figure 1. *Euphorbia* diterpene esters and tiglane molecular networks from the two species extracts. (A) diterpene esters comparison between *E. pithyusa* (blue) and *E. cupanii* (red). (B) analogue search and reference compound annotations with the previously isolated compounds (1–10) from *E. pithyusa*, and the MN-guided isolated compounds from *E. cupanii* (11a–h, 12, 14b–c and 15a–c), and their clustering (I–V).

To refine the annotation of the observed diterpene esters, we performed unsupervised substructure annotation of the fragmentation spectra with the *MS2LDA* web app.²⁶ This unsupervised approach used the latent Dirichlet allocation (LDA)

algorithm³¹ to discovery frequently co-occurring fragments and neutral losses (called

motifs or Mass2Motifs) from the fragmentation spectra, which indicates the ~~existence~~ ^{occurrences}

of a structural features. The results of *MS2LDA* were then mapped into the diterpene

esters/ molecular network for subsequent interpretation (Figure 2 and Table S5,

Supporting Information). ^{Significantly} Remarkably, *MS2LDA* recovered efficiently Mass2Motifs

corresponding to fragmentation already described for 14-oxopremyrsinol

(Mass2Motifs 240 and 129) and myrsinol esters (Mass2Motifs 67 and 219).¹⁴ Further

inspection of the results showed specific motifs (Mass2Motifs 16, 193, and 271) in

molecules observed in *E. cupanii*, suggesting derivatives with uncharacterized

structural features. While ~~the~~ Mass2Motif 276 shared the backbone fragment ions

(*m/z* 247, 277, 265, and 295) with ~~the~~ Mass2Motif 67 indicating a myrsinol ester

derivatives. In addition, ~~we also observed~~ a motif (Mass2Motif 262) ^{ed (the presence of} corresponding to

the loss of 124 Da ^{was also observed} indicating the presence of nicotinoylated derivatives. Overall, the

mapping of the annotated Mass2Motifs into the molecular networks was found to be

consistent with the network's topology, alongside ~~with~~ the annotation obtained in

library spectral matching in analogue search mode. Moreover, the use of

1
2
3
4 unsupervised substructure annotation completed nicely molecular networking, by
5
6
7 revealing some of the spectral features involved in the clustering of the fragmentation
8
9
10 spectra by highlighting the presence of structural motifs.
11
12

13
14 Based on the interpretation of the MS/MS data with molecular networks and
15
16 ~~the next aim was to elucidate~~
17 unsupervised substructure annotation, ~~we aimed at identifying~~ some of the putative
18 ~~~~~~~~~
19 ~~this~~
20 new metabolites produced by *E. cupanii*. For that purpose, a color coding was
21
22
23
24 employed in the molecular network with the piechart nodes showing the compound
25
26
27
28 intensity across the different fractions (Figure S6, Supporting Information), and it was journal
29
30
31 indicated that many of these ions were concentrated in fractions F9–12 (pink and
32
33
34
35 green nodes).
36
37

38
39 These fractions were selected for Molecular-Networking-guided (MN-guided)
40
41
42 isolation using flash chromatography and/or preparative HPLC, allowing the isolation
43
44
45
46 of 20 diterpene esters derivatives (11a–h, 12, 13, 14a–c, 15a–c, 16a–c, and 17),
47
48 for which nd
49 whose 16 have never been described previously.
50 ^^
51
52
53
54
55
56
57
58
59
60

Figure 2. Diterpene esters molecular network from *E. pithyusa* and *E. cupanii* extracts with their Mass2Motifs annotation and their associated fragment ions, generated with *MS2LDA* web-app and visualized with Cytoscape.

um
CAY YAM
INDICATE
MOLE COUNTRY
IN THE
CAPTION
WHICH
COMPOUND
CAME FROM
WHICH
SPECIES?

Identification of diterpene esters isolated from *E. cupanii* extract.

TEXT STARTS ON
THIS LINE

The comprehensive 1D- and 2D-NMR data analysis, mass spectrometry and comparison with literature data were used to identify ten myrsinol (11a-h, 12, and 13) and six derivatives. premysins (14a-c and 15a-c). Four phorbol esters (16a-c, and 17) were obtained in the present study in the course of the isolation procedure, and were investigated for their antiviral potential.^{14,15,26,32,33}

The HRESIMS (High Resolution Electrospray Ionisation Mass Spectrometry) of compound 11a showed a cation at m/z 681.2698 $[M + Na]^+$, corresponding to the molecular formula $C_{38}H_{42}O_{10}$, indicating 18 hydrogen deficiencies. From this formula and its proton NMR spectroscopic data ($CDCl_3$, Tables S7 and S9, Supporting Information), this compound seems to belong to the myrsinane class as the compounds

previously isolated in *Euphorbia myrsinites*,³⁴ *E. seguieriana*,³⁵ and *E. prolifera*.³⁶ The

analysis of its ^{13}C and ^1H NMR spectra revealed the presence of four ester carbonyls

which were assigned to two acetate groups (δ_{C} 172.2 and 167.8; δ_{H} 1.69 and 2.02)

and two benzoate groups [δ_{C} 165.6, 132.6, 129.1, 128.9 and 128.1; δ_{H} 8.01, 7.53,

7.38) and (δ_{C} 165.7, 132.8, 130.3, 129.4 and 129.7; δ_{H} 7.94, 7.55, 7.41)].

Examination of the HSQC spectrum and the carbon chemical shifts of compound **11a**

allowed the determination of three methyl groups, four oxymethines, and three

carbons without HSQC correlations. One proton at δ_{H} 3.85 did not depict correlation

in the HSQC spectrum, was assigned to a free hydroxyl at the C-5 position as

determined by COSY correlation with the oxymethine H-5. From COSY spectrum,

two spin systems were assigned from H-1 to H-5 [δ_{H} 2.96 (dd, J = 16.2, 10.3 Hz, H-

1 α), δ_{H} 2.68 (dd, J = 16.2, 9.2 Hz, H-1 β), 2.23 (m, H-2), two methine at δ_{H} 5.25 (t, J =

3.5 Hz, H-3) and δ_{H} 4.23 (dd, J = 10.1, 2.7 Hz, H-5) linked by H-4 at δ_{H} 2.96 (dd, J =

10.1, 3.5 Hz)], and from H-7 to H-12, as depicted in Figure 3. The second spin

system begin by one oxymethine at δ_{H} 5.63 (d, J = 6.6 Hz, H-7) connected at one *cis*-

substituted double bond at δ_{H} 6.21 (ddd, J = 9.7, 6.6, 1.4 Hz, H-8) and 5.84 (dd, J =

9.7, 5.0 Hz, H-9), which was coupled to two adjacent protons at δ_{H} 3.28 (dd, $J = 5.0$, 4.1 Hz, H-11), and 3.25 (d, $J = 4.1$ Hz, H-12). The HMBC correlation showed an exomethylene group [δ_{C} 112.7, δ_{H} 4.92 and 4.87, br s (H-18a and H-18b, respectively) and δ_{H} 1.91, s, H-19] connected to H-11. This compound was characterized by the presence of an ether bond between C-17 and C-13 [δ_{C} 68.6, δ_{H} 4.20 and 3.48, d, $J = 12.0$ Hz (H-17a and H-17b, respectively)]. As depicted in Figure 3, HMBC correlations between oxymethines and ester carbonyl carbons allowed the attachment of one acetate group and two benzoate groups at C-3, C-7 and C-14, respectively. The last acetate group was assumed to be at the C-15 position. The relative configuration of compound **11a** was established by ROESY experiment (Figure 3). Cross peaks between H-2/H-3, H-4/H-14, H₂-17/H-7/H₃-20 on the one hand, and H-5/H-12 on the other hand indicated that these protons are [⊖]cofacial, and they were oriented as α - and β -, arbitrarily α -oriented and β -oriented, respectively. Therefore, compound **11a** was assigned identified as (3*R*,4*S*,5*R*,7*R*,11*R*,12*R*,13*S*,14*R*,15*R*)-3,15-*O*-diacetyl-7-14-*O*-dibenzoyl-5-hydroxymyrsinol.

Figure 3. Key COSY (bold, left), HMBC (blue arrows, left) and ROESY (red arrows, right) correlations of compound **11a**.

The comparison of NMR and HRESIMS data for the compounds **11b**, **11c**, and **11h** isolated from *E. cupanii*, with the myrsinol ester **11a** and the previously isolated compounds **11d-g**,^{36,37} suggested that these molecules ~~had~~ ^{have} also a myrsinane skeleton. The structure of the ~~compounds~~ **11b**, **11c**, and **11h** were elucidated with the same approach described above for compound **11a**, and it was found that they were also myrsinol esters differing by their acylation pattern. The latter was solved by analyzing ~~inspecting~~ the results of HMBC and ROESY experiments. In addition, compounds **11d**, **11f**, and **11h** were obtained in crystalline form and their crystal X-ray diffraction analysis allowed their absolute configuration to be assigned as shown in ~~the~~ Figure 4 (Table S10, Supporting Information). Based on the results of X-ray diffraction

crystallography for **11d**, **11f**, **11h**, and from their co-occurrence, myrsinane esters

11a–h are assumed to have the same absolute configuration.

Figure 4. ORTEP view of the X-ray structure of compounds **11d**, **11f**, **11h**, **12**, and **15b**.

The structural elucidation of the new compounds **12** and **13** was established using the same approach used for the myrsinol esters **11a–c**. Compound **12** possessed a

molecular formula of $C_{37}H_{46}O_{13}$ ($[M + Na]^+$ at m/z 721.2843), and the 1D- and 2D-

NMR spectroscopic data interpretation confirmed that **12** ^{is} was also a myrsinol ester,

but possessed ^s an additional acetate group at ~~the~~ position C-10. The structure and

absolute configuration of **12** ^{were} was confirmed by X-ray diffraction analysis (Figure 4 and

Table S10, Supporting Information). Thus, the structure of **12** was assigned ^s as

(3*R*,4*S*,5*R*,7*R*,11*R*,12*R*,13*S*,14*R*,15*R*)-3,5,7,10,15-*O*-pentacetyl-14-*O*-benzoyl-

^{In turn, the} 10,18-dihydromyrsinol. The MS spectra of compound **13** showed a protonated

molecular ion at m/z 735.2649 $[M + H]^+$, corresponding to the molecular formula

$C_{39}H_{43}O_{14}$. Its NMR data were comparable to those of compound **12** with four acetate

^s group at C-3, C-6, C-7, C-10 and C-15, but a carbonyl function at C-14 was assumed

instead of a benzoyl group ^{as} for compound **12**. The deshielded carbon at C-2 (88.0)

suggested that ^a the benzoyl group ^{is} was located at this position. Compound **13** was

^{elucidated} identified as an unprecedented 14-oxomyrsinol derivative. ^{the} Give NAME

The examination of NMR data (Tables S8 and S9, Supporting Information) of

compounds **14a-c** and **15a-c** indicated that they ^{possess} were closely comparable to ^{structures}

myrsinanes **11a-h**, bearing however a ^{gem} gem-dimethylcyclopropane moiety at C-9/C-

11 instead of an isopropenyl group in the case of **11a-h**. These compounds,

possessing such a 5/7/6/3-tetracyclic framework, are premyrsinane esters. Their 2D

structure and relative configuration were solved by interpretation of 1D- and 2D- NMR

spectroscopy data. All compounds **14a-c** and **15a-c** possess an ether bond between

C-13 and C-17, similarly to the myrsinol esters **11a-h**, **12**, and **13**. In addition, HMBC

correlations from H-1 and H-12 to the carbonyl C-14 suggested that compounds

14a-c were 14-oxopremyrsinols, whereas for compounds **15a-c**, HMBC correlations

from H-1 and H-12 to the oxymethine C-14 indicated that an acyl function is located

at this position. The location of the acyloxy groups were determined

interpretation of HMBC correlations and the relative configurations of all compounds

were determined by interpretation of ROESY correlations. Additionally, X-ray

diffraction analysis was performed on compound **15b** crystal (Figure 4 and Table

S10, Supporting Information), confirming its absolute configuration. Thus, it can be

proposed that premyrsinol esters **15a** and **15c**, and by extension, the premyrsinol

esters **14a-c** share the same absolute configuration.

1
2
3 exhibited
4 Compound **16a** showed a sodium adduct ion at m/z 607.3246 $[M + Na]^+$ in its
5
6
7 HRESIMS data, consistent with the molecular formula $C_{34}H_{48}O_8$. Examination of its
8
9
10 NMR data ($CDCl_3$, Table S11, Supporting Information) showed the presence of one
11
12
13 ketone carbonyl (δ_C 209.1), two ester carbonyls (δ_C 167.0 and 179.7), and two olefinic
14
15
16 signals attributed to two vicinal-monosubstituted double bonds (δ_C 161.0 (C-1), 133.0
17
18
19 (C-2), and δ_H 7.59 (br d, $J = 10.4$ Hz, H-1)], and [δ_C 140.5 (C-6), 129.4 (C-7), and δ_H
20
21
22 5.68 (d, $J = 4.9$ Hz, H-7)]. Further examination of the 2D-NMR data (COSY and
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

that **16a** was a phorbol diester. The two acyloxy substituents were assigned to a deca-2*E*,4*Z*-dienoyl group (δ_C 167.0, 121.2, 140.1, 126.7, 121.2, 28.5, 29.2, 31.6, 22.9, and 11.8; δ_H 5.84, 7.58, 6.11, 5.86, 2.26, 1.39, 1.27, 1.26, and 0.86) and an isobutyrate moiety (δ_C 179.7, 34.4, 18.8, and 18.7; δ_H 2.57, 1.18, and 1.15). The attachment of the deca-2*E*,4*Z*-dienoyl moiety at C-12 was established thanks to HMBC correlation from the oxymethine proton to the ester carbonyl carbon. From these data, it was deduced that the isobutyrate group was necessarily located at C-13 (Figure 5). The relative configuration of compound **16a** was established based on

Thus, this
the examination of ROESY correlations as depicted in Figure 5. The compound was
assigned
identified as ~~12~~ 12β -O-[deca-2*E*,4*Z*-dienoyl]-13 α -isobutyryl-4 β -phorbol.

Figure 5. Key COSY (bold, left), HMBC (blue arrows, left) and ROESY (red arrows, right) correlations of compound **16a**.

The same general approach was used to elucidate the structure of phorbol esters

16b, **16c**, and **17** (Table S11, Supporting Information). Compounds **16b** and **16c**
with
possessed the same 4 β -phorbol skeleton as compound **16a**. The phorbol ester **16b**

only differed from compound **16a** by the acylation pattern at position C-13, with a 2-

methylobutyryl group instead of an isobutyryl group in compound **16a**. Compound **16c**

was a stereoisomer of 4 β -phorbol **16a**, possessing a deca-2*Z*,4*E*-dienoyl moiety at C-

13 instead of a deca-2*E*,4*Z*-dienoyl moiety for compound **16a**. Diterpene esters **17**

and **16c** shared the same acylation pattern. However, ~~the structure~~ of compound **17**

~~possessed~~ ^{exhibited} the rare 5-ene-7-oxo moiety, previously observed for the trigowiin A,³⁸

^{procedures.} which could be formed by autoxidation during the isolation step.³⁹ Compound **17** was

^{elucidated} identified as 12 β -O-[deca-2Z,4E-dienoyl]-13 α -isobutyryl-5-ene-7-oxo-4 β -phorbol.

^{have been} Several phorboids isolated from *Euphorbia* spp. were shown to potently inhibit the

^{Accordingly,} chikungunya virus replication (CHIKV).^{14,15,26,32} ~~For that reason, we have evaluated~~

^{were evaluated} compounds **16a-c** and **17** for their ability to inhibit the CHIKV replication (Table S12,

^{The results obtained} Supporting Information). ~~Results~~ showed that 4 β -phorbol esters **16a-c** possessed a

^{the} selective anti-CHIKV activity in the sub-micromolar range (EC_{50} CHIKV < 0.8 μ M, SI

> 12). Further evaluation of the bioactivity could not be achieved due to the limited

^{in each case.} amount of material obtained. Compound **17** showed a lower inhibition [EC_{50} CHIKV =

4.5 $\pm$ 0.6 μ M (SI = 6.3)] than **16a-c**, indicating that the presence of an α,β

^{is} unsaturated carbonyl function at C-7 ~~proves~~ deleterious for the anti-CHIKV activity.

This finding was also observed for trigowiin A,³⁸ which possesses the same carbon

backbone as **17** substituted by a C₁₂-chain fatty acid at C-12 and an acetyl moiety at

C-13.

Fragmentation pathway of diterpene esters and Mass2Motif annotation.

The MS/MS spectra of the isolated compounds from *E. cupanii* (**11a-h**, **12**, **13**, **14a-c**, **15a-c**, **16a-c**, and **17**) were deposited ^{into} to the GNPS spectral public library and the molecular networks annotation was updated as shown in the Figure 1B. Examination of the MS/MS spectra for the premyrsinol ester ^{analogues} found in cluster IV and V (**14a-c** and **15a-c**, respectively) ^{premyrsinol ester} allowed the establishment of their characteristic diterpene backbone fragment ions. The premyrsinol esters found in cluster IV (**14a**, **14b**, and **14c**, click for the MS/MS spectra) showed consistent backbone fragment ions at m/z 295, 267, 265, and 237, while the premyrsinane esters **15a**, **15b**, and **15c** (click for the spectra) showed fragment ions at m/z 297, 279, 267, and 249 (Figures S13 and S14, respectively, Supporting Information). The difference of 27.99 Da between fragment ions at m/z 295.1703 ($[C_{20}H_{23}O_2]^+$, $\Delta_{m/z} = 1.67$ ppm) and m/z 267.1743 ($[C_{19}H_{23}O]^+$, $\Delta_{m/z} = 2.21$ ppm) observed in the MS/MS spectra of the compound **14c** (Figure S13, Supporting Information), indicated the neutral loss of a CO ^{as} already observed for the previously isolated premyrsinol esters **(1-7)**.¹⁴ This neutral loss was not observed for the premyrsinol esters **15a-c**.

As shown in Figures S13 and S14 (Supporting Information), the MS/MS spectra of premyrsinol esters **14a-c** and **15a-c** exhibited a characteristic neutral loss of CH₂O (30.01 Da, for m/z 267 → 237, and m/z 279 → 249, respectively), resulting from the cleavage of the ether link between C-13 and C-17. Interestingly, unsupervised substructure annotation (Figure 2) was able to recover these fragmentation patterns, with the Mass2Motif 193 (main fragments at m/z 265 and 237) proposed for compounds **14a-c** and the putative related derivatives. While the Mass2Motifs 16 and 76 were observed for compounds **15a-c**, with backbone fragment ions at m/z 297, 279, and 249, and an additional neutral loss corresponding to an acid benzoic (122 Da) in the Mass2Motif 16. As depicted in the Figures 1B and 2, the 13,17-oxy-14-oxopremyrsinols (**14a-c**) and 13,17-oxypremyrsinols (**15a-c**) were found in the clusters IV and V, respectively.

The cluster II included compounds **11a-h** and along with the other main nodes, exhibited were annotated by the Mass2Motifs 67 and 219, which presented in their fragmentation spectra similar ions as the myrsinane ester **8**¹⁴ (m/z 295, 277, 265, and 247). The compound **12**, a 10,18-dihydromyrsinol ester, was also found in the cluster

II (Figure 1B), and was characterized by the Mass2Motifs 271, also showing fragment

ions at m/z 295, 277, 265, and 247 (Figure 1B). Significantly, Remarkably, the use of MS2LDA

made possible differentiation the 10,18-dihydromyrsinol ester 12 from other

myrsinane esters (11a-h), although they clustered together in the molecular network.

ketone carbonyl group
The compound 13 possesses a ketone at C-14 instead of an ester group and was

not observed in the cluster I-IV with the other myrsinane derivatives (11a-h, and 12).

The acquisition of its MS/MS spectra ([click here](#)) showed the presence of ions at m/z

309, 291, 281, 263, 251, and 233 (Figure S15, Supporting Information), showing a

neutral loss of CH_2O (30.01 Da, for m/z 281 $\rightarrow$ 251 and m/z 263 $\rightarrow$ 233) and of CO

(27.99 Da, for m/z 309 $\rightarrow$ 281 and m/z 291 $\rightarrow$ 263).

Overall, the retrospective analysis of the molecular network with MS2LDA

confirmed that it is mainly composed of premyrsinane and myrsinane derivatives.

many
This interpretation was consistent with most phytochemical investigations conducted

on ~~a~~ *Euphorbia* species, where the diterpene esters discovered in a single species

are polyester ~~s~~ derivatives of few scaffolds of *Euphorbia*'s diterpenes, differing by their

acylation pattern or their oxidation degree.

Finally, the phorbol esters (**16a-c** and **17**) isolated from *E. cupanii* extract were not detected in the molecular networks. Inspection of the LC-MS/MS data showed that these ions were not selected for MS/MS fragmentation in untargeted mode, due to low intensity and/or co-elution with other ions. Yet, the MS/MS spectra of **16a-c** and **17** were acquired from the isolated compounds and are accessible by clicking on the following number: ^s**16a**, **16b**, **16c**, and **17**.

Annotation of diterpene esters using Combinatorial Network Annotation Propagation

(C-NAP)

The use of molecular networking, with unsupervised ~~structure annotation~~ and the partial annotation a reference MS/MS spectra, made possible to ~~partially annotate~~ the fragmentation spectra by providing information on some structural features of these putative premyrsinane and myrsinane derivatives. This partial annotation corresponds to a level 3 annotation based on Metabolomics Standard Initiative.⁴⁰ Multiple computational approaches, such as MetFrag,⁴¹ MaGMA,⁴² Sirius/CSI:FingerID,⁴³ Dereplicator+,⁴⁴ can rank chemical structure candidates for each fragmentation

spectra, which corresponded to a level 2 putative annotation. ^{(In} By essence, these

approaches are limited to structures present in the database(s) employed. Recently,

Network Annotation Propagation (NAP) was introduced for the propagation of

annotations in molecular networks.⁴⁵ In brief, NAP starts by using the MetFusion⁴⁶

concept to re-rank a list of candidate structures proposed by MetFrag for each nodes^X

^(was used to) in the molecular network. Then, NAP re-ranked^X the candidate lists using the ^(by employing)

molecular network topology based on expected structural similarity between

connected components, propagating spectral library annotations (*Fusion* score) or

the consensus of candidates (*Consensus* score) for connected nodes.

^{INDEX} The use of NAP with chemical structures from the Dictionary of Natural Products ^{annotation}

(DNP) and SuperNatural II (SN) databases, allowed the annotation of 69% of the

nodes (244 out of 266 nodes) from the diterpene esters^S molecular network of *E.*

pithyusa and *E. cupanii* (Table 1). Amongst^X the structures ranked at the first position

(TOP1) by MetFrag, only 2.5% corresponded to *Euphorbia* diterpene esters, and only

1.2% to premyrsinane or myrsinane esters. The results for NAP with the *Consensus*

score showed that 10.2% of the structures ranked at TOP1 were *Euphorbia* diterpene

esters, and that 0.8% were premyrsinane or myrsinane esters. Only 180 nodes had a *Fusion* annotation and amongst ~~them~~^{same}, 37.8% had a *Euphorbia* diterpene ester ranked at TOP1, and 16.7% ^{Since} were premyrsinane or myrsinane esters (Figure S16, Supporting Information). ^{was} ^a ~~Because~~ the *Fusion* score is calculated only for node with at least one direct neighboring node with a spectral library annotation (performed here ^{had} in analogue search mode), less nodes ~~have~~ a *Fusion* score annotations than *MetFrag* and *Consensus* score. The results of NAP with the DNP and SN databases were not concordant with the the annotation derived from molecular networking and unsupervised substructure derived annotation ^{that} ~~which~~ indicated that they are premyrsinane and myrsinane derivatives. The low number of likely annotations obtained with NAP was due to the absence of corresponding premyrsinane and myrsinane candidates in the databases employed. Indeed, *Euphorbia* diterpene esters are relatively rare molecules, with approximately 600 entries in the DNP (239,726 total entries), and only 63 and 42 entries for the premyrsinane and myrsinane derivatives, respectively.

Table 1. Annotations Obtained with NAP and C-NAP for the Molecular Network.

Database ?
new annotation

	Nodes annotated (% of nodes)	TOP1 that are diterpene esters (% of annotation)	TOP1 that are premyrsinane or myrsinane esters (% of annotation)
MetFrag (NAP)	244 (93.2%)	6 (2.5%)	3 (1.2%)
Fusion (NAP)	180 (69.0%)	68 (37.8%)	30 (16.7%)
Consensus (NAP)	244 (93.2%)	25 (10.2%)	2 (0.8%)
MetFrag (CNAP)	244 (93.2%)	7 (2.9%)	5 (2.0%)
Fusion (CNAP)	180 (69.0%)	180 (37.8%)	173 (96.1%)
Consensus (CNAP)	244 (93.2%)	217 (88.9%)	216 (88.5%)
Hybrid Fusion Consensus (CNAP)	244 (93.2%)	221 (90.6%)	214 (87.7%)

NO BUILDING
= IN
ANNOTATION
OR COPY
OF THIS
TABLE

To overcome this limitation, ~~we~~^{is} introduced here the concept of the Combinatorial
Network Annotation Propagation (C-NAP). By leveraging the fact that one single

Euphorbia species produces a limit set of diterpene backbone that are polyesterified
Euphorbia with a usual set of short acyl substituents,^{47,48} a Combinatorial Structures DataBase

(CSDB) of premyrsinane and myrsinane theoretical derivatives was created with the

SmiLib software and used for NAP.^{49,50} The generation of CSDB had been

successfully applied for the annotation of the fragmentation spectra with compound

classes that are made of limited building blocks, including lipids⁵¹ and

phenylcoumarin⁵² derivatives. The present CSDB was generated using two

premyrsinane and three myrsinane main scaffolds, along with 10 common acyls

1
2
3
4 "building blocks" found in *Euphorbia* diterpenes (See Figures S17 and S18 in

5
6 Supporting Information). The CSDB initially obtained, consisted of 5.7 millions

7
8
9 structures. ^{Since} Because a large proportion of the structures consisted of positional

10
11
12 isomers differing ⁱⁿ by the location of the acyl substituents (21,306 unique positional

13
14
15 isomers, with 4,340 unique molecular formula), that ^s ^{could} cannot be discriminated in

16
17
18 tandem mass spectrometry without reference standards, ^{was maintained} we kept only one

19
20
21 representative structure per unique positional isomers. The CSDB, alongside with the

22
23
24 DNP and SN databases (more than 30,000 and 79,273 unique molecular formulas

25
26
27 each, respectively), were used for C-NAP (Figure 6). The results showed that for the

28
29
30 nodes annotated with ^{the} *Fusion* score (180 nodes), 96.1% of the structure obtained at ^s ^{the}

31
32
33 TOP1 position were a premyrsinane or a myrsinane derivative (173 nodes). Although

34
35
36 more annotations were obtained with the *Consensus* annotation (244 nodes), the

37
38
39 premyrsinane or myrsinane derivatives were less represented with 88.5% of TOP1

40
41
42 candidates (217 nodes). Nonetheless, a detailed examination of the annotations

43
44
45 proposed by the *Fusion* and the *Consensus* score showed that, despite a higher

46
47
48 number of premyrsinane/myrsinane derivatives obtained with the *Consensus* score,

the annotations obtained with the *Fusion* score were more consistent with the annotations derived from molecular networking and *MS2LDA*. For that reason, ~~we~~ ^(this) have opted for a hybrid approach that ~~we~~ ^{was taken} named *Hybrid-Fusion Consensus*, where ~~we~~ ^{was preferred} for each node, ~~we preferred~~ ^{only} the candidate from the *Fusion* score (Figure 6) and used the candidate from *Consensus* score when no *Fusion* score was available. The summary of the annotations obtained with the different scores is presented in Figure 7.

~~Remarkably, amongst~~ ^{Among} the premyrsinanes and myrsinanes obtained as TOP1 rank with the *Fusion* score, 74% were not present in the DNP or the SN databases. By inspecting the fragmentation spectra for ~~some annotations obtained with C-NAP~~ ^{Supporting Information} (Figures 6A-C and Figures S19-S21), ~~we observed~~ ^{it was} that the fragmentation pattern (diterpene backbone fragment ions and acyls neutral losses) agreed with the structure proposed. Nonetheless, the structure proposed by C-NAP could correspond to an isomer, differing by the position of ~~the~~ oxidation on the diterpene backbone, or the location of the acyl substituents.

These results showed that, because of their scarcity in the databases due to their exceptional diversity, the structural annotation of *Euphorbia* diterpene esters by tandem mass spectrometry with computational methods using regular structure databases is limited and prone to errors. Herein, it was shown that the generation of a *utilization* CSDB and its employment with C-NAP can rapidly annotate diterpene esters, including compounds that have not been described yet. Nonetheless, because C-NAP is a supervised method, it can be applied only in cases where a minimum prior knowledge on the sample composition (isolated compounds, chemotaxonomy, e.g. lvs and frs) is available. The annotation of new compound annotations in MS/MS spectra can also be achieved effectively by using a set of biotransformations to predict theoretical candidates.⁵⁴ Methods such as *iMet*,⁵⁵ *Hybrid Similarity Search*,⁵⁶ have been described for the annotation of new molecules in MS/MS spectral data. However, these computational methods do not have their code source available which prevents their general use. This is a major problem. For that reason, open and collaborative tools that can propose structural annotation of putative new derivatives from inputted biotransformations, such as the recent MetWork web-server⁵⁷ are needed. Recently published

Figure 6. Diterpene esters molecular network from *E. pithyusa* and *E. cupanii* samples, annotated with Combinatorial Network Annotation Propagation (C-NAP), using a Combinatorial Structure Database (CSDB) of premyrsinane and myrsinane esters. The annotated spectra for the first rank structure (TOP1) proposed by C-NAP are presented for three nodes (Box 6A–6C and Figures S19–S21, in Supporting Information). The mass spectra of these nodes can be consulted here: [6A](#), [6B](#), and [6C](#).

Figure 7. The first rank (TOP1) annotation obtained with different scores (*MetFrag*, *Fusion*, *Consensus*, and *Hybrid Fusion Consensus*) with both NAP and C-NAP.

The analysis of *E. pithyusa* and *E. cupanii* extracts with untargeted LC-MS/MS allowed us to establish that the extract of *E. pithyusa* was dominated by triterpenoids.

while the *E. cupanii* was characterized by a large diversity of diterpene esters. The analysis of their diterpene esters profiles showed considerable differences with only

7% of the diterpene esters shared between the two species. The use of unsupervised structure annotation with *MS2LDA* enriched the diterpene ester molecular networks by highlighting some of the fragmentation pattern dictating the spectral clustering. Like observed in a previous study,¹⁷ our results demonstrated that this integration has a strong potential for the annotation of *Euphorbia* diterpene's fragmentation spectra, that are characterized by the presence numerous isomers, including positional isomers and macrocyclic functionalization, which often led to similar fragment ions.⁵³ In the present study, *MS2LDA* highlighted the presence of putatively new premyrsinane- and myrsinane derivatives in *E. cupanii* extract, which prompted us to carry out a guided isolation procedure that led to the discovery of twenty diterpene esters belonging to the myrsinane (11a-h, 12, 13), premyrsinane (14a-c and 15a-c), and phorbol groups (16a-c). The phorbol esters isolated were found to be selective inhibitors of the CHIKV replication.

The retrospective annotation of the molecular networks using the reference MS/MS spectra and computational methods showed that the diterpenoid profile of *E. pithyusa* was characterized by a diversity of 14-oxopremyrsinol, while *E. cupanii's* analogues.

the occurrence of *E. cupanii* profile was dominated by myrsinol and 13,17-oxy-premyrsinol derivatives. Moreover, the profile was studied, characterized, and wide range of *E. pithyusa* extracts was marked by the presence of a large diversity of triterpenoids.

In conclusion, our results showed that *E. pithyusa* and *E. cupanii* collected in Corsica are producing different specialized metabolites, which indicates that the speciation process of these two taxa was mediated via a change in their biosynthetic chemical

arsenal. The main chemical difference in the specialized metabolite profiles of *E.*

pithyusa and *E. cupanii* (triterpenoids vs. diterpenoids, respectively) could have been

driven by their biotopic distribution (coastal vs. inland, respectively) in relation to the

natural enemies present (insects or microbes or other wild animals), but also in

relation to livestock pressure in an island that has been shaped by agro-pastoral

practices for centuries. Overall, our results are supporting the taxonomical

classification of *E. cupanii* at the species rank (rather than the subspecies rank) as

already proposed from the morphological and botanical analysis of the two taxa,²⁰

and further supported by their different number of chromosomes.²⁵ Nonetheless, *E.*

cupanii collected in Corsica showed a different chemical composition compared to the

specimens collected in Sardinia and studied by Appendino and colleagues,¹⁹ and that

very
unusual!
I suggest
YW
omit
this

characterized by the presence of lathyrane and 14-oxopremyrsinane derivatives.

Thus, this observation suggests the existence of multiple populations across the two

islands and maybe within Sardinia, where *E. cupanii* is present on multiple sites.

From a methodological standpoint, we ~~introduced~~ ^{was introduced} C-NAP and showed that the use

of CSDB can rapidly propose *Euphorbia* diterpene ester structure in tandem mass

spectrometry data, including new compounds, which ~~is impossible~~ ^{has not proven possible} with usual

computational approaches depending on structure databases. C-NAP has the

potential to annotate other classes of metabolites constituted by a limited number of

building blocks and can be readily used through the NAP web app available on the

GNPS web-platform. We anticipate that the development of hybrid methods using

both CSDB and chemical transformations ^{using} are key to further accelerate ^{ing} the discovery

of new molecules from mass spectral data.

EXPERIMENTAL SECTION

General Experimental Procedures. See ^S supporting ^I information for details on the general experimental procedure (S22).

Plant Material. The aerial parts of *Euphorbia cupanii* (Guss. ex Bertol.) were collected in October 2014 on the road D71 in Morosaglia (Corsica, France) by M. N.-E. and L.-F. N. following a maintenance of the road by the regional services. The voucher specimen (ME-3-151-A) was deposited at the Herbarium of the University of Corsica (Laboratoire de Chimie des Produits Naturels, Corte) and identified by M. N.-E., L.-F. N., and Marie-José Battesti.

Extraction and Isolation. The same protocol used for *E. pithyusa* was applied to *E. cupanii* as previously described.¹⁴ In brief, the diterpenoids enriched extract (9.3 g) was concentrated under vacuum and mixed with Celite (10 g), then separated by flash chromatography on a prepacked silica column (GraceResolv, 120 g) to obtain 14 fractions (F1–F14). The gradient used was heptane–CHCl₃ (40:60 to 0:100 in 40 minutes), then CHCl₃–MeOH (100:0 to 80:20 in 60 min). Based on the diterpene esters molecular network interpretation, F10 to F12 were selected for the isolation of targeted diterpenoid compounds. See Supporting Information for purification details of for the compounds **11a–h**, **12**, **13**, **14a–c**, **15a–c**, **16a–c**, and **17** (S23).

(3R,4S,5R,7R,11R,12R,13S,14R,15R)-3,15-O-diacetyl-7,14-O-dibenzoyl-5-

Cmp 5

hydroxymyrsinol (**11a**): amorphous powder; $[\alpha]_{\text{D}}^{25} -150.0$ (c 1, EtOH); UV (EtOH) λ_{max}

(log ϵ) 273 (3.74), 230 (4.54), 205 (4.32) nm; for ^1H and ^{13}C NMR spectroscopic data,

(Supporting Information)

see Tables S7 and S9; HRESIMS m/z 681.2698 $[\text{M} + \text{Na}]^+$ (calcd for $\text{C}_{38}\text{H}_{42}\text{O}_{10}\text{Na}$,

semicolumn $\Rightarrow$ ③

681.2676), MS/MS spectrum is deposited in the GNPS spectral library:

CCMSLIB00000840355.

(3R,4S,5R,7R,11R,12R,13S,14R,15R)-15-O-acetyl-7,14-O-dibenzoyl-5-hydroxy-3-

Cmp 5

O-propanoylmyrsinol (**11b**): amorphous powder; $[\alpha]_{\text{D}}^{25} -138.0$ (c 1, EtOH); UV (EtOH)

λ_{max} (log ϵ) 273 (3.47), 230 (4.48), 204 (4.25) nm; For ^1H and ^{13}C NMR spectroscopic

(Supporting Information)

data, see Tables S7 and S9; HRESIMS m/z 695.2836 $[\text{M} + \text{Na}]^+$ (calcd for

semicolumn $\Rightarrow$ ③

$\text{C}_{39}\text{H}_{44}\text{O}_{10}\text{Na}$, 695.2832), MS/MS spectrum is deposited in the GNPS spectral library:

CCMSLIB00000840359.

(3R,4S,5R,7R,11R,12R,13S,14R,15R)-3,7-O-diacetyl-5,14-O-dibenzoylmyrsinol

Cmp 5

(**11c**): amorphous powder; $[\alpha]_{\text{D}}^{25} -6.0$ (c 1, EtOH); UV (EtOH) λ_{max} (log ϵ) 274 (4.49),

231 (4.59), 204 (4.34) nm; for ^1H and ^{13}C NMR spectroscopic data, see Tables S7 and

(Supporting Information)

SEMICON

S9; HRESIMS m/z 681.2679 $[M + Na]^+$ (calcd for $C_{38}H_{42}O_{10}Na$, 681.2676) MS/MS

spectrum is deposited in the GNPS spectral library: [CCMSLIB00000840358](#).

(3R,4S,5R,7R,11R,12R,13S,14R,15R)-5,15-O-diacetyl-14-O-benzoyl-7-O-nicotinyl-

3-O-propanoylmyrsinol (11h): crystals; $[\alpha]^{25}_D$ -42.0 (c 1, EtOH); UV (EtOH) λ_{max} (log ϵ)

263 (3.70), 226 (4.37), 204 (4.27) nm; for 1H and ^{13}C NMR spectroscopic data, see

(Supporting Information)

Tables S7 and S9; HRESIMS m/z 738.2886 $[M + Na]^+$ (calcd for $C_{40}H_{45}NO_{11}Na$,

738.2890) MS/MS spectrum is deposited in the GNPS spectral library:

[CCMSLIB00000840362](#).

(3R,4S,5R,7R,11R,12R,13S,14R,15R)-3,5,7,10,15-O-pentacetyl-14-O-benzoyl-

10,18-dihydromyrsinol (12): crystals; $[\alpha]^{25}_D$ -408.0 (c 1, EtOH); UV (EtOH) λ_{max} (log ϵ)

274 (3.18), 231 (4.21), 203 (3.85) nm; for 1H and ^{13}C NMR spectroscopic data, see

(Supporting Information)

Tables S7 and S9; HRESIMS m/z 721.2843 $[M + Na]^+$ (calcd for $C_{37}H_{46}O_{13}Na$,

721.2836) MS/MS spectrum is deposited in the GNPS spectral library:

[CCMSLIB00000840363](#).

(2R,3R,4S,5R,7R,11R,12R,13S,15R)-3,5,7,10,15-O-pentacetyl-2-O-benzoyl-14-

oxo-10,18-dihydromyrsinol (13): amorphous powder; $[\alpha]^{25}_D$ -202.0 (c 1, EtOH); UV

(EtOH) $\lambda_{\max}$ (log ϵ) 230 (4.30), 203 (4.14) nm; for ^1H and ^{13}C NMR spectroscopic data,

(Supporting Information)

see Tables S7 and S9; HRESIMS m/z 735.2649 $[\text{M} + \text{H}]^+$ (calcd for $\text{C}_{39}\text{H}_{43}\text{O}_{14}$,

structure $\rightarrow$ ⑤

735.2653) MS/MS spectrum is deposited in the GNPS spectral library:

CCMSLIB00000840364.

(3R,4S,5R,7R,9R,11S,12R,13S,15R)-3,7-O-diacetyl-5-O-benzoyl-13,17-oxy-14-

CNPS

oxopremyrinsol (**14a**): amorphous powder; $[\alpha]_{\text{D}}^{25} -50.5$ (c 1, EtOH); UV (EtOH) $\lambda_{\max}$

(log ϵ) 230 (4.29), 206 (3.95); for ^1H and ^{13}C NMR spectroscopic data, see Tables S8

(Supporting Information)

and S9; HRESIMS m/z 577.2434 $[\text{M} + \text{Na}]^+$ (calcd for $\text{C}_{31}\text{H}_{38}\text{O}_9\text{Na}$, 577.2414) MS/MS

structure $\rightarrow$ ⑤

spectrum is deposited in the GNPS spectral library: CCMSLIB00000840349.

(3R,4S,5R,7R,9R,11S,12R,13S,15R)-3,7,15-O-triacetyl-5-O-benzoyl-13,17-oxy-14-

oxopremyrinsol (**14b**): amorphous powder; $[\alpha]_{\text{D}}^{25} -11.2$ (c 1, EtOH); IR $\nu_{\max}$ 1721, 1370,

CNPS

1266, 1227, 1097, 1070, 1024, 835, 710 cm^{-1} ; UV (EtOH) $\lambda_{\max}$ (log ϵ) 231 (3.60) nm;

(Supporting Information)

for ^1H and ^{13}C NMR spectroscopic data, see Tables S8 and S9; HRESIMS m/z

619.2515 $[\text{M} + \text{Na}]^+$ (calcd for $\text{C}_{33}\text{H}_{40}\text{O}_{10}\text{Na}$, 619.2519) MS/MS spectrum is deposited

structure $\rightarrow$ ⑤

in the GNPS spectral library: CCMSLIB00000840350.

(3*R*,4*S*,5*R*,7*R*,9*R*,11*S*,12*R*,13*S*,15*R*)-7,15-*O*-diacetyl-5-*O*-benzoyl-3-propanoyl-

13,17-oxypremyrinsinol (**14c**): amorphous powder; $[\alpha]_D^{25}$ -6.0 (*c* 1, EtOH); UV

(EtOH) $\lambda_{\max}$ (log ϵ) 231 (3.33) nm; for ^1H and ^{13}C NMR spectroscopic data, see Tables

(Supporting Information)

S8 and S9; HRESIMS m/z 633.2697 $[\text{M} + \text{Na}]^+$ (calcd for $\text{C}_{34}\text{H}_{42}\text{O}_{10}\text{Na}$, 633.2676)

MS/MS spectrum is deposited in the GNPS spectral library: [CCMSLIB00000840351](#).

(3*R*,4*S*,5*R*,7*R*,9*R*,11*S*,12*R*,13*S*,14*R*,15*R*)-3,7-*O*-diacetyl-5,14-*O*-dibenzoyl-13,17-

oxypremyrinsinol (**15a**): amorphous powder; $[\alpha]_D^{25}$ -6.5 (*c* 1, EtOH); UV (EtOH) $\lambda_{\max}$ (log

ϵ) 231 (4.35), 202 (4.27) nm; for ^1H and ^{13}C NMR spectroscopic data, see Tables S8

(Supporting Information)

and S9; HRESIMS m/z 683.2819 $[\text{M} + \text{Na}]^+$ (calcd for $\text{C}_{38}\text{H}_{44}\text{O}_{10}\text{Na}$, 683.2832) MS/MS

spectrum is deposited in the GNPS spectral library: [CCMSLIB00000840352](#).

(3*R*,4*S*,5*R*,7*R*,9*R*,11*S*,12*R*,13*S*,14*R*,15*R*)-3,14-*O*-diacetyl-5-7-*O*-dibenzoyl-13,17-

oxypremyrinsinol (**15b**): crystals; $[\alpha]_D^{25}$ -4.0 (*c* 1, EtOH); UV (EtOH) $\lambda_{\max}$ (log ϵ) 230

(4.54), 205 (4.27) nm; for ^1H and ^{13}C NMR spectroscopic data, see Tables S8 and S9

HRESIMS m/z 683.2843 $[\text{M} + \text{Na}]^+$ (calcd for $\text{C}_{38}\text{H}_{44}\text{O}_{10}\text{Na}$, 683.2832) MS/MS

spectrum is deposited in the GNPS spectral library: [CCMSLIB00000840353](#).

(3*R*,4*S*,5*R*,7*R*,9*R*,11*S*,12*R*,13*S*,14*R*,15*R*)-7-*O*-^Aacetyl-5,14-*O*-dibenzoyl-3-propanoyl-13,17-oxypremyrsinol (**15c**): ^Aamorphous powder; $[\alpha]_D^{25}$ -24.3 (*c* 1, EtOH); IR $\nu_{\max}$ 3325, 2973, 2884, 1379, 1327, 1273, 1087, 1046, 879 cm^{-1} ; UV (EtOH) $\lambda_{\max}$ (log ϵ) 274 (3.26), 231 (4.28), 204 (3.79) nm; for ^1H and ^{13}C NMR spectroscopic data, see (Supporting Information) Tables S8 and S9; HRESIMS m/z 697.2983 $[\text{M} + \text{Na}]^+$ (calcd for $\text{C}_{39}\text{H}_{46}\text{O}_{10}\text{Na}$, 697.2989), MS/MS spectrum is deposited in the GNPS spectral library: CCMSLIB000008403524.

12 β -*O*-[deca-2*E*,4*Z*-dienoyl]-13 α -isobutyl-4 β -phorbol (**16a**): ^Aamorphous powder; $[\alpha]_D^{25}$ +0.6 (*c* 1, EtOH); UV (EtOH) $\lambda_{\max}$ (log ϵ) 264 (4.57), 205 (4.33) nm; For ^1H and ^{13}C NMR spectroscopic data, see Table S11; HRESIMS m/z 607.3246 $[\text{M} + \text{Na}]^+$ (calcd for $\text{C}_{34}\text{H}_{48}\text{O}_8\text{Na}$, 607.3247), MS/MS spectrum is deposited in the GNPS spectral library: CCMSLIB00000840365.

12 β -*O*-[deca-2*E*,4*Z*-dienoyl]-13 α -(2-methylbutyl)-4 β -phorbol (**16b**): ^Aamorphous powder; $[\alpha]_D^{25}$ +2.0 (*c* 1, EtOH); UV (EtOH) $\lambda_{\max}$ (log ϵ) 264 (4.23), 206 (3.97) nm; for ^1H and ^{13}C NMR spectroscopic data, see Table S11; HRESIMS m/z 621.3422 $[\text{M} +$

Na]⁺ (calcd for C₃₅H₅₀O₈Na, 621.3403) MS/MS spectrum is deposited in the GNPS

spectral library: CCMSLIB00000840366.

12β-O-[deca-2Z,4E-dienoyl]-13α-isobutyl-4β-phorbol (**16c**): amorphous powder;

[α]_D²⁵ +5.2 (c 1, EtOH); UV (EtOH) λ_{max} (log ε) 262 (3.93), 206 (3.75) nm; For ¹H and

¹³C NMR spectroscopic data, see Table S11; HRESIMS m/z 607.3264 [M + Na]⁺

(calcd for C₃₄H₄₈O₈Na, 607.3247) MS/MS spectrum is deposited in the GNPS spectral

library: CCMSLIB00000840367.

12β-O-[deca-2Z,4E-dienoyl]-13α-isobutyl-5-ene-7-oxo-4β-phorbol (**17**): amorphous

powder; [α]_D²⁵ +16.0 (c 1, EtOH); UV (EtOH) λ_{max} (log ε) 261 (4.34) nm; for ¹H and

¹³C NMR spectroscopic data, see Table S11; HRESIMS m/z 621.3040 [M + Na]⁺

(calcd for C₃₄H₄₆O₉Na, 621.3040) MS/MS spectrum is deposited in the GNPS

spectral library: CCMSLIB00000840368.

X-ray structure determination. Colorless crystals suitable for single-crystal X-ray

analysis were obtained after incubation of concentrated Et₂O solutions saturated with

hexane vapors at low temperature for the five compounds (**11d**, **11f**, **11h**, **12**, and

15b) analyzed herein. All the data were collected on a Rigaku diffractometer

1
2
3 constituted by a MM007 HF rotating-anode generator, equipped with Osmic CMF
4
5
6
7 confocal optics, and a Rapid II curved Image Plate, operating at the Cu-K α radiation
8
9
10 ($\lambda = 1.54187 \text{ \AA}$). Samples were irradiated at room temperature. The ORTEP views of
11
12
13 the compounds **11d**, **11f**, **11h**, **12**, and **15b** are reported in Figure 4. See Table S10
14
15
16 for the X-ray crystallography data collection parameters and structure refinement
17
18
19 statistics and S24 ^(Supporting Information) for a detailed discussion on X-ray structure determination, ^(Supporting Information) in the
20
21
22 supporting information. Crystallographic data have been deposited with the
23
24
25
26
27
28 Cambridge Crystallographic Data Centre (deposit No. CCDC 1518590–1518594).
29
30
31 Copies of the data can be obtained, free of charge, on application to the Director,
32
33
34
35 CCDC, 12 Union Road, Cambridge CB2 1EZ, UK (fax: +44-(0)223-336033 or e-mail:
36
37
38 deposit@ccdc.cam.ac.uk).
39
40

41
42 **LC-MS/MS Analysis.** The LC-ESI-MS/HRMS analysis was performed as
43
44
45 previously describe ^d for *E. pithyusa*.¹⁴ The mass spectrometry data were deposited on
46
47
48 the public MassIVE repository under the accession number MSV000082850, and the
49
50
51
52 MS/MS spectrum of the previous isolated compounds **11d–g** were deposited in the
53
54
55
56
57
58
59
60

GNPS spectral library: CCMSLIB00000840356, CCMSLIB00000840360,
CCMSLIB00000840357, and CCMSLIB00000840361, respectively.

Virus Cell-Based Anti-Chikungunya Assay. The protocol performed on Vero
(African green monkey kidney) cells with the Chikungunya virus strain 899 was
described in a previous publication.¹⁴

Feature-Based Molecular Networking. Acquisition files were converted to mzML
format using the MSConvert part of the ProteoWizard suite.⁵⁸ Then, these files were
processed with MZmine2²⁸ and the feature-based molecular networking workflow²⁶
available on the GNPS platform²⁹ (Global Natural Products Social Molecular
Networking: <http://gnps.ucsd.edu>) was employed [<https://ccms-ucsd.github.io/GNPSDocumentation/featurebasedmolecularnetworking/>]. The
molecular networks were generated from the .MGF file exported in MZmine2 with the
“GNPS export module”. The detailed MZmine2 parameters are described in the

Supporting Information (S25). Cytoscape 3.6.1³⁰ software was used for the
visualization. The ^{reference} molecular networking ^{information} job can be accessed ^{as follows:} here:

<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=a6ff2e43eb014c7ab53630805f0e>

32a3

^S
^A
Unsupervised substructure annotation. The fragmentation spectra were mined for co-occurring fragments and neutral losses using the latent Dirichlet allocation

algorithm³¹ on the *MS2LDA* web app.⁵⁹ The following parameters were used for the

task: job: isolation_window (0.5), min_ms2_int (90), n_its (1000), K (300), ms2_bin (0.05

Da), doc_m2m_prob_threshold (0.05). The result of the *MS2LDA* annotation is publicly

available at the following link: (<http://ms2lda.org/basicviz/summary/669/>) The table "All

Fragmentation Spectra and Mass2Motifs matching details" was converted with an in-

house *python* script into a format that can be readily mapped Mass2Motifs

annotations onto the molecular networks generated using the feature-based

molecular networking workflow ([https://ccms-](https://ccms-ucsd.github.io/GNPSDocumentation/featurebasedmolecularnetworking/)

[ucsd.github.io/GNPSDocumentation/featurebasedmolecularnetworking/](https://ccms-ucsd.github.io/GNPSDocumentation/featurebasedmolecularnetworking/)). The script

and its documentation are available on GitHub

(https://github.com/lfnathias/MS2LDA_to_GNPS_molecular_networks). The inspection

and the annotation of the Mass2Motifs was realized on the *MS2LDA.org* web-app.

CAPD

USL
SAME
FONT
STYLE

Combinatorial Database Generation for Network Annotation Propagation. The

combinatorial database of premyrsinane and myrsinane esters was generated with

Smilib v2.0,^{49,50} available as a Java executable from [\[http://melolab.org/smilib/\]](http://melolab.org/smilib/) and

enabling the combinatorial generation of

enables to combinatorially generate structures from inputted scaffolds, building

blocks. The input and output structures were encoded as SMILES.⁶⁰ The list of

structures obtained were formatted in a NAP compatible custom database using a

dedicated web-app <http://dorresteinappshub.ucsd.edu:5002/upload>. Note that

alternatively, the NAP database can generate locally using a python script, available

as a *Jupyter notebook* [\[http://jupyter.org/\]](http://jupyter.org/). The script and the files used to generate

the CSDB, as well as step-by-step tutorial are available on GitHub

[\[https://github.com/lfnathias/Combinatorial_Network_Annotation_Propagation\]](https://github.com/lfnathias/Combinatorial_Network_Annotation_Propagation).

Network Annotation Propagation. The molecular networks were annotated with the

Network Annotation Propagation (NAP)⁵⁸ workflow available on GNPS

(<https://gnps.ucsd.edu/ProteoSAFe/static/gnps-theoretical.jsp>). The result of NAP was

mapped into the molecular networks by importing the NAP output file into the molecular

networking job in Cytoscape. The chemical structures were visualized in Cytoscape with the

ChemViz2 plug-in [\[http://www.cgl.ucsf.edu/cytoscape/chemViz/\]](http://www.cgl.ucsf.edu/cytoscape/chemViz/). The NAP parameters were

set as follows: 25 first candidates, exact mass searches within 30 ppm, [M + H], [M + Na], [M

+ NH₄] as adduct types, cosine of 0.5 to sub-select inside a cluster, and 10 maximum of

USE SAME
FUNK
JITLW

1
2
3 candidate structures in the graph. The NAP ^{ins} jobs can be accessed here with the following
4
5 databases: *Dictionary of Natural Products* (DNP) and *SuperNatural II*,⁶¹
6
7
8 ([https://proteomics2.ucsd.edu/ProteoSAFe/status.jsp?task=047e9a665c904479b4c9](https://proteomics2.ucsd.edu/ProteoSAFe/status.jsp?task=047e9a665c904479b4c9a33e84daacc8)
9
10 [a33e84daacc8](https://proteomics2.ucsd.edu/ProteoSAFe/status.jsp?task=047e9a665c904479b4c9a33e84daacc8)); DNP, SN, and the premyrsinane and myrsinane combinatorial
11
12
13
14
15 database (CSDB)
16
17
18 ([https://proteomics2.ucsd.edu/ProteoSAFe/status.jsp?task=9df8a887c66649ab8524d](https://proteomics2.ucsd.edu/ProteoSAFe/status.jsp?task=9df8a887c66649ab8524dc93367e8370)
19
20 [c93367e8370](https://proteomics2.ucsd.edu/ProteoSAFe/status.jsp?task=9df8a887c66649ab8524dc93367e8370))
21
22
23
24
25
26
27
28

29 ASSOCIATED CONTENT

30 Supporting Information.

31
32
33
34
35
36 ¹H NMR spectra for the compounds **11d–g**; ¹H, ¹³C, HMBC, ROESY, and HRESIMS
37
38
39
40 data for the compounds **11a–h**, **12**, **13**, **14a–c**, **15a–c**, **16a–c**, and **17**; MS/MS
41
42
43
44 spectra of the compounds **14c**, **15b**, and **13**; The detailed discussion on X-ray
45
46
47 structure determination for **11d**, **11f**, **11h**, **12**, and **15b**.
48
49

50 This material is available free of charge via the ACS Publications website at
51
52
53
54 <http://pubs.acs.org>.
55
56

57 AUTHOR INFORMATION

Corresponding Author

* Tel: + 1 858-534-6607. Fax: + 1 858-822-0041. E-mail: pdorrestein@ucsd.edu.

* Tel: + 33 1 69 82 30 85. Fax: + 33 1 69 07 72 47. E-mail: marc.litaudon@cnrs.fr.

ORCID

Louis-Félix Nothias: 0000-0001-6711-6719

Fanny Roussi: 0000-0002-5941-9901

Pascal Retailleau: 0000-0003-3995-519X

Julien Paolini: 0000-0002-3109-1430

Pieter C. Dorrestein: 0000-0002-3003-1030

David Touboul: 0000-0003-2751-774X

Marc Litaudon: 0000-0002-0877-8234

Melissa Nothias-Esposito: 0000-0002-7221-6033

REMOVE UNDERLINE

Notes

All other authors declare no competing financial interest.

ACKNOWLEDGMENTS

MOEO LOFONO

not in mic
↓

ME and LFN were funded by a "Contrat Doctoral" awarded by the University of

Corsica and would like to thank the Prof. Jean Costa for his continuous support. This

not in mic
↓

work was supported by an "Investissement d'Avenir" grant managed by Agence

Nationale de la Recherche (CEBA, ANR-10-LABX-25-01). We thank the NIH for

U.S. National Institutes

of

Health

supporting this work under NIH-UCSD Center for Computational Mass Spectrometry

(P41 GM103484) and the NIH grant on reuse of metabolomics data (R03 CA211211)

The authors would like to thank Justin van der Hooft (Wageningen Univ.) and Simon

University of

University

Rogers (Univ. Glasgow) for their help with MS2LDA. The authors also would like to

acknowledge the help of Marie-José Battesti for with the plant identification, and also

of Jean-Francois Gallard (CNRS) for support with the NMR spectra acquisition. The

authors are grateful to C. Collard, N. Verstraeten, and C. Vanderheydt from the

Laboratory for Virology and Experimental Chemotherapy at the Rega Institute for

Medical Research in Leuven, for the evaluation of antiviral activity.

REFERENCES

- (1) Halle, F.; Oldeman, R. A. A.; Tomlinson, P. B. *Tropical Trees and Forests: An Architectural Analysis*; Springer Science & Business Media, 2012.
- (2) Govaerts, R.; Frodin, D. G.; Radcliffe-Smith, A.; Carter, S. *World Checklist and Bibliography of Euphorbiaceae (with Pandaceae)*; Royal Botanic Gardens, Kew, 2000. *UK*
- (3) Horn, J. W.; Xi, Z.; Riina, R.; Peirson, J. A.; Yang, Y.; Dorsey, B. L.; Berry, P. E.; Davis, C. C.; Wurdack, K. J. Evolutionary Bursts in Euphorbia (Euphorbiaceae) Are Linked with Photosynthetic Pathway. *Evolution* **2014**, *68*, 3485–3504.
- (4) Salmaki, Y.; Zarre, S.; Esser, H.-J.; Heubl, G. ~~Seed and Gland Morphology in Euphorbia (Euphorbiaceae) with Focus on Their Systematic and Phylogenetic Importance, a Case Study in Iranian Highlands~~. *Flora - Morphol. Distrib. Funct. Ecol. Plants* **2011**, *206*, 957–973.
- (5) Bruyns, P. V.; Mapaya, R. J.; Hedderson, T. A. ~~New Subgeneric Classification for Euphorbia (Euphorbiaceae) in Southern Africa Based on ITS and PsbA-TrnH Sequence Data~~. *Taxon* **2006**, *55*, 397–420.
- (6) Park, K.-R.; Jansen, R. K. ~~A Phylogeny of Euphorbiaceae Subtribe Euphorbiinae (Euphorbiaceae) Based on Molecular Data~~. *J. Plant Biol.* **2007**, *50*, 644–649.
- (7) Zimmermann, N. F. A.; Ritz, C. M.; Hellwig, F. H. ~~Further Support for the Phylogenetic Relationships within Euphorbia L. (Euphorbiaceae) from nrITS and TrnL-TrnF IGS Sequence Data~~. *Plant Syst. Evol.* **2010**, *286*, 39–58.
- (8) Steinmann, V. W.; Porter, J. M. ~~Phylogenetic Relationships in Euphorbiaceae (Euphorbiaceae) Based on ITS and NdhF Sequence Data~~. *Ann. Mo. Bot. Gard.* **2002**, *453–490*. *ADD VOLUME NUMBERS*
- (9) Steinmann, V. W. ~~La Inclusión de Pedilanthus En Euphorbia (Euphorbiaceae)~~. *Acta Bot. Mex.* **2003**, 45–50.
- (10) Steinmann, V. W.; Van Ee, B.; Berry, P. E.; Gutiérrez, J. ~~The Systematic Position of Cubanthus and Other Shrubby Endemic Species of Euphorbia (Euphorbiaceae) in Cuba~~. In *Anales del Jardín Botánico de Madrid*; Real Jardín Botánico, 2007; Vol. 64, pp 123–133.

Madrid

Berlin and Heidelberg

- (11) Horn, J. W.; van Ee, B. W.; Morawetz, J. J.; Riina, R.; Steinmann, V. W.; Berry, P. E.; Wurdack, K. J. ~~Phylogenetics and the Evolution of Major Structural Characters in the Giant Genus Euphorbia L. (Euphorbiaceae). *Mol. Phylogenet. Evol.* **2012**, *63*, 305–326.~~
- (12) López-Maury, L.; Marguerat, S.; Bähler, J. ~~Tuning Gene Expression to Changing Environments: From Rapid Responses to Evolutionary Adaptation. *Nat. Rev. Genet.* **2008**, *9*, 583–593.~~
- (13) Sedio, B. E. ~~Recent Breakthroughs in Metabolomics Promise to Reveal the Cryptic Chemical Traits That Mediate Plant Community Composition, Character Evolution and Lineage Diversification. *New Phytol.* **2017**, *214*, 952–958.~~
- (14) Esposito, M.; Nothias, L.-F.; Retailleau, P.; Costa, J.; Roussi, F.; Neyts, J.; Leyssen, P.; Touboul, D.; Litaudon, M.; Paolini, J. ~~Isolation of Premyrsinane, Myrsinane, and Tiglane Diterpenoids from Euphorbia Pithyusa Using a Chikungunya Virus Cell-Based Assay and Analogue Annotation by Molecular Networking. *J. Nat. Prod.* **2017**, *80*, 2051–2059.~~
- (15) Nothias, L.-F.; Boutet-Mercey, S.; Cachet, X.; De La Torre, E.; Laboureur, L.; Gallard, J.-F.; Retailleau, P.; Brunelle, A.; Dorrestein, P. C.; Costa, J.; et al. ~~Environmentally Friendly Procedure Based on Supercritical Fluid Chromatography and Tandem Mass Spectrometry Molecular Networking for the Discovery of Potent Antiviral Compounds from Euphorbia Semiperfoliata. *J. Nat. Prod.* **2017**, *80*, 2620–2629.~~
- (16) Olivon, F.; Allard, P.-M.; Koval, A.; Righi, D.; Genta-Jouve, G.; Neyts, J.; Apel, C.; Pannecouque, C.; Nothias, L.-F.; Cachet, X. et al. ~~Bioactive Natural Products Prioritization Using Massive Multi-Informational Molecular Networks. *ACS Chem. Biol.* **2017**, *12*, 2644–2651.~~
- (17) Ernst, M.; Nothias-Scaglia, L.-F.; Hooft, J. van der; Silva, R. R.; Saslis-Lagoudakis, C. H.; Grace, O. M.; Martinez-Swatson, K.; Hassemer, G.; Funez, L.; Simonsen, H. T.; et al. ~~Did a Plant-Herbivore Arms Race Drive Chemical Diversity in Euphorbia? *bioRxiv* **2018**, 323014.~~
- (18) Jeanmonod, D.; Gamisans, J. *Flora Corsica*, Edisud 921; Aix-en-Provence, 2007; Vol. 1.

ADD ALL
CO-MUTATION

MUTATION
ENTRY:
HAS THREE
BAND A
FULL PAPER
CITING
SUBSEQUENT
TO THE
PREPRINT
TITLE WITH
BE
CITING
FROM

STING
PRELIMINARY NAME
↓

- (19) Appendino, G.; Belloro, E.; Tron, G. C.; Jakupovic, J.; Ballero, M. ~~Diterpenoids from Euphorbia pithyusa Subsp. cupanii~~. *J. Nat. Prod.* **1999**, 62, 1399–1404.
- (20) Valsecchi, F. ~~Le Piante Endemiche Della Sardegna: 80-83~~. *Boll. Della Soc. Sarda Sci. Nat.* **1980**, 19, 323–342.
- (21) Euphorbia pithyusa subsp. cupanii (Guss. ex Bertol.) Radcl.-Sm. — The Plant List. } Give website address
- (22) ITIS Standard Report Page: Euphorbia pithyusa ssp. cupanii
- (23) Viviani, D. *Florae Corsicae Specierum Novarum, Vel Minus Cognitarum*. *Diagnosis: Quam in Florae Italicae Fragmenti Alterius Prodrum* [Gennae]. 1824.
- (24) eFlore – Tela Botanica. ← ADD WEBSITE ADDRESS
- (25) Cardona, M. A. Contribucion Al Estudio Citotaxonomico de La Flora de Baleares. *Colloq. Soc. Catal. Biol.* **1978**, 10–11. ADD VOLUME NUMBER 10/11
- (26) Nothias, L.-F.; Nothias-Esposito, M.; da Silva, R.; Wang, M.; Protsyuk, I.; Zhang, Z.; Sarvepalli, A.; Leyssen, P.; Touboul, D.; Costa, J.; (et al.) Bioactivity-Based Molecular Networking for the Discovery of Drug Leads in Natural Product Bioassay-Guided Fractionation. *J. Nat. Prod.* **2018**, 81, 758–767. ADD NAME OF ALL CO-AUTHORS
- (27) Feature Based Molecular Networking - GNPS Documentation. - WEBSITE ADDRESS?
- (28) Pluskal, T.; Castillo, S.; Villar-Briones, A.; Orešič, M. MZmine 2: Modular Framework for Processing, Visualizing, and Analyzing Mass Spectrometry-Based Molecular Profile Data. *BMC Bioinformatics* **2010**, 11, 395. SPACE
- (29) Wang, M.; Carver, J. J.; Phelan, V. V.; Sanchez, L. M.; Garg, N.; Peng, Y.; Nguyen, D. D.; Watrous, J.; Kaponov, C. A.; Luzzatto-Knaan, T.; (et al.) Sharing and Community Curation of Mass Spectrometry Data with Global Natural Products Social Molecular Networking. *Nat. Biotechnol.* **2016**, 34, 828–837.
- (30) Smoot, M. E.; Ono, K.; Ruscheinski, J.; Wang, P.-L.; Ideker, T. Cytoscape 2.8: New Features for Data Integration and Network Visualization. *Bioinformatics* **2011**, 27, 431–432.
- (31) Hooft, J. J. J. van der; Wandy, J.; Barrett, M. P.; Burgess, K. E. V.; Rogers, S. Topic Modeling for Untargeted Substructure Exploration in Metabolomics. *Proc. Natl. Acad. Sci.* **2016**, 113, 13738–13743.

USA

- (32) Nothias-Scaglia, L.-F.; Pannecouque, C.; Renucci, F.; Delang, L.; Neyts, J.; Roussi, F.; Costa, J.; Leyssen, P.; Litaudon, M.; Paolini, J. ~~Antiviral Activity of Diterpene Esters on Chikungunya Virus and HIV Replication.~~ *J. Nat. Prod.* **2015**, *78*, 1277–1283.
- (33) Remy, S.; Olivon, F.; Desrat, S.; Blanchard, F.; Eparvier, V.; Leyssen, P.; Neyts, J.; Roussi, F.; Touboul, D.; Litaudon, M. ~~Structurally Diverse Diterpenoids from *Sandwithia Guyanensis*.~~ *J. Nat. Prod.* **2018**, *81*, 901–912.
- (34) Öksüz, S.; Gürek, F.; Gil, R. R.; Pengsuparp, T.; Pezzuto, J. M.; Cordell, G. A. ~~Four Diterpene Esters from *Euphorbia Myrsinites*.~~ *Phytochemistry* **1995**, *38*, 1457–1462.
- (35) Öksüz, S.; Gürek, F.; Qiu, S.-X.; Cordell, G. A. ~~Diterpene Polyesters from *Euphorbia Seguieriana*.~~ *J. Nat. Prod.* **1998**, *61*, 1198–1201.
- (36) Xu, J.; Guo, Y.; Xie, C.; Li, Y.; Gao, J.; Zhang, T.; Hou, W.; Fang, L.; Gui, L. ~~Bioactive Myrsinot Diterpenoids from the Roots of *Euphorbia Prolifera*.~~ *J. Nat. Prod.* **2011**, *74*, 2224–2230. *SPKLC*
- (37) Jeske, F.; Jakupovic, J.; Berendsohn, W. ~~Diterpenes from *Euphorbia Seguieriana*.~~ *Phytochemistry* **1995**, *40*, 1743–1750.
- (38) Bourjot, M.; Delang, L.; Nguyen, V. H.; Neyts, J.; Guéritte, F.; Leyssen, P.; Litaudon, M. ~~Prostratin and 12-O-Tetradecanoylphorbol 13-Acetate Are Potent and Selective Inhibitors of Chikungunya Virus Replication.~~ *J. Nat. Prod.* **2012**, *75*, 2183–2187.
- (39) Schmidt, R.; Hecker, E. ~~Autoxidation of Phorbol Esters under Normal Storage Conditions.~~ *Cancer Res.* **1975**, *35*, 1375–1377.
- (40) Sumner, L. W.; Amberg, A.; Barrett, D.; Beale, M. H.; Beger, R.; Daykin, C. A.; Fan, T. W.-M.; Fiehn, O.; Goodacre, R.; Griffin, J. L. ~~Proposed Minimum Reporting Standards for Chemical Analysis.~~ *Metabolomics* **2007**, *3*, 211–221.
- (41) Wolf, S.; Schmidt, S.; Müller-Hannemann, M.; Neumann, S. ~~In Silico Fragmentation for Computer Assisted Identification of Metabolite Mass Spectra.~~ *BMC Bioinformatics* **2010**, *11*, 148.
- (42) Ridder, L.; Hooft, J. J. J. van der; Verhoeven, S.; Vos, R. C. H. de; Schaik, R. van; Vervoort, J. ~~Substructure-Based Annotation of High-Resolution Multistage MSn Spectral Trees.~~ *Rapid Commun. Mass Spectrom.* **2012**, *26*, 2461–2471.

- (43) Böcker, S.; Dührkop, K. ~~Fragmentation Trees Reloaded~~. *J. Cheminformatics* **2016**, *8*, 1–26.
- (44) Mohimani, H.; Gurevich, A.; Shlemov, A.; Mikheenko, A.; Korobeynikov, A.; Cao, L.; Shcherbin, E.; Nothias, L.-F.; Dorrestein, P. C.; Pevzner, P. A. ~~Dereplication of Microbial Metabolites through Database Search of Mass Spectra~~. *Nat. Commun.* **2018**, *9*, 4035. ← RMY CLOSING PAGE NUMBER ?
- (45) Silva, R. R. da; Wang, M.; Nothias, L.-F.; Hooft, J. J. J. van der; Caraballo-Rodríguez, A. M.; Fox, E.; Balunas, M. J.; Klassen, J. L.; Lopes, N. P.; Dorrestein, P. C. ~~Propagating Annotations of Molecular Networks Using in Silico Fragmentation~~. *PLOS Comput. Biol.* **2018**, *14*, e1006089.
- (46) Gerlich, M.; Neumann, S. ~~MetFusion: Integration of Compound Identification Strategies~~. *MetFusion*. *J. Mass Spectrom.* **2013**, *48*, 291–298.
- (47) Shi, Q.-W.; Su, X.-H.; Kiyota, H. ~~Chemical and Pharmacological Research of the Plants in Genus Euphorbia~~. *Chem Rev* **2008**, *108*, 4295–4327.
- (48) Vasas, A.; Rédei, D.; Csupor, D.; Molnár, J.; Hohmann, J. ~~Diterpenes from European Euphorbia Species Serving as Prototypes for Natural Product-Based Drug Discovery~~. *Eur. J. Org. Chem.* **2012**, *2012*, 5115–5130.
- (49) Schüller, A.; Hähnke, V.; Schneider, G. ~~SMILib v2.0: A Java-Based Tool for Rapid Combinatorial Library Enumeration~~. *QSAR Comb. Sci.* **2007**, *26*, 407–410.
- (50) Schüller, A.; Schneider, G.; Byvatov, E. ~~SMILIB: Rapid Assembly of Combinatorial Libraries in SMILES Notation~~. *QSAR Comb. Sci.* **2003**, *22*, 719–721.
- (51) Kind, T.; Liu, K.-H.; Yip Lee, D.; DeFelice, B.; Meissen, J. K.; Fiehn, O. ~~LipidBlast – in-Silico Tandem Mass Spectrometry Database for Lipid Identification~~. *Nat. Methods* **2013**, *10*, 755–758.
- (52) Qiu, F.; Fine, D. D.; Wherritt, D. J.; Lei, Z.; Sumner, L. W. ~~PlantMAT: A Metabolomics Tool for Predicting the Specialized Metabolic Potential of a System and for Large-Scale Metabolite Identifications~~. *Anal. Chem.* **2016**, *88*, 11373–11383.
- (53) Nothias-Scaglia, L.-F.; Schmitz-Afonso, I.; Renucci, F.; Roussi, F.; Touboul, D.; Costa, J.; Litaudon, M.; Paolini, J. ~~Insights on Profiling of Phorbol,~~

- Deoxyphorbol, Ingenol and Jatrophane Diterpene Esters by High Performance Liquid Chromatography Coupled to Multiple Stage Mass Spectrometry. *J. Chromatogr. A* **2015**, 1422, 128–139. *SMILE*
- (54) Ridder, L.; van der Hooft, J. J. J.; Verhoeven, S.; de Vos, R. C. H.; Vervoort, J.; Bino, R. J. *In Silico Prediction and Automatic LC-MSn Annotation of Green Tea Metabolites in Urine*. *Anal. Chem.* **2014**, 86, 4767–4774.
- (55) iMet: A Network-Based Computational Tool to Assist in the Annotation of Metabolites from Tandem Mass Spectra - Analytical Chemistry (ACS Publications). *GIVE A PROPER CITATION*
- (56) Moorthy, A. S.; Wallace, W. E.; Kearsley, A. J.; Tchekhovskoi, D. V.; Stein, S. E. *Combining Fragment-Ion and Neutral-Loss Matching during Mass Spectral Library Searching: A New General Purpose Algorithm Applicable to Illicit Drug Identification*. *Anal. Chem.* **2017**, 89, 13261–13268.
- (57) Beauxis, Y.; Genta-Jouve, G. *Network: A Web Server for Natural Products Anticipation*. *Bioinformatics*. *COMPLETE THE*
- (58) Adusumilli, R.; Mallick, P. *Data Conversion with ProteoWizard MsConvert*. In *Proteomics: Methods and Protocols*; Comai, L., Katz, J. E., Mallick, P., Eds.; Methods in Molecular Biology; Springer New York: New York, NY, 2017; pp 339–368. *in msc* *vol. 93*
- (59) Wandy, J.; Zhu, Y.; van der Hooft, J. J. J.; Daly, R.; Barrett, M. P.; Rogers, S. *Me2lda.Org: Web-Based Topic Modelling for Substructure Discovery in Mass Spectrometry*. *Bioinformatics* **2018**, 34, 317–318.
- (60) SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules - Journal of Chemical Information and Modeling (ACS Publications). *GIVE A PROPER CITATION*
- (61) Banerjee, P.; Erehman, J.; Gohlke, B.-O.; Wilhelm, T.; Preissner, R.; Dunkel, M. *Super Natural II--a Database of Natural Products*. *Nucleic Acids Res.* **2015**, 43, D935-939.

GRAPHICAL ABSTRACT