



HAL
open science

Isolation of Picrotoxanes from *Austrobuxus carunculatus* Using Taxonomy-Based Molecular Networking

Florent Olivon, Pascal Retailleau, Sandy Desrat, David Touboul, Fanny Roussi, Cecile Apel, Marc Litaudon

► **To cite this version:**

Florent Olivon, Pascal Retailleau, Sandy Desrat, David Touboul, Fanny Roussi, et al.. Isolation of Picrotoxanes from *Austrobuxus carunculatus* Using Taxonomy-Based Molecular Networking. *Journal of Natural Products*, 2020, 83 (10), pp.3069-3079. 10.1021/acs.jnatprod.0c00636 . hal-02992365

HAL Id: hal-02992365

<https://hal.science/hal-02992365>

Submitted on 6 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Isolation of Picrotoxanes from *Austrobuxus*
carunculatus Using Taxonomy-based Molecular
Networking

Florent Olivon, Pascal Retailleau, Sandy Desrat, David Touboul, Fanny Roussi,

Cécile Apel,* and Marc Litaudon*

Université Paris-Saclay, CNRS, Institut de Chimie des Substances Naturelles, UPR 2301,

91198, Gif-sur-Yvette, France

ABSTRACT: A unique collection of 292 extracts from 107 New Caledonian Euphorbiaceae species *sensu lato* was profiled by LC-MS² and the metabolite content organized by molecular networking. Based on the assumption that taxon-specific molecules are more likely to be structurally novel, taxonomic data were mapped on spectral networks to detect genus-specific clusters. Using this approach, a group of compounds unique to the genus *Austrobuxus* was highlighted. The subsequent MS-guided purification of the fruit EtOAc extract of *Austrobuxus carunculatus* led to the isolation of 13 new monolactone and “norditerpene” picrotoxanes (**2-14**), along with the known tutin (**1**). The structures of the new compounds were elucidated by HRESIMS and NMR spectroscopic data analysis, and the absolute configurations of compounds **1**, **3**, **7**, **11**, **12** and **14** were determined by single-crystal X-ray diffraction analysis. The relative and absolute configurations of compounds **4** and **5** were ascertained by chemical transformation of compound **3**. The absolute configurations of other members of the series have been proposed on the basis of biogenetic considerations and specific rotation values of similar sign and magnitude. Compounds **1-14** were evaluated for their antiproliferative activities against HCT116 colon, U87-MG glioblastoma and A549 lung human cancer cell lines. Compounds bearing an acyl chain at C-2 (i.e. **2**, **4** and **13**) showed IC₅₀ values in the micromolar range for the three cell lines used.

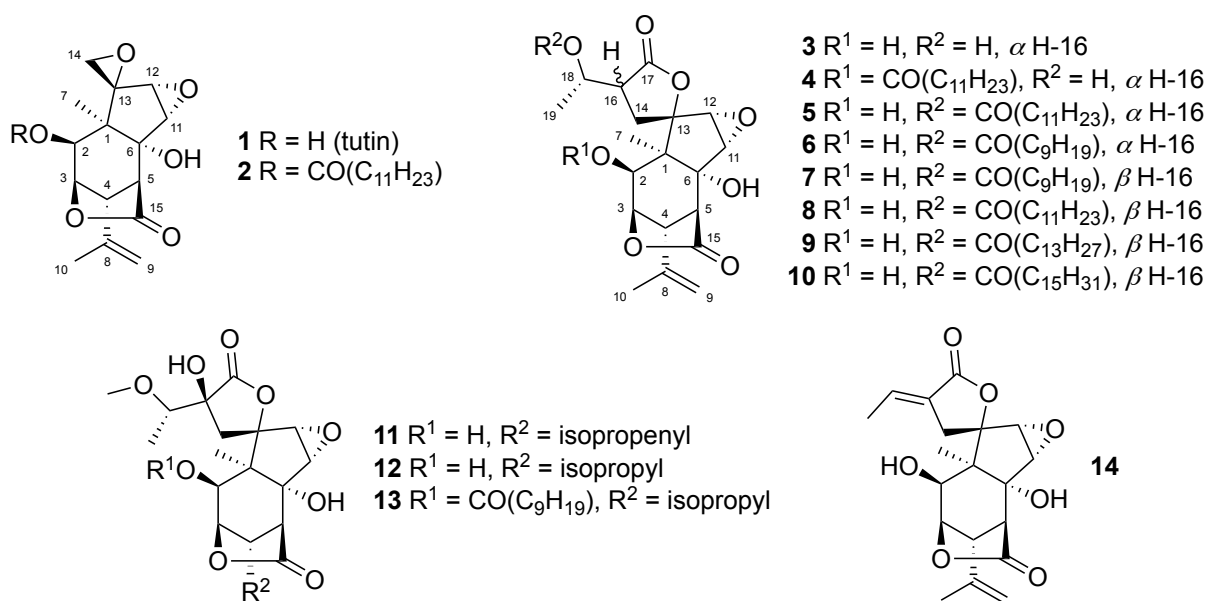
It is now widely accepted that the occurrence of specialized metabolites (SM) within the plant kingdom and their exceptional chemical diversity is the result of evolutionary processes.^{1,3} Millions of years of environmental screening have endowed natural products with structures and functions shaped and optimized to overcome abiotic stress and fight natural enemies. Chemosystematics aims at studying the distribution of metabolites within the plant kingdom and has always relied on botanical classification. However, it has been shown that within the same botanical family, certain taxa biosynthesize SM that are or become specific to their producer. Based on the postulate that different groups of SM may have an identical ecological role, their occurrence within a clade is sometimes mutually exclusive, in order to maintain an advantageous cost/benefit ratio.⁴ As a result, a single and structurally homogeneous group of SM may dominate within a given taxon. This unequal distribution between tribes or genera has even been used to build the classification of certain plant families.⁵ Based on this observation, a methodology identifying genus- or species-specific metabolites has been developed, with the ultimate goal of discovering novel structures or scaffolds. The present prioritization strategy consists of adding taxonomic data to a spectral network resulting from a tandem mass spectrometry data organizational approach.^{6,9} The rapid detection of taxon-specific clusters and subsequent dereplication have helped guide further work on the most promising extracts.

The work described herein is in continuation of the chemical exploration of an extract collection of 292 New Caledonian Euphorbiaceae species *sensu lato* initiated in the above-mentioned studies.^{6,8} The currently accepted classification of Euphorbiaceae *sensu lato* by the Angiosperm Phylogeny group recognizes five different families: Euphorbiaceae Juss. *sensu stricto*, Pandaceae Engl. & Gilg, Phyllanthaceae Martynov, Picrodendraceae Small and Putranjivaceae Endl..¹⁰

From this collection, the approach used enabled the detection of ion clusters specific to the genus *Austrobuxus* Miq.. The corresponding compounds were particularly abundant in the fruit extract of the species *Austrobuxus carunculatus* (Baill.) Airy Shaw. *A. carunculatus*, for which no chemical investigation has been reported so far, belongs to the family Picrodendraceae and is endemic to New Caledonia. The MS-guided purification of its fruit EtOAc extract led to the isolation of 13 new picrotoxanes: 1 monolactone (**2**) and 12 structurally related “norditerpenes” (**3-14**), along with the known tutin (**1**).

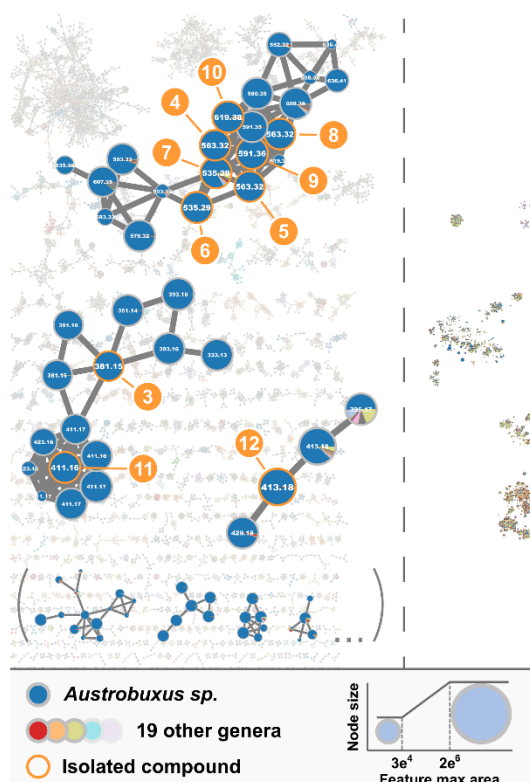
Picrotoxanes are found mainly as tetra- or pentacyclic complex structures containing up to 12 stereogenic centers.¹¹ Most members of this chemical class are neurotoxins acting as GABA-receptor antagonists,¹²⁻¹⁹ and are used as tools in neurobiological research. So far, the “norditerpene” structural variant of the picrotoxanes is restricted to the family Picrodendraceae.¹¹ With reference to compounds isolated from *Austrobuxus swanii*,²⁰ the presently isolated compounds were named austrobuxusins E-M (**2**, **4-6**, **9**, **10**, **12-14**) and 16-*epi*-austrobuxusins B, G and H (**11**, **8** and **7**, respectively). Compound **3** was named 16-*epi*-pretoxin. The present report describes their isolation, structure elucidation and cytotoxic activity evaluation using the HCT116, A549 and U87-MG cell lines.

RESULTS AND DISCUSSION



A collection of 292 EtOAc extracts was prepared and profiled by UHPLC-HRMS² as previously described.⁶⁸ The set comprised EtOAc extracts of different plant parts prepared from 107 New Caledonian species (93% are endemic) coming from 20 genera, of which 40% were Euphorbiaceae *sensu stricto*, 46% Phyllanthaceae, 13% Picrodendraceae and 1% Putranjivaceae. The Pandaceae family is not represented in New Caledonia and is therefore not part of the sample series used.

A. Molecular Networks



B. t-SNE projection

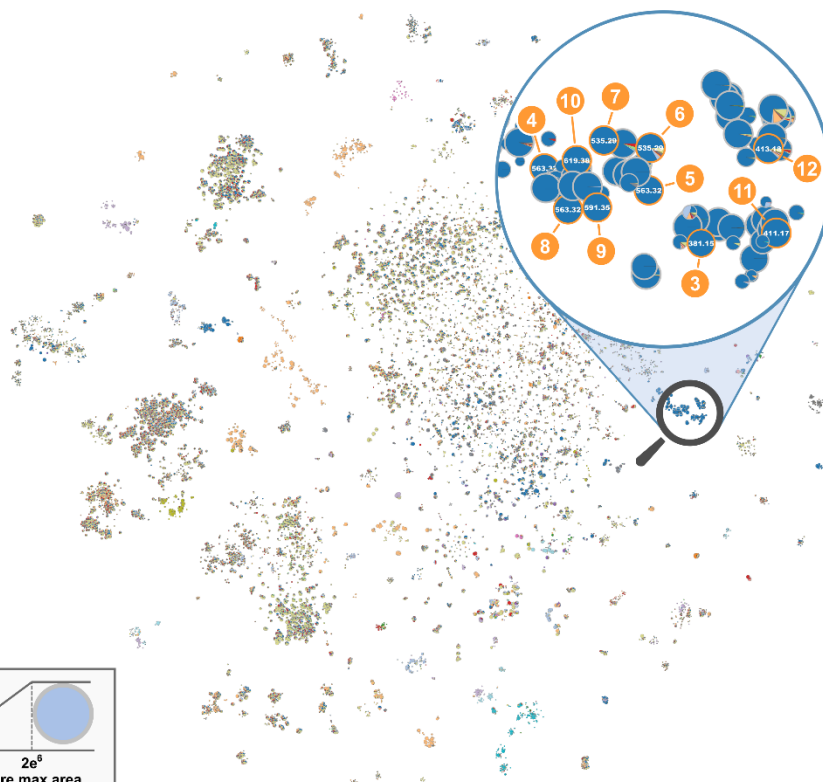


Figure 1. Multi-informative molecular maps generated with MetGem. Networks were created with the preprocessed .mgf file exported from the Euphorbiaceae dataset analysis in MZmine 2. The 20 genera represented in the collection were mapped with a color tag. Genus-specific molecular families were detected by spotting single-colored clusters. (A) Standard molecular networks view, $[M+H]^+$ clusters specific to *Austrobuxus* species (with the whole network in background); (B) t-SNE projection of the Euphorbiaceae dataset, and a zoom in on the area bringing together the $[M+H]^+$ nodes specific to *Austrobuxus* species. *Austrobuxus* clusters depicted in Figure 1A are separated and scattered in the classical representation of molecular networks while they are gathered on the t-SNE visualization. The node size is proportional to the MS¹ peak area value obtained from the sample in which the feature's abundance is the highest.

Spectral data were processed via MZmine 2²¹ before being organized as molecular networks using MetGem.²² A multi-informative molecular map was created by overlaying color-coded taxonomic data on the initial networks. For this purpose, the extracts were grouped according to the genus they belong. The color mapping consisted of assigning a specific color to each group, so as to easily differentiate multicolored ion clusters (for which the corresponding ubiquitous compounds come from multiple taxa), from single-colored clusters, which are therefore genus or species-specific. Using this approach, several clusters bringing together protonated ions specific to the genus *Austrobuxus* were highlighted (Figure 1). These clusters were separated and scattered in the classical representation of molecular networks (Figure 1A), while they were well gathered on the t-SNE visualization along with other isolated nodes (Figure 1B). The latter preserves all relations between groups of spectra and allows the gathering of sub-structurally related compounds.

When analyzing a large sample set, t-SNE parameters can be modulated for easier browsing within the graph (i.e. to avoid overloading and obtain a clearer delimitation of the different clusters). Nodes can be filtered to keep only those that interact with a representative number of other nodes: in the representation Figure 1B, nodes were retained when sharing at least 6 cosine scores above 0.5. Spectral similarity information is maintained since the filter applies to nodes, not cosine scores.

None of the compounds within the specific clusters could be annotated using the GNPS libraries.²³ However, database searching from predicted molecular formulas, together with taxonomic considerations, suggested the presence of pretoxin or its isomers.²⁴ Despite a moderate ranking (22nd candidate considering the CSI:FingerID bio database), this information was corroborated by using Sirius.^{25,26}

The new compounds of interest were detected as present mainly in the fruit EtOAc extract of *Austrobuxus carunculatus*. Its MS-guided chemical investigation led to the isolation of 14

microtoxanes for which the structures were elucidated by HRESIMS and NMR spectroscopic data analysis (Tables 1-4, Table S1 and S3, Supporting Information). Their absolute configurations were determined by single-crystal X-ray diffraction (S2, Supporting Information), semi-synthesis and on the basis of biogenetic considerations. Compounds **1**, **2**, **13** and **14** were not detected within the clusters of interest but co-purified during the targeted isolation of compounds **3-12**.

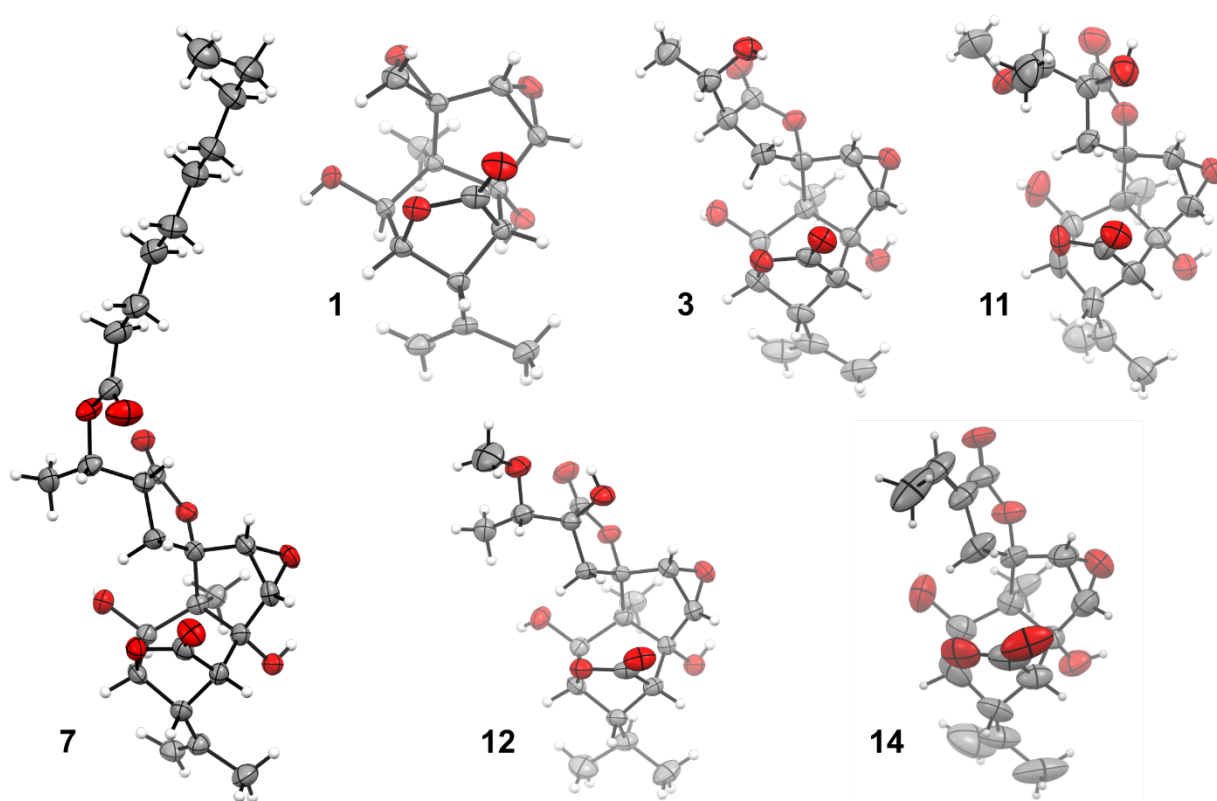


Figure 2. ORTEP views of compounds **1**, **3**, **7**, **11**, **12** and **14**. (Ellipsoids drawn at 50% (30 % for **14**) of probability and hydrogens are shown with spheres of arbitrary size).

Compound **1** was identified as the known tutin²⁷ microtoxane on the basis of the data obtained. Its HRESIMS data showed a $[M-H_2O+H]^+$ ion at m/z 277.1107 (calcd for $C_{15}H_{17}O_5^+$, 277.1081) corresponding to the molecular formula $C_{15}H_{18}O_6$. The combination of 1D- and 2D-

NMR data, as well as X-ray diffraction providing some evidence regarding the absolute configuration, allowed the confirmation of its characteristic five-fused rings and their substituents (See Table S1, Supporting Information for NMR data and Figures 2 and S2.2 Supporting Information for X-ray ORTEP view^{28, 29} and overlay diagram for the original KEGBAZ³⁰ and the present model of **1**).

Compound **2** gave the molecular formula C₂₇H₄₀O₇, as indicated by HRESIMS of the [M+H]⁺ ion peak at *m/z* 477.2828 (calcd for C₂₇H₄₁O₇⁺, 477.2847). Its NMR spectra displayed similar proton and carbon resonances and cross correlations with those of tutin indicating the presence of the same α -isopropenyl group at C-4 on the cyclohexane ring, the ester bridge between C-3 and C-5 forming the γ -lactone ring, the α -methyl group at C-1 as well as the α -hydroxy group at C-6. Compound **2** also contained a *cis*-fused α -epoxy function located in the C-11/C-12 positions and a spiro-epoxy moiety at C-13 on the cyclopentane ring. The major structural difference between compound **2** and tutin was the presence of a C₁₂-saturated-carbon ester side chain instead of a hydroxy group at C-2. This was supported by the presence of an ester carbonyl resonance at δ 173.8, which correlated with the H-2 methine proton at δ_{H} 5.33 (s, H-2) and two methylene resonances at δ_{H} 2.33 (m, H₂-2') and 1.61 (m, H₂-3') in the HMBC spectrum. The small deshielding effect observed on H-2 was consistent with the substitution of tutin C-2 hydroxy group by an ester moiety. The relative configuration of C-2 was deduced from the ROESY correlations observed between H-2 and H-3, H₃-7 and the olefinic proton H-9b of the isopropenyl group, indicating the β -orientation of the long-chain ester moiety. The absolute configuration of **2** is assumed to be the same as that of tutin on the basis of biogenetic considerations and their specific rotation values of similar sign and magnitude.

Table 1. ¹H and ¹³C NMR Data for Compounds 2-5 in CDCl₃ (3 in CD₃OD)

position	2		3		4		5	
	δ _c	δ _a (J in Hz)	δ _c	δ _a (J in Hz)	δ _c	δ _a (J in Hz)	δ _c	δ _a (J in Hz)
1	44.8	-	55.5	-	52.9	-	54.1	-
2	74.3	5.33 s	74.2	4.23 brs	75.9	5.45 brs	73.5	4.28 brs
OH-2	-	-	-	-	-	-	-	-
3	80.5	4.81 brd (4.8)	84.8	4.80 m	79.4	4.87 brd (4.8)	82.7	4.82 brd (4.8)
4	50.0	3.20 m	51.1	3.26 m	50.2	3.21 m	49.8	3.18 m
5	49.0	3.13 brd (4.4)	50.9	3.10 d (4.3)	48.5	3.13 d (4.4)	48.7	3.10 d (4.1)
6	78.0	-	78.0	-	77.3	-	77.0	-
OH-6	-	-	-	-	-	2.78 brs	-	2.54 brs
7	20.1	1.26 s	24.6	1.20 s	23.6	1.16 s	23.9	1.24 s
8	140.1	-	143.7	-	140.1	-	141.0	-
9-a	113.1	5.01 brs	111.2	4.71 brs	112.8	4.97 brs	111.9	4.92 brs
9-b	-	5.14 brs	-	4.88 brs	-	5.13 brs	-	4.69 brs
10	23.0	1.91 s	23.4	1.93 s	23.0	1.88 s	23.1	1.88 s
11	60.6	3.74 d (3.0)	63.0	3.69 m	61.8	3.73 d (2.7)	61.9	3.74 d (2.7)
12	60.1	3.16 d (3.0)	64.3	3.52 brd (2.7)	63.7	3.63 d (2.7)	64.0	3.34 d (2.7)
13	65.0	-	92.7	-	90.5	-	90.7	-
14-α	52.5	3.74 d (4.9)	33.0	3.68 dd (13.9, 7.6)	31.7	3.51 dd (14.6, 11.4)	31.3	3.61 dd (14.3, 11.2)
14-β	-	2.82 d (4.9)	-	1.96 dd (13.9, 11.1)	-	1.95 dd (14.6, 6.6)	-	1.81 dd (14.3, 7.9)
15	174.4	-	176.9	-	174.0	-	174.4	-
16	-	-	49.3	2.95 ddd (11.1, 7.6, 3.5)	47.5	2.70 ddd (11.4, 6.6, 5.8)	45.4	3.16 ddd (11.2, 7.9, 3.9)
OH-16	-	-	-	-	-	-	-	-
17	-	-	179.6	-	177.0	-	176.4	-
18	-	-	69.1	3.96 qd (6.6, 3.5)	69.2	3.94 qd (6.3, 5.8)	70.3	5.05 qd (6.5, 3.9)
OMe-18	-	-	-	-	-	-	-	-
19	-	-	20.7	1.31 d (6.6)	21.4	1.34 d (6.3)	17.1	1.36 d (6.5)
1'	173.8	-	-	-	173.5	-	173.5	-
2'	34.4	2.33 m	-	-	34.3	2.30 m	34.7	2.26 t (7.5)
3'	24.9	1.61 m	-	-	25.0	1.58 m	25.1	1.57 m
4'-7'	29.1-29.9	1.17-1.34 m	-	-	29.1-29.7	1.15-1.35 m	29.1-29.9	1.15-1.30 m
8'	29.1-29.9	1.17-1.34 m	-	-	29.1-29.7	1.15-1.35 m	29.1-29.9	1.15-1.30 m
9'	29.1-29.9	1.17-1.34 m	-	-	29.1-29.7	1.15-1.35 m	29.1-29.9	1.15-1.30 m
10'	32.1	1.17-1.34 m	-	-	32.0	1.15-1.35 m	32.1	1.15-1.30 m
11'	22.9	1.17-1.34 m	-	-	22.8	1.15-1.35 m	22.9	1.15-1.30 m
12'	14.3	0.85 t (6.9)	-	-	14.3	0.83 t (6.9)	14.3	0.85 t (6.7)

Table 2. ¹H and ¹³C NMR Data for Compounds 6-9 in CDCl₃

position	6		7		8		9	
	δ _c	δ _n (J in Hz)	δ _c	δ _n (J in Hz)	δ _c	δ _n (J in Hz)	δ _c	δ _n (J in Hz)
1	54.2	-	52	-	52	-	52	-
2	73.7	4.29 brs	74.2	4.34 brs	74.2	4.34 brd (4.4)	73.7	4.30 brs
OH-2	-	-	-	-	-	-	-	-
3	82.6	4.82 brd (4.7)	82.5	4.84 brd (4.9)	82.5	4.84 brd (4.8)	82.8	4.80 brd (4.1)
4	49.8	3.19 m	49.8	3.19 m	49.8	3.19 m	49.7	3.15 m
5	48.7	3.11 d (4.4)	48.9	3.10 d (4.2)	48.9	3.10 d (4.3)	48.9	3.07 d (3.6)
6	77	-	76.9	-	76.9	-	76.8	-
OH-6	-	2.48 brs	-	2.49 brs	-	2.73 brs	-	2.76 s
7	23.9	1.23 s	23.6	1.34 s	23.6	1.34 s	23.5	1.30 s
8	141	-	141	-	141	-	141	-
9-a	111.9	4.93 brs	112.2	4.94 brs	112.2	4.94 brs	112	4.90 brs
9-b	-	4.69 brs	-	4.69 brs	-	4.69 brs	-	4.68 brs
10	23.1	1.88 s	23	1.88 s	23	1.88 s	23	1.85 s
11	61.9	3.75 d (2.7)	61.9	3.73 d (2.8)	61.9	3.74 d (2.8)	61.8	3.70 brs
12	64.1	3.35 d (2.7)	61.8	3.44 d (2.8)	61.8	3.44 d (2.8)	61.6	3.43 brs
13	90.6	-	90.5	-	90.5	-	90.7	-
14-α	-	3.61 dd (14.3, 11.2)	30.3	3.62 dd (13.9, 11.5)	30.2	3.62 dd (13.9, 11.5)	30.4	3.61 dd (13.5, 11.6)
14-β	31.3	1.82 dd (14.3, 7.9)	-	2.32 dd (13.9, 9.6)	-	2.32 dd (13.9, 9.6)	-	2.29 dd (13.5, 9.6)
15	174.2	-	174.5	-	174.5	-	174.9	-
16	45.4	3.15 ddd (11.2, 7.9, 3.9)	45.9	2.87 ddd (11.5, 9.6, 5.5)	45.9	2.87 ddd (11.5, 9.6, 5.5)	45.8	2.87 ddd (11.6, 9.6, 5.3)
OH-16	-	-	-	-	-	-	-	-
17	176.2	-	174.7	-	174.7	-	174.9	-
18	-	5.06 qd (6.5, 3.9)	68.8	5.20 qd (6.4, 5.5)	68.8	5.20 qd (6.4, 5.5)	69	5.13 qd (6.4, 5.3)
OMe-18	70.3	-	-	-	-	-	-	-
19	17.1	1.37 d (6.5)	18.4	1.36 d (6.4)	18.4	1.36 d (6.4)	18.3	1.32 d (6.4)
1'	173.5	-	173.4	-	173.4	-	173.6	-
2'	34.7	2.26 m	34.7	2.25 t (7.6)	34.7	2.25 t (7.6)	34.6	2.22 t (7.2)
3'	25.1	1.58 m	25.1	1.58 m	25.1	1.58 m	25.1	1.55 m
4'-7'	29.3-29.7	1.20-1.32 m	29.3-29.7	1.15-1.35 m	29.3-29.9	1.20-1.35 m	29.3-29.9	1.15-1.35 m
8'	32.1	1.20-1.32 m	32.1	1.15-1.35 m	29.3-29.9	1.20-1.35 m	29.3-29.9	1.15-1.35 m
9'	22.9	1.20-1.32 m	22.9	1.15-1.35 m	29.3-29.9	1.20-1.35 m	29.3-29.9	1.15-1.35 m
10'	14.3	0.85 t (6.9)	14.3	0.85 t (6.9)	32.1	1.20-1.35 m	29.3-29.9	1.15-1.35 m
11'	-	-	-	-	22.8	1.20-1.35 m	29.3-29.9	1.15-1.35 m
12'	-	-	-	-	14.3	0.85 t (6.9)	32.1	1.15-1.35 m
13'	-	-	-	-	-	-	22.8	1.15-1.35 m
14'	-	-	-	-	-	-	14.2	0.83 t (6.9)

Table 3. ¹H and ¹³C NMR Data for Compounds 10-13 in CDCl₃ (12 in CD₃OD)

position	10		11		12		13	
	δ _c	δ _n (J in Hz)	δ _c	δ _n (J in Hz)	δ _c	δ _n (J in Hz)	δ _c	δ _n (J in Hz)
1	52	-	52.1	-	53.3	-	51.3	-
2	74.1	4.33 brs	74.4	4.31 brs	75.4	3.77 brs	75.4	5.00 brs
OH-2	-	-	-	3.03 brs	-	-	-	-
3	82.6	4.84 brd (4.9)	82.6	4.86 brd (4.9)	85.4	4.58 brs	80.5	4.61 brd (4.2)
4	49.8	3.18 m	49.8	3.18 m	54.7	2.17 m ^r	53.7	2.18 m ^r
5	48.9	3.10 d (4.3)	48.9	3.10 d (4.2)	50.4	2.84 brs	48	2.93 d (3.3)
6	76.9	-	77.1	-	78.4	-	77.8	-
OH-6	-	2.52 s	-	2.57 s	-	-	-	-
7	23.6	1.34 s	23.6	1.37 s	24.1	1.40 s	23.1	1.34 s
8	141	-	141	-	28.1	2.17 m ^r	27.3	2.18 m ^r
9-a	112.1	4.94 brs	112.1	4.94 brs	22.4	1.06 d (4.0)	22.1	1.04 m ^r
9-b	-	4.69 brs	-	4.68 brs	-	-	-	-
10	23	1.88 s	23.0	1.88 s	20.6	0.92 d (4.0)	20.5	1.04 m ^r
11	61.9	3.73 d (2.7)	61.8	3.72 d (2.8)	63.1	3.64 brs	62.1	3.72 d (2.8)
12	61.7	3.44 d (2.7)	62.6	3.95 d (2.8)	62.9	3.93 brs	62.5	4.20 d (2.8)
13	90.5	-	91.1	-	93.7	-	91.4	-
14-a	30.3	3.62 dd (13.9, 11.5)	36.3	3.65 d (15.1)	37.4	3.94 d (14.3)	36	3.39 d (14.4)
14-b	-	2.32 dd (13.9, 9.5)	-	2.16 d (15.1)	-	2.00 d (14.3)	-	2.18 m ^r
15	174.7	-	174.7	-	177.8	-	174.9	-
16	45.9	2.87 ddd (11.5, 9.5, 5.7)	79.1	-	80.4	-	79.6	-
OH-16	-	-	-	3.33 brs	-	-	-	-
17	174.6	-	175.9	-	178.6	-	175.5	-
18	68.8	5.19 qd (6.3, 5.7)	78.2	3.50 q (6.3)	78.9	3.53 q (6.3)	77.9	3.53 q (6.3)
OMe-18	-	-	57.8	3.36 s	57.8	3.31 s	57.4	3.33 s
19	18.4	1.35 d (6.3)	13.1	1.23 d (6.3)	13.6	1.19 d (6.3)	13.3	1.25 d (6.3)
1'	173.3	-	-	-	-	-	174.1	-
2'	34.7	2.24 t (7.6)	-	-	-	-	34	2.34 t (7.6)
3'	25.1	1.58 m	-	-	-	-	25	1.57 m
4'-7'	29.3-29.9	1.15-1.35 m	-	-	-	-	29.3-29.9	1.15-1.35 m
8'	29.3-29.9	1.15-1.35 m	-	-	-	-	32.1	1.15-1.35 m
9'	29.3-29.9	1.15-1.35 m	-	-	-	-	22.9	1.15-1.35 m
10'	29.3-29.9	1.15-1.35 m	-	-	-	-	14.4	0.85 t (6.9)
11'	29.3-29.9	1.15-1.35 m	-	-	-	-	-	-
12'	29.3-29.9	1.15-1.35 m	-	-	-	-	-	-
13'	29.3-29.9	1.15-1.35 m	-	-	-	-	-	-
14'	32.1	1.15-1.35 m	-	-	-	-	-	-
15'	22.9	1.15-1.35 m	-	-	-	-	-	-
16'	14.3	0.85 t (6.9)	-	-	-	-	-	-

Overlapped.

Table 4. ^1H and ^{13}C NMR Data for Compound **14** in CDCl_3

position	14	
	δ_c	δ_n (J in Hz)
1	52.8	-
2	74.1	4.33 brd (4.0)
OH-2	-	3.38 brd (5.1)
3	82.6	4.86 brd (4.8)
4	49.8	3.21 m
5	48.9	3.12 d (4.3)
6	77.1	-
OH-6	-	2.52 s
7	23.7	1.35 s
8	141	-
9-a	112.1	4.70 brs
9-b	-	4.93 brs
10	23.1	1.89 s
11	61.9	3.74 d (2.8)
12	63.2	3.34 d (2.8)
13	90.2	-
14-a	32.6	4.18 brd (17.8)
14-b	-	2.67 brd (17.8)
15	174.8	-
16	128.2	-
OH-16	-	-
17	170.3	-
18	136	6.69 qdd (7.1, 3.3, 2.9)
OMe-18	-	-
19	16	1.83 brd (7.1)

Overlapped.

Analysis of its data allowed compound **3** to be identified as an epimer of the known pretoxin.²⁴ In contrast to monolactone picrotoxanes such as compounds **1** and **2**, the so-called “norditerpene” picrotoxanes possess an extended carbon skeleton. The main difference between compounds **1** and **3** was the presence of a substituted spiro γ -butyrolactone at C-13 in **3** instead of the spiro-epoxy moiety in **1**. The structure and absolute configuration of **3**, which was recrystallized from MeOH, was confirmed unambiguously by X-ray diffraction data analysis. (See Table 1 for the NMR data and Figures 2 and S2.2, Supporting Information for X-ray ORTEP representation of **3**). Compound **3** is 16-*epi*-pretoxin. In the literature, the absence of crystallographic data and complete NMR data for pretoxin did not allow their comparison with those of compound **3**. Only the melting points of the two compounds could be compared and proved to be different, confirming their structural difference (Mp = 128-130°C²⁴ and 208-210°C for pretoxin and 16-*epi*-pretoxin, respectively).

The HRESIMS data of compound **4** showed a $[M+H]^+$ ion peak at m/z 563.3176 (calcd for $C_{31}H_{47}O_9^+$, 563.3215), corresponding to the molecular formula, $C_{31}H_{46}O_9$. The NMR spectroscopic data showed the presence of similar structural features to those of 16-*epi*-pretixin. However, compound **4** was found to possess a C_{12} -saturated-carbon ester side chain at C-2 instead of a hydroxy group as in 16-*epi*-pretixin. The MS² spectrum of the sodium-cationized molecule showed a fragment ion corresponding to a neutral loss of a C_{12} fatty acyl group followed by the loss of a molecule of water, which confirmed the nature of the side chain. As for compound **2**, bearing the same acyl chain at C-2, key COSY, HMBC and ROESY correlations were used to validate the structure as depicted. However, there was some doubt about the relative configuration of C-18 in the side chain of **4** connected to C-16. To resolve this uncertainty, the semi-synthesis of compound **4** was carried out from compound **3**, for which the absolute configuration is known. Esterification of **3** by lauric acid afforded the semi-synthetic compound of interest, with the ¹H NMR spectrum along with the specific rotation value ($[\alpha]_D^{20} +12$, c 0.1, MeOH) being identical to those of compound **4**, confirming their identical relative and absolute configurations (See S3, Supporting Information for comparison of the ¹H NMR spectra).

Compound **5** exhibited the same molecular formula and shared very similar data to **4**. The only structural difference between **5** and **4** was the attachment of the ester side-chain to C-18 instead of C-2 in **4**, as suggested by the HMBC correlation between H-18 and C-1'. The presence of an acyl chain at C-18 resulted in the downfield shift of protons H-16 and H-18 resonances (δ_H 2.70 to 3.15 and 3.94 to 5.06, respectively). As for compound **4**, the relative and absolute configurations of compound **5** were determined from the identical ¹H NMR spectroscopic data and optical rotation values between the semi-synthetic and natural compounds (See S3, Supporting Information for comparison of the ¹H NMR spectra).

The structure of compound **6** only differed from compound **5** by the possession of a shorter acyl chain at C-18, comprising 10 instead of 12 carbon atoms. Its absolute configuration may be assumed to be the same as that of compounds **3-5** on the basis of biogenetic considerations and their specific rotation values of similar sign and magnitude.

The analysis of NMR spectroscopic data of compound **7** indicated its structure as being very close to those of **5** and **6**. However, slight differences in ¹H NMR resonances of the methines H-12, H-16 and H-18, the methylene protons H₂-14, and the methyl protons H₃-7 suggested a different stereochemistry. Indeed, the relative configuration of C-16 was established on the basis of strong ROESY correlations between H-16, H-12 and H-14-β, all oriented above the mean plane. Thus, compound **7** was assigned as the epimer at C-16 of compound **6**. Crystals suitable for single X-ray analysis, obtained by vapor diffusion from methanol against hexane, were used to confirm its structure and absolute configuration (See Figure 2 and S2.2, Supporting Information for the X-ray ORTEP view of **7**).

The structures of compounds **8-10** were shown to only differ from that of **7** by the presence of longer acyl chains at C-18, with respectively, 2, 4 or 6 additional carbon atoms. The length of the chains was determined from the comparison of molecular formulas and analysis of the MS² spectra obtained from [M+Na]⁺ ions of **8**, **9** and **10**. They showed neutral losses of 200, 228 and 256 atomic mass units, respectively, corresponding to the loss of C₁₂, C₁₄ and C₁₆ acyl chains. Single crystals of **9** of relatively poor quality, only confirmed the relative configuration of the molecule (See S2.2, Supporting Information for the X-ray ORTEP view of **9**). On the basis of biogenetic considerations and specific rotation values all having similar signs and magnitude ($[\alpha]_D^{20}$ values of -17, -23, -24 and -18 for **7**, **8**, **9** and **10**, respectively), the same absolute configuration as that for compound **7** could be proposed for **8-10**.

The molecular formula of compound **11** was established as C₂₀H₂₆O₉ by HRESIMS, which showed a [M+H]⁺ ion peak at *m/z* 411.1680 (calcd for C₂₀H₂₇O₉⁺, 411.1661). The NMR

spectroscopic data of compound **11** were comparable to those of compounds **7-10**, but revealed the presence of a tertiary oxygenated carbon at C-16 (δ_c 79.1) instead of a methine carbon, and an additional singlet for a methoxy group resonating at δ_c 57.8 ppm, instead of the long acyl chains as in **7-10**. HMBC correlations from H-18 (δ_H 3.50), H-14- β (δ_H 2.16) and H₃-19 (δ_H 1.23) to C-16, and from the hydroxy proton at δ_H 3.33 (brs, OH-16) to C-14 on the one hand, and the COSY correlation between OH-16 and H-14- α (δ_H 3.65) on the other hand, indicated that a hydroxy group is attached at C-16. The HMBC correlation from the methyl singlet (δ_H 3.36) to C-18 (δ_c 78.2) and the ROESY correlations between the methoxy group protons and H-18 and H₃-19, confirmed the location at C-18 of the methoxy group. X-ray crystallographic analysis of compound **11**, crystallized by slow evaporation in MeOH, confirmed the relative and absolute configurations of **11** (Figure 2 and S2.2 Supporting Information). Compound **11** is a C-16 epimer of austrobuxusin B.²⁰

Compound **12** showed the molecular formula C₂₀H₂₈O₉, as indicated by HRESIMS. Its NMR spectroscopic data were similar to those of compound **11** but revealed the presence of an isopropyl group (δ_H 2.17, m, H-8; 1.06, d, $J = 4.0$ Hz, H₃-9 and 0.92, d, $J = 4.0$ Hz, H₃-10 and δ_c 28.1, 22.4 and 20.6, C-8, C-9 and C-10, respectively) at C-4 instead of an isopropenyl group. The location of this group at C-4 was confirmed by HMBC correlations from H₃-9 and H₃-10 to C-4 (δ_c 54.7) and C-8. The molecular structure and absolute configuration were reliably confirmed by X-ray crystallography analysis of a single-crystal obtained by slow evaporation in diethyl ether (Figure 2 and S2.2 Supporting Information).

The HRESIMS data of compound **13** showed a [M+H]⁺ ion peak at m/z 567.3195 (calcd for C₃₀H₄₇O₁₀⁺, 567.3175) corresponding to the molecular formula, C₃₀H₄₆O₁₀. The NMR spectroscopic data of **13** generally were comparable to those of compound **12**, but revealed the presence of a C₁₀-saturated-carbon ester side chain at C-2 instead of the hydroxy group present in **12**. The same key COSY and HMBC correlations that supported the assignments

of the ester chains at C-2 in **2** and **4** were also observed in **13** and were used to establish the structure as depicted. Crystallographic analysis confirmed the relative configuration of the molecule but the single crystals obtained were of insufficient quality to determine its absolute configuration (see S2.2, Supporting Information for the ORTEP view of **13**). However, this is assumed to be identical as that of compounds **11** and **12** based on their common biogenesis.

The HRESIMS data of compound **14** showed a protonated molecular ion at m/z 363.1467 $[M+H]^+$ (calcd 363.1449), corresponding to the molecular formula, $C_{19}H_{22}O_7$. The molecular structure consisted of the “norditerpene” picrotoxane skeleton comprising an ethylidene γ -butyrolactone. The 1H - 1H COSY correlation between a methyl doublet (δ_H 1.83, brd, $J = 7.1$ Hz, H₅-19) and a vinylic proton at δ_H 6.69 (qdd, $J = 7.1, 3.3, 2.9$ Hz, H-18) as well as the HMBC correlation from H₅-19 to C-18 (δ_C 136.0) indicated the presence of a CH_3 -CH= unit. Its position at C-16 was confirmed by the presence of HMBC correlations from H₅-19 to the quaternary carbon at δ_C 128.2 (C-16) and the carbonyl carbon at δ_C 170.3 (C-17), and from H-18 to carbons C-14, C-17 and C-16. Single crystals of **14** were grown by vapor diffusion from acetone against chloroform. X-ray crystallographic analysis confirmed the structure and absolute configuration of **14** as depicted in Figures 2 and S2.2 (Supporting Information). In view of the “norditerpene” picrotoxane biosynthesis scheme proposed by Corbella et al.,³¹ compound **14** could be formed from dehydration of 16-*epi*-pretoxin (**3**).

Compounds **1-14** were evaluated at 1 and 10 μ M for their antiproliferative activities against the HCT116 colon, U87-MG glioblastoma, and A549 lung human cancer cell lines. No compound showed cytotoxicity at 1 μ M. However, compounds **2**, **4** and **13** exhibited discernible cytotoxic activities at 10 μ M for all three cell lines. Their subsequent evaluations at 10 different concentrations allowed IC_{50} values to be determined and were found to be in the micromolar range. In particular, compounds **4** and **13** showed IC_{50} values of 5.8 ± 0.5 μ M for A549 cell line and 0.7 ± 0.1 μ M for U87-MG cell line, respectively. Interestingly,

compounds **2**, **4** and **13** are the only ones in the series that share a common acyl chain at C-2, which appears to be essential for cytotoxic activity for the cell lines evaluated.

The application of the presently used methodology in combining spectral and taxonomic data facilitated the isolation of 13 new picrotoxane-type compounds. The MetGem dual display offers two complementary visualizations, and is even more useful when studying large sets of samples. While the standard molecular networks representation may appear clearer and specific clusters can be easily located/defined, the t-SNE organization provides additional information by joining clusters and isolated nodes related to a larger group. The detection of taxon-specific clusters should be associated with a comprehensive dereplication step because specificity does not necessarily imply compound novelty or originality. In this regard, it is helpful to share MS² information in order to enrich databases with new experimental spectra. MS² data from most of the compounds isolated in this work were deposited in the GNPS public library accordingly (GNPS IDs can be found in the Experimental Section). These data are valuable for direct annotation of features via library searching, to train and refine computational metabolite annotation algorithms and to optimize the efficiency of annotation workflows that also consider orthogonal information (molecular network topology, predicted retention times, taxonomy, etc.).³²⁻³⁴ This is all the more important for structural classes not represented in spectral databases, as seems to be the case for picrotoxanes.

Furthermore, the efficiency and relevance of the methodology used greatly depend on the constitution of the sample series studied. When studying a botanical family or a phylogenetically related group, as was the case in the present study, it is beneficial to bring together as many different genera (or subsets) as possible. By covering the largest molecular extent within these groups, this allows for more reliable observations of their ubiquitous

chemical spaces and thus, avoids the emergence of “falsely taxon-specific” clusters. The present collection contains a non-homogeneous distribution of Picrodendraceae representatives (*Austrobuxus*, *Longetia* and *Scagea*, i.e, the only three Picrodendraceae genera present in New Caledonia). Consequently, the picrotoxanes isolated in this work were detected as specific to the genus *Austrobuxus*, but their occurrence within several other Picrodendraceae genera is known.¹¹ These compounds are nonetheless structurally fascinating and it would have been interesting to have greater sampling allowing a finer distinction to be made between the different types of picrotoxanes and their distribution within the Picrodendraceae.

Taxa originally belonging to the Euphorbiaceae have been reclassified into five different families. In particular, based on *rbcL* gene sequences analysis, Savolainen et al. elevated the Euphorbiaceae subfamily Oldfieldioideae to the Picrodendraceae (as the Pseudanthaceae).³⁵ The isolation of “norditerpene” picrotoxanes in *Austrobuxus carunculatus* appears to be consistent with this recent taxonomic revision, as this scaffold has only been reported in species in the Picrodendraceae. By visualizing chemical contents and highlighting specific molecular families, the molecular networking approach has facilitated the detection of chemotaxonomic markers, and thus appears as a convenient tool to support botanical classifications.

EXPERIMENTAL SECTION

General Experimental Procedures. Optical rotations were measured at 20 °C on an Anton Paar MCP 300 polarimeter. Melting points were measured with a Büchi Melting Point B-540 apparatus. NMR spectra were recorded on a Bruker 300 MHz instrument (Avance 300) for compounds **5** and **8** and on a Bruker 500 MHz instrument (Avance 500) for

compounds **1-4**, **6**, **7** and **9-14**. Chemical shifts (relative to CDCl_3 or CD_3OD for **1**, **3** and **12**) are in ppm and coupling constants are in Hz. Kinetex analytical and semi-preparative C_{18} columns (250×4.6 mm and 250×10 mm; $5 \mu\text{m}$ Phenomenex) were used for HPLC separations using a Waters autopurification system equipped with a sample manager (Waters 2767), a column fluidics organizer, a binary pump (Waters 2525), a UV-vis diode array detector (190–600 nm, Waters 2996), and a PL-ELS 1000 ELSD Polymer Laboratory detector. HRMS data were acquired using an Acquity Waters UPLC coupled to a Waters LCT Premier XE mass spectrometer. The UPLC system was equipped with a Waters Acquity PDA detector. Separation was achieved on a BEH C_{18} column ($1.7 \mu\text{m}$, 2.1 mm \times 50 mm) at a flow rate of $0.6 \text{ mL}\cdot\text{min}^{-1}$. Elution was conducted with a $\text{H}_2\text{O}-\text{CH}_3\text{CN} + 0.1\%$ formic acid gradient as follows: 95:5 to 0:100 in 5.5 min. The ionization was carried out using an electrospray ionization source in the positive mode (range m/z 80–1500). All solvents were purchased from Carlo Erba (France) or SDS (Peypin, France) and analytical plates (Si gel 60 F254) were from Merck (France). Prepacked GraceResolv silica cartridges were used for flash chromatography using a Teledyne Isco Combiflash Rf 200i.

Plant Material. Fruits of *Austrobuxus carunculatus* were collected at Port Boisé (New Caledonia) in May 1998 and authenticated by M. Litaudon. A voucher specimen (LIT-0538) has been deposited at the Herbarium IRD de Nouméa, New Caledonia. Fruits were cut before being dried using a hot air generator (the temperature did not exceed $40 \text{ }^\circ\text{C}$) and stored at room temperature.

Extraction and Isolation. The plant material (1.3 kg, dry wt.) was extracted three times for 3 hours at room temperature with EtOAc (2 L each). The EtOAc extract was concentrated in vacuo at $40 \text{ }^\circ\text{C}$. This extract was dissolved in 0.5 L of CH_3CN and subjected to a liquid/liquid

partition process with *n*-heptane (2 × 0.5 L) to afford 7.2 g of a CH₃CN-soluble fraction. This residue was subjected to silica gel column flash chromatography (120 g) using a gradient of *n*-heptane-EtOAc-MeOH of increasing polarities (100:0:0 to 0:100:0 to 0:80:20, 50 mL.min⁻¹) to afford 33 fractions, F1–F33, according to their TLC profiles. F12 (98 mg) was purified by silica gel column flash chromatography (12 g) using a gradient of *n*-heptane-EtOAc (95:5 to 60:40, in 15 min) to afford compound **2** (71 mg). F14 (86.5 mg), F15 (760.2 mg) and F16 (231.8 mg) were subjected to semi-preparative HPLC (Kinetex C₁₈, H₂O-CH₃CN 15:85 (F14), H₂O-CH₃CN 25:75 for 10 min to H₂O-CH₃CN 0:100 in 10 min (F15), H₂O-CH₃CN 20:80 (F16) at 4.5 mL.min⁻¹) to afford compound **13** (8.5 mg, *t_r* 6.9 min), compounds **5-10** (101.5 mg, *t_r* 11.3 min; 15.5 mg, *t_r* 7.3 min; 16.8 mg, *t_r* 6.7 min; 159.0 mg, *t_r* 9.6 min; 108.7 mg, *t_r* 16.2 min; 26.3 mg, *t_r* 20.8 min) and compound **4** (79.7 mg, *t_r* 9.2 min), respectively. F18 (124.4 mg) was purified by silica gel column flash chromatography (12 g) using a gradient of *n*-heptane-EtOAc (80:20 to 50:50, in 20 min) leading to four fractions, F18-1 to F18-4. F18-2 (69.2 mg) was separated by semi-preparative HPLC (Kinetex C₁₈, H₂O-CH₃CN 75:25 at 4.5 mL.min⁻¹) to yield compounds **1** (13.0 mg, *t_r* 5.0 min) and **14** (14.5 mg, *t_r* 12.2 min). Purification of F20 (215.8 mg) by semi-preparative HPLC (Kinetex C₁₈, H₂O-CH₃CN 75:25 at 4.5 mL.min⁻¹) furnished compounds **11** (18.9 mg, *t_r* 6.5 min) and **12** (21.0 mg, *t_r* 7.9 min). F22 (736.2 mg) was purified by silica gel column flash chromatography (80 g) using a gradient of *n*-heptane-EtOAc (60:40 to 50:50, in 20 min) leading to fractions F22-1 and F22-2. F22-1 (468.2 mg) was separated by semi-preparative HPLC (Kinetex C₁₈, H₂O-CH₃CN 75:25 at 4.5 mL.min⁻¹) to yield compound **3** (207.3 mg, *t_r* 6.4 min).

(1R,2S,3R,4R,5S,6R,11S,12R,13R)-Tutin (1): white needles; $[\alpha]_D^{20} +14$ (*c* 1, MeOH); ^1H and ^{13}C NMR, see Table S1, Supporting Information; HRESIMS m/z 277.1107 $[\text{M}-\text{H}_2\text{O}+\text{H}]^+$ (calcd for $\text{C}_{15}\text{H}_{17}\text{O}_5^+$, 277.1081). X-ray crystallographic data, see S2, Supporting Information. GNPS ID: CCMSLIB00005723341 $[\text{M} + \text{H}]^+$.

(1R,2S,3R,4R,5S,6R,11S,12R,13R)-Austrobuxusin E (2): white amorphous powder; $[\alpha]_D^{20} +11$ (*c* 0.1, MeOH); ^1H and ^{13}C NMR, see Table 1; HRESIMS m/z 477.2828 $[\text{M}+\text{H}]^+$ (calcd for $\text{C}_{27}\text{H}_{41}\text{O}_7^+$, 477.2847). GNPS ID: CCMSLIB00005723342 $[\text{M} + \text{H}]^+$, CCMSLIB00005723343 $[2\text{M} + \text{H}]^+$.

(1R,2S,3R,4R,5S,6R,11S,12R,13S,16S,18S)-16-epi-Pretoxin (3): white needles; mp 208-210°C; $[\alpha]_D^{20} +18$ (*c* 0.1, MeOH); ^1H and ^{13}C NMR, see Table 1; HRESIMS m/z 381.1550 $[\text{M}+\text{H}]^+$ (calcd for $\text{C}_{19}\text{H}_{25}\text{O}_8^+$, 381.1555). X-ray crystallographic data, see S2, Supporting Information. GNPS ID: CCMSLIB00005723344 $[\text{M}+\text{H}]^+$.

(1R,2S,3R,4R,5S,6R,11S,12R,13S,16S,18S)-Austrobuxusin F (4): white amorphous powder; $[\alpha]_D^{20} +12$ (*c* 0.1, MeOH); ^1H and ^{13}C NMR, see Table 1; HRESIMS m/z 563.3176 $[\text{M}+\text{H}]^+$ (calcd for $\text{C}_{31}\text{H}_{47}\text{O}_9^+$, 563.3215). GNPS ID: CCMSLIB00005723345 $[\text{M}+\text{H}]^+$, CCMSLIB00005723346 $[\text{M}+\text{Na}]^+$.

(1R,2S,3R,4R,5S,6R,11S,12R,13S,16S,18S)-Austrobuxusin G (5): white amorphous powder; $[\alpha]_D^{20} +6$ (*c* 0.1, MeOH); ^1H and ^{13}C NMR, see Table 1; HRESIMS m/z 563.3223 $[\text{M}+\text{H}]^+$ (calcd for $\text{C}_{31}\text{H}_{47}\text{O}_9^+$, 563.3215). GNPS ID: CCMSLIB00005723347 $[\text{M}+\text{H}]^+$, CCMSLIB00005723348 $[\text{M}+\text{Na}]^+$.

(1R,2S,3R,4R,5S,6R,11S,12R,13S,16S,18S)-Austrobuxusin H (6): white amorphous powder; $[\alpha]_D^{20} +11$ (*c* 0.5, MeOH); ^1H and ^{13}C NMR, see Table 2; HRESIMS m/z 535.2928 $[\text{M}+\text{H}]^+$ (calcd for $\text{C}_{29}\text{H}_{43}\text{O}_9^+$, 535.2913). GNPS ID: CCMSLIB00005723349 $[\text{M}+\text{H}]^+$, CCMSLIB00005723350 $[\text{M}+\text{Na}]^+$.

(1R,2S,3R,4R,5S,6R,11S,12R,13S,16R,18S)-16-epi-Austrobuxusin H (7): white needles; $[\alpha]_D^{20}$ -17 (c 0.1, MeOH); ^1H and ^{13}C NMR, see Table 2; HRESIMS m/z 535.2908 $[\text{M}+\text{H}]^+$ (calcd for $\text{C}_{29}\text{H}_{43}\text{O}_9^+$, 535.2913). X-ray crystallographic data, see S2, Supporting Information. GNPS ID: CCMSLIB00005723351 $[\text{M}+\text{H}]^+$, CCMSLIB00005723352 $[\text{M}+\text{Na}]^+$.

(1R,2S,3R,4R,5S,6R,11S,12R,13S,16R,18S)-16-epi-Austrobuxusin G (8): white amorphous powder; $[\alpha]_D^{20}$ -23 (c 0.1, MeOH); ^1H and ^{13}C NMR, see Table 2; HRESIMS m/z 563.3218 $[\text{M}+\text{H}]^+$ (calcd for $\text{C}_{31}\text{H}_{47}\text{O}_9^+$, 563.3215). GNPS ID: CCMSLIB00005723353 $[\text{M}+\text{H}]^+$, CCMSLIB00005723354 $[\text{M}+\text{Na}]^+$.

(1R,2S,3R,4R,5S,6R,11S,12R,13S,16R,18S)-Austrobuxusin I (9): white crystalline powder; mp 131-133°C; $[\alpha]_D^{20}$ -24 (c 0.1, MeOH); ^1H and ^{13}C NMR, see Table 2; HRESIMS m/z 591.3533 $[\text{M}+\text{H}]^+$ (calcd for $\text{C}_{33}\text{H}_{51}\text{O}_9^+$, 591.3539). X-ray crystallographic data, see S2, Supporting Information. GNPS ID: CCMSLIB00005723355 $[\text{M}+\text{H}]^+$, CCMSLIB00005723356 $[\text{M}+\text{Na}]^+$.

(1R,2S,3R,4R,5S,6R,11S,12R,13S,16R,18S)-Austrobuxusin J (10): white amorphous powder; $[\alpha]_D^{20}$ -18 (c 0.1, MeOH); ^1H and ^{13}C NMR, see Table 3; HRESIMS m/z 619.3870 $[\text{M}+\text{H}]^+$ (calcd for $\text{C}_{35}\text{H}_{55}\text{O}_9^+$, 619.3852). GNPS ID: CCMSLIB00005723357 $[\text{M}+\text{H}]^+$, CCMSLIB00005723358 $[\text{M}+\text{Na}]^+$.

(1R,2S,3R,4R,5S,6R,11S,12R,13S,16S,18S)-16-epi-Austrobuxusin B (11): white needles; $[\alpha]_D^{20}$ -21 (c 0.1, MeOH); ^1H and ^{13}C NMR, see Table 3; HRESIMS m/z 411.1680 $[\text{M}+\text{H}]^+$ (calcd for $\text{C}_{20}\text{H}_{27}\text{O}_9^+$, 411.1661). X-ray crystallographic data, see S2, Supporting Information. GNPS ID: CCMSLIB00005723359 $[\text{M}+\text{H}]^+$.

(1R,2S,3R,4R,5S,6R,11S,12R,13S,16S,18S)-Austrobuxusin K (12): white needles; $[\alpha]_D^{20}$ -18 (c 0.1, MeOH); ^1H and ^{13}C NMR, see Table 3; HRESIMS m/z 413.1850 $[\text{M}+\text{H}]^+$ (calcd for $\text{C}_{20}\text{H}_{29}\text{O}_9^+$, 413.1817). X-ray crystallographic data, see S2, Supporting Information. GNPS ID: CCMSLIB00005723360 $[\text{M}+\text{H}]^+$.

(1R,2S,3R,4R,5S,6R,11S,12R,13S,16S,18S)-Austrobuxusin L (13): white amorphous powder; $[\alpha]_{\text{D}}^{20}$ (c 0.1, MeOH); ^1H and ^{13}C NMR, see Table 3; HRESIMS m/z 567.3195 $[\text{M}+\text{H}]^+$ (calcd for $\text{C}_{30}\text{H}_{47}\text{O}_{10}^+$, 567.3175). X-ray crystallographic data, see S2, Supporting Information.

(1R,2S,3R,4R,5S,6R,11S,12R,13S)-Austrobuxusin M (14): white needles; $[\alpha]_{\text{D}}^{20}$ +15 (c 0.1, MeOH); ^1H and ^{13}C NMR, see Table 4; HRESIMS m/z 363.1467 $[\text{M}+\text{H}]^+$ (calcd for $\text{C}_{19}\text{H}_{23}\text{O}_7^+$, 363.1449). X-ray crystallographic data, see S2, Supporting Information.

X-ray Crystallography. X-ray analyses experimental details, ORTEP views, Crystal data, Structure refinement and Conventional hydrogen bonds of compounds **1**, **3**, **7**, **9**, **11**, **12**, **13** and **14** are available in S2 Supporting Information. Crystallographic data have been deposited in the Cambridge Crystallographic Data Center as entries 1996118-1996126 (compounds **1**, **3**, **7**, **9**, **11**, **12**, **13**, and **14** (both polymorphs) respectively) and can be obtained, free of charge, on application to the Director, CCDC, 12 Union Rd., Cambridge CB21EZ, U.K. [fax, +44(0)-1233-336033; email, deposit@ccdc.cam.ac.uk].

Conversion of Compound 3 into Compounds 4 and 5. To a solution of DMAP (1 mg, 11 μmol , 0.2 equiv.), EDCI.HCl (12 mg, 60 μmol , 1.1 equiv.) and lauric acid (12 mg, 58 μmol , 1.05 equiv.) in anhydrous dichloromethane (2 mL) at 0 °C under an argon atmosphere, was added compound **3** (21 mg, 55 μmol , 1 equiv.). The reaction mixture was stirred at room temperature for 24 h, transferred to a separatory funnel and subsequently washed three times with an aqueous solution of NaHCO_3 . The organic layer was dried with MgSO_4 , then filtered and concentrated in vacuo. ^1H NMR analysis confirmed the presence of the two ester products **4** and **5** in an 8:2 ratio, respectively. Purification by semi-preparative HPLC (Kinetex C_{18} ,

H₂O-MeCN 25:75 at 4.5 mL.min⁻¹) afforded the semi-synthetic compounds **4** (5.6 mg, 10 μmol, 18%, *t_r* 13.4 min) and **5** (0.3 mg, 0.5 μmol, 1%, *t_r* 12.1 min).

Data Dependent LC-ESI-HRMS² Analysis. LC analyses were performed with a Dionex Ultimate 3000 RSLC system equipped with an Accucore C₁₈ column (2.1 × 100 mm; 2.6 μm, ThermoScientific). The mobile phase consisted of water-acetonitrile (H₂O-CH₃CN) with 0.1% formic acid 20:80 held for 5 min, then a gradient from 20:80 to 100:0 in 20 min, at a flow rate of 350 μL.min⁻¹. The temperature of the column oven was set at 30 °C and the injection volume at 5 μL.

LC-HRESIMS² analysis was achieved by coupling the LC system to a hybrid quadrupole time-of-flight mass spectrometer Agilent 6540 (Agilent Technologies, Massy, France) equipped with an ESI dual source, operating in the positive-ion mode. Source parameters were set as follows: capillary temperature 325 °C, source voltage 3500 V, sheath gas flow rate 7 L.min⁻¹, nebulizer pressure 30 psi, drying-gas flow rate 10 L.min⁻¹, drying-gas temperature 350 °C, stealth-gas temperature 350 °C, skimmer voltage 45 V, fragmentor voltage 150 V, and nozzle voltage 500 V.

MS scans were operated in the full-scan mode from *m/z* 100 to 1000 (0.1 sec scan time) with a mass resolution of 20,000 at *m/z* 922. MS¹ scan was followed by MS² scans of the five most intense ions above an absolute threshold of 3000 counts. Selected parent ions were fragmented with a collision energy fixed at 35 eV and an isolation window of 1.3 amu.

Use of a calibration solution, containing two internal reference masses (purine, C₅H₄N₄, *m/z* 121.0509, and HP-921 [hexakis-(1H,1H,3H-tetrafluoropentoxy)phosphazene], C₁₈H₁₈O₆N₃P₃F₂₄, *m/z* 922.0098), routinely led to mass accuracy below 2 ppm. LC-UV data were analyzed with Chromeleon software (Dionex) and MS data acquisition and processing were performed using MassHunter Workstation software (Agilent Technologies, Massy, France).

MZmine 2 and Sirius Parameters. The 292 MS² data files were converted from the .d (Agilent) standard data-format to .mzXML format using MSConvert software, part of the ProteoWizard package.³⁶ All .mzXML were then processed using MZmine 2 v31.²¹ The mass detection was conducted using a noise level of 500 counts for MS and 50 counts for MSMS dimension. The ADAP chromatogram builder was used with a minimum group size of scans of 5, a group intensity threshold of 1000, a minimum highest intensity of 2000 and a m/z tolerance of 10 ppm.³⁷ The ADAP wavelets deconvolution algorithm was used with the following standard settings: S/N threshold = 7, minimum feature height = 2500, coefficient/area threshold = 25, peak duration range 0.05 – 0.5 min, RT wavelet range 0.01 - 0.1. However, considering the diverse molecular contents of extracts included within this wide-ranging collection, these parameters needed to be re-optimized for a few specific samples showing divergent chromatographic profiles and peak shapes. MS² scans were paired using a m/z tolerance range of 0.01 Da and RT tolerance range of 0.1 min. Isotopologues were grouped using the isotopic peaks grouper algorithm, with a m/z tolerance of 5 ppm and a RT tolerance of 0.1 min. Peak alignment was performed using the join aligner module (m/z tolerance = 5 ppm, weight for m/z = 2, weight for RT = 1, absolute RT tolerance 0.2 min). The peak list was gap-filled with the same RT and m/z range gap filler module (m/z tolerance of 5 ppm. Eventually, the .mgf spectral data file and its corresponding .csv metadata file (containing RT and peak areas) were exported using the dedicated “Export to GNPS” built-in modules.³⁸ The .mgf file was imported into Sirius 4.4.17 and spectra of interest were processed using the Molecular Formula Identification and the CSI:FingerID modules.^{25,26} The MS2 MassDev parameter was set to 5 ppm and only the bio database was considered for structure prediction.

Molecular Network Analysis. The two files mentioned above were imported into MetGem.²² MS² spectra were window-filtered by choosing only the top six peaks within the ± 50 Da window throughout the spectrum. The data were filtered by removing all peaks in the ± 17 Da range around the precursor m/z . The m/z tolerance windows used to find the matching peaks was set to 0.01 Da, and cosine scores were kept under consideration for spectra sharing 5 matching peaks at least. The standard molecular networks were created where edges were filtered to have a cosine score above 0.7. Further edges between two nodes were kept in the network only if each of the nodes appeared in each other's respective top 10 most similar nodes.

Regarding the t-SNE output, nodes were kept under consideration only if they shared at least 6 cosine scores above 0.5 with others ("at least 6 cosine scores above 0.5" parameter). The number of iterations, perplexity, learning rate, and early exaggeration parameters were set to 10000, 10, 150, and 12, respectively. The Barnes-Hut approximation was activated using a θ angle of 0.2°.

Both standard molecular networks and t-SNE outputs were imported in Cytoscape 3.7.2 for advanced network visualization and image export.³⁹

Data availability. Raw .d and .mzXML files of the EtOAc fruit extract of *Austrobuxus carunculatus* were deposited in the MassIVE data repository (MassIVE ID: MSV000085524). Raw .d files of the 292 extracts, .mgf and .csv output files, and molecular networks files are also available on MassIVE (MassIVE ID: MSV000086113).

Cell Culture and Proliferation Assay. Cancer cell lines were obtained from the American Type Culture Collection (Rockville, MD, USA) and were cultured according to the supplier's instructions. Human HCT-116 colorectal carcinoma cells were grown in Gibco McCoy's

medium 5A supplemented with 10% fetal calf serum (FCS) and 1% L-glutamine. Human A549 lung cancer cells were grown in RPMI 1640 supplemented with 10% FCS and 1% L-glutamine. Human U87-MG glioblastoma cells were grown in DMEM supplemented with 10% FCS and 1% L-glutamine. Cells were maintained at 37 °C in a humidified atmosphere containing 5% CO₂.

Cell viability was determined by a luminescent assay according to the manufacturer's instructions (Promega, Madison, WI, USA). For IC₅₀ determinations, the cells were seeded in 96-well plates (3 × 10³ cells/well) containing 90 μL of growth medium. After 24 h of culture, the cells were treated with the test compounds at 10 different final concentrations. Each concentration was obtained from serial dilutions in culture medium starting from the stock solution. Control cells were treated with the vehicle. Experiments were performed in triplicate.

After 72 h of incubation, 100 μL of CellTiter Glo Reagent were added for 15 min before recording luminescence with a spectrophotometric plate reader PolarStar Omega (BMG LabTech). The dose-response curves were plotted with Graph Prism software and the IC₅₀ values were calculated using the Graph Prism software from polynomial curves (four or five-parameter logistic equations).

Doxorubicin was used as a positive control with IC₅₀ values of 56.6 ± 0.8, 99.6 ± 2.3 and 90.0 ± 3.0 nM for HCT116, U87-MG and A549 cell lines, respectively.

ASSOCIATED CONTENT

Supporting Information

¹H and ¹³C NMR chemical shifts of compound **1**, X-ray analyses experimental details, ORTEP views, Crystal data, Structure refinement and Conventional hydrogen bonds of

compounds **1, 3, 7, 9, 11, 12, 13** and **14**, ^1H , ^{13}C , HSQC, HMBC, COSY, ROESY NMR spectra of compounds **1-4, 6, 9, 11-14**, ^1H , ^{13}C of compounds **5, 8, 10**, ^1H , ^{13}C , HMBC of compound **7** and HRESIMS data for compounds **1-14**. (PDF)

The Supporting Information is available free of charge on the ACS Publications website.

AUTHOR INFORMATION

Corresponding authors

* Tel: + 33 1 69 82 30 38. E-mail: cecile.apel@cnrs.fr

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

ORCID

Florent Olivon: 0000-0002-3662-5390

Pascal Retailleau: 0000-0003-3995-519X

Sandy Desrat: 0000-0002-6544-9242

David Touboul: 0000-0003-2751-774X

Fanny Roussi: 0000-0002-5941-9901

Cécile Apel: 0000-0003-3785-9620

Marc Litaudon: 0000-0002-0877-8234

Notes

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

The authors are very grateful to the South Province of New Caledonia that has facilitated our field investigation and B. Fogliani for his contribution to species collection. We thank H. Lévaique and J. Bignon, who performed the cytotoxic assays on the three cancer cell lines. This work has benefited from an “Investissement d’Avenir” grant managed by Agence Nationale de la Recherche (CEBA, ref ANR-10-LABX-25-01).

REFERENCES

- (1) Firm, R. D.; Jones, C. G. *Nat. Prod. Rep.* **2003**, *20*, 382-391.
- (2) Futuyma, D. J.; Agrawal, A. A. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 18054.
- (3) Kessler, A.; Kalske, A. *Annu. Rev. Ecol. Evol. Syst.* **2018**, *49*, 115-138.
- (4) Wink, M. *Phytochemistry* **2003**, *64*, 3-19.
- (5) Tetenyi, P. *Ann. Missouri Bot. Gard.* **1987**, *74*, 600-608.
- (6) Olivon, F.; Allard, P.-M.; Koval, A.; Righi, D.; Genta-Jouve, G.; Neyts, J.; Apel, C.; Pannecouque, C.; Nothias, L.-F.; Cachet, X.; Marcourt, L.; Roussi, F.; Katanaev, V. L.; Touboul, D.; Wolfender, J.-L.; Litaudon, M. *ACS Chem. Biol.* **2017**, *12*, 2644-2651.
- (7) Olivon, F.; Apel, C.; Retailleau, P.; Allard, P. M.; Wolfender, J. L.; Touboul, D.; Roussi, F.; Litaudon, M.; Desrat, S. *Org. Chem. Front.* **2018**, *5*, 2171-2178.
- (8) Olivon, F.; Remy, S.; Grelier, G.; Apel, C.; Eydoux, C.; Guillemot, J.-C.; Neyts, J.; Delang, L.; Touboul, D.; Roussi, F.; Litaudon, M. *J. Nat. Prod.* **2019**, *82*, 330-340.
- (9) 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.; Porto, C.; Bouslimani, A.; Melnik, A. V.; Meehan, M. J.; Liu, W.-T.; Crüsemann, M.; Boudreau, P. D.; Esquenazi, E.; Sandoval-Calderón, M.; Kersten, R. D.; Pace, L. A.; Quinn, R. A.; Duncan, K. R.; Hsu, C.-C.; Floros, D. J.; Gavilan, R. G.; Kleigrew, K.; Northen, T.; Dutton, R. J.; Parrot, D.; Carlson, E. E.; Aigle, B.; Michelsen, C. F.; Jelsbak, L.; Sohlenkamp, C.; Pevzner, P.; Edlund, A.; McLean, J.; Piel, J.; Murphy, B. T.; Gerwick, L.; Liaw, C.-C.; Yang, Y.-L.; Humpf, H.-U.; Maansson, M.; Keyzers, R. A.; Sims, A. C.; Johnson, A. R.; Sidebottom, A. M.; Sedio, B. E.; Klitgaard, A.; Larson, C. B.; Boya P, C. A.; Torres-Mendoza, D.; Gonzalez, D. J.; Silva, D. B.; Marques, L. M.; Demarque, D. P.; Pociute, E.; O'Neill, E. C.; Briand, E.; Helfrich, E. J. N.; Granatosky, E. A.; Glukhov, E.; Ryffel, F.; Houson, H.; Mohimani, H.; Kharbush, J. J.; Zeng, Y.; Vorholt, J. A.; Kurita, K. L.; Charusanti, P.; McPhail, K. L.; Nielsen, K. F.; Vuong, L.; Elfeki, M.; Traxler, M. F.; Engene, N.; Koyama, N.; Vining, O. B.; Baric, R.; Silva, R. R.; Mascuch, S. J.; Tomasi, S.; Jenkins, S.; Macherla, V.; Hoffman, T.; Agarwal, V.; Williams, P. G.; Dai, J.; Neupane, R.; Gurr, J.; Rodríguez, A. M. C.; Lamsa, A.; Zhang, C.; Dorrestein, K.; Duggan, B. M.; Almaliti, J.; Allard, P.-M.; Phapale, P.; Nothias, L.-F.; Alexandrov, T.; Litaudon, M.; Wolfender, J.-L.; Kyle, J. E.; Metz, T. O.; Peryea, T.; Nguyen, D.-T.; VanLeer, D.; Shinn, P.; Jadhav, A.; Müller, R.; Waters, K. M.; Shi, W.; Liu, X.; Zhang, L.; Knight, R.; Jensen, P. R.; Palsson, B. Ø.; Pogliano, K.; Linington, R. G.; Gutiérrez, M.; Lopes, N. P.; Gerwick, W. H.; Moore, B. S.; Dorrestein, P. C.; Bandeira, N. *Nat. Biotechnol.* **2016**, *34*, 828-837.

- (10) The Angiosperm Phylogeny Group. *Bot. J. Linn. Soc.* **2003**, *141*, 399-436.
- (11) Gössinger, E., Picrotoxanes. In *Fortschritte der Chemie organischer Naturstoffe / Progress in the Chemistry of Organic Natural Products*; Kinghorn, A. D.; Falk, H.; Kobayashi, J., Eds.; Springer Verlag: Vienna, 2010, *93*, 71-210.
- (12) Hosie, A. M.; Ozoe, Y.; Koike, K.; Ohmoto, T.; Nikaido, T.; Sattelle, D. B. *Br. J. Pharmacol.* **1996**, *119*, 1569-1576.
- (13) Newland, C. F.; Cull-Candy, S. G. *J. Physiol.* **1992**, *447*, 191-213.
- (14) Ozoe, Y.; Akamatsu, M. *Pest Manag. Sci.* **2001**, *57*, 923-931.
- (15) Ozoe, Y.; Akamatsu, M.; Higata, T.; Ikeda, I.; Mochida, K.; Koike, K.; Ohmoto, T.; Nikaido, T. *Biorg. Med. Chem.* **1998**, *6*, 481-492.
- (16) Ozoe, Y.; Hasegawa, H.; Mochida, K.; Koike, K.; Suzuki, Y.; Nagahisa, M.; Ohmoto, T. *Biosci. Biotechnol. Biochem.* **1994**, *58*, 1506-1507.
- (17) Porter, L. A. *Chem. Rev.* **1967**, *67*, 441-464.
- (18) Smart, T. G.; Constanti, A. *Proc. R. Soc. Lond, B* **1986**, 191-216.
- (19) Twyman, R. E.; Rogers, C. J.; Macdonald, R. L. *Neurosci. Lett.* **1989**, *96*, 89-95.
- (20) Demirkiran, O.; Campitelli, M.; Wang, C.; Feng, Y. *Tetrahedron* **2016**, *72*, 8400-8405.
- (21) Pluskal, T.; Castillo, S.; Villar-Briones, A.; Orešič, M. *BMC Bioinform.* **2010**, *11*, 395.
- (22) Olivon, F.; Elie, N.; Grelier, G.; Roussi, F.; Litaudon, M.; Touboul, D. *Anal. Chem.* **2018**, *90*, 13900-13908.
- (23) <https://gnps.ucsd.edu/ProteoSAFe/libraries.jsp>
- (24) Corbella, A.; Jommi, G.; Rindone, B.; Scolastico, C. *Tetrahedron* **1969**, *25*, 4835-4841.
- (25) Dührkop, K.; Fleischauer, M.; Ludwig, M.; Aksenov, A. A.; Melnik, A. V.; Meusel, M.; Dorrestein, P. C.; Rousu, J.; Böcker, S. *Nat. Methods* **2019**, *16*, 299-302.
- (26) Dührkop, K.; Shen, H.; Meusel, M.; Rousu, J.; Böcker, S. *Proc. Natl. Acad. Sci. U.S.A.* **2015**, *112*, 12580.
- (27) Easterfield, T. H.; Aston, B. C. *J. Chem. Soc., Trans.* **1901**, *79*, 120-126.
- (28) Burnett, M. N.; Johnson, C. K. *ORTEP-III: Oak Ridge Thermal Ellipsoid Plot Program for Crystal Structure Illustrations*; Oak Ridge, TN, United States, 1996-07-01, 1996.
- (29) Macrae, C. F.; Edgington, P. R.; McCabe, P.; Pidcock, E.; Shields, G. P.; Taylor, R.; Towler, M.; van de Streek, J. *J. Appl. Crystallogr.* **2006**, *39*, 453-457.
- (30) Gulbis, J. M.; Mackay, M. F.; Wong, M. G. *Aust. J. Chem.* **1989**, *42*, 1881-1896.
- (31) Corbella, A.; Gariboldi, P.; Jommi, G.; Scolastico, C. *J. Chem. Soc. D* **1969**, 634-635.
- (32) Bonini, P.; Kind, T.; Tsugawa, H.; Barupal, D. K.; Fiehn, O. *Anal. Chem.* **2020**, *92*, 7515-7522.
- (33) da Silva, R. R.; Wang, M.; Nothias, L.-F.; van der Hooft, J. J. J.; Caraballo-Rodríguez, A. M.; Fox, E.; Balunas, M. J.; Klassen, J. L.; Lopes, N. P.; Dorrestein, P. C. *PLoS Comput. Biol.* **2018**, *14*, e1006089.
- (34) Rutz, A.; Dounoue-Kubo, M.; Ollivier, S.; Bisson, J.; Bagheri, M.; Saesong, T.; Ebrahimi, S. N.; Ingkaninan, K.; Wolfender, J.-L.; Allard, P.-M. *Front. Plant Sci.* **2019**, *10*, 1329.
- (35) Savolainen, V.; Fay, M. F.; Albach, D. C.; Backlund, A.; Bank, M. v. d.; Cameron, K. M.; Johnson, S. A.; Lledó, M. D.; Pintaud, J. C.; Powell, M.; Sheahan, M. C.; Soltis, D. E.; Soltis, P. S.; Weston, P.; Whitten, W. M.; Wurdack, K. J.; Chase, M. W. *Kew Bull.* **2000**, *55*, 257-309.
- (36) Chambers, M. C.; Maclean, B.; Burke, R.; Amodei, D.; Ruderman, D. L.; Neumann, S.; Gatto, L.; Fischer, B.; Pratt, B.; Egertson, J.; Hoff, K.; Kessner, D.; Tasman, N.; Shulman, N.; Frewen, B.; Baker, T. A.; Brusniak, M.-Y.; Paulse, C.; Creasy, D.; Flashner, L.; Kani, K.; Moulding, C.; Seymour, S. L.; Nuwaysir, L. M.; Lefebvre, B.; Kuhlmann, F.; Roark, J.; Rainer, P.; Detlev, S.; Hemenway, T.; Huhmer, A.; Langridge, J.; Connolly, B.; Chadick, T.;

Holly, K.; Eckels, J.; Deutsch, E. W.; Moritz, R. L.; Katz, J. E.; Agus, D. B.; MacCoss, M.; Tabb, D. L.; Mallick, P. *Nat. Biotechnol.* **2012**, *30*, 918-920.

(37) Myers, O. D.; Sumner, S. J.; Li, S.; Barnes, S.; Du, X. *Anal. Chem.* **2017**, *89*, 8696-8703.

(38) Nothias, L. F.; Petras, D.; Schmid, R.; Dührkop, K.; Rainer, J.; Sarvepalli, A.; Protsyuk, I.; Ernst, M.; Tsugawa, H.; Fleischauer, M.; Aicheler, F.; Aksenov, A.; Alka, O.; Allard, P.-M.; Barsch, A.; Cachet, X.; Caraballo, M.; Da Silva, R. R.; Dang, T.; Garg, N.; Gauglitz, J. M.; Gurevich, A.; Isaac, G.; Jarmusch, A. K.; Kameník, Z.; Kang, K. B.; Kessler, N.; Koester, I.; Korf, A.; Gouellec, A. L.; Ludwig, M.; Christian, M. H.; McCall, L.-I.; McSayles, J.; Meyer, S. W.; Mohimani, H.; Morsy, M.; Moyne, O.; Neumann, S.; Neuweger, H.; Nguyen, N. H.; Nothias-Esposito, M.; Paolini, J.; Phelan, V. V.; Pluskal, T.; Quinn, R. A.; Rogers, S.; Shrestha, B.; Tripathi, A.; van der Hooft, J. J. J.; Vargas, F.; Weldon, K. C.; Witting, M.; Yang, H.; Zhang, Z.; Zubeil, F.; Kohlbacher, O.; Böcker, S.; Alexandrov, T.; Bandeira, N.; Wang, M.; Dorrestein, P. C. *bioRxiv* **2019**, 812404.

(39) Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N. S.; Wang, J. T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. *Genome Res.* **2003**, *13*, 2498-2504.

Graphical Abstract:

