

Supplementary Information

Comparative Analyses of Metabolomic Fingerprints and Cytotoxic Activities of Soft Corals from the Colombian Caribbean

Liliana Santacruz ¹, Olivier P. Thomas ^{2,*}, Carmenza Duque ³, Mónica Puyana ⁴
and Edison Tello ^{1,*}

¹ Bioprospecting Research Group and Bioscience Doctoral Program, Faculty of Engineering, Campus Puente del Común, Universidad de La Sabana, Chía, Colombia

² Marine Biodiscovery, School of Chemistry and Ryan Institute, National University of Ireland Galway (NUI Galway), University Road, H91 TK33 Galway, Ireland

³ Departamento de Química, Universidad Nacional de Colombia, Carrera 30 # 45-03, Bogotá, Colombia; cduqueb@unal.edu.co

⁴ Departamento de Ciencias Biológicas y Ambientales, Universidad Jorge Tadeo Lozano, Carrera 4 # 22-61, Bogotá, Colombia; monica.puyana@utadeo.edu.co

* Correspondence: olivier.thomas@nuigalway.ie (L.S.); edisson.tello@unisabana.edu.co (E.T.);
Tel.: +353-(0)9-149-3563 (L.S.); Tel.: +57-32-0699-5696 (E.T.)

Index of Figures

Figure S1. Principal component analysis (PCA); in the triangle the pools (QC) are shown. 3

Index of Tables

Table S1. Name of species of soft corals used in this research a Collection Code (ICN), assigned by Collection of the Institute of Natural Sciences of the National University of Colombia, (Bogotá, Colombia). 2

Table S2. Percentage of cytotoxic activity against three cancerous cell lines. (1 if it is above 50% 0 if it is below) 4

Table S3. Scrip using Galaxy software to UPLC-MS data extraction..... 5

Table 1. Name of species of soft corals used in this research a Collection Code (ICN), assigned by Collection of the Institute of Natural Sciences of the National University of Colombia, (Bogotá, Colombia).

Name Assigned	Name of the Species	Accession	Code ICN-UN
Cali 2	<i>Plexaura kukenthali</i>	C2P	ICN-MHN (Po)-CO-271
Cali 4	<i>Pseudopterogorgia albatrossae</i>	C4P	ICN-MHN (Po)-CO-272
Cali 6	<i>Eunicea succinea</i>	C6E	ICN-MHN (Po)-CO-273
Cali 8	<i>Eunicea succinea plantaginea</i>	C8E	ICN-MHN (Po)-CO-274
Cali 9	<i>Eunicea clavigera</i>	C9E	ICN-MHN (Po)-CO-275
Cali 11	<i>Eunicea flexuosa</i>	C11E	ICN-MHN (Po)-CO-276
Cali 12	<i>Plexaura cf. nina</i>	C12P	ICN-MHN (Po)-CO-279
Cali 13	<i>Eunicea fusca</i>	C13E	ICN-MHN (Po)-CO-277
Cali 15	<i>Pseudopterogorgia albatrossae</i>	C15Pst	ICN-MHN (Po)-CO-278
Cali 16	<i>Eunicea fusca</i>	C16E	ICN-MHN (Po)-CO-277
Cali 17	<i>Eunicea clavigera thin form</i>	C17E	ICN-MHN (Po)-CO-280
Cali 19	<i>Eunicea flexuosa</i>	C19E	ICN-MHN (Po)-CO-276
Cali 24	<i>Plexaurella nutans</i>	C24Plr	ICN-MHN (Po)-CO-281
Gra 1	<i>Plexaurella fusifera</i>	G1P Plr	ICN-MHN (Po)-CO-282
Gra 2	<i>Plexaura homomalla</i>	G2P	ICN-MHN (Po)-CO-283
Gra 3	<i>Eunicea clavigera thin form</i>	G3E	ICN-MHN (Po)-CO-284
Gra 4	<i>Eunicea asperula</i>	G4E	ICN-MHN (Po)-CO-285
Gra 5	<i>Eunicea clavigera</i>	G5E	ICN-MHN (Po)-CO-275
Gra 6	<i>Plexaura kukenthali</i>	G6P	ICN-MHN (Po)-CO-292
Gra 8	<i>Eunicea clavigera</i>	G8E	ICN-MHN (Po)-CO-275
Gra 9	<i>Plexaura sp.</i>	G9P	ICN-MHN (Po)-CO-286
Gra 10	<i>Eunicea knightii</i>	G10E	ICN-MHN (Po)-CO-287
Gra 13	<i>Eunicea clavigera</i>	G13E	ICN-MHN (Po)-CO-275
Gra 14	<i>Eunicea cf. calyculata</i>	G14E	ICN-MHN (Po)-CO-288
Gra 17	<i>Pseudoplexaura flagellosa</i>	G17Ps	ICN-MHN (Po)-CO-289
Gra 18	<i>Plexaura kukenthali</i>	G18P	ICN-MHN (Po)-CO-292
Gra 19	<i>Eunicea tayrona</i>	G19E	ICN-MHN (Po)-CO-290
Gra 22	<i>Plexaurella sp.</i>	G22P	ICN-MHN (Po)-CO-291

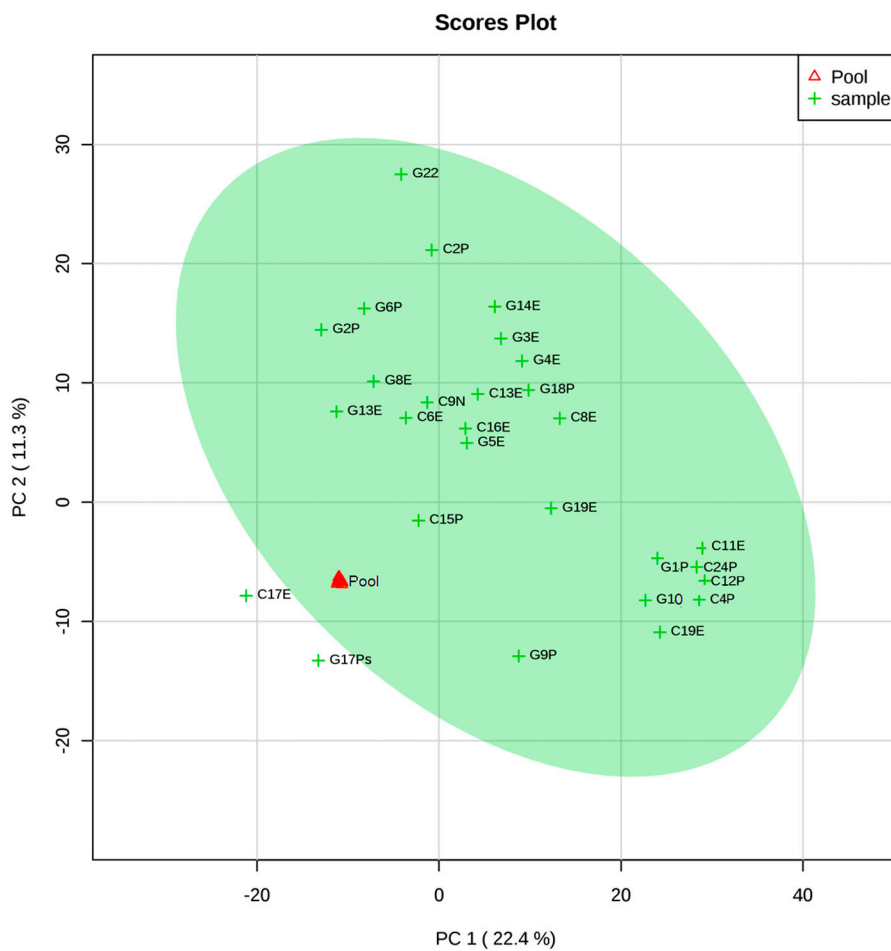


Figure 1. Principal component analysis (PCA); in the triangle the pools (QC) are shown.

Table 2. Percentage of cytotoxic activity against three cancerous cell lines. (1 if it is above 50% 0 if it is below).

Obs	Name	SiHa (%)	PC3 (%)	A549 (%)	L929 (%)	SiHa	PC3	A549
1	G9Px	84.5	20.4	31.0	49.8	1	0	0
2	C6Es	69.0	61.7	35.1	28.9	1	1	0
3	G17Ef	66.8	40.3	71.5	65.9	1	0	1
4	G14Ec	63.5	22.0	65.7	19.2	1	0	1
5	C17Ec	61.3	17.0	27.5	13.8	1	0	0
6	C16Ee	60.5	30.0	39.1	21.3	1	0	0
7	G3Ec	59.9	33.1	45.3	5.8	1	0	0
8	G6Pk	57.0	40.3	48.5	23.0	1	0	0
9	G4Ea	53.9	35.5	67.6	14.6	1	0	1
10	C9NE	53.3	32.1	14.7	17.3	1	0	0
11	C8Ec	48.4	35.9	43.6	13.4	0	0	0
12	C2Pk	46.2	7.8	62.4	16.2	0	0	1
13	G10Ek	46.1	1.3	45.4	14.5	0	0	0
14	G19Et	45.9	39.7	67.8	18.8	0	0	1
15	G8Ec	45.9	38.5	49.9	23.4	0	0	0
16	Cali3	45.9	40.9	44.0	18.4	0	0	0
17	G5Ec	39.8	50.6	46.2	12.6	0	1	0
18	G1Pf	39.4	23.8	33.0	11.5	0	0	0
19	G22P	37.6	33.4	41.9	20.7	0	0	0
20	G13Ec	36.9	30.9	49.7	6.8	0	0	0
21	C11Ef	36.8	12.1	49.3	21.7	0	0	0
22	C13EA	36.4	27.2	43.8	3.5	0	0	0
23	G2Po	34.8	37.8	21.1	14.3	0	0	0
24	G18Pk	34.0	44.0	52.5	31.2	0	0	1
25	C4Pa	27.1	21.7	33.1	14.2	0	0	0
26	C12Pn	23.7	16.0	43.3	21.7	0	0	0
27	C19Ef	15.1	16.1	36.1	20.7	0	0	0
28	C24Pa	13.1	65.1	39.0	13.3	0	1	0
29	Doxorubicin	60.1	46.9	58.2	16.0			

Table 3. Scrip using Galaxy software to UPLC-MS data extraction.

<p>Step 1: Input dataset (soft corals flow)</p>
<p>Step 2: xcms.xcmsSet</p> <p>Extraction method for peaks detection: centWave, ppm=15 max tolerated ppm m/z deviation in consecutive scans in ppm 15.</p> <p>peakwidth=c(5,20), min,Max peak width in seconds, 5,20</p> <p>snthresh=3, signal/Noise threshold</p> <p>prefilter=c(3,5000))</p>
<p>Step 3: xcms.group</p> <p>xset RData file, Output dataset 'xsetRData' from step 2</p> <p>Method to use for grouping:</p> <p>xset<-group(xset,bw=2,mzwid=0.015,minfrac=0.3)</p> <p>density Bandwidth: 2</p> <p>Minimum fraction of samples necessary: 0.3</p> <p>Width of overlapping m/z slices: 0.015</p> <p>Maximum number of groups to identify in a single m/z slice (50)</p>
<p>Step 4: xcms.retcor, Method to use for retention time correction</p> <p>xset RData file, Output dataset 'xsetRData' from step 3</p> <p>xset2<-retcor(xset, method="obiwarp", plotype ="deviation")</p> <p>peakgroups</p> <p>Smooth method</p> <p>loess</p> <p>Number of extra peaks to allow in retention time correction correction groups: 1</p> <p>Number of missing samples to allow in retention time correction groups: 1</p> <p>Degree of smoothing for local polynomial regression fitting: 0.2</p> <p>Family, gaussian, plotype, deviation.</p>
<p>Step 5: xcms.group, Method to use for grouping</p> <p>xset RData file, output dataset 'xsetRData' from step 4</p> <p>xset3<-group(xset2,bw=2,mzwid=0.015,minfrac=0.3)</p> <p>density Bandwidth: 2</p> <p>Minimum fraction of samples necessary: 0.3</p> <p>Width of overlapping m/z slices: 0.015</p> <p>Maximum number of groups to identify in a single m/z slice: 50</p>
<p>Step 6: xcms.fillPeaks, Filling method</p> <p>xset RData file</p> <p>Output dataset 'xsetRData' from step 5</p>

```
xset = fillPeaks(chrom)
```

Step 7: CAMERA.annotate

RData file,

Output dataset 'xsetRData' from step 6

num_digits: 4

groupFWHM: multiplier of the standard deviation: 6

groupFWHM: percentage of FWHM width: 0.6

findIsotopes: max. ion charge: 3

findIsotopes: max. number of expected isotopes: 4

findIsotopes: The percentage number of samples, which must satisfy the C12/C13 rule for isotope annotation:
0.5

General ppm error: 5

Mzabs: 0.015

General used intensity value

findAdducts: Which polarity mode was used for measuring of the ms sample negative

groupCorr: correlation threshold (0.1), 0.75

groupCorr: Method selection for grouping peaks after correlation analysis into pseudospectra

hcs, groupCorr: significant correlation threshold (0.05)

findAdducts: If no ruleset is provided, calculate ruleset with max. number n of [nM+x]

diffreport: number of the most significantly different analytes to create EICs for, (50)

diffreport: width (in seconds) of EICs produced, (200)

diffreport: logical indicating whether the reports should be sorted by p-value, (True)