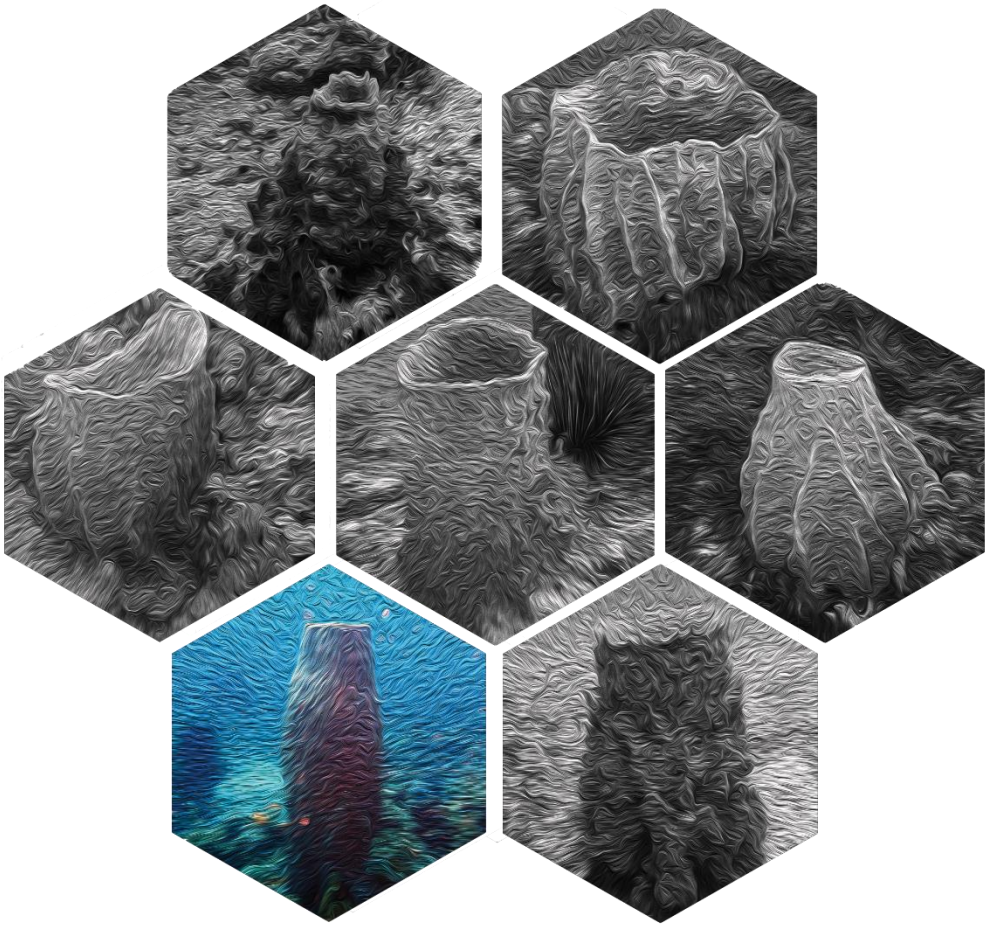
Giant barrel sponges in diverse habitats: a story about the metabolome

Bayona Maldonado, L.M.



Chapter 5

Metabolic variation in Caribbean giant barrel sponges: influence of age and sea-depth

Lina M. Bayona[1], Min-Sun Kim[2], Thomas Swierts[3], Geum-Sook Hwang[4], Nicole J. de Voogd[3,5], Young Hae Choi[1]

[1] Natural Products Laboratory, Institute of Biology, Leiden University, 2333 BE Leiden, The Netherlands.

[2] Food Analysis Center, Korea Food Research Institute, Wanju, Korea.

[3] Naturalis Biodiversity Center, Marine Biodiversity, 2333 CR Leiden, The Netherlands.

[4] Integrated Metabolomics Research Group, Western Seoul Center, Korea Basic Science Institute, Seoul, Korea.

[5] Institute of Environmental Sciences, Leiden University, 2333 CC Leiden, The Netherlands.

Manuscript submitted to marine environmental research

Abstract

The biochemical differentiation of widely distributed long-living marine organisms according to their age or the depth of waters in which they grow is an intriguing topic in marine biology. Especially sessile life forms such as sponges, could be expected to actively regulate biological processes and interactions with their environment through chemical signals in a multidimensional manner. Conventional targeted analysis is too limited to be able to give comprehensive answers to this interesting, yet complex question. In recent years, the development of chemical profiling methods such as metabolomics provided an approach that has encouraged the investigation of the chemical interactions of these organisms. In this study, LC-MS based metabolomics followed by feature based molecular networking (FBMN) was used to explore the effects of both biological and environmental factors on the metabolome of giant barrel sponges, chosen as model organisms as they are distributed throughout a wide range of sea-depths. Samples of giant barrel sponges belonging to two different genetic groups and of different ages, were collected along a significant depth gradient (7-43 m). Using molecular networking, metabolites were deduced mainly for the families of compounds present in the samples. To identify the metabolites influenced by each factor, further multivariate data analyses were performed including the correlations between the metabolome and the considered factors, i.e., age and depth. It was found that depth only caused significant changes in the metabolome of one of the genetic groups. The effect of sponge age, on the other hand, had a significant effect on the metabolome of sponges, independently of their genetic group or depth of growth. Among the identified chemical components of the sponges, the phospholipid profile was found to be most affected by age and depth but each of these factors affected specific phospholipids. As a result, two previously unreported metabolites that were found to be related to depth, compounds **1** and **2**, were chemically elucidated using NMR and MS.

Keyword: *Xestospongia muta*, metabolomics, phospholipids, molecular networking, environmental condition, sponge growth.

1. Introduction

In recent years, coral reefs around the world have experienced substantial changes in their biodiversity as a response to changes in the environment resulting from both natural and anthropogenic phenomena (Hughes et al. 2007; Pandolfi et al. 2003). Global warming causes an increase in sea surface temperatures that leads to massive bleaching events of corals and are the primary driver of recent global reef decline (Hughes et al. 2017). The decay of corals has a profound influence on other organisms in the reef, including sponges and algae, which are taking over the substrate space freed by the coral decline (Bell et al. 2013, 2018; McManus and Polsenberg 2004). Sponges in particular, play a vital role in reef life, not only because they are the most abundant taxa in Caribbean reefs, but also because they are the responsible for the transfer of nutrients along the trophic chain (Diaz and Rützler 2001; Pawlik and McMurray 2020; Rix et al. 2018).

Giant barrel sponges, which belong to the genus *Xestospongia*, have been identified as a important reef player. Due to their large size and barrel-shape appearance, these sponges provide a habitat for many other animals (Swierts, et al. 2018a). In fact, the Caribbean giant barrel sponges can cover up to 9% of the available surface (Zea 1993), and recent studies have shown that the population growth rate of *Xestospongia muta* has been rapidly increasing (McMurray et al. 2015), to the point that they are one of the most abundant organisms in reefs (Loh and Pawlik 2014). They also play an active part in the regulation of the nitrogen and carbon cycles in their habitat (Fiore et al. 2013; Southwell et al. 2008)

Xestospongia muta is characterized as a high microbial abundance (HMA) sponge with a slow growth rate and fluctuating chemical defenses (Gloeckner et al. 2014; Loh and Pawlik 2014; Pawlik et al. 1995). The interaction of these sessile organisms with their environment, as well as their biological processes, must be mediated essentially, through chemical signals. It can be anticipated, thus, that changes in environmental conditions will be reflected at a metabolic level as an alteration of the metabolome (Hay 2009; Paul et al. 2006; Pawlik 2011). As expected, a wide range of metabolites have been reported in *Xestospongia* spp. sponges, most of which have shown significant responses to bioassays, confirming that they might have a role in various biological and physiological processes (Zhou et al. 2010). However, the biotic or abiotic factors driving the production of these compounds remain unclear mainly due to the lack of methods able to deal with their chemical complexity. The traditional single-targeted method has proved to be too limited to dereplicate such chemically complex situations, and it was not till the advent of a more holistic approach such as that provided by metabolomics, that these problems could be successfully revisited.

The use of "omics" platforms to gain insight into the chemical ecology of marine sponges has changed the perspectives for their exploration (Paul et al. 2019). In particular metabolomics, defined as the study of all metabolites present in an organism under specific conditions (Viant 2007), has been very useful for environmental studies (Bundy et al. 2008). Although targeted analyses have revealed significant information on changes in metabolites resulting from environmental factors (Noyer et al. 2011; Page et al. 2005), the use of metabolomics as an unbiased method enables a much broader approach that can lead to the identification of responses associated with a variety of previously unreported compounds.

Giant barrel sponges used in this study as model organisms, have been considered as the "redwood of the reef" due to their unparalleled longevity. For instance, a specimen found in Curaçao was estimated to be up to 2300 years old, positioning *Xestospongia* as one of the oldest animals alive (Van Soest et al. 2012). With the development of a growth model for *Xestospongia*, determining the age of these sponges has become relatively easy (McMurray et al. 2008). However, the effect that age might have on the lifecycle and physiological variation of the sponge is still far from being clearly understood. In particular, age could be expected to affect the metabolome, as has been reported in studies on most other organisms, such as several terrestrial plants (Lee et al. 2019; Yoon et al. 2019), humans (Jové et al. 2014; Yu et al. 2012) and marine sponge associated bacteria (Ng et al. 2013).

Biological development of marine sponges could also very likely be greatly affected by another factor: sea-depth. Marine habitats can vary substantially depending on their distance from the sea surface due to great variations in environmental conditions such as the availability of sunlight, pH, pressure, temperature and presence of predators. Therefore, giant barrel sponges, which can be found at depths ranging from less than 10 to beyond 100 m (Olson and Gao 2013; Van Soest et al. 2014), should unavoidably be influenced by these environmental variations. Several studies on the influence of depth on the microbiome of this sponge suggested that depth gradients could cause great changes in some of the microorganisms associated with the sponge (Morrow et al. 2016; Olson and Gao 2013; Villegas-Plazas et al. 2019). Considering that *X. muta* is an HMA sponge, changes in the microbiome could be expected to have a significant effect on the metabolome of the sponge. Notably, notwithstanding the importance of this relationship in the field of marine ecology, the studies on the effect of depth on the metabolic production of giant barrel sponges are still limited.

Recent studies have shown that the species previously known as *X. muta* is in fact, a species complex that can be separated into three distinct genetic groups (cryptic species) according to their mitochondrial DNA (Swierts et al., 2017). In the present study, samples belonging to

X. muta of two different genetic groups, different ages and depths were collected in Curacao and chemically profiled in a holistic manner to detect possible variations in their metabolome. For this, liquid chromatography coupled with mass spectrometry (LC-MS) was used as an analytical platform, combined with multivariate data analysis (MVDA) and a molecular networking workflow. The high sensitivity of LC-MS, one of its outstanding features, allowed the detection of metabolites present in very low concentrations (Goulitquer et al. 2012). After applying MVDA to the obtained data, a molecular networking workflow allowed it interpretation from a broader perspective.

2. Result and Discussion

Sponges are so far one of the most studied sources of marine natural products (MNP). Among these, giant barrel sponges have been most extensively studied since their widespread distribution and relatively large size facilitates their identification and collection. Their chemical composition includes sterols, brominated fatty acids and terpenoids (Zhou et al. 2010). However, the qualitative and quantitative composition of their metabolome was found to vary depending on environmental conditions (Bayona et al. 2020; Villegas-Plazas et al. 2019). The correlations between changes in the metabolome of giant barrel sponges and two possible driving factors of its variation (age and depth) were investigated using LC-MS, given its capacity to detect very low concentrations of most metabolites. Samples were collected in the Caribbean Sea off the coast of Curaçao. With the recent findings of Swierts and co-workers (Swierts et al. 2017) regarding the existence of three genetic groups among what had been previously considered to be only *X. muta* in mind, the collected samples were studied and classified accordingly. Thus, among the 69 samples collected, 41 belonged to genetic group 7 and 28 to genetic group 8 (Swierts et al. 2017). Considering that these genetic groups could have different responses to biotic or abiotic factors such as age or depth, the interaction between genetic groups and these two factors was also included in the study.

The effect of age, depth and genetic group on the chemical profile (LC-MS data) was firstly tested using a PERMANOVA analysis (Table 5.1). Among the factors evaluated, age and genetic group were found to significantly influence the variation of the sponge metabolome, whereas depth correlation was only marginally significant. The effect of the interaction between the studied factors was also calculated. In this case, no correspondence was found between metabolomic changes related to sample age and depth and/or genetic groups, that is, the changes in the metabolome caused by the aging processes did not appear to be affected by the genetic group nor by the depth of the sample collection. On the other hand, while the effect of depth was marginally significant when considering the whole sample set, it did seem

to vary according to the genetic group. This is supported by the fact that the interaction between the depth of collection and the genetic groups was found to be significant.

Table 5.1: Permutational multivariate analysis of variance (PERMANOVA) testing the effect of depth of sample collection, age, and their genetic group of the samples on the metabolome.

	F-value	R^2	P value
Age	$F_{(1,69)} = 2.45$	0.024	0.014*
Depth	$F_{(1,69)} = 2.04$	0.020	0.050
Genetic group	$F_{(1,69)} = 30.16$	0.300	<0.001*
Age*Depth	$F_{(1,69)} = 0.78$	0.007	0.593
Age* Genetic Group	$F_{(1,69)} = 0.64$	0.006	0.781
Depth* Genetic Group	$F_{(1,69)} = 2.57$	0.026	0.020*
Age*Depth*Genetic group	$F_{(1,69)} = 0.79$	0.008	0.596

*Significant factors confirmed by PERMANOVA.

In view of the significance of their interaction, the effect of depth on each genetic group was analyzed individually. The PERMANOVA results showed the metabolome of genetic group 7 to be largely affected by the depth ($F_{(1,45)} = 3.018$, $R^2 = 0.064$, $p = 0.003$) and this was also confirmed by the results of a canonical principal coordinate analysis (Figure 5.1). In contrast, the metabolome of genetic group 8 was not clearly correlated with the depth factor (PERMANOVA, $F_{(1,27)} = 1.457$, $R^2 = 0.053$, $p = 0.112$). This finding is in accordance with previous studies that identified depth as one factor that could influence the chemical composition of *X. muta* (Villegas-Plazas et al. 2019). However, the dependence of a particular genetic group on changes due to depth has never been studied.

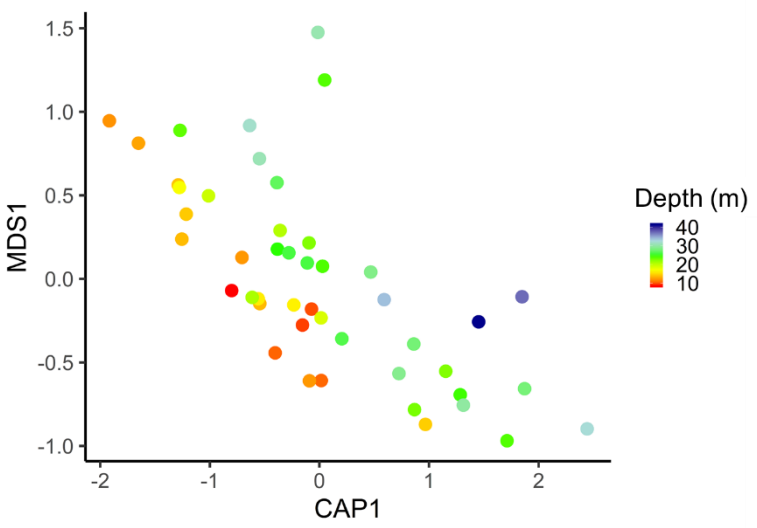


Figure 5.1: Canonical principal coordinate analysis between depth (m) factor and metabolome of the genetic group 7. Constrain canonical variable 1 (CAP 1) vs non constrain variable out of multidimensional scaling (MDS1).

The principal coordinate analysis (PCoA), an unsupervised multivariate data analysis, revealed differences between the metabolomes of the two phylogenetic groups (Figure 5.2a) Samples belonging to group 7 were found to have more unique features than group 8, consisting mostly of secondary metabolites (Figure 5.2b). The higher chemical richness of group 7 can be associated with its higher potential for producing metabolites implying that sponges belonging to group 7 are more likely to produce different secondary metabolites as a response to changes in environmental factors along the depth gradient. Furthermore, considering that *X. muta* is an HMA sponge, the variations observed in its metabolome could be associated with changes in the metabolism of the host itself (sponge), in the metabolism or the composition of its symbionts (microbiome) or changes in the metabolic interaction between the sponge and the symbionts. The alteration of the microbiome of giant barrel sponges in the Caribbean due to environmental factors related to differences in depth has been reported (Morrow et al. 2016; Villegas-Plazas et al. 2019) as was the influence of the phylogenetic group of the host on the composition of the microbiome of *Xestospongia* spp (Swierts 2019; Swierts et al. 2018b). From this perspective, the changes in the metabolome due to depth could also be attributed not only to different responses according to the phylogenetic group but also to differences in their specific symbionts

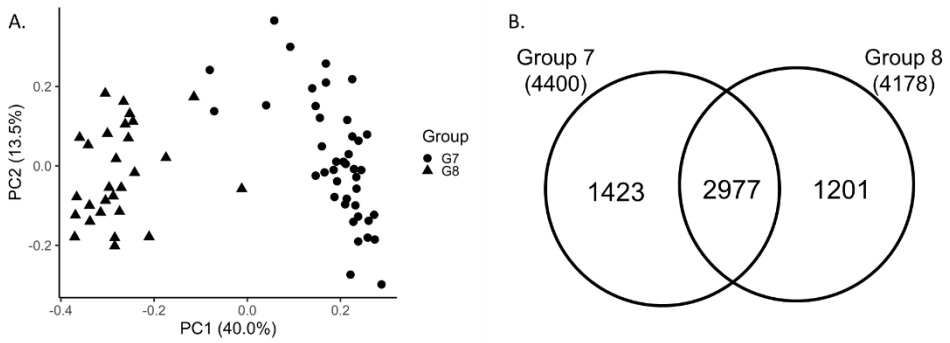


Figure 5.2: A: Principal coordinate analysis of 69 *Xestospongia* spp. samples using LC-MS data, principal coordinate 1 (PC1) Vs principal coordinate 2 (PC2). B: Venn-Diagram of the features obtained from the LC-MS data matrix in the intercept shows features present in both genetic groups.

The LC-MS data obtained from samples belonging to genetic group 7 was analyzed using the molecular networking workflow in the GNPS platform (Nothias et al. 2020; Wang et al. 2016) followed by *in silico* dereplication methods such as network annotation propagator (da Silva et al. 2018) and MolNetEnhancer (Ernst et al. 2019). This resulted in 40 molecular networks that contain 3 or more nodes, 20 of which could be classified into a molecular family according to the ClassyFire chemical ontology (Djoumbou Feunang et al. 2016). Most of the classified networks corresponded to lipids that include both fatty acids and glycerophospholipids or lipophilic compounds such as terpenoids and steroids while others were classified as polyketides and aromatic compounds such as benzoic acid derivates or quinones. To gain insight into the metabolites influenced by the depth gradient, the 15 features that most contributed in each direction of the first axis in the canonical principal coordinate analysis - positive (deep) and negative (shallow) - were selected for further investigation. As is shown in Figure 5.3, most of the features related to samples collected at shallow depths were grouped in one cluster and classified as glycerophospholipids. A closer look at the MS spectra of these compounds revealed that they all contain a phosphatidylcholine moiety. In addition, considering their molecular weight, it was deduced that they only contain one fatty acid bonded to the glycerol unit, implying that they belong specifically to the *lyso*-phosphatidylcholine lipids (*lyso*-PC) family. Lastly, the isotopic pattern showed the presence of bromine atoms in some of the compounds.

Table 5.2: ^{1}H-NMR and ^{13}C-NMR (600 MHz, CH_3OH-d_4) of compounds **1** and **2**

	Compound 1			Compound 2		
	δ_H (ppm)	mult, *J* (Hz)	δ_C (ppm)	δ_H (ppm)	mult, *J* (Hz)	δ_C (ppm)
1'	---	---	174.8		---	175.1
2'	2.48	t, 7.4	33.8	2.39	t, 7.5	34.2
3'	1.81	quint, 7.2	25.2	1.74	quint, 7.4	25.1
4'	2.35	td, 7.0, 1.9	19.4	2.35	qd,7.4, 1.1	30.5
5'	---	---	88.1	5.89	dt, 10.7, 7.5	142.6
6'	---	---	80.9	5.61	brd, 10.7	111.3
7'	5.45	dm, 15.8	111.3		---	85.6
8'	5.99	dt, 15.8, 7.1	144.0		---	93.7[a]
9'	2.08	m	33.6	5.66	dq, 15.8, 1.8	111.3
10'	1.40	m	29.5	6.10	dt, 15.8, 7.1	144.9
11'	1.44	m	29.1	2.17	qd, 7.0, 1.4	33.5
12'	2.04	m	32.0	1.53	m	29.1[b]
13'	6.07	td, 7.7, 1.4	137.4	1.53	m	28.9[b]
14'			114.8	2.30	m	19.8
15'	6.78	dm, 7.6	132.3		---	93.6[a]
16'	6.56	d, 7.6	113.4		---	78.3
17'	---	---	---	6.24	dt, 14.0, 2.3	119.2
18'	---	---	---	6.70	d, 14.0	118.0
1	4.19, 4.12	dd, 11.4, 4.5, dd, 11.4, 6.2	66.3	4.17, 4.12	dd, 11.4, 4.7, dd, 11.4, 6.1	66.2
2	3.98	m	69.8	3.98	m	69.8
3	3.90	m	67.8	3.90	m	67.8
1''	4.30	m	60.4	4.10	m	61.6
2''	3.64	m	67.5	3.24	tm, 4.8	50.9
N-Me	3.23	s	54.7	2.73	s	33.5

[a] Interchangeable carbons, [b] interchangeable carbons

The dereplication of the selected features led to the identification of one brominated polyacetylene fatty acid and a *lyso*-phosphatidylcholine lipid that contains palmitic acid as the acyl moiety. Further studies of the features of the glycerophospholipids network that showed different brominating patterns led to the identification of a previously unreported compound, 1-*O*-((7*E*,13*E*,15*Z*)-14,16-dibromohexadeca-7,13,15-trien-5-ynoyl)-sn-glycero-3-phosphocholine (Compound **1**). This compound was isolated as a white powder. The full elucidation of this compound was done using both NMR (Table 5.2) and HRMS. Its (+)-HRMS

spectrum showed the proton adduct $[M+H]^+$ ions at *m/z* 642.0835, 644.0836, and 646.0803 corresponding to the molecular formula, $C_{24}H_{38}Br_2NO_7P^+$ (calc. 644.0810, 642.0831, 646.0790); the presence of the two bromine atoms was deduced from the isotopic pattern. The ^{13}C-NMR spectrum showed the presence of a carbonyl carbon at δ_C 174.8 in addition to six sp^2 carbons and 2 *sp* carbons corresponding to three alkenes and one alkyne, respectively, indicating the presence of an acyl moiety. The carbons attached to heteroatoms (oxygen and nitrogen) in the region between δ_C 54 and 70 suggested the presence of one glycerol and one phosphatidylcholine moiety. The two bromine atoms observed in the mass spectra were found to be bonded to carbons involved in double bonds as shown by the shift in the chemical shift of these carbons. The full structure of compound **1** was determined using the COSY and HMBC spectra. The brominated fatty acid and phosphatidylcholine were found to be bonded to C1 and C3 of the glycerol molecule, respectively. Using the coupling constant, the stereochemistry of the double bonds was established as 7*E*,15*Z.* The stereochemistry of the double bond in C'-13 was determined to be *E* due to the correlation observed in the NOESY between protons in position 12' and 15'.

Lyso-phosphatidylcholines *(lyso-PC)* account for only 0.3 to 4% of the total lipid content in sponge extracts (Genin et al. 2008), coinciding with studies conducted in animal cell membranes (Torkhovskaya et al. 2007). Lysophospholipids have been recognized as playing a role beyond being part of the membrane structure as molecular signaling molecules (Birgbauer and Chun 2006; D'Arrigo and Servi 2010; Torkhovskaya et al. 2007). These types of molecules have been related to different mechanism of cell proliferation , such as the induction of growing factors synthesis or increasing the sensitivity of cells to growth factors (Torkhovskaya et al. 2007). In addition, similar molecules have been reported to be involved in the self-recognition and activation of immune systems though pro-inflammation signaling in sponges and corals (Müller and Müller 2003; Quinn et al. 2016). Therefore, the increase in the production of *lyso*-PC in shallow depth sponges could be related to a defense mechanism developed as a response to higher predatory stress in shallower waters (Chanas and Pawlik 1997). A similar phenomenon has been observed in sponges of the genus *Oscarella* that displayed a seasonal variation of two lysophospholipids with an increased production that coincided with the embryogenesis and the larval development period (Ivanisevic et al. 2011). This shows that these types of molecules can perform a variety of roles in the development of the sponge. However, the variation in their structures makes it difficult to establish specific roles for each of them. Lastly, even though lysophospholipids in general, and *lyso*-PC in particular, are universally present among metazoans, some chemical features in the acyl chain present in compounds from *Xestospongia*, such as bromine atoms and acetylenic groups, are

characteristic of these specific organisms (Zhou et al. 2010; Zhou et al. 2015). The fatty acids from *Xestospongia* have been reported to have a wide range of biological activities, suggesting that it is not only the lysophospholipid that has a biological role in this case, but the free fatty acid which could *per se* play an additional role in the sponge.

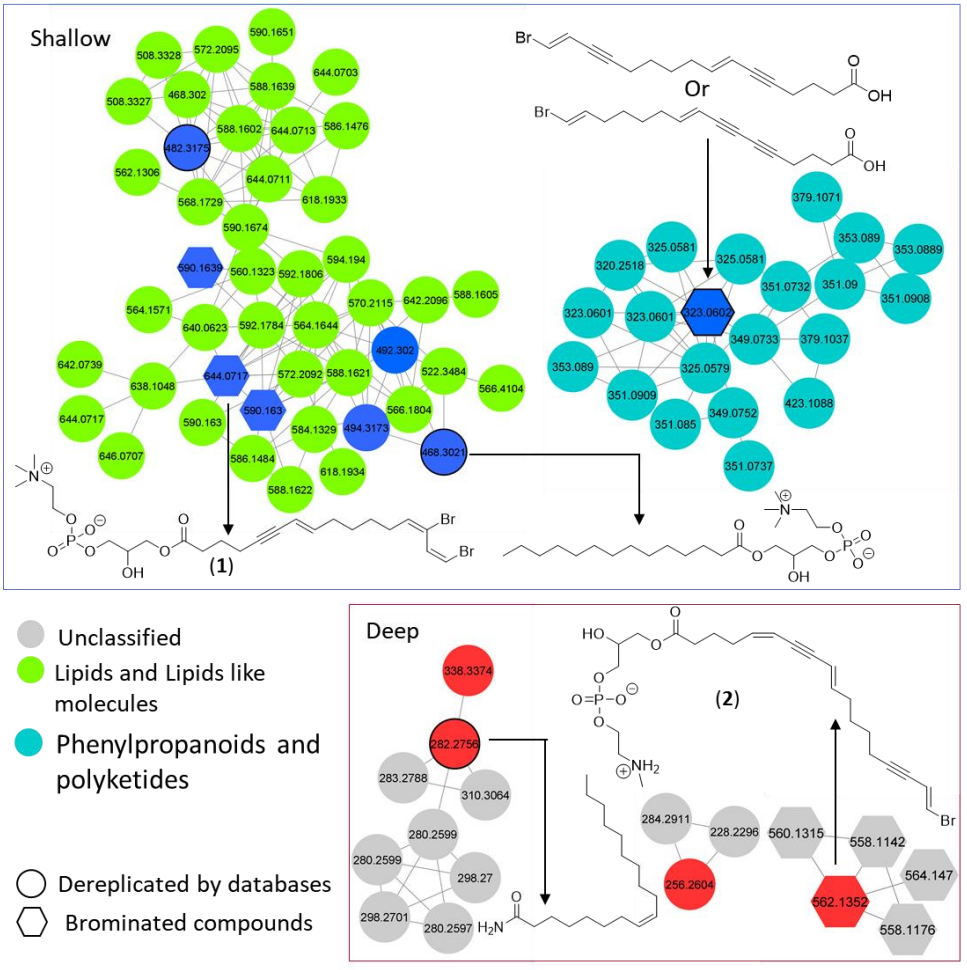


Figure 5.3: Molecular network showing the clusters of the features identified to be changing due to depth. The nodes corresponding to important features obtained from the canonical analysis of principal coordinate are colored in blue (shallow) and red (deep)

For samples collected at increasing depth, the features that were identified as important were not grouped in one single cluster. Instead, the significant compounds were found in small clusters within the molecular networking analysis, as shown in Figure 5.3. This indicates that the environmental changes along the depth gradient do not cause a change in the production

of a specific group of molecules, but rather, that the changes observed in the metabolome are compound specific. One of the reasons why the environmental changes along the depth gradient produce a response involving specific compounds could be that not all environmental and biotic factor variations are directly related to depth. While abiotic factors such as pressure and temperature are related to depth, other factors such as predatory stress can vary along the depth gradient. This can cause heterogenicity in the compounds whose production is altered at larger depths. Among the compounds that were identified as associated with greater depths it was possible to isolate another previously unreported compound, 1-*O*-((5*Z*,9*E*,17*E*)-18-bromooctadeca-5,9,17-trien-7,15-diynoyl)-sn-glycero-3-*N*-Methylethanolamine (compound **2**).

Compound **2** was also isolated as a white powder. Its (+)-QTOF-ESI-MS spectrum showed the proton adduct $[M+H]^+$ ions at *m/z* 560.1394 and 562.1379, with a relative intensity of1:1, indicating the presence of one bromine atom in the molecule. Consequently, the molecular formula was established as $C_{24}H_{36}BrNO_7P$ (calc. 560.1413 and 562.1392). Its NMR spectroscopic data (Table 5.2) was used to characterize the structure of this compound. The carbonyl carbon at δ_C 175.1 together with the *sp* and sp^2 carbons confirmed the presence of an acyl chain in the molecule. In addition, the absence of a terminal methyl or methylene group together with the unusual chemical shift of the carbons in terminal double bond at δ_c 119.2 and 118.0 confirmed the presence of a bromine atom in this position. The full structure of compound **2** was determined using the COSY and HMBC. The compound was determined to be formed by three moieties, a brominated unsaturated fatty acid, a glycerol molecule and *N*-methylethanolamine**.** The brominated fatty acid and *N*-methylethanolamine are bonded to C1 and C3 of the glycerol molecule, respectively. Using the coupling constant, the stereochemistry of the double bounds was established as 5*Z*,11*E*,17*E*.

Similar to the compounds found in shallower waters, compound **2** also belongs to the lysophospholipids family. However, the presence of a *N*-methylethanolamine moiety is very unusual. Although ethanolamine and choline are moieties commonly found in animal cells, *N*-methylethanolamine has been mostly reported to be present in the membrane of several by microorganisms (Dahal and Kim 2017; Goldfine and Ellis 1963; Schubotz et al. 2011). This suggests that compound **2** is produced totally or partially by an associated microorganism. As was discussed previously, the *Xestospongia* microbiome experiences changes due to depth. The fact that one of the metabolites associated with changes in depth contains a moiety produced by microorganisms is an example of the way in which changes in the microbiome can be reflected in changes in the metabolome of the holobiont.

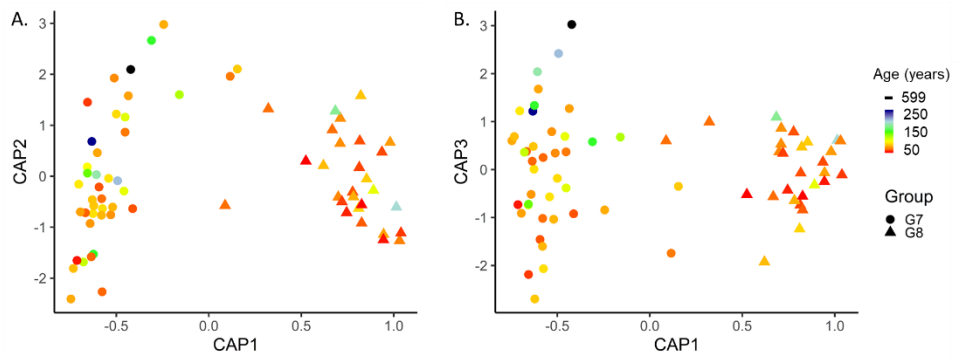


Figure 5.4: Canonical principal coordinate analysis using age (years), depth (m) and genetic group as factors for samples of genetic group 7 and 8. A: CAP 1 vs 2 and B: CAP 1 vs 3.

In the second part of the study the effect of age on the metabolome of the sponge was investigated (Figure 5.4). The metabolome of sponges proved to be affected by age, providing an important insight into their biological and chemical processes. From a biological perspective, this could imply that in sponges, the biological stage of development of a sponge is reflected by the metabolome as observed in other organisms (Yoon et al. 2019; Yu et al. 2012). For sponges in general, and particularly in the case of giant barrel sponges, the finding that the metabolome can be an indicator for the development of the sponge is important, as very little is known about the biological development of these sponges. Moreover, these findings suggest that external factors could also affect the metabolome, as older sponges have had to adapt over time to more variations in the conditions of the water column, for example in the microbiome, the temperature or type and amount of nutrients. Therefore, it is possible that over the years, older sponges have, for example, recruited specific symbionts into their metabolome that afforded an adaptive advantage (Ribes et al. 2016; Turon et al. 2018).

The fact that the interaction between age and depth does not affect the metabolome is in agreement with the lack of any influence of depth on growth rate of *X. muta* (McMurray et al. 2008). Similarly, the interaction between age and genetic groups showed no significant effect on the metabolome, possibly because growth rates have been calculated for *X. muta,* but not for the individual genetic groups. Each genetic group could have different growth rates, as suggested by previous reports of differences in growth rate of *X. muta* and its sister species *X. testudinaria* (McGrath et al. 2018).

Considering that the effect of age was reflected in the metabolome, the second aim of this research was to identify those metabolites most affected by this process. The canonical principal coordinated analysis revealed the effect of age mainly CAP 2 and 3 as can be observed

in Figure 5.4. The features that were found to contribute the most to this axis were selected as characteristic metabolites associated with age. Subsequently, a molecular network was built using the GNPS platform followed by *in silico* tools to increase the annotation of the features. This showed that most of the clusters in which the nodes corresponding to the compounds related to age were located were not classified. However, a closer look into the MS spectra of the selected compounds allowed us to draw some conclusions about the chemical structures of these compounds.

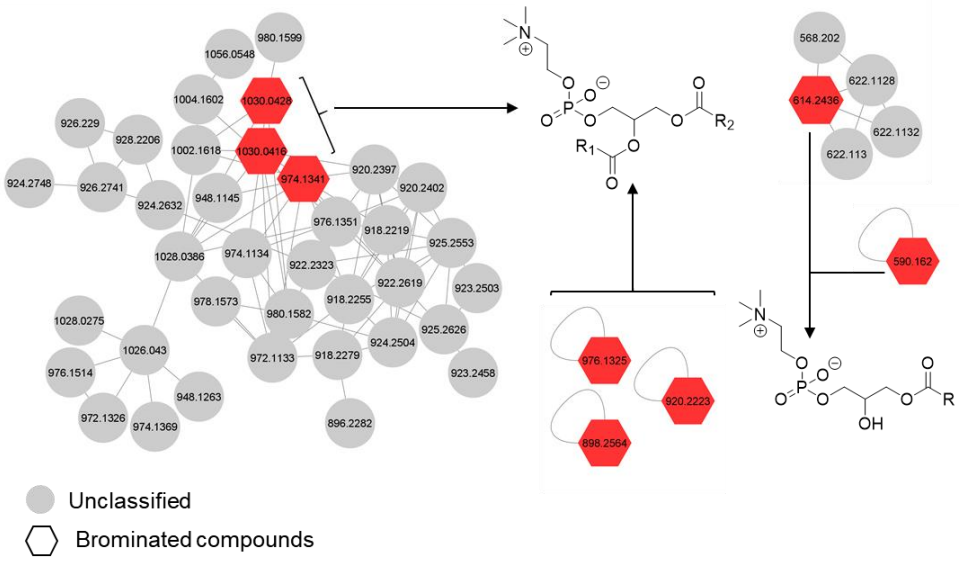


Figure 5.5: Molecular network showing the clusters of the features identified to be changing depending on the age of the sponge. The nodes corresponding to features related to older sponges obtained from the canonical analysis of principal coordinate are colored in and red.

As shown in Figure 5.5, most of the compounds that were identified to be related to older sponges displayed *m/z* values between 800 and 1000. This, together with their fragmentation pattern, indicated that the structure of these compounds corresponded to glycerophospholipids with one phosphatidylcholine and two acyl chains joined to the glycerol moiety. Further, the isotopic pattern of these compounds revealed the presence of 1 to 5 bromine atoms. A possible explanation for their presence could be that were cellular division increased in older sponges these compounds could also be increased as they are a constitutive element in membranes. Considering that it has been reported that the growth rate of giant barrel sponges decreases with the age of the sponge (McMurray et al. 2008), it is possible that the increment in the cellular division associated with these membrane lipids is related to higher cell turnover rates rather than a growing state (Alexander et al. 2014). Cell turnover

rates have proven to be particularly fast in sponges compared to other animals and it is believed to be a strategy to prevent cellular damage (De Goeij et al. 2009). Another alternative could be that these molecules act as a sort of reservoir of the bioactive polybrominated fatty acids that are bonded to the glycerol. Polybrominated fatty acids have been reported to exhibit a wide range of biological activities, such as antiviral, antimicrobial and cytotoxic activity. Therefore, besides a mere functional role, the glycerophospholipids could act as a bank of these fatty acids that are involved in defense mechanisms and can be released in the event of situations that require their participation. However, further research in this field is needed in order to prove this hypothesis.

3. Conclusions

The metabolome of giant barrel sponges belonging to two genetic groups revealed their different response to changes along a depth gradient. Genetic group 7 showed a higher content of lysophosphatidylcholine lipids at shallower depths. These compounds have been reported to play several biological roles in animal cells indicating that environmental conditions such as light, temperature and predatory stress present in shallower waters might trigger the production of this family of compounds. Samples collected at greater depths did not reveal the overexpression of any specific group of compounds, but one of the compounds that was found to be related to these conditions contains a *N*- methylethoanolamine unit that has been reported only in bacteria, suggesting that it is the result of the interaction between the metabolism of the sponges and their associated microorganisms. Apart from depth, age was also found to affect the metabolome of giant barrel sponges with older specimens having increased amounts of glycerophospholipids. This is an important contribution as it is the first time that age has been evaluated as factor for change in the metabolome of these sponges and could be indicative of an increase in cell replication in older sponges.

4. Materials and Methods

4.1 Sample collection

Giant barrel sponges (*Xestospongia* spp) were collected by SCUBA diving in Curaçao from multiple locations with different depths around the island (appendix 2 Table S1). Collected sponge samples were immediately stored on site in 98% ethanol (w/w) at -20 °C. The samples were identified by DNA sequencing, using the I3-M11 and partition of the CO1 mitochondrial gene according to Swierts and co-workers (Swierts et al. 2013). The DNA analysis allowed the samples to be identified as belonging to three main genetic classes for giant barrel sponges in the Caribbean (groups 7 and 8) (Swierts et al. 2017). Although a few samples could be

identified as belonging to group 9, these were not included in the study as there was not a sufficient quantity to carry out statistical analysis. The age of the specimens was determined through on-field measurements according to previous work (McMurray et al., 2008). The depth at which each specimen was collected was also recorded and that value was used for the analyses.

4.2 Sample preparation for metabolomic analysis

The *Xestospongia* spp. samples were ground and extracted with ethanol and sonicated for 20 min three times. The 3 extracts were combined and desalted on 500mg C-18 SPE cartridges of 45 µm particle size (Supelco Supelclean, Bellefonte, PA, USA). For this, 50 mg of each extract were loaded onto the cartridge and eluted with water (F1), methanol (F2) and dichloromethane/methanol 1:1 (F3). The methanol fractions were analyzed by LC-MS.

4.3 Liquid chromatography- Mass spectrometry (LC-MS) analysis

The methanol fractions obtained were taken to dryness with a Centrivap vacuum concentrator (Labconco, Kansas City, MO, USA) and the obtained residue was redissolved in 50% methanol (v/v) to a final concentration of 1 mg/mL. Samples were injected in randomized sequence within each genetic group. The analysis was carried out using an UHPLC-DAD-ESI-MS system consisting of a UPLC Acquity I-Class (Waters, Milford, MA, USA) hyphenated to a Bruker Impact HD MS spectrometer (Bremen, Germany) with electrospray ionization (ESI). The separation was performed on a Waters C18 (2.1 x 100 mm, 2.1 µm) column and eluted with gradient of 0.1% formic acid in water (A) and 0.1% formic acid in acetonitrile (B) of 10% to 100% B in 30 min, and 100% B for 5 min; 100% to 10% B in 3 minutes and 2 minutes equilibration at 10% B. The flow rate was 0.3 mL/min, column temperature was 40 °C and injection volume was 2 µL. The mass spectrometer parameters were set as follows: 1.5 bar of nebulizer gas, 6.0 L/min of drying gas, drying gas was 350 °C, and capillary voltage was 4000 V. The mass spectrometer was operated in positive mode in a range of 50 to 1200 m/z.

4.4 Statistical analysis

Data files obtained from the LC–MS analyses were converted to mzXML format using Brucker Daltonics DataAnalysis (version 4.1, Bremen, Germany). LC-MS data were processed using MZMine2 (Pluskal et al. 2010). To build the feature matrix, the mass detection was performed using centroid data. The noise level was set at 10000 for MS and 100 for MS/MS. The chromatograms were built using the ADAP chromatogram builder (Myers et al. 2017) with a minimum number of scans of 3, group intensity threshold 1000, minimum highest intensity 10000 and m/z tolerance of 0.05. Chromatograms were deconvoluted using a baseline cutoff

algorithm with the following parameters: minimum peak height of 10000, peak width range of 0.02−1.00 min, and baseline level of 1000. Chromatograms were deisotoped using an isotopic peak grouper algorithm with an *m/z* tolerance of 0.05 and retention time tolerance of 0.1 min. The features of each sample were aligned using a join alignment algorithm with the following parameters: 0.05 *m/z* tolerance and 0.1 retention time tolerance. Using these parameters, three matrixes were created, one including all samples and one matrix with samples of each genetic group.

All the statistical analyses were performed using RStudio 1.2.1335 (RStudio team,2018). The PERMANOVA (Permutational analysis of variance) analyses were performed using adonis() function in the 'vegan' package (Oksanen J., 2018). This was applied to different age, sea-depth and genetic group samples, for which 9999 permutation and Bray-Curtis as dissimilarity measurement were used to investigate mutual interactions between metabolites and the factors. Canonical principal coordinate analyses were performed for the data matrixes obtained from LC-MS analysis and the factors found to be significant in the PERMANOVA capscale() function with Bray-Courtois dissimilarity as a distance parameter was used. Variation in the metabolome of genetic groups was assessed with a principal coordinate analysis (PCoA) using the cmdscale() function and the Bray-Courtois dissimilarity matrix as an input.

4.5 Molecular networking and dereplication

Feature base molecular networking (FBMN) workflow (Nothias et al. 2020) was carried out using the GNPS online platform (http://gnps.ucsd.edu) (Wang et al. 2016). Quantitative data matrix and MS/MS obtained from MZmine2 are shown in section 4.4. Files were uploaded to GNPS platform and the ions in the region between +/- 17 Da around the precursor *m/z* were removed to filter the data. Additionally, only the top six fragment ions of the MS/MS data were considered for further use to reduce data size. The mass tolerance for both precursor and the MS/MS fragment ions were set to 0.03 Da. With these tolerance and filter parameters, a molecular network was constructed, for which minimum cosine value for the edges and the matched fragments were set at 0.7 and 6, respectively. In addition, edges between two nodes were only included when both nodes were in common to their respective top 10 matching nodes. The maximum size of a molecular family was set to 100. The generated spectra in the network were then dereplicated using GNPS spectral libraries (Horai et al. 2010; Wang et al. 2016), and visualized using Cytoscape software (Shannon et al. 2003).To improve the annotation of the features in the MNs Network Annotation Propagation (NAP) an *in silico* tool was used (da Silva et al. 2018). The following NAP parameters were employed for the

annotation: exact mass error for database search, 10 ppm; $[M + H]^+$ as adduct type:, cosine = 0.7 to select inside a cluster; 10 maximum candidate structures in the graph. The result from both FBMN and NAP were introduced in MolNetEnhancer workflow (https://ccms-ucsd.github.io/GNPSDocumentation/molnetenhancer/) (Ernst et al. 2019)to classify the features into chemical classes according to ClassyFire chemical ontology (Djoumbou Feunang et al. 2016).The results of the molecular networks for the matrixes with all the samples and the samples from genetic group 7 can be found in https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=cd2ffac11ce24c32ad6b8f84ce85240e and https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=48bd32ff37cc45a98575a78121bb33d2, respectively.

4.6 Isolation and identification of compounds

Two compounds were isolated from two of the ethanol extracts of the sponge samples collected at different depths. The crude extracts were fractionated using 20-mL SPE LC-18 cartridges (Supelco) eluted successively with 100 mL each of 100% water, 80% 60%, 40% 20% methanol/water mixtures, 100 % methanol and methanol-dichloromethane (1:1, v/v). The SPE fractionation yielded seven fractions labelled FS1-FS7 and FD1-FD7 for sample1 and 2, respectively. Using LC-MS profiling, the target compounds, corresponding to the top 15 most contributing features (negative and positive) to the CAP 1 from the canonical principal coordinate analysis (appendix 2 Table S2), were identified in FS5 and FD5. Successive semi-preparative HPLC using an Agilent 1200 series system (Santa Clara, CA, USA) and a Luna 5 µm, C-18, 250 mm x 10 mm column (Phenomenex, Torrance, CA, USA) eluted at a flow rate of 3.5 mL/min allowed the isolation of compound **1** from FS5 and compound **2** from FD5.

The structure of these compounds was elucidated by NMR analyses at 25°C with an AV-600 MHz NMR spectrometer (Bruker, Karlsruhe, Germany), operating at the ^{1}H-NMR frequency of 600.13 MHz, equipped with a TCI cryoprobe and Z gradient system. Deuterated methanol was used as the internal lock. Their 1-D ^{1}H-NMR and ^{13}C attached proton test (APT) spectra were recorded initially and 2-D ^{1}H-NMR techniques, COSY, HSQC, and HMBC were used to confirm the connectivity of the atoms in compound **1** and **2**

Appendix 2

Table S1: Information of depth, age, genetic group and geographical location of *Xestospongia* the samples collected in Curaçao

Code	Depth	Age	Genetic group
LMB360	28.7	28	Group 7
LMB362	16.2	82	Group 7
LMB366	9.7	18	Group 7
LMB367	8.6	37	Group 7
LMB368	23.0	60	Group 7
LMB369	26.0	25	Group 7
LMB370	29.8	33	Group 7
LMB374	27.9	40	Group 7
LMB375	28.0	54	Group 7
LMB376	20.6	47	Group 7
LMB377	13.8	141	Group 7
LMB379	14.6	599	Group 7
LMB380	23.7	9	Group 7
LMB381	13.7	36	Group 7
LMB382	30.8	47	Group 7
LMB383	22.6	125	Group 7
LMB384	9.4	18	Group 7
LMB385	12.1	43	Group 7
LMB386	12.3	62	Group 7
LMB387	24.1	63	Group 7
LMB388	24.7	52	Group 7
LMB389	37.3	28	Group 7
LMB390	31.8	40	Group 7
LMB391	32.4	44	Group 7
LMB393	24.0	53	Group 7
LMB394	12.8	23	Group 7
LMB396	26.3	42	Group 7
LMB397	21.0	209	Group 7
LMB398	22.6	13	Group 7
LMB399	16.0	21	Group 7
LMB400	12.3	65	Group 7
LMB402	10.5	266	Group 7
LMB403	14.4	50	Group 7
LMB405	10.4	179	Group 7
LMB408	18.8	72	Group 7
LMB409	24.4	104	Group 7
LMB410	24.0	141	Group 7
LMB411	17.1	96	Group 7
LMB412	34.0	85	Group 7
LMB414	26.9	61	Group 7
LMB415	14.2	23	Group 7
LMB416	14.8	10	Group 8
LMB418	24.1	8	Group 8
LMB419	24.1	15	Group 8
LMB422	19.8	25	Group 8
LMB423	15.4	17	Group 8
LMB425	20.7	25	Group 8
LMB426	24.0	25	Group 8
LMB427	28.7	171	Group 8
LMB428	32.2	45	Group 8
LMB429	22.7	35	Group 8
LMB430	18.8	9	Group 8
LMB432	34.2	49	Group 8
LMB433	14.8	18	Group 8
LMB434	9.1	32	Group 8
LMB436	18.1	34	Group 8
LMB438	9.4	195	Group 8
LMB440	17.3	41	Group 8
LMB441	33.8	20	Group 8
LMB442	15.2	16	Group 8
LMB444	22.3	52	Group 8
LMB445	15.9	13	Group 8
LMB446	17.7	82	Group 8
LMB448	17.2	58	Group 8
LMB452	21.0	14	Group 8
LMB453	10.0	30	Group 8
LMB455	19.4	8	Group 8
LMB456	16.1	38	Group 8
LMB457	19.5	55	Group 8

Table S2: Top 15 most contributing features (negative and positive) to the CAP 1 from the canonical principal coordinate analysis

m/z value	Retention time (min)	CAP 1 value
590.1722	14.9	0.066929
1030.038	24.0	0.072301
576.4033	19.3	0.074658
664.454	19.2	0.077708
616.4698	30.2	0.077844
892.1917	25.1	0.084319
562.1453	15.0	0.085022
620.4288	19.3	0.08547
751.5224	33.9	0.086833
338.3374	29.3	0.118655
896.218	25.5	0.121921
793.4967	30.8	0.126926
256.2603	23.5	0.147761
593.3947	30.2	0.182476
282.2756	24.3	0.707957
590.1629	15.4	-0.17504
496.3693	19.5	-0.16223
644.0717	14.1	-0.146
494.3173	16.3	-0.13164
468.3021	14.5	-0.11018
482.3175	15.0	-0.08795
482.3174	16.5	-0.07844
323.0601	19.8	-0.07703
590.1638	14.8	-0.07477
492.3019	15.2	-0.06393
663.0504	15.5	-0.06306
510.3483	18.6	-0.05543
721.4981	33.6	-0.05309
482.3174	15.5	-0.05242

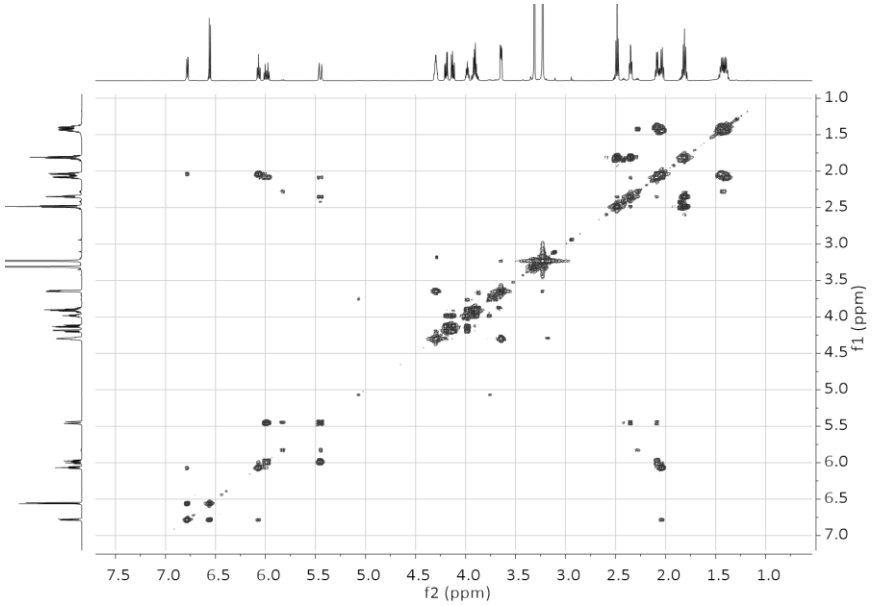


Figure S1: ^{1}H-^{1}H Correlation spectroscopy (COSY) spectrum (CH_3OH-d_4, 600 MHz) of Compound (**1**).

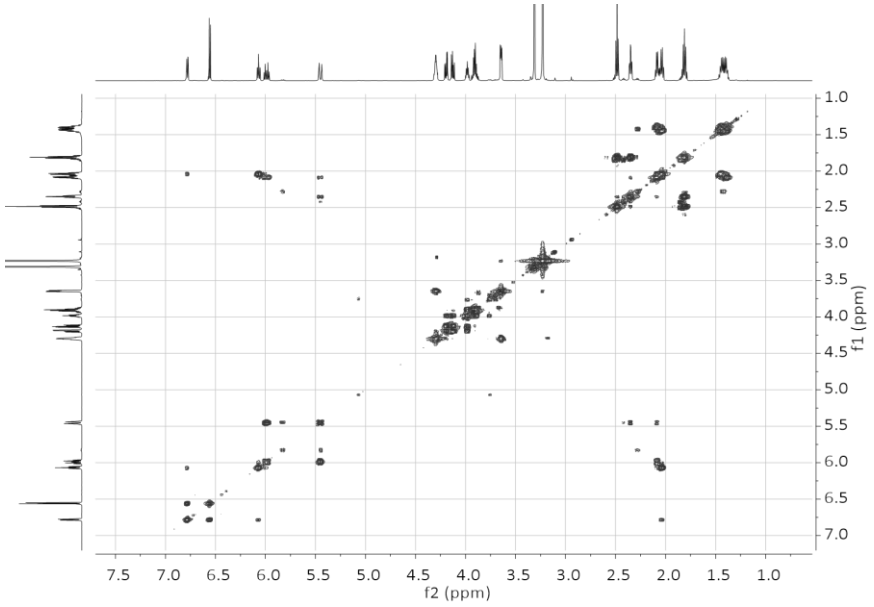


Figure S2: Heteronuclear multiple bond correlation (HMBC) spectrum (CH_3OH-d_4, 600 MHz) of Compound (**1**).

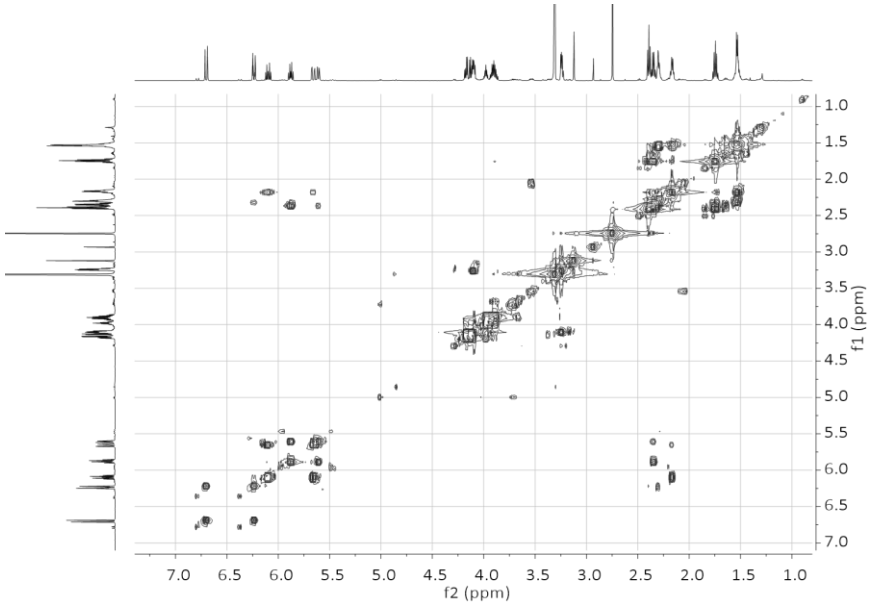


Figure S3: ^{1}H-^{1}H Correlation spectroscopy (COSY) spectrum (CH_3OH-d_4, 600 MHz) of Compound (**2**).

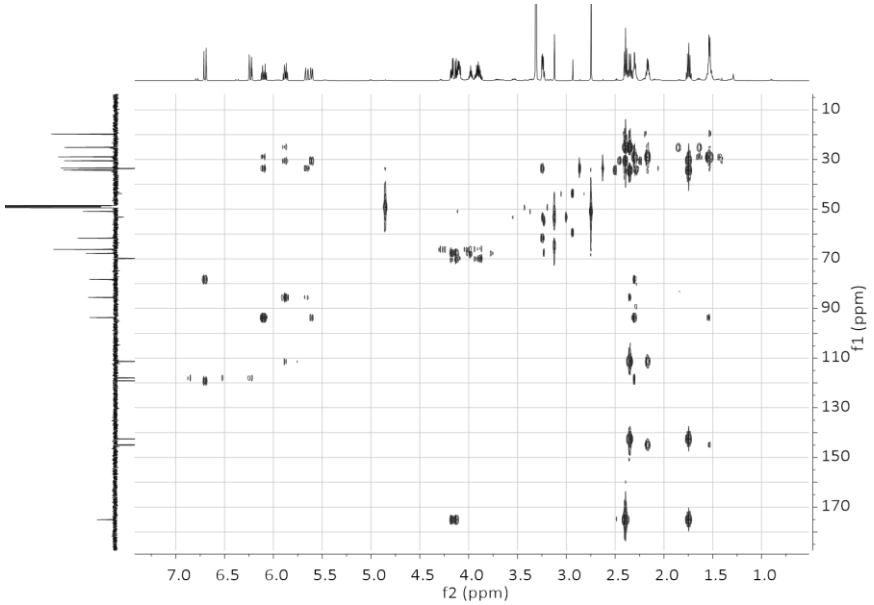


Figure S4: Heteronuclear multiple bond correlation (HMBC) spectrum (CH_3OH-d_4, 600 MHz) of Compound (**2**).