Figure S1. The phylogenetic tree of ascidians using 18S rRNA gene sequences. Sequences obtained from this study are labeled with location (CA, FL, PNG, VN) and sample code. Other sequences were obtained from NCBI with the following accession numbers: *Megalocercus huxleyi* FM244868.1, *Herdmania litoralis* AB564300.1, *Pyura haustor* AY903926.1, *Pyura spinifera* JF961826.1, *Ecteinascidia turbinata* FM244848.1, *Phallusia fumigata* FM244844.1, *Ascidia ceratodes* L12378.2, *Aplidium pliciferum* AB211067.1, *Eudistoma gilboviride* AB211069.1, *Cystodytes* sp_TG-2008 FM244842.1, *Lissoclinum badium* AB211078.1, *Lissoclinum bistratum* AB211084.1, *Lissoclinum patella* AB211085.1, *Didemnum* sp DidSB AB211073.1, *Tridemnum paracyclops* AB211077.1, *Didemnum molle* AB211071.1, *Didemnum* sp DidSA/57 AB211072.1

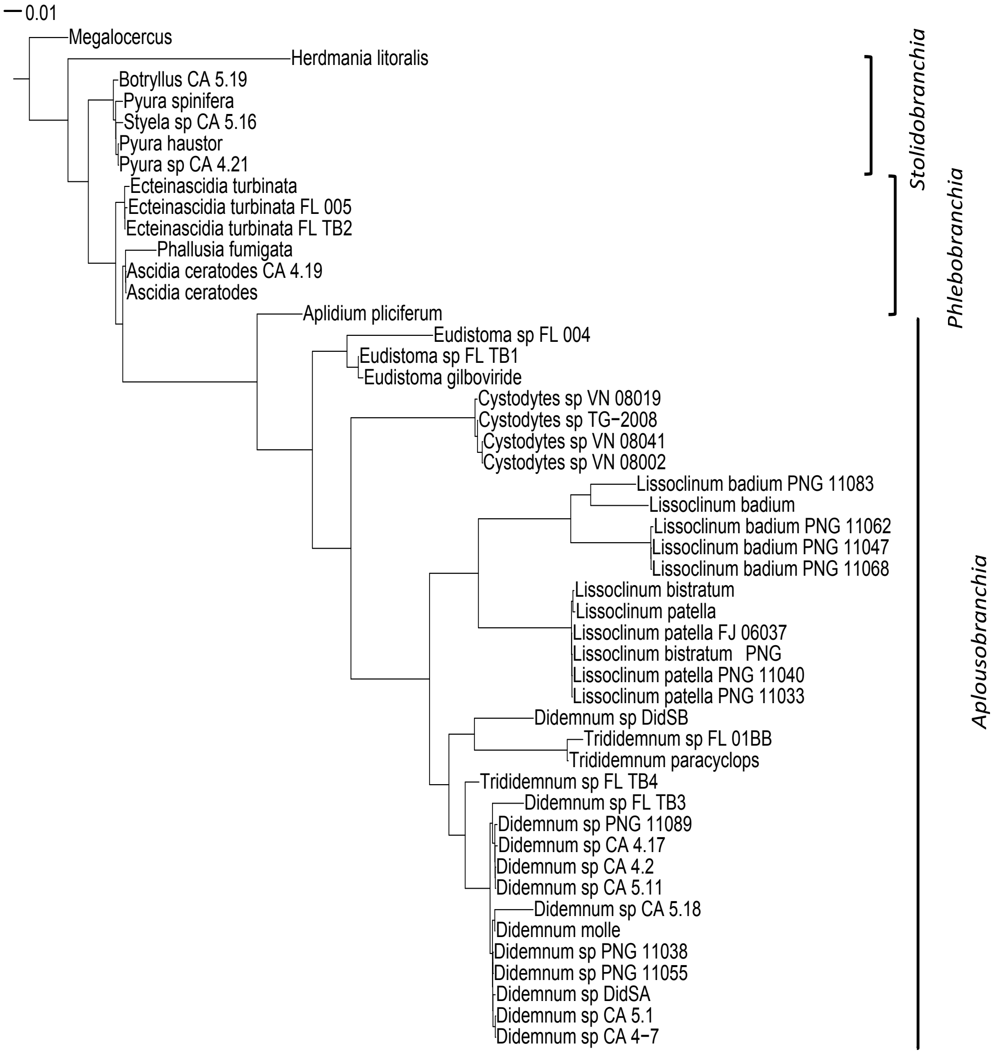


Figure S2. Abundance and order level composition of Deltaproteobacteria (A) and Gammaproteobacteria (B) showing ascidian species specificity of the most abundant groups. (C) Phylogeny of the top ten most abundant OTUs in *L. patella* and *E. turbinata* samples, arrows indicate known natural products producers. Bar plots and phylogenetic trees were generated using Phyloseq R package.

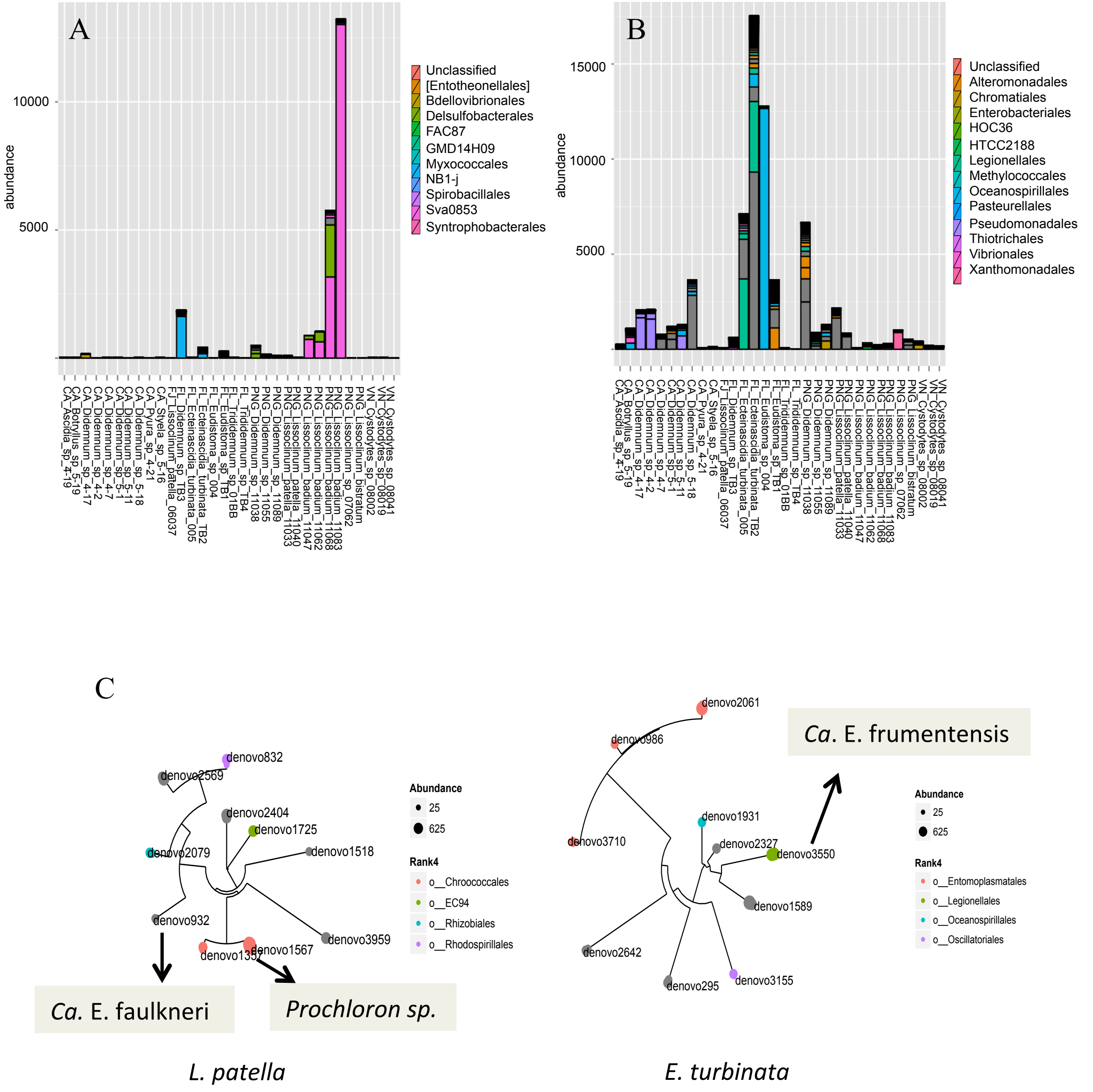


Figure S3. Effects of seasonal variation on the composition of ascidian microbiomes. The graph shows that the major phyla (filtered to top 20 percent) are maintained in ascidians collected from fall and spring of 2011 (x-axis), while the abundance of different species of bacteria varies.

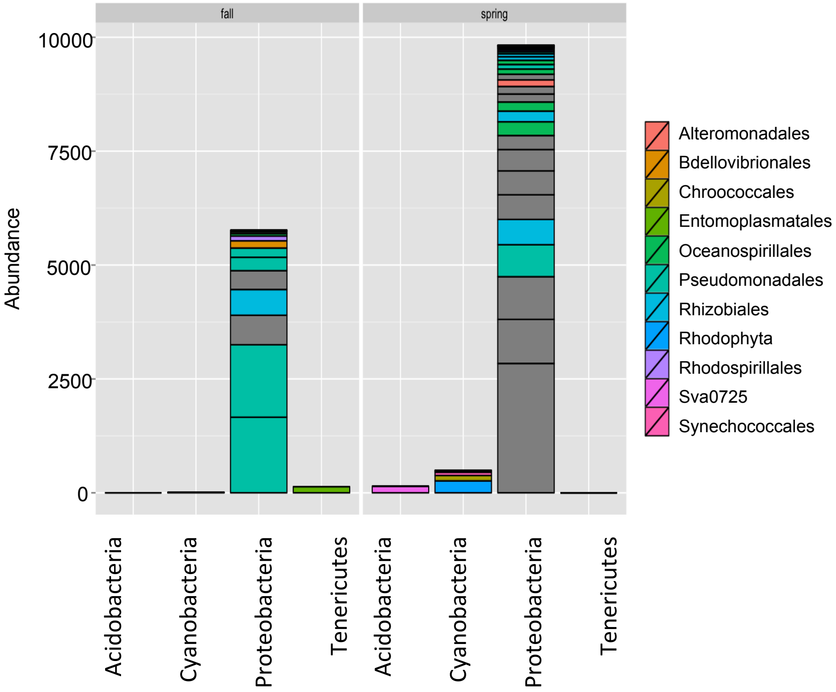


Figure S4. UniFrac-based 3D PCoA plots of all samples showing the clustering of samples according to (A) type of organisms (black: ascidians, red: mollusks) and (B) location (black: CA, green: FL, blue: PNG, light blue: VN, red: FJ).

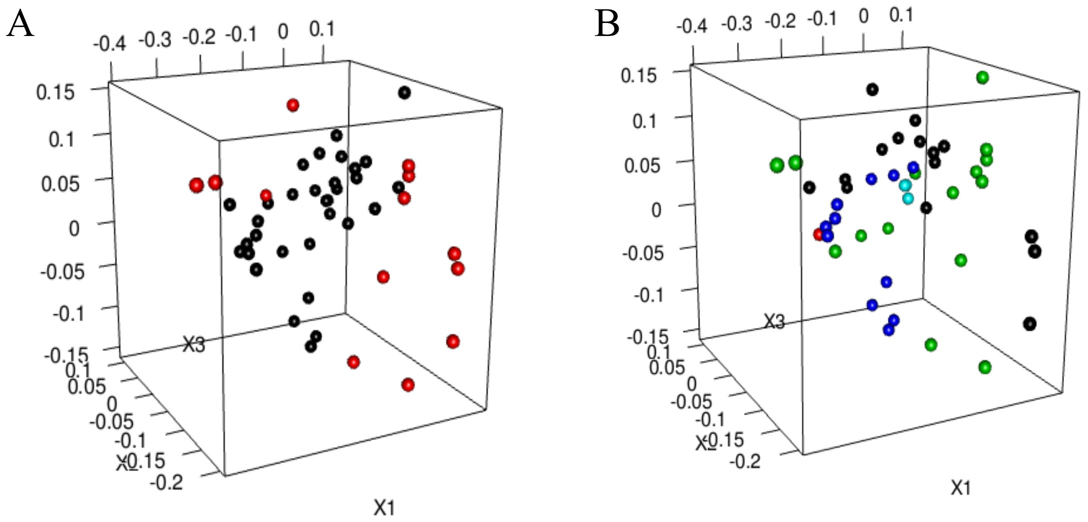


Figure S5. PCA biplots of (A) ascidian organic extracts showing correlation of samples from the same location (red arrows) with specific extract components (black print indicating retention time and m/z (B) resulting PCA biplot of assigned lipids after assignment of the LCMS data through Lipid Maps database showing the same pattern as A.

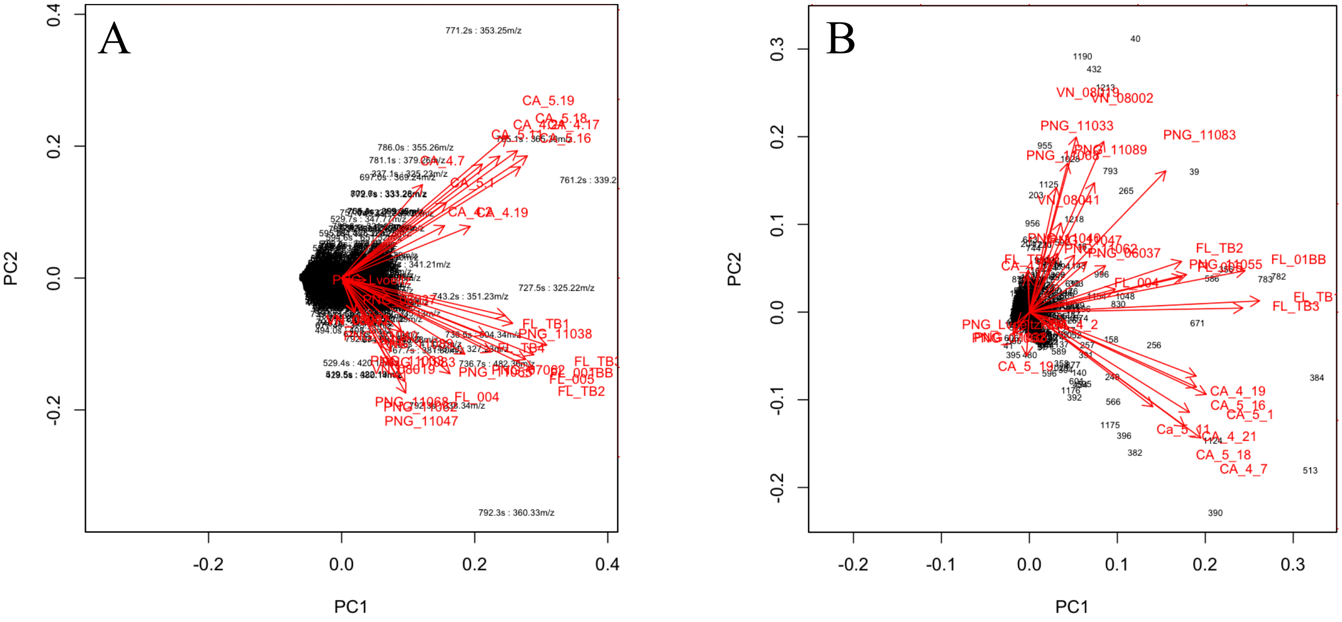
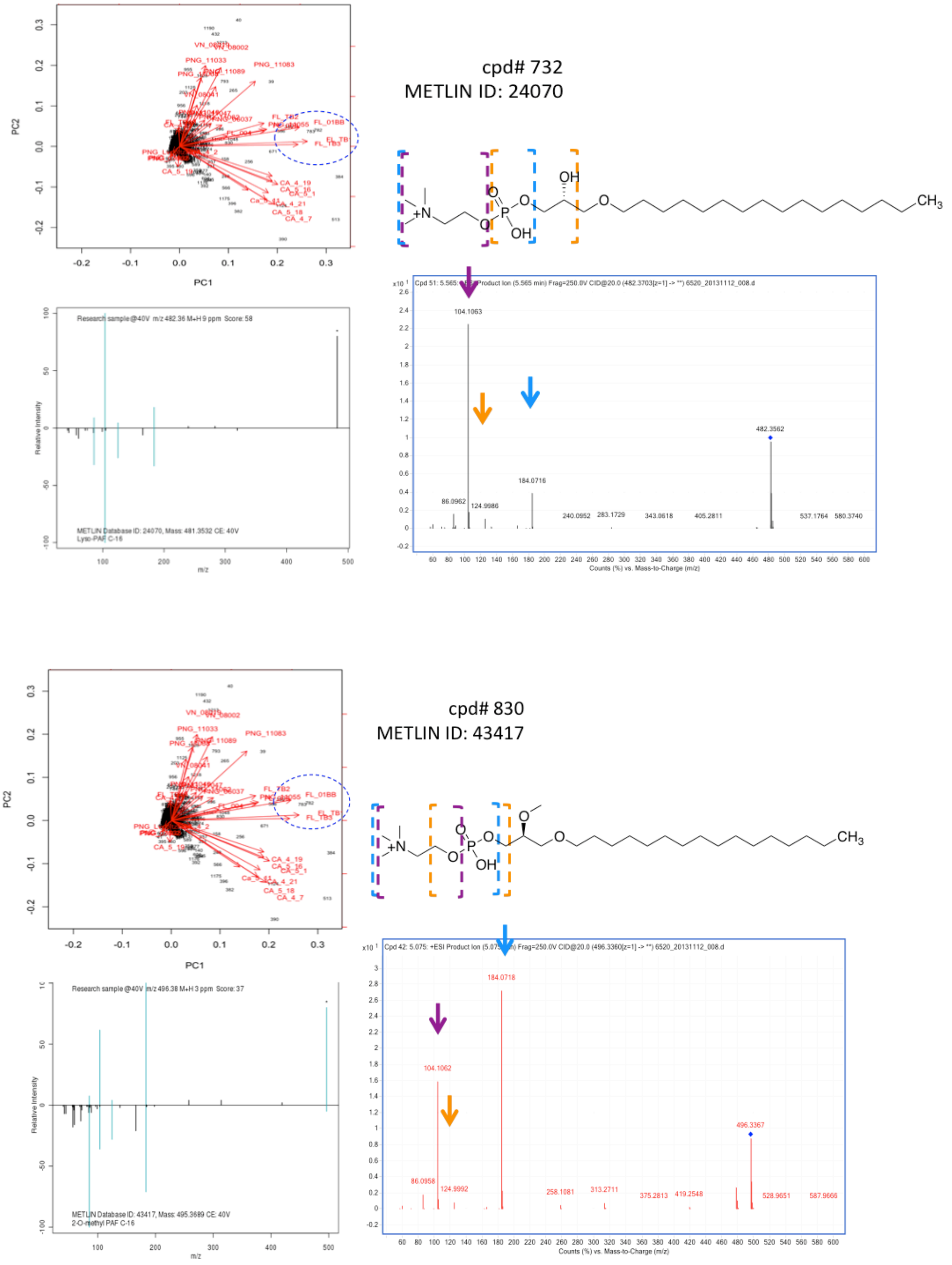


Figure S6. Example LC-MS/MS spectra of lipids that were correlated with extracts from different locations (PCA biplots are shown on the left panels). Compounds or compound families were identified by comparison of MS/MS fragmentation patterns and accurate masses to compounds in METLIN database (right panels).



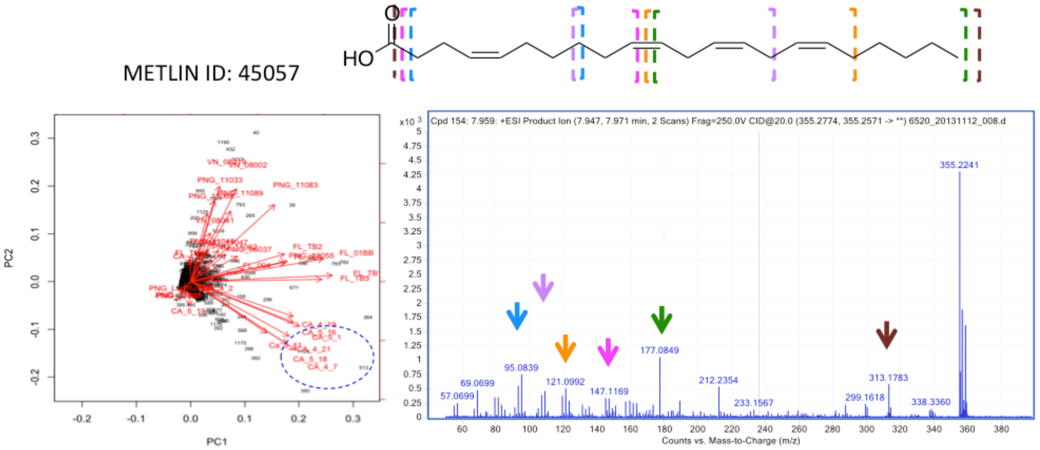

METLIN ID: 45057
O
HO
PC1
PC2
-0.2
-0.1
0.0
0.1
0.2
0.3
VN_08002
PNG_11033
PNG_11089
PNG_11083
VN_08041
PNG_06037
FL_TB2
FL_01BB
FL_TB3
CA_4_19
CA_5_16
CA_5_1
CA_4_21
CA_5_18
CA_4_7
CA_5_19
Cpd 154: 7.959: +ESI Product Ion (7.947, 7.971 min, 2 Scans) Frag=250.0V CID@20.0 (355.2774, 355.2571 -> **) 6520_20131112_008.d
x10 3
5
4.75
4.5
4.25
4
3.75
3.5
3.25
3
2.75
2.5
2.25
2
1.75
1.5
1.25
1
0.75
0.5
0.25
0
355.2241
57.0699
69.0699
95.0839
121.0992
147.1169
177.0849
212.2354
233.1567
299.1618
313.1783
338.3360
60
80
100
120
140
160
180
200
220
240
260
280
300
320
340
360
380
Counts vs. Mass-to-Charge (m/z)

Figure S7A. PCA of samples from Papua New Guinea showing a clustering driven by secondary metabolites.

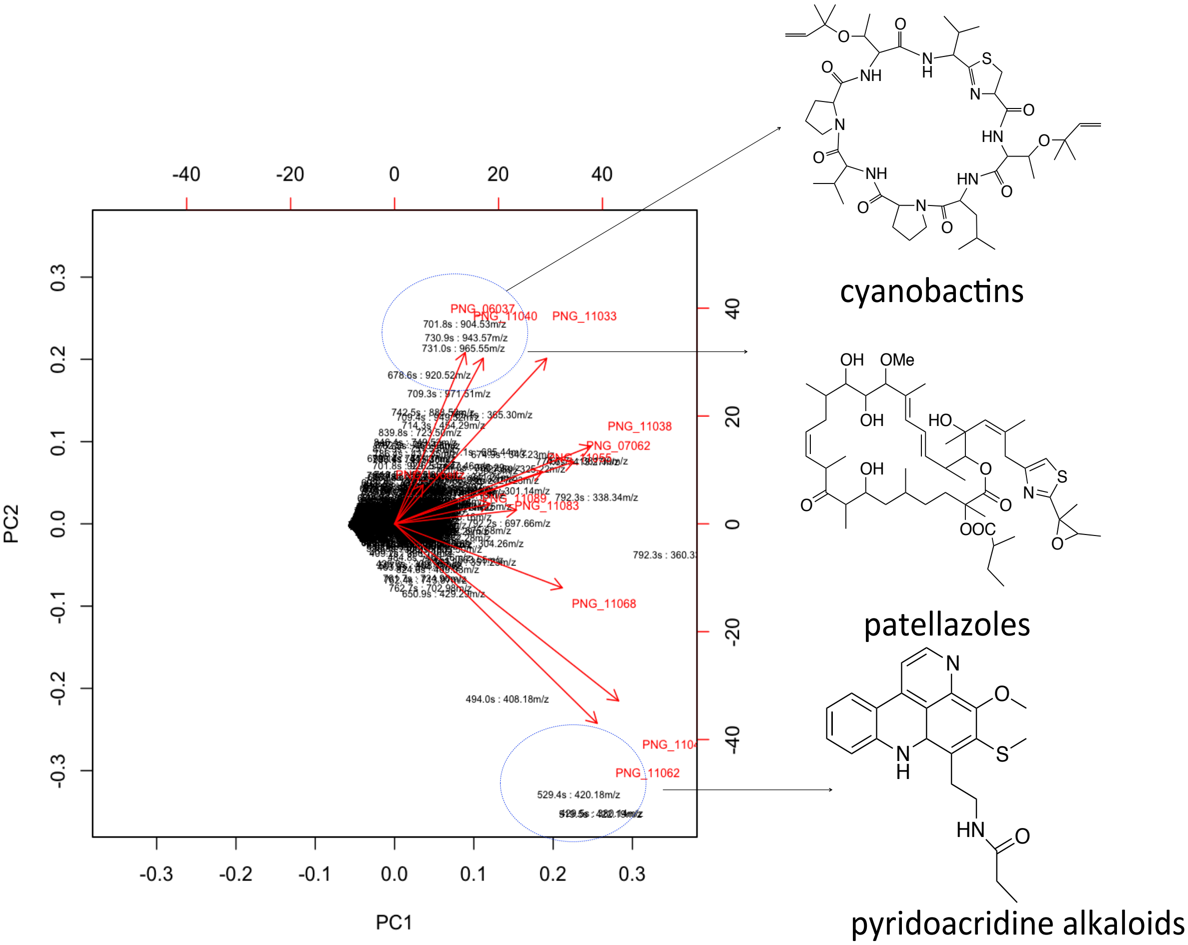


Figure S7B. Spectrum and selected ion chromatogram of varamine A from *Lissoclinum badium* samples.

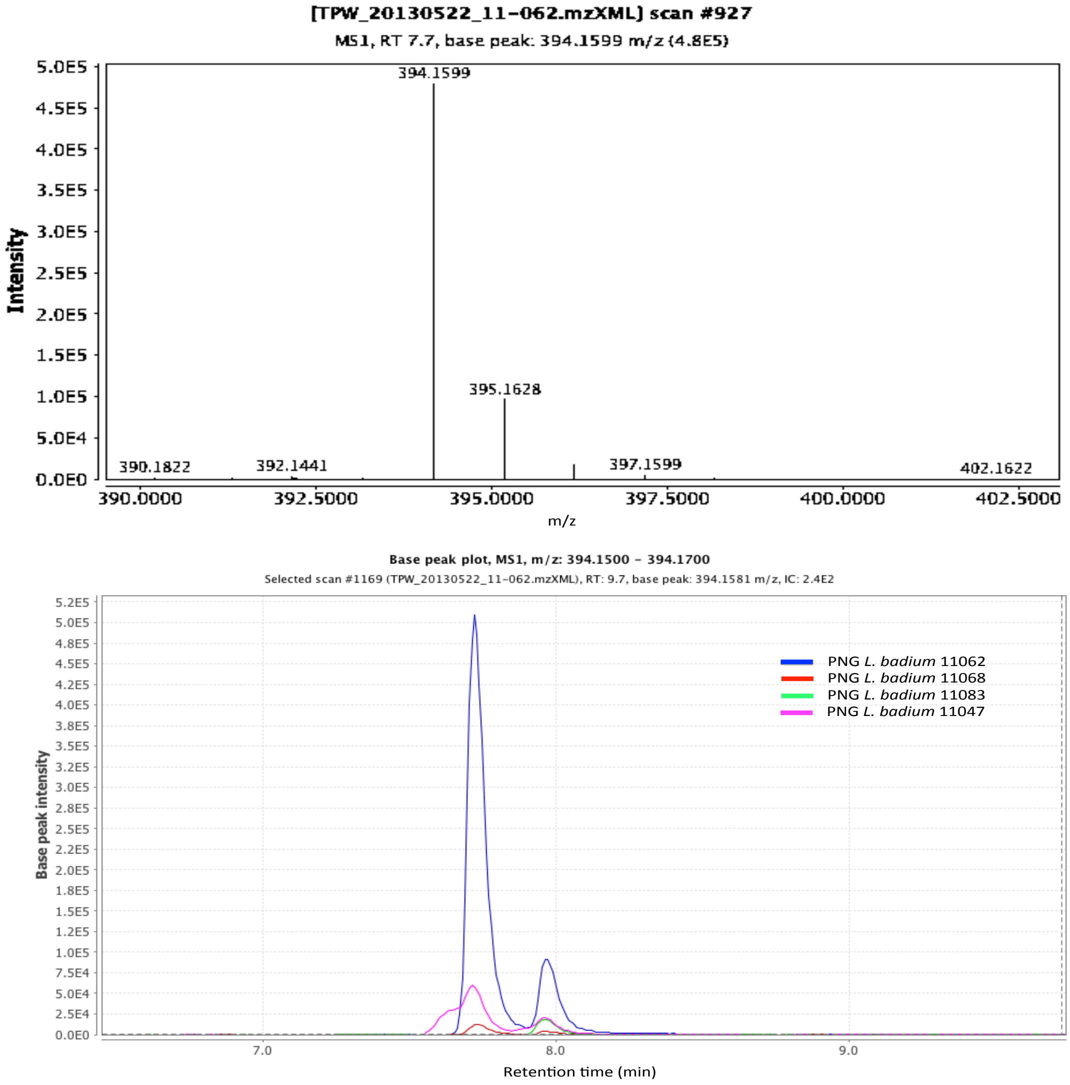


Figure S7C. Spectrum and selected ion chromatogram of varamine B from *Lissoclinum badium* samples.

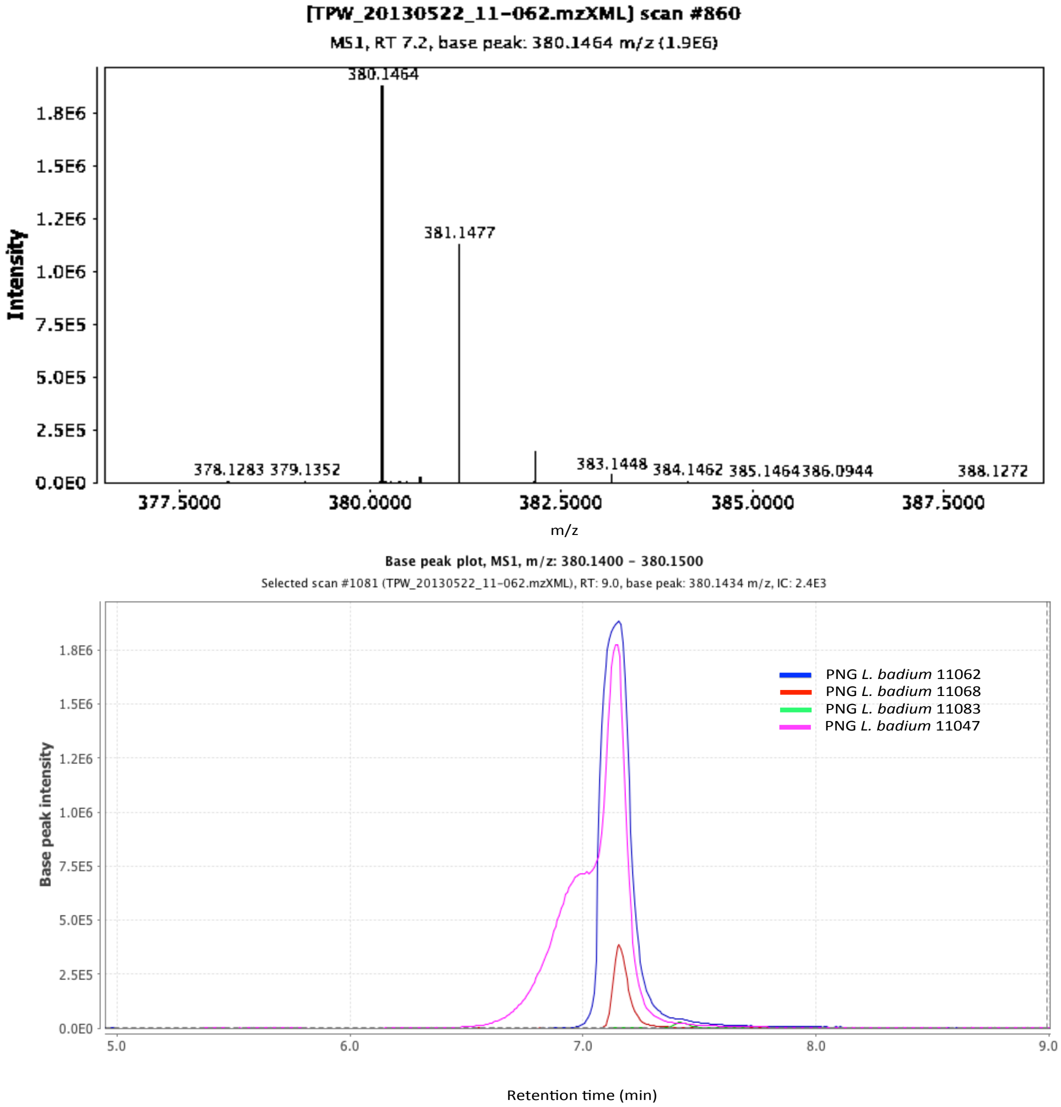


Figure S7D. Spectrum and selected ion chromatogram of diplamine from *Lissoclinum badium* samples.

\

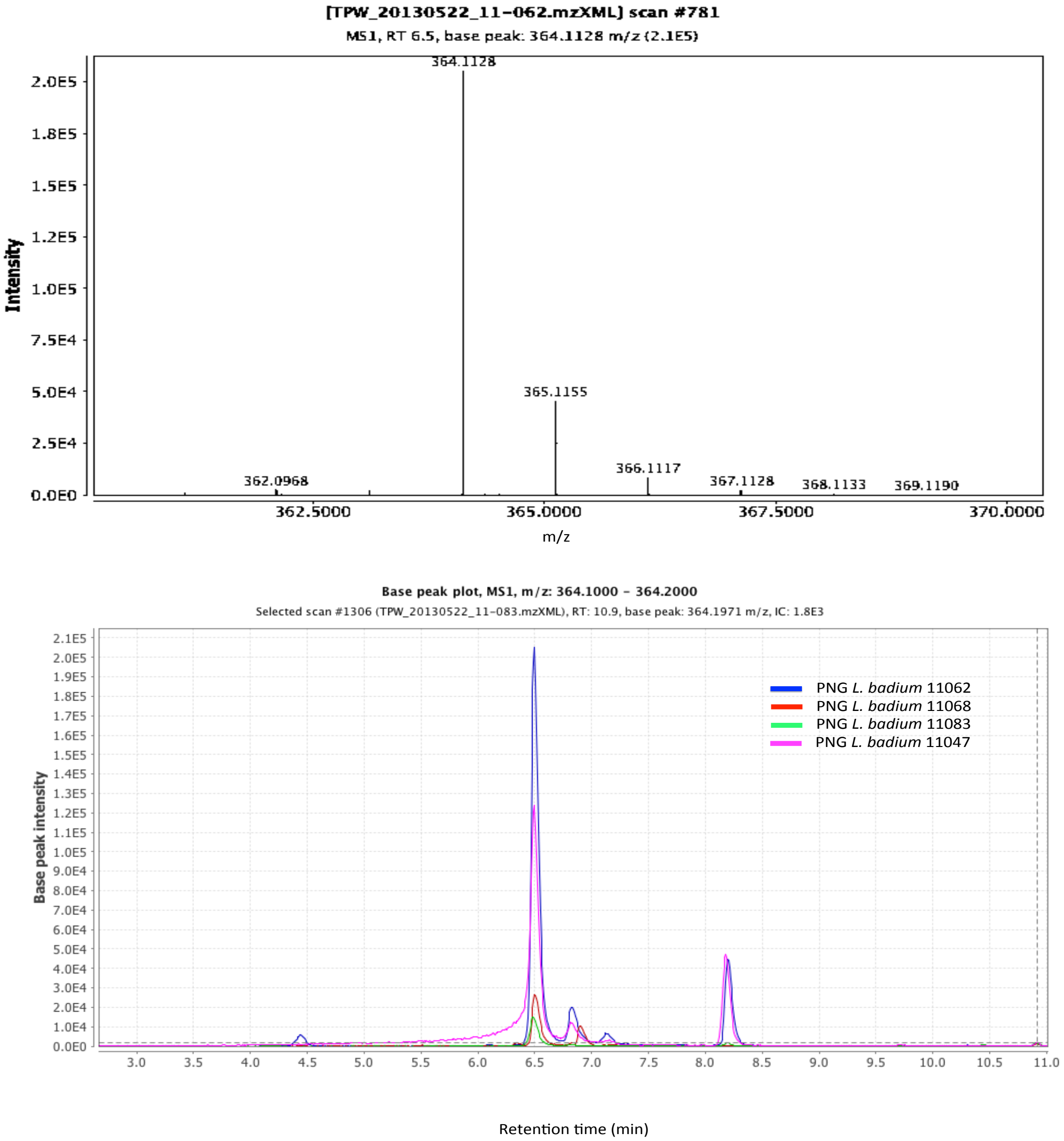


Figure S7E. Spectrum and selected ion chromatogram of lissoclin B from *Lissoclinum badium* samples.

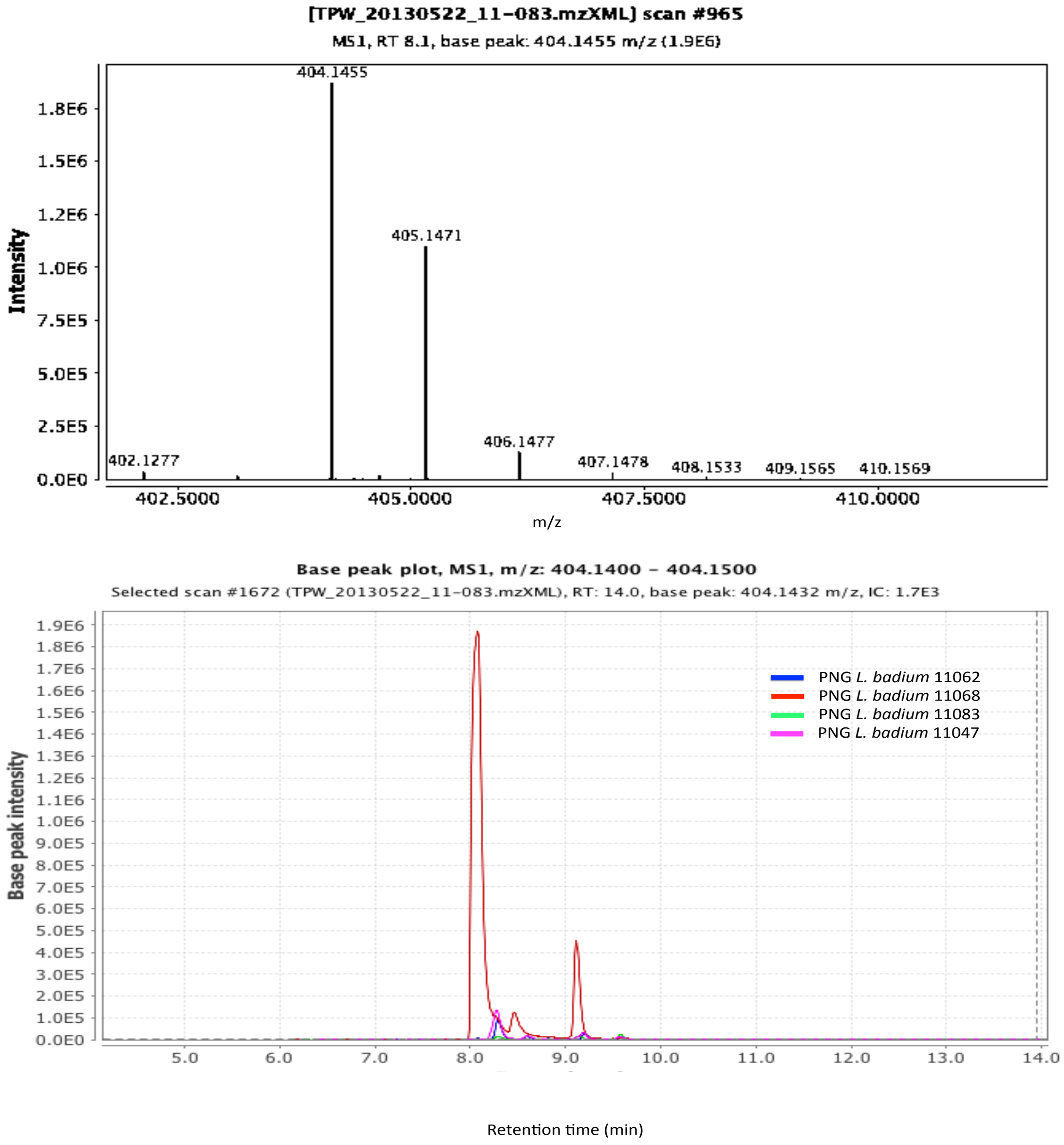


Figure S7F. Spectrum and selected ion chromatogram of patellazole A from *Lissoclinum patella* samples.

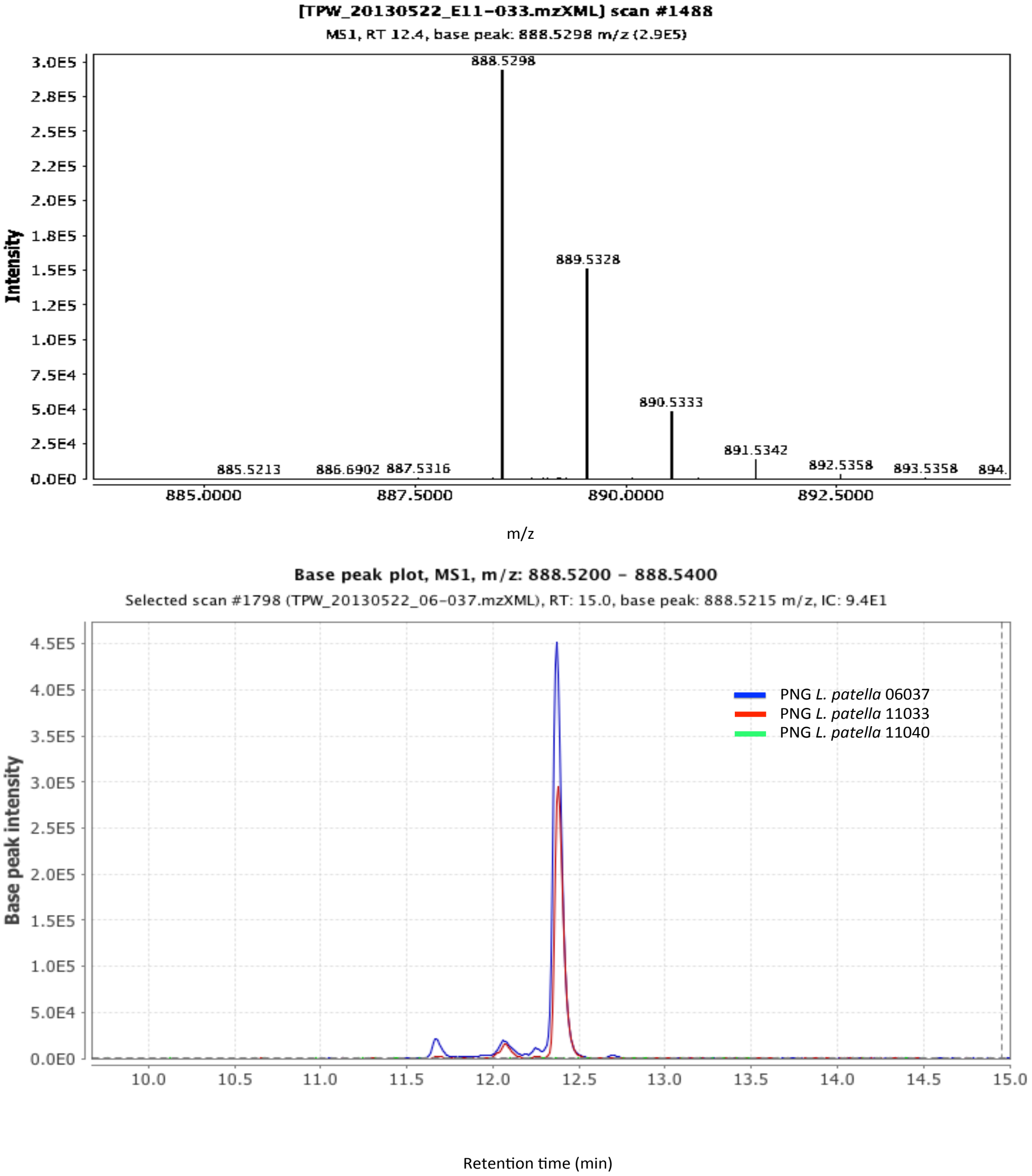


Figure S7G. Spectrum and selected ion chromatogram of patellin 3 from *Lissoclinum patella* samples.

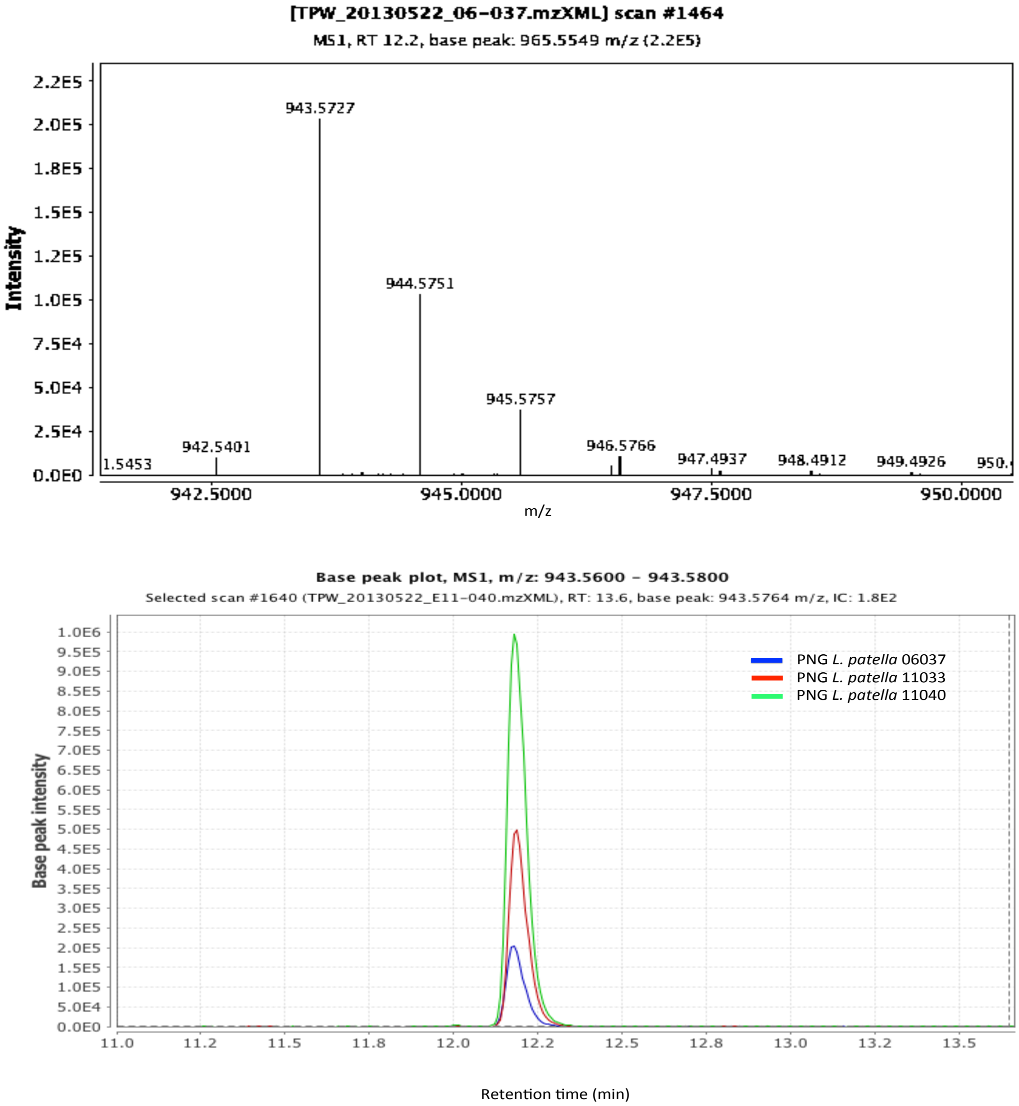


Figure S7H. Spectrum and selected ion chromatogram of cystodytin A from *Cystodytes* sp. samples.

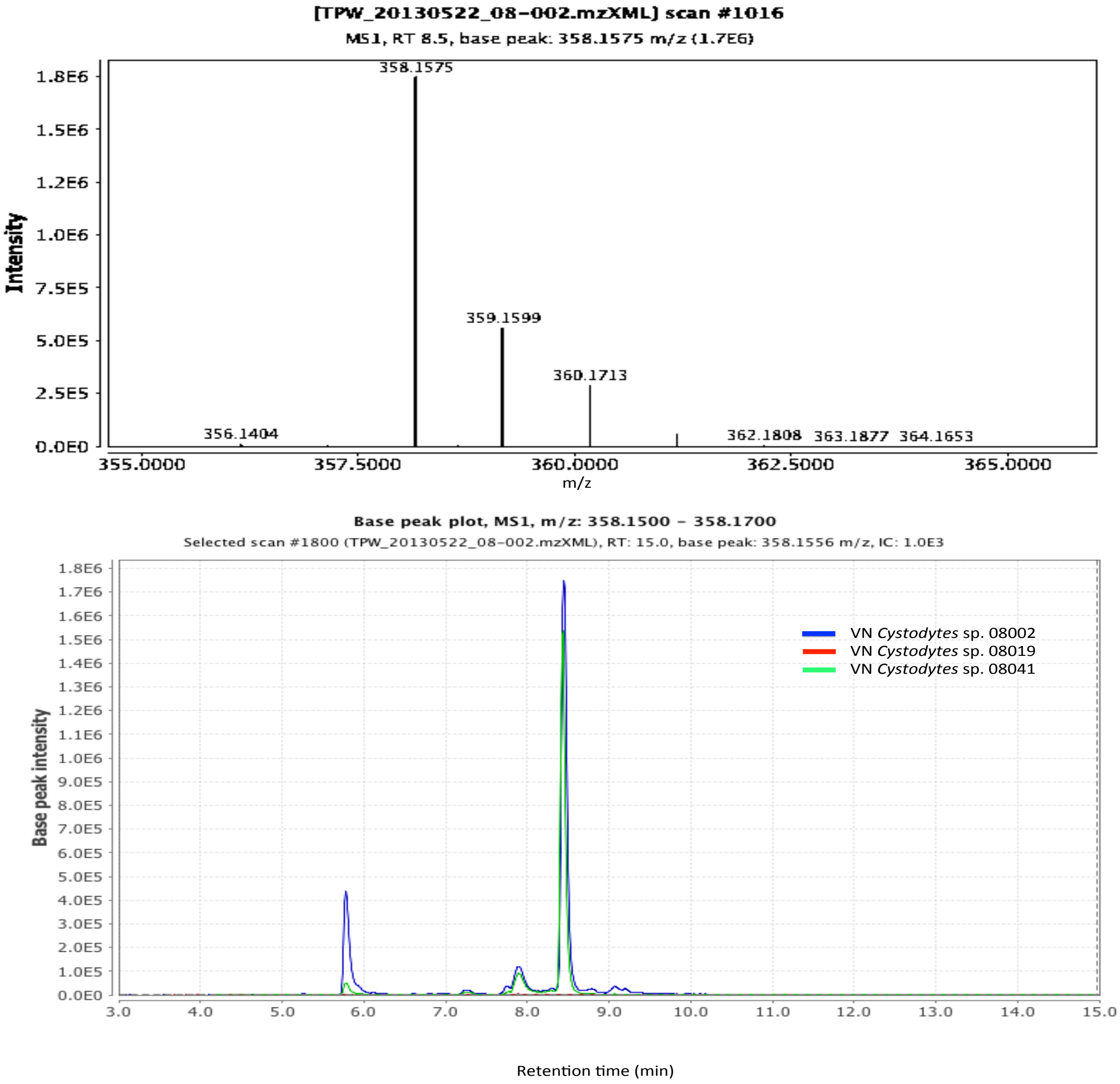


Figure S7I. Spectrum and selected ion chromatogram of cystodytin A from *Cystodytes sp* samples

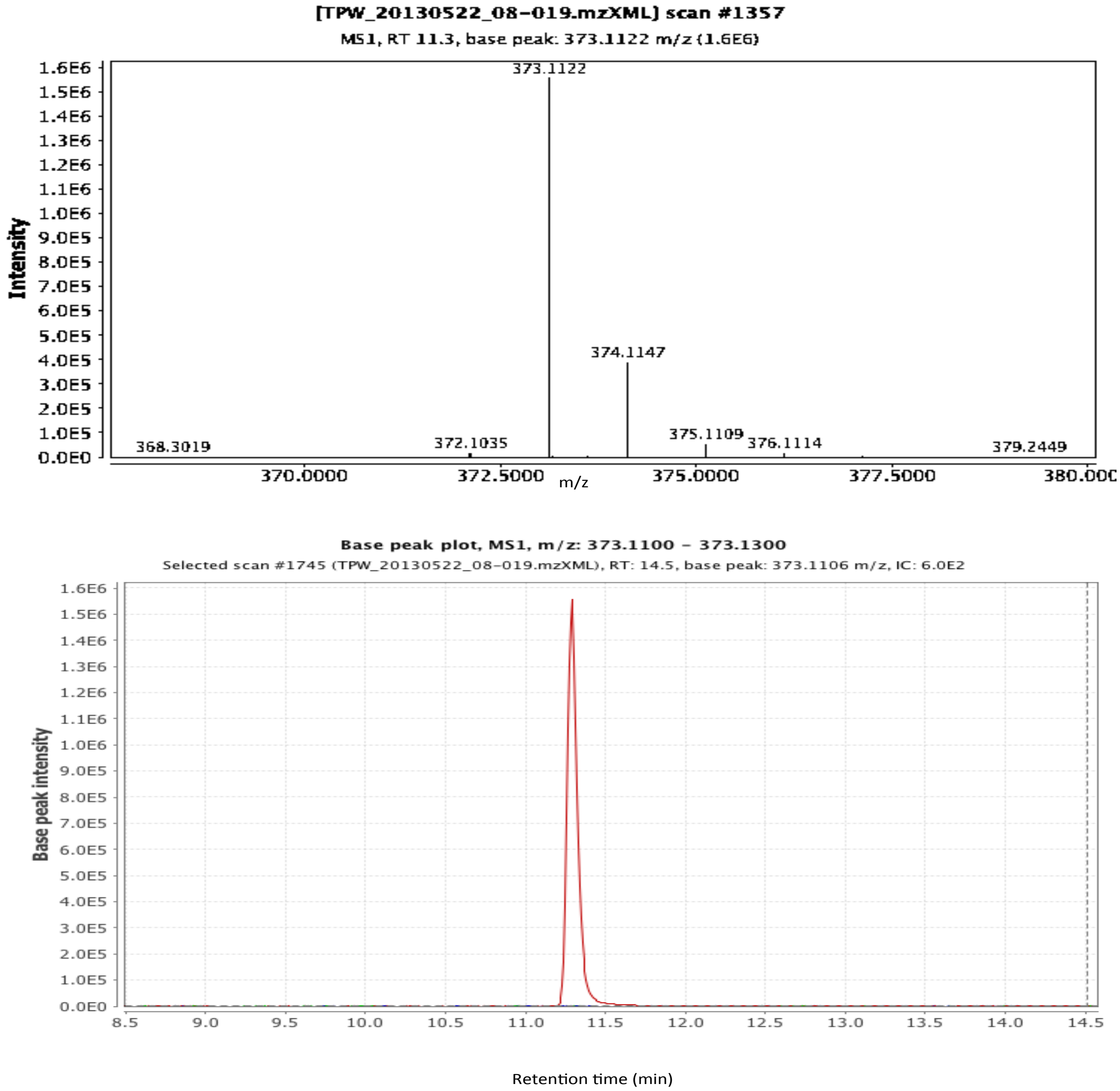


Figure S8. Bioactivity profile of ascidian extracts showing the distribution of bioactivity across different locations. Activities are shown as percent growth inhibition relative to standard compounds (see methods section).

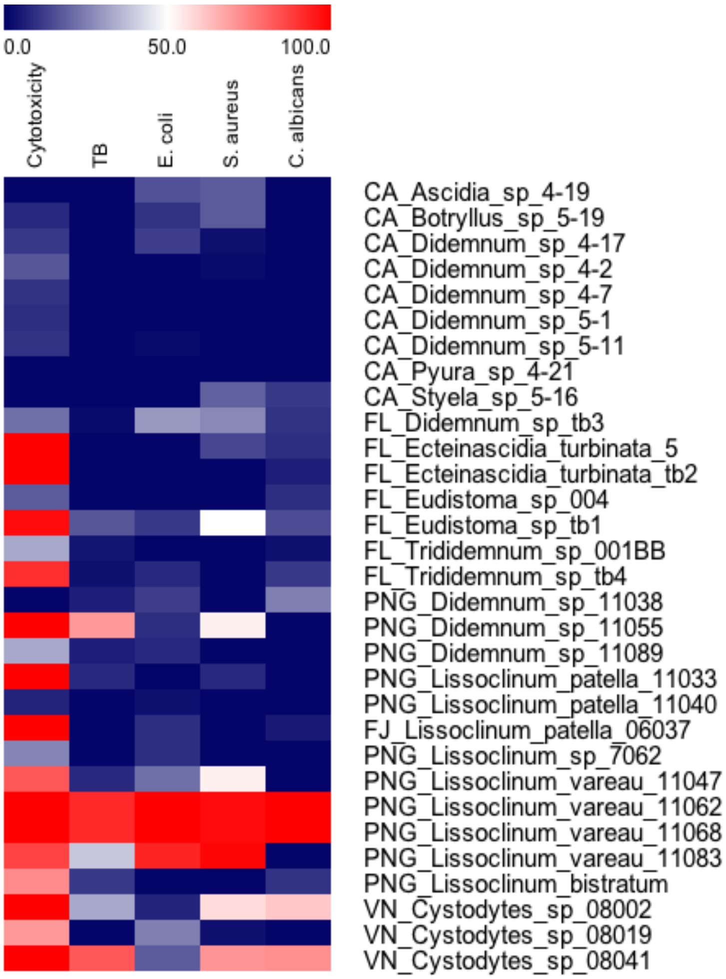


Figure S9. Group diversity analysis using the Shannon and Simpson indices. Calculations were made using the mcpHill function in the simboot R package. This function reports multiplicity adjusted p-values from distribution free tests (see methods for details).

Location (number of samples)	Pvalue(Shannon)	Pvalue(Simpson)
CA (10)	0.9958	0.9972
FL (7)	0.9564	0.6014
PNG (15)	0.9980	0.8104
Host (number of samples)		
Cystodytes sp. (3)	1.0000	0.9994
Didemnum sp. (6)	0.7678	0.6856
*Didemnum sp.*_2 (3)	0.0082	0.0072
Ecteinascidia turbinata (2)	1.0000	0.9824
Lissoclinum badium (4)	0.6970	0.8482
Lissoclinum patella (3)	0.8678	0.7086