Supplementary Information for **Biosignatures of Diverse Eukaryotic Life from a Snowball Earth Analogue Environment in Antarctica**

Fatima Husain*[1], Jasmin L. Millar[2], Anne D. Jungblut[3], Ian Hawes[4], Thomas W. Evans[1,^], Roger E. Summons[1]

[1]Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts, United States of America

[2]School of Earth and Environmental Sciences, Cardiff University, Cardiff, United Kingdom

[3]Department of Science, Natural History Museum, London, United Kingdom

[4]Coastal Marine Field Station, University of Waikato, Tauranga, New Zealand

*Corresponding author: fhusain@mit.edu

^Present Address: Shell Global Solutions International B.V., The Hague, The Netherlands

Supplementary Figures contain the following sections:

1 | Systematic and Common Sterol Names
2 | Co-elution of (24R)-ethyl-5α-cholestan-3β-ol and 24-ethylcholesta-5,24(28)Z-dien-3β-ol
3 | Mass Spectra of Unidentified Steroids
4 | Possible C_{31} Steranes Generated from the Pond Seventy Mat
5 | dMRM Chromatograms of the Relict Microbial Mat Hydrocarbon Fraction
6 | Relative Abundances of 18S rRNA Gene Communities obtained from Meltwater Pond Microbial Mats from the McMurdo Ice Shelf, Antarctica

Supplementary Information

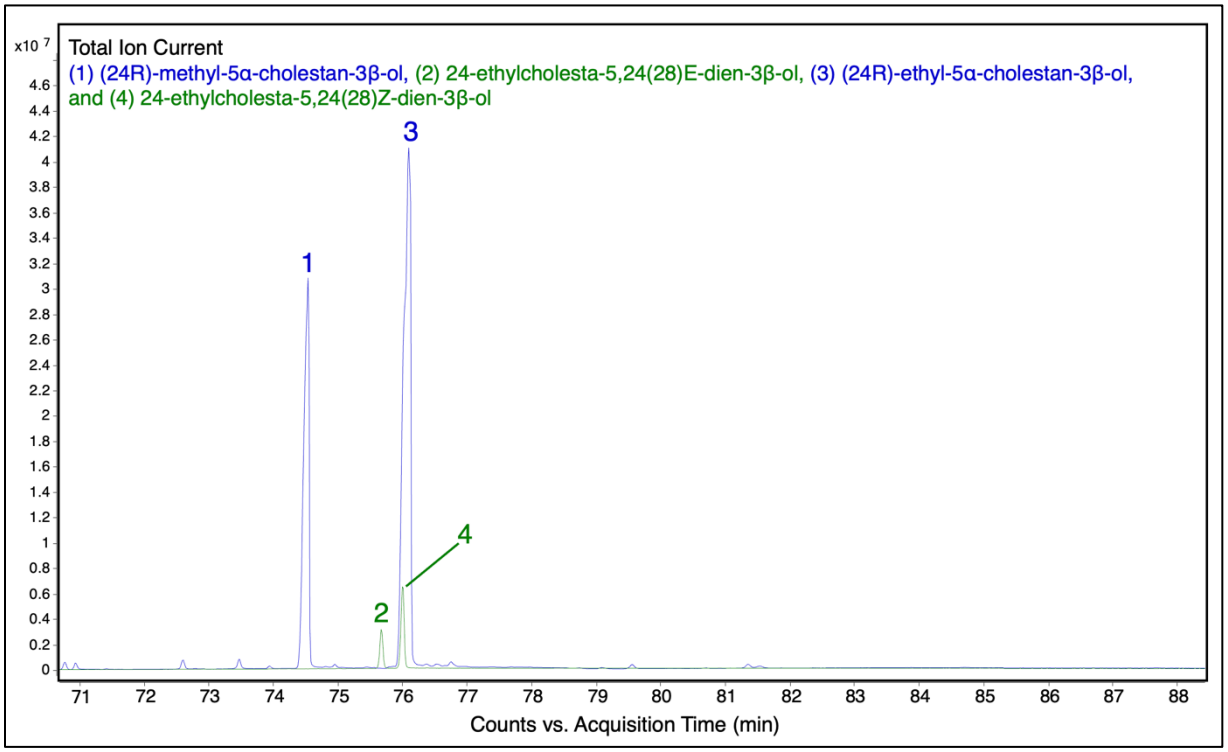
1 | Systematic and Common Sterol Names

Peak	Systematic Name	Common Name
1	22-trans-24-norcholesta-5,22E-dien-3β-ol	24-Norcholestadienol
2	5α-cholest-22E-en-3β-ol	5α-Cholest-22E-en-3β-ol
3	5β-cholestan-3β-ol	Coprostanol
4	Unknown Steroid	Unknown Steroid
5	27-nor-(24S)-cholesta-5,22E-dien-3β-ol	Occelasterol
6	Unknown Steroid	Unknown Steroid
7	cholesta-5,22E-dien-3β-ol	22-Dehydrocholesterol
8	Unknown Steroid	Unknown Steroid
9	Unknown Steroid	Unknown Steroid
10	cholest-5-en-3β-ol(d7)	Cholesterol-D7
11	cholest-5-en-3β-ol	Cholesterol
12	5α-cholestan-3β-ol	Cholestanol
13	Unknown Steroid	Unknown Steroid
14	24-methylcholesta-5,22E-dien-3β-ol	Brassicasterol + Diatomsterol
15	5α-cholest-7-en-3β-ol	Lathosterol
16	24-methylcholesta-5,24(28)-dien-3β-ol	24-Methylenecholesterol
17	(24R)-ethyl-5β-cholestan-3β-ol	5β-Stigmastanol
18	(24R)-methylcholest-5-en-3β-ol	Campesterol
19	(24R)-methyl-5α-cholestan-3β-ol	Campestanol
20	(24S)-ethylcholesta-5,22E-dien-3β-ol	Stigmasterol
21	23,24-dimethyl-5α-cholest-22E-en-3β-ol	4α-Desmethyl-dinosterol
21	Unknown Compound	Unknown Compound
22	24-methyl-5α-cholest-7-en-3β-ol	Ergosta-7-en-3β-ol
23	Unknown Steroid	Unknown Steroid
24	(24R)-ethylcholest-5-en-3β-ol	β-Sitosterol
25	(24R)-ethyl-5α-cholestan-3β-ol	Stigmastanol
25	24-ethylcholesta-5,24(28)Z-dien-3β-ol	Isofucosterol
26	4α,23,24-trimethylcholest-22E-en-3β-ol	Dinosterol
26	see table description	α-Tocopherol
27	Unknown Steroid	Unknown Steroid
28	(24S)-ethyl-5α-cholesta-7,22E-dien-3β-ol	Chondrillasterol
29	Unknown Steroid	Unknown Steroid
30	Unknown Steroid	Unknown Steroid

Supplementary Table 1. Systematic and Common Sterol Names. Peak numbers correspond to individual peaks identified in **Table 2** of the main text. The systematic name for peak 26 is 2,5,7,8-Tetramethyl-2-(4,8,12-trimethyltridecyl)-6-chromanol.

2 | Co-elution of (24R)-ethyl-5α-cholestan-3β-ol and 24-ethylcholesta-5,24(28)Z-dien-3β-ol

On a standard 60m DB-5MS gas chromatography column, some sterols have been shown to co-elute [1]. While such co-elutions tend to affect dinosteroids, phytosterols such as 24-ethylcholesta-5,24(28)Z-dien-3β-ol and (24R)-ethyl-5α-cholestan-3β-ol may also co-elute. To demonstrate the co-elution of these two compounds, authentic standards were co-injected and run using the same program applied to the sample set. Co-elution was observed around minute 76 of the run, during which 24-ethylcholesta-5,24(28)Z-dien-3β-ol elutes entirely within (24R)-ethyl-5α-cholestan-3β-ol (**Supplementary Fig. 1**). While it is possible that at low concentrations, co-elution may be minimized, the dominance of the C_{29} sterols (24S)-ethylcholesta-5,22E-dien-3β-ol and (24R)-ethylcholest-5-en-3β-ol in this sample set results in (24R)-ethyl-5α-cholestan-3β-ol concentrations which significantly exceed those of 24-ethylcholesta-5,24(28)Z-dien-3β-ol.



Supplementary Fig. 1. Partial gas chromatogram of peaks from sterol standards. Peaks are coloured and numbered according to standard. Peak 1 corresponds to (24R)-methyl-5α-cholestan-3β-ol; Peak 2 corresponds to 24-ethylcholesta-5,24(28)E-dien-3β-ol; Peak 3 corresponds to (24R)-ethyl-5α-cholestan-3β-ol; Peak 4 corresponds to 24-ethylcholesta-5,24(28)Z-dien-3β-ol.

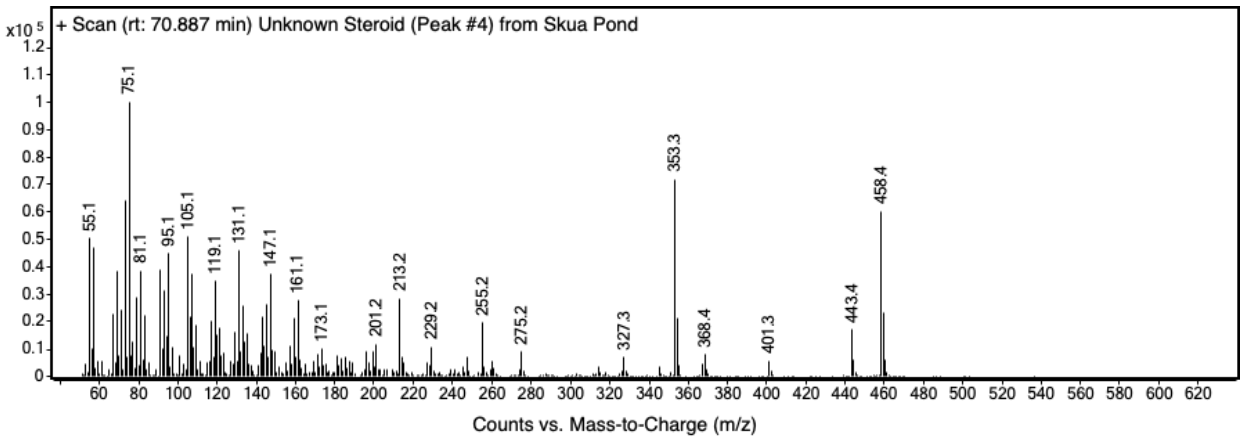
2 | Mass Spectra of Unidentified Steroids

Altogether, nine compounds with mass spectral features characteristics of sterols could not be definitively identified via co-injection of standards or against reference spectra. Their background-subtracted mass spectra are shown in **Figs. S2-S10**.

It is possible that co-elution with non-steroidal compounds, such as carboxylic acids, long chain alcohols, or other polar molecules complicates these identifications. Many unknown sterols were identified in Skua Pond in particular, which is perhaps indicative of a markedly diverse eukaryotic assemblage and active sterol biosynthesis at work. Due to the late elution time of polar compounds such as sterols with this study's run program, instrument backgrounds which tended to be high in column bleed (207 Da) were subtracted from the mass spectra below. We have also included some interpretations that were suggested

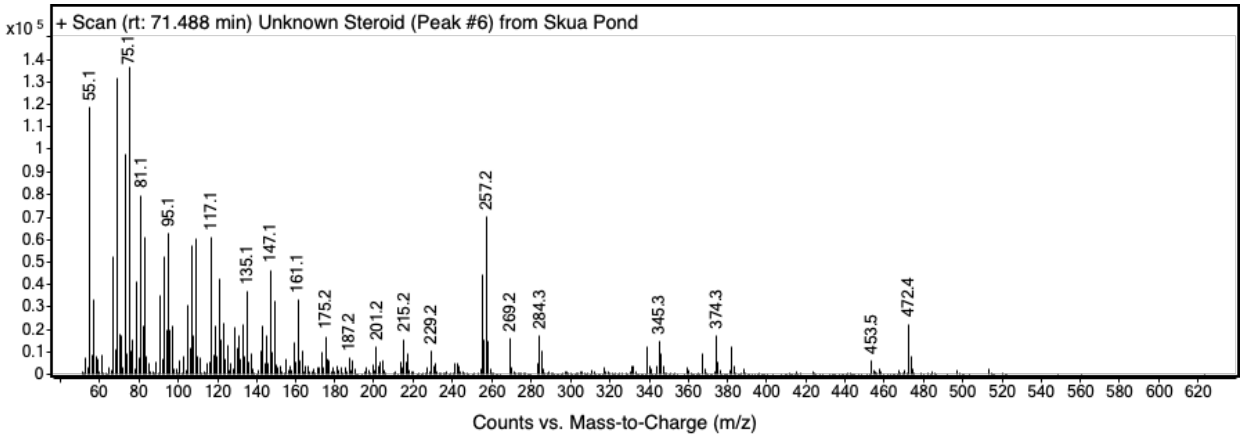
during peer review of this manuscript; we are grateful to the reviewers for furthering efforts to characterize these molecules.

The mass spectrum of peak #4 (**Supplementary Fig. 2**) contained a 458 Da molecular ion, followed by M-CH_3 fragments 443 Da, 353 Da, and 255 Da. The base peak ion, 75 Da, is characteristic of hydroxysteroids [2] such as cholest-5-en-3β-ol and 5α-cholestan-3β-ol, which elute after this unknown compound. A reviewer suggested that the elution time of this molecule makes it more likely to have a Δ8 unsaturation rather than a Δ7, unless this molecule has a 5β hydrogen and Δ7 unsaturation.



Supplementary Fig. 2. Mass spectrum of Peak #4 from Skua Pond. Ions from this unidentified compound are shown and labelled with their mass-to-charge ratios.

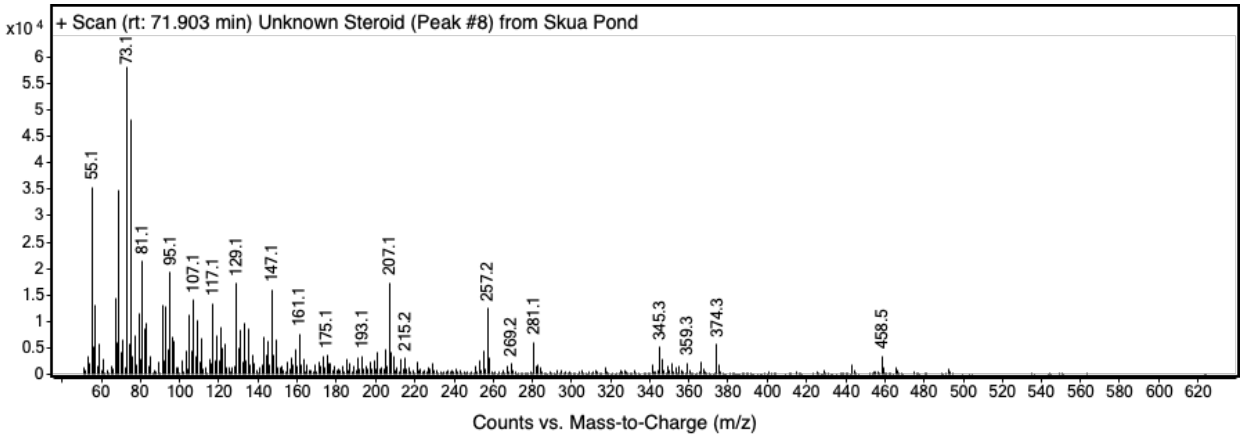
The mass spectrum of peak #6 (**Supplementary Fig. 3**) contained a 472 Da molecular ion followed by 257 Da, 255 Da, 215 Da, and 129 Da fragment ions often found in steroids, perhaps indicative of unsaturations at C-5 or C-7. A 75 Da base peak was observed as in Peak #4. However, the presence of a strong 117 Da ion peak suggests the presence of a carboxylic acid, indicating possible co-elution.



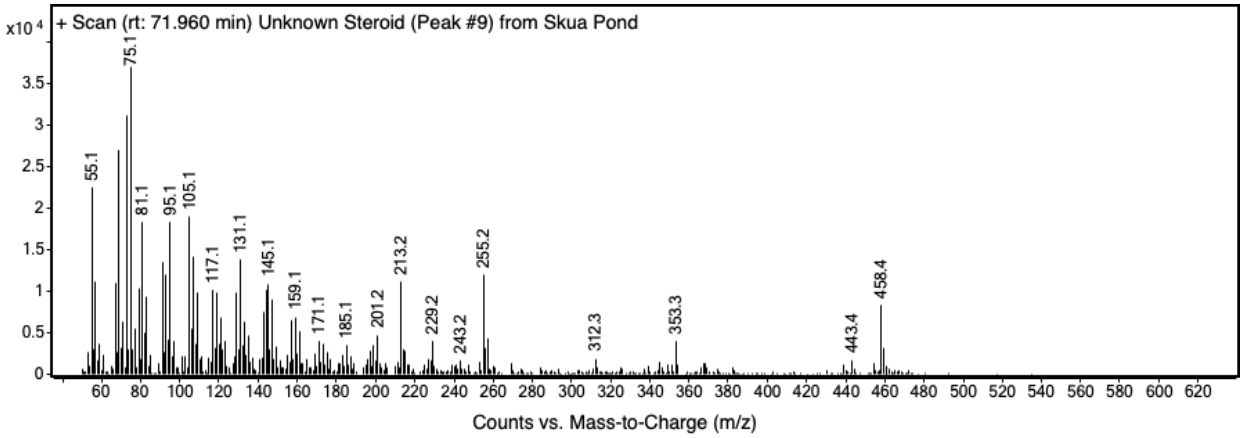
Supplementary Fig. 3. Mass spectrum of Peak #6 from Skua Pond. Ions from this unidentified compound are shown and labelled with their mass-to-charge ratios.

Peaks #8 and #9 (**Figs. S4** and **S5**) in Skua Pond elute as a doublet. The spectra for these peaks contain a 458 Da molecular ion and a small 443 Da M-15 fragment. However, the first peak, peak #8, contains a 257 Da mass fragment ion, while peak #9 contains a 255 Da mass fragment ion, suggesting that the two compounds differ in the positions of their unsaturations. Peak #8 contains 257 Da, 215 Da, and 129 Da

fragments which tend to occur in Δ22 5α-stanols. A reviewer suggested that peak #9 is unlikely to contain a Δ8 unsaturation based on its retention time.

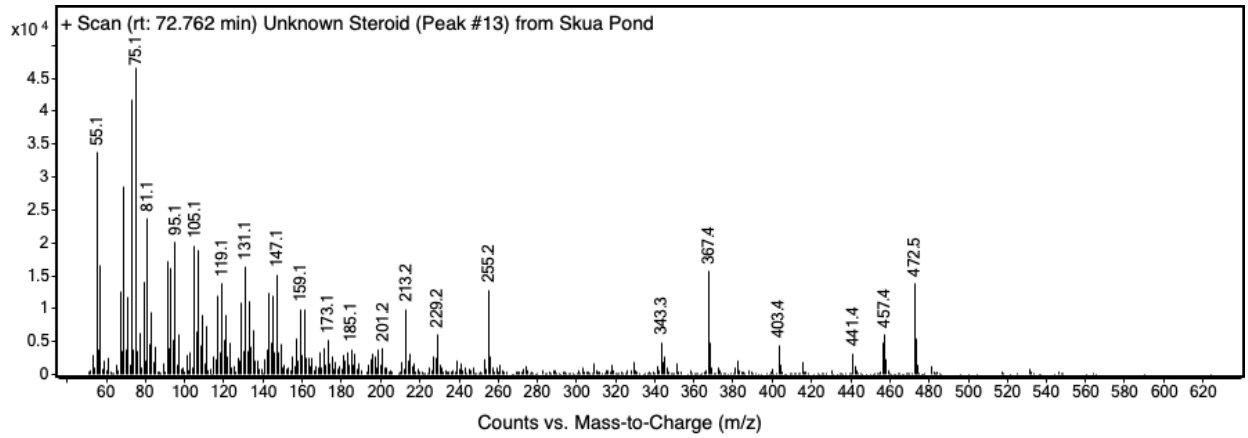


Supplementary Fig. 4. Mass spectrum of Peak #8 from Skua Pond. Ions from this unidentified compound are shown and labelled with their mass-to-charge ratios.



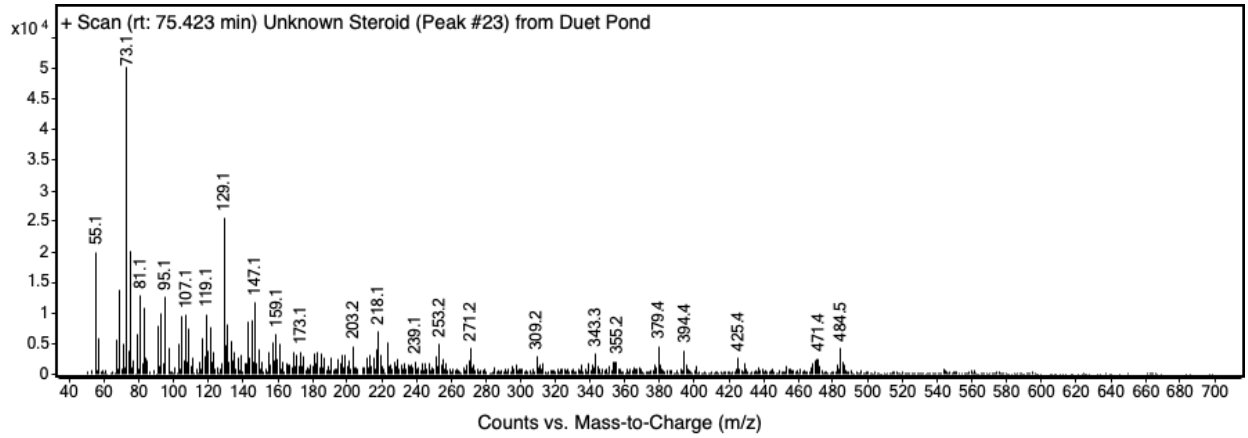
Supplementary Fig. 5. Mass spectrum of Peak #9 from Skua Pond. Ions from this unidentified compound are shown and labelled with their mass-to-charge ratios.

The mass spectrum of peak #13 (**Supplementary Fig. 6**) contains a 472 Da molecular ion, followed by a 457 Da M-15 fragment, as well as 255 Da, 213 Da, 131 Da fragments, and a 75 Da base peak. The 343 Da fragment may arise from M-129; a 129 Da fragment is present in less intensity than the 131 Da fragment. The simultaneous presence of these fragments suggests that this molecule contains a possible Δ5 unsaturation. However, as a reviewer noted, these interpretations may be complicated by co-elution with another molecule with a 456 Da molecular ion.



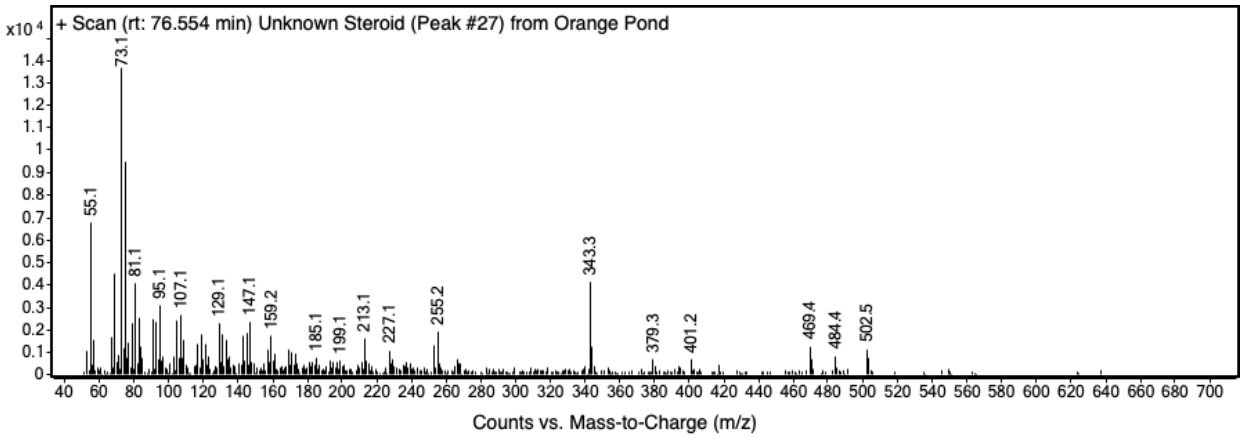
Supplementary Fig. 6. Mass spectrum of Peak #13 from Skua Pond. Ions from this unidentified compound are shown and labelled with their mass-to-charge ratios.

The mass spectrum of peak #23 (**Supplementary Fig. 7**) contains a 484 Da molecular ion with 355 Da, 343 Da, 271 Da, 253 Da, and 129 Da fragments, as well as a 73 Da base peak. The presence of 253 Da and 271 Da mass fragments suggests unsaturations at C-5 and C-8, C-5 and C-7, or at C-8 and C-14.



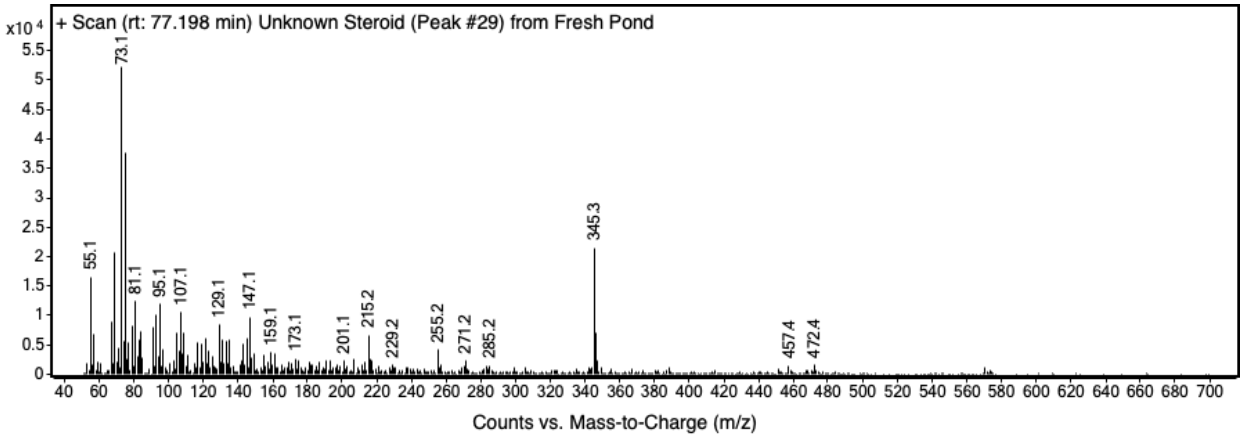
Supplementary Fig. 7. Mass spectrum of Peak #23 from Duet Pond. Ions from this unidentified compound are shown and labelled with their mass-to-charge ratios.

The mass spectrum of peak #27 (**Supplementary Fig. 8**) contains a 502 Da fragment, though it is unclear whether this ion represents the molecular ion or is the result of co-elution with another molecule. 484 Da, 469 Da, 343 Da, 255 Da, 213 Da, 129 Da, and 73 Da fragments were also detected. The dual presence of 255 Da and 213 Da fragments suggest that the molecule may contain an unsaturation at C-5.



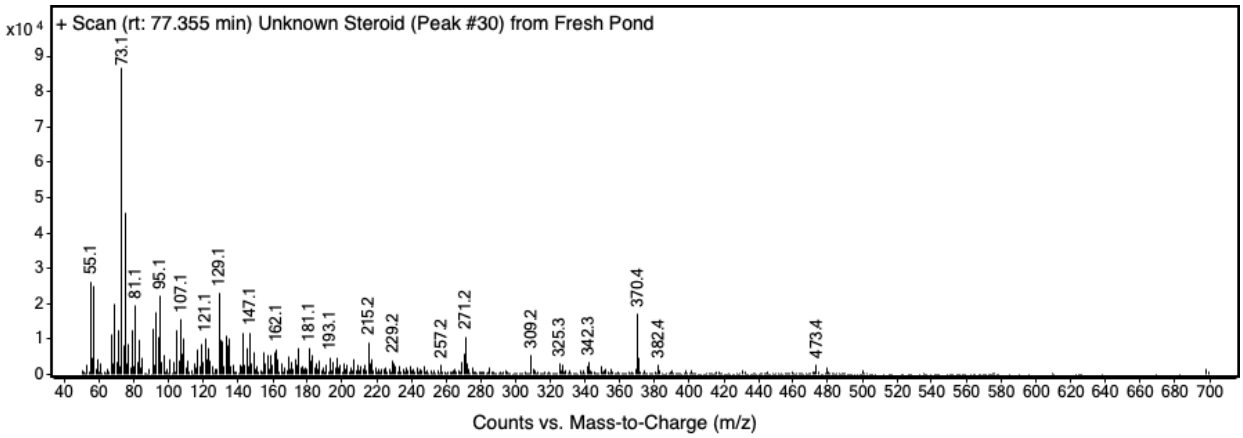
Supplementary Fig. 8. Mass spectrum of Peak #27 from Orange Pond. Ions from this unidentified compound are shown and labelled with their mass-to-charge ratios.

The mass spectrum of peak #29 (**Supplementary Fig. 9**) contains 472 Da, 457 Da, 345 Da, 271 Da, 255 Da, 229 Da, 215 Da, 129 Da fragments with a 73 Da base peak. The presence of 255 Da and 229 Da fragments together are commonly associated with the presence of an unsaturation at C-7. A reviewer notes that the 345 Da ion present corresponds to a loss of a 127 Da ion, consistent with a C_9H_{19} side chain.



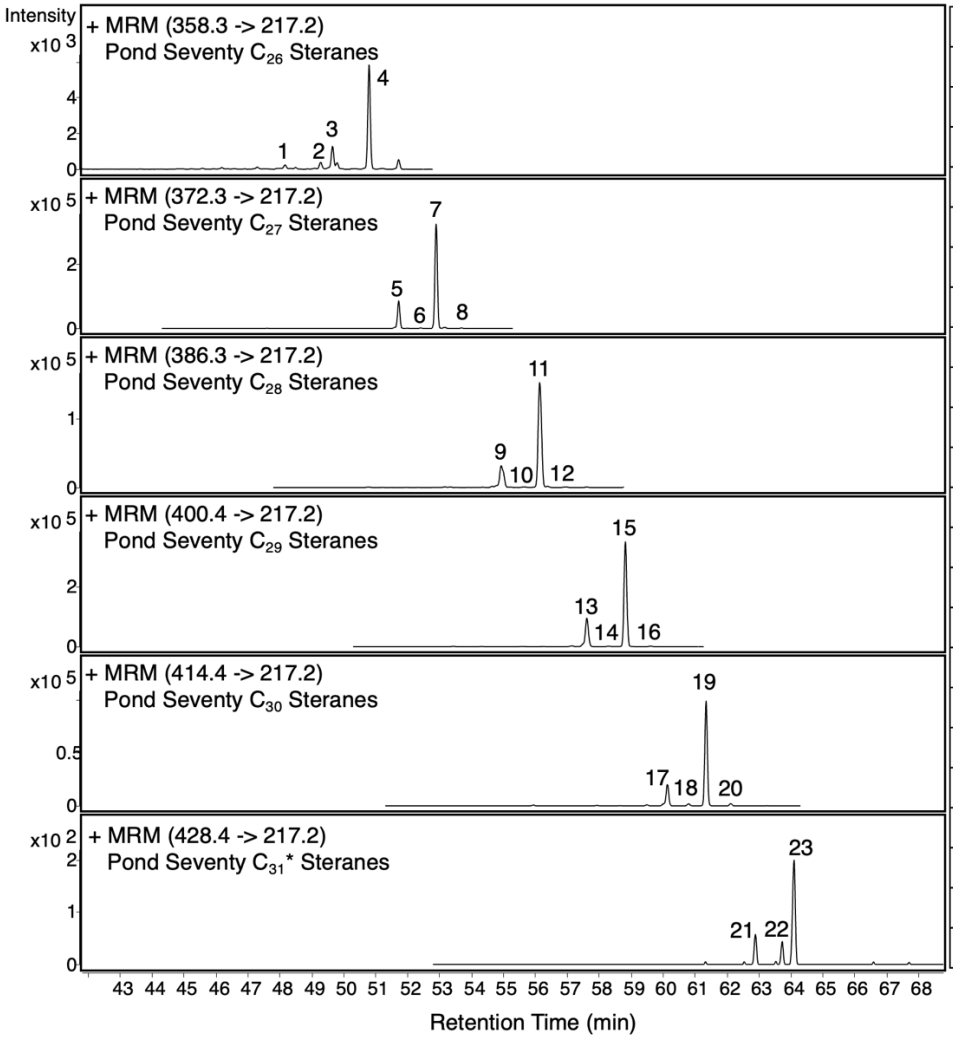
Supplementary Fig. 9. Mass spectrum of Peak #29 from Fresh Pond. Ions from this unidentified compound are shown and labelled with their mass-to-charge ratios.

The mass spectrum of peak #30 (**Supplementary Fig. 10**) contains several fragment ions commonly associated with steroids, including 73 Da, 129 Da, 215 Da, 229 Da, 257 Da, and 271 Da, though its higher mass-to-charge ratio peaks do not readily lend themselves to interpretation without a clear molecular ion present.



Supplementary Fig. 10. Mass spectrum of Peak #30 from Fresh Pond. Ions from this unidentified compound are shown and labelled with their mass-to-charge ratios.

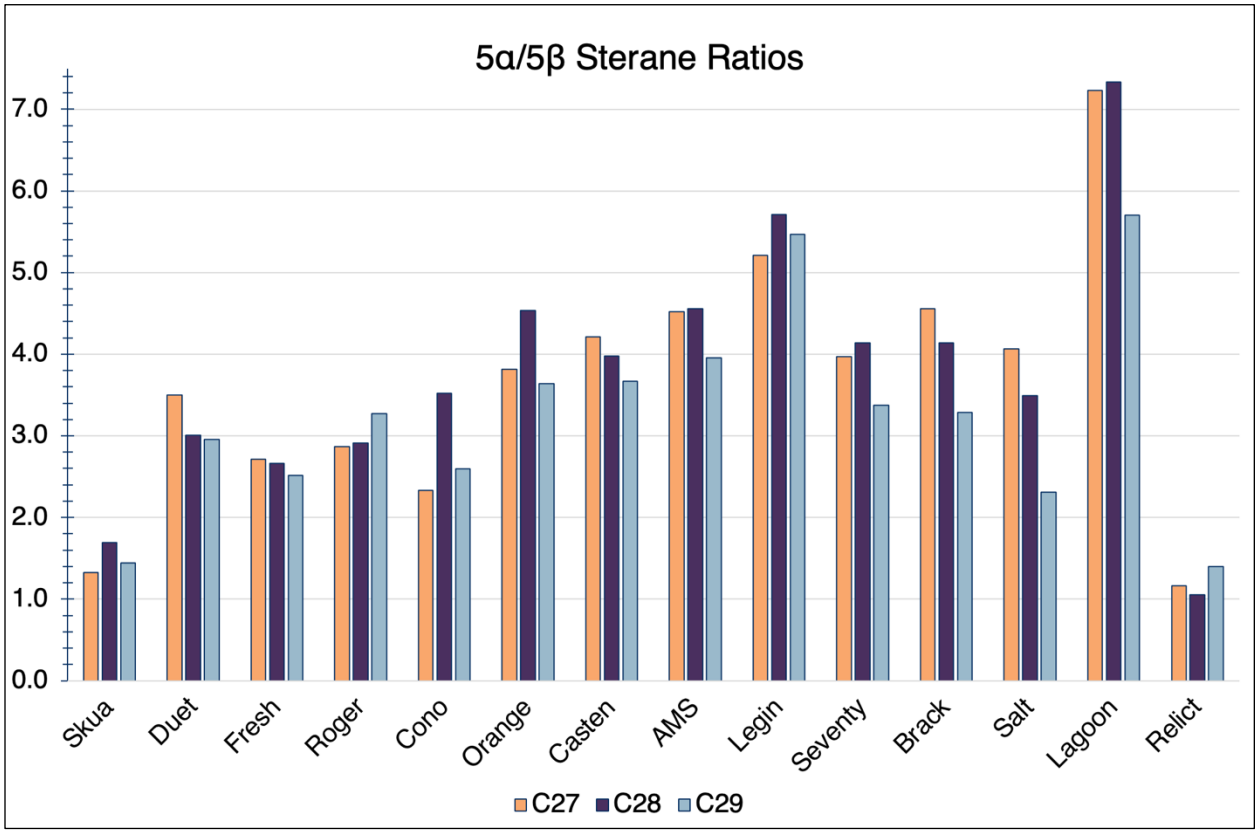
3 | Possible C_{31} Steranes Generated from Pond Seventy Mat



Number	Peak Identity
1	24-norcholestane βαα20R
2	24-norcholestane ααα20R
3	27-norcholestane βαα20R
4	27-norcholestane ααα20R
5	cholestane βαα20R
6	C_{27} norsterane βαα20R
7	cholestane ααα20R
8	C_{27} norsterane ααα20R
9	24-methylcholestane βαα20R
10	C_{28} norsterane βαα20R
11	24-methylcholestane ααα20R
12	C_{28} norsterane ααα20R
13	24-ethylcholestane βαα20R
14	C_{29} norsterane βαα20R
15	24-ethylcholestane ααα20R
16	C_{29} norsterane ααα20R
17	24-isopropylcholestane βαα20R
18	putative C_{30} norsterane βαα20R
19	24-isopropylcholestane ααα20R
20	putative C_{30} norsterane ααα20R
21	putative C_{31} sterane βαα20R
22	putative C_{31} norsterane βαα20R
23	putative C_{31} sterane ααα20R

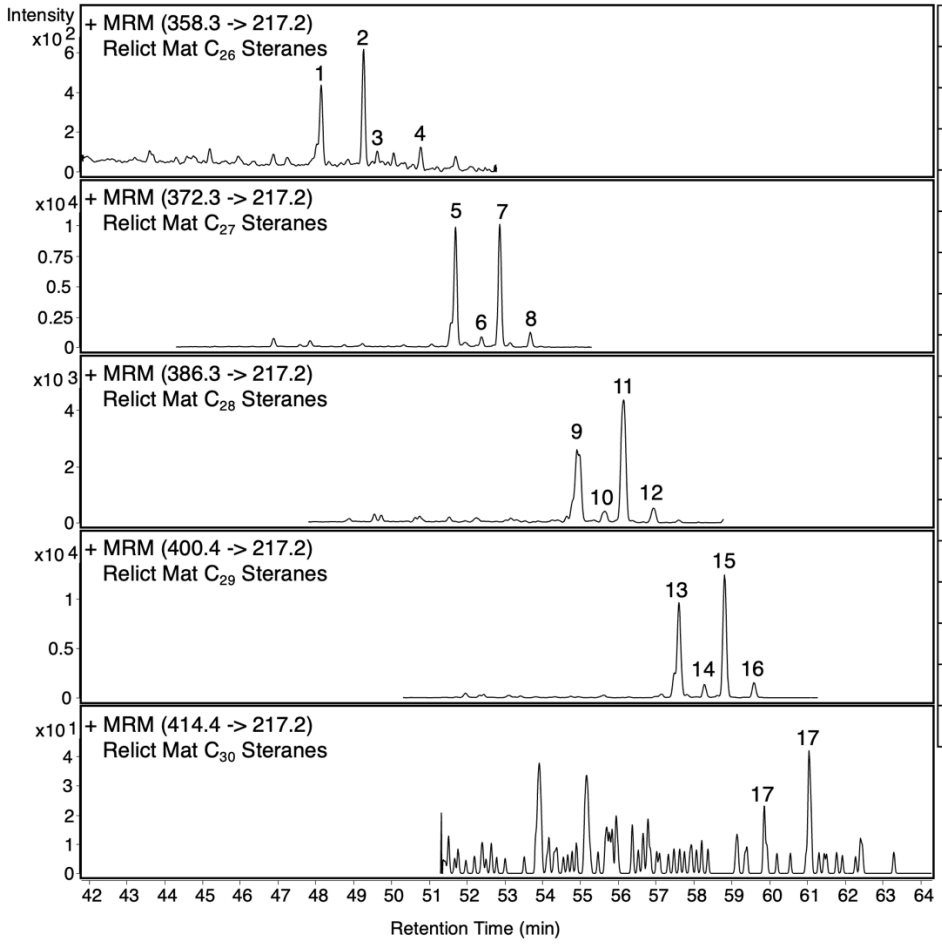
Supplementary Fig. 11. **dMRM chromatograms showing steranes from Pond Seventy sterol hydrogenation.** Numerically labelled peaks are accompanied by peak identities.

4 | ααα20R/βαα20R ratios for C_{27}, C_{28}, and C_{29} steranes



Supplementary Fig. 12. ααα20R/βαα20R ratios for C_{27}, C_{28}, and C_{29} steranes. Ratios calculated using dMRM responses of the ααα20R isomers. Certain names are abbreviated: Cono is short for Conophyton, and Casten is short for Castenholz.

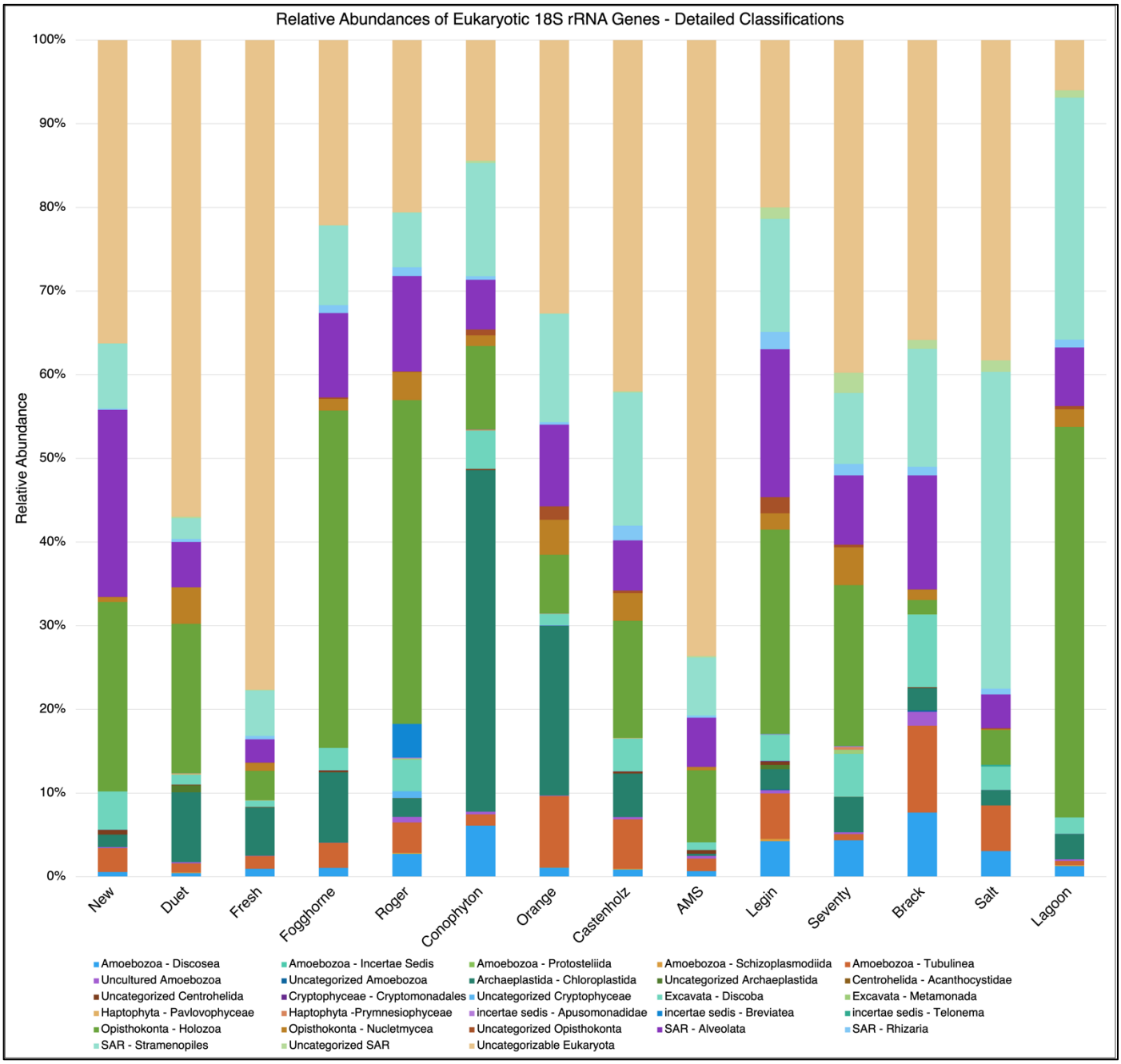
5 | dMRM Chromatograms of the Relict Microbial Mat Hydrocarbon Fraction



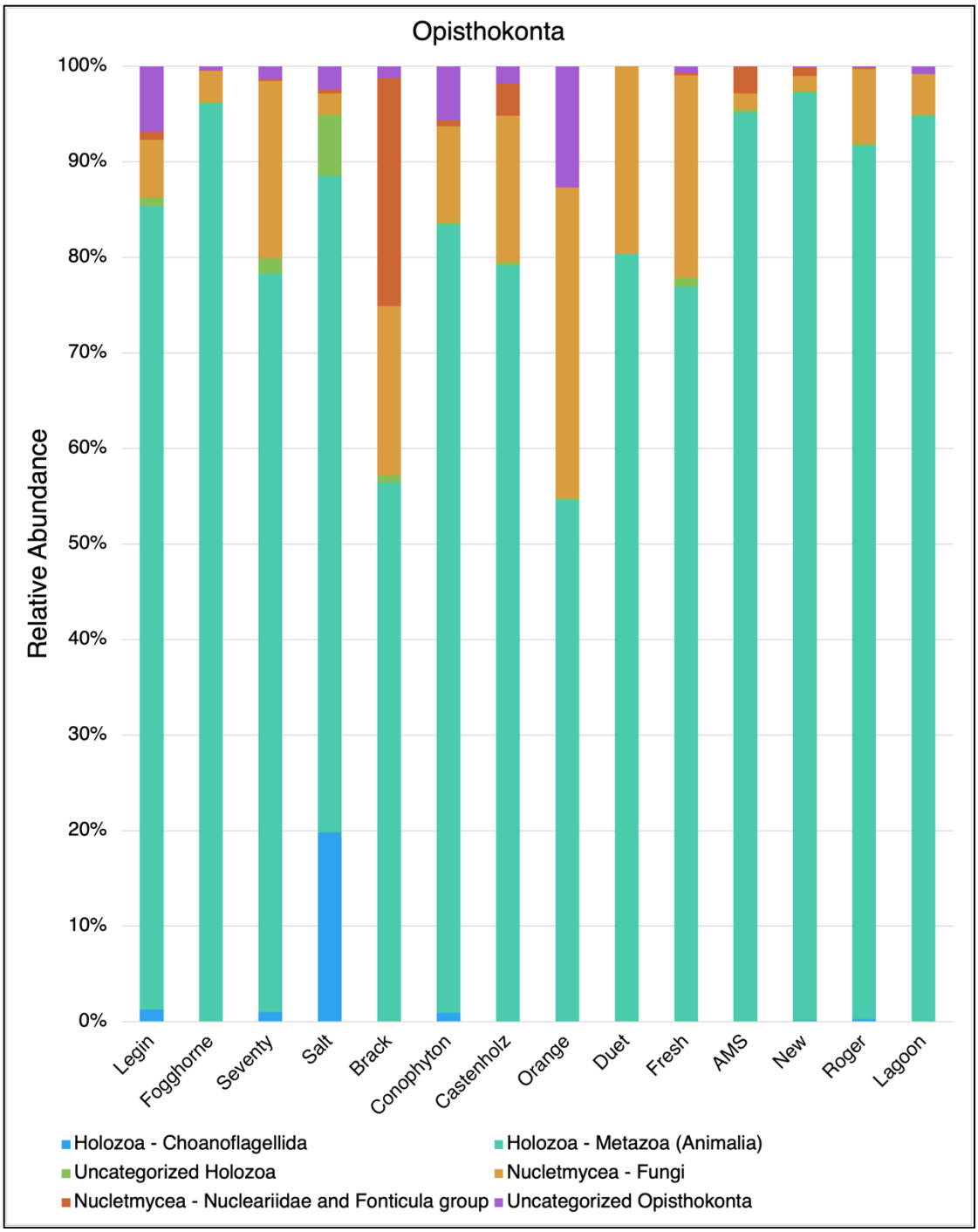
Number	Peak Identity
1	24-norcholestane βαα20R
2	24-norcholestane ααα20R
3	27-norcholestane βαα20R
4	27-norcholestane ααα20R
5	cholestane βαα20R
6	C_{27} norsterane βαα20R
7	cholestane ααα20R
8	C_{27} norsterane ααα20R
9	24-methylcholestane βαα20R
10	C_{28} norsterane βαα20R
11	24-methylcholestane ααα20R
12	C_{28} norsterane ααα20R
13	24-ethylcholestane βαα20R
14	C_{29} norsterane βαα20R
15	24-ethylcholestane ααα20R
16	C_{29} norsterane ααα20R
17	4-methylsteranes

Supplementary Fig. 13. dMRM chromatograms showing steranes from the relict microbial mat hydrocarbon fraction. Numerically labelled peaks are accompanied by peak identities.

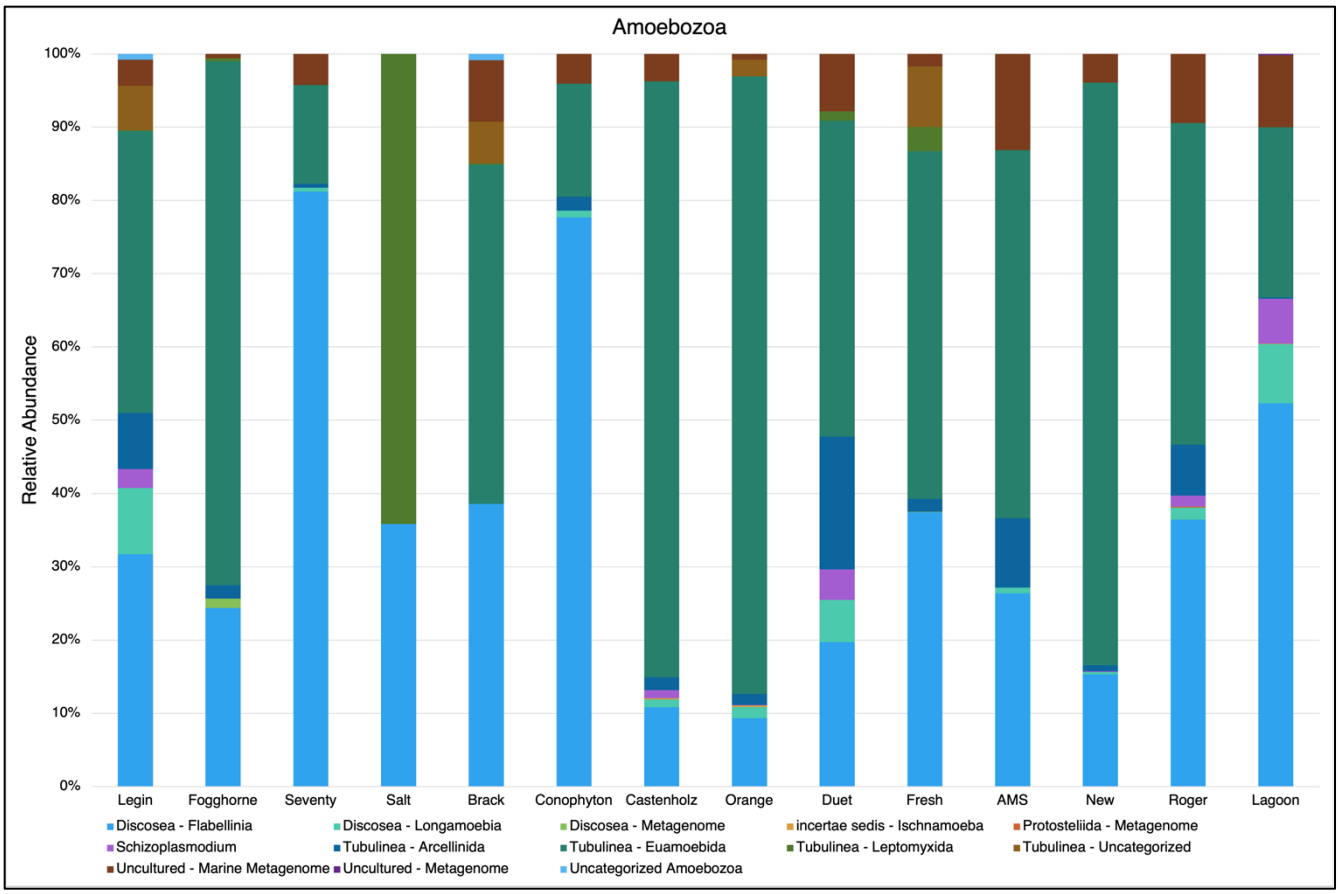
6 | Relative Abundances of 18S rRNA Gene Communities obtained from Meltwater Pond Microbial Mats from the McMurdo Ice Shelf, Antarctica



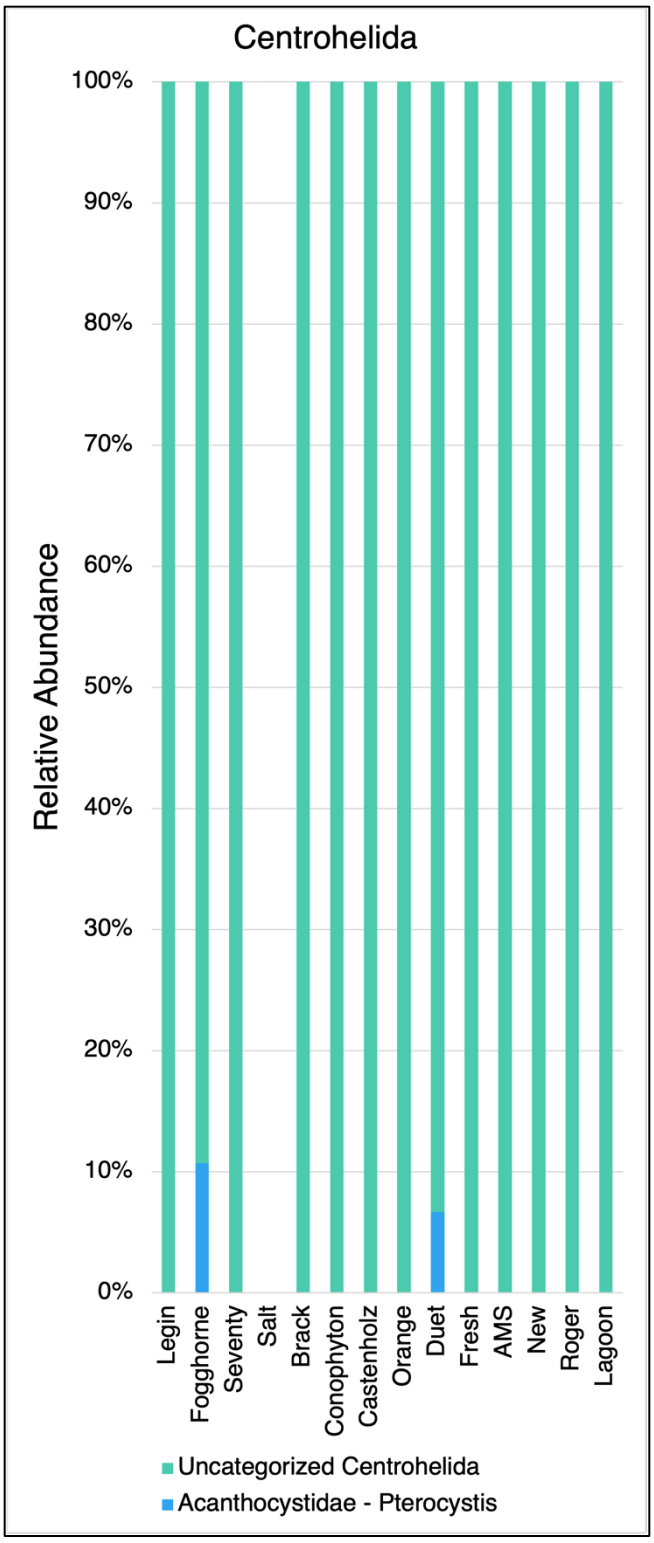
Supplementary Fig. 14. Relative abundances of 18S rRNA gene communities with detailed classifications. The distributions of 18S rRNA gene sequences from eukaryotic groups are shown and labelled for the ponds and the Bratina Lagoon. Source data are provided as a Source Data file.



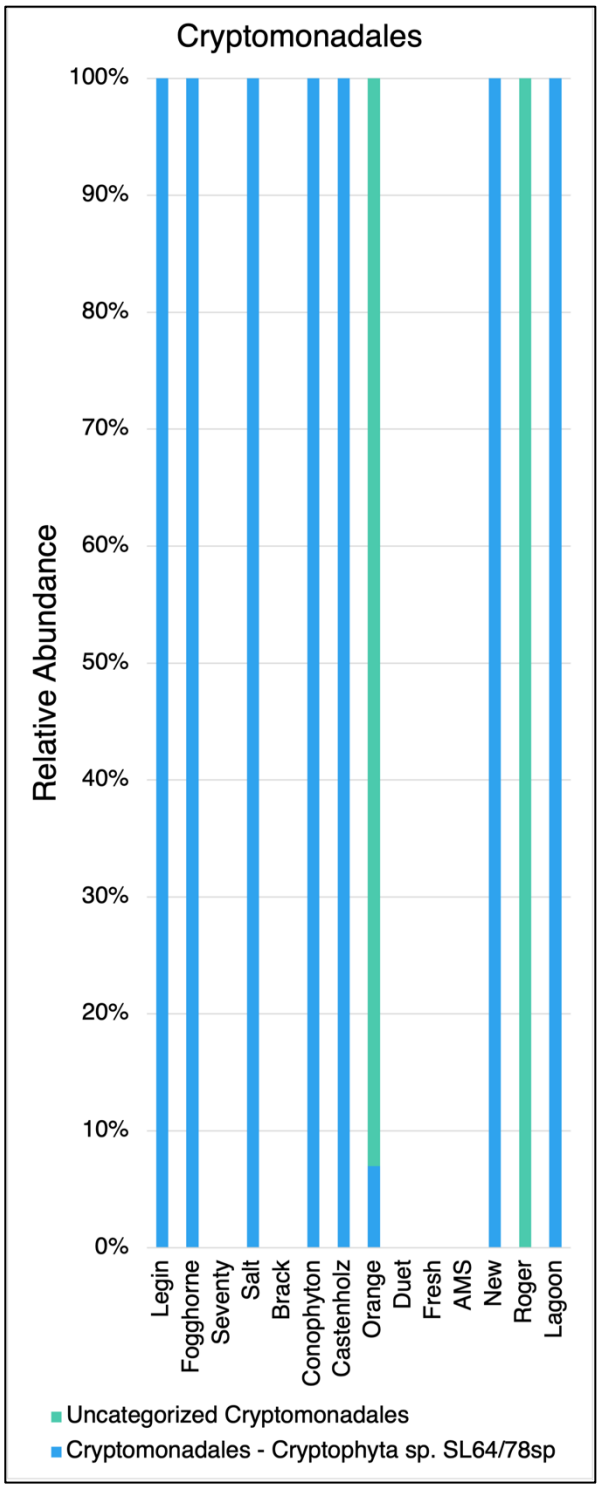
Supplementary Fig. 15. Relative abundances of microbial mat Opisthokonta communities. The distributions of 18S rRNA gene sequences assigned to Opisthokonta are shown and labelled for the ponds and the Bratina Lagoon. Source data are provided as a Source Data file.



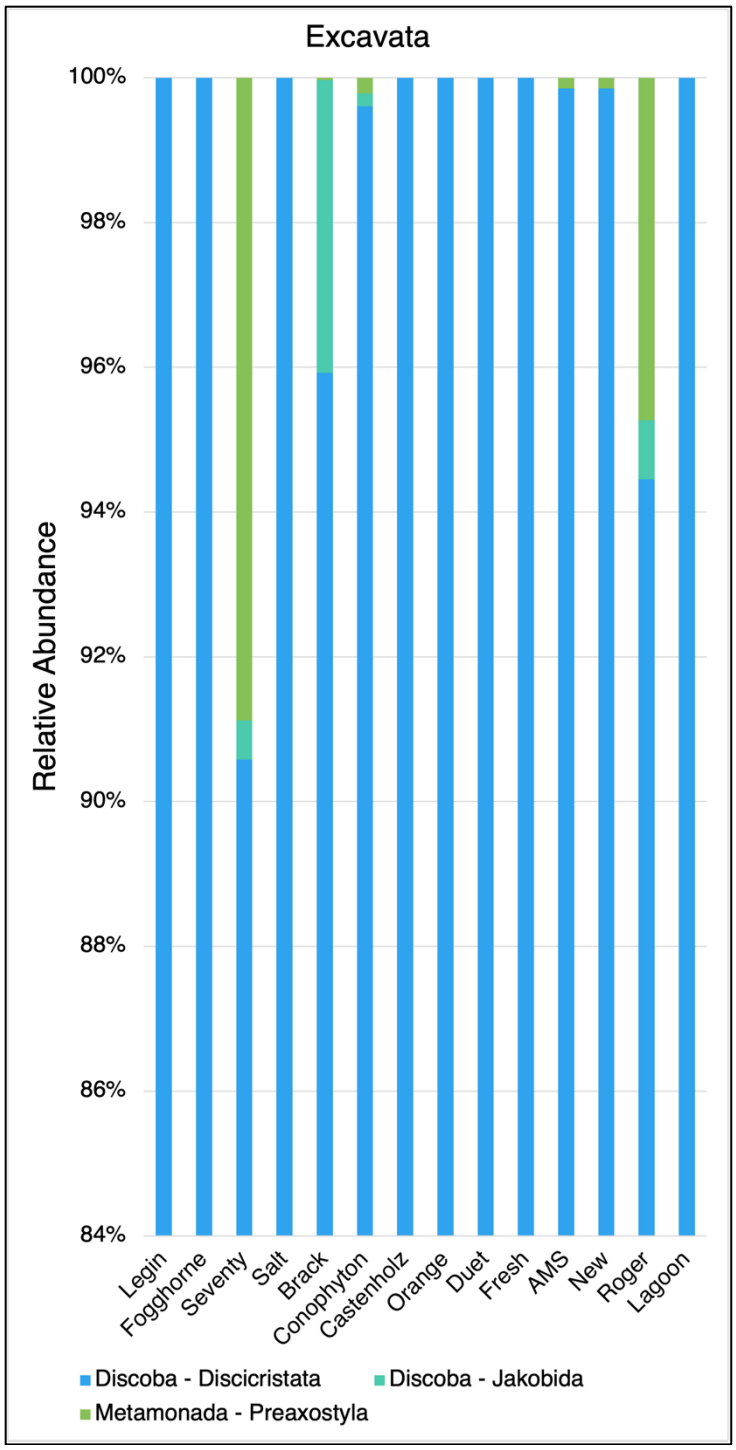
Supplementary Fig. 16. Relative abundances of microbial mat Amoebozoa communities. The distributions of 18S rRNA gene sequences assigned to Amoebozoa are shown and labelled for the ponds and the Bratina Lagoon. Source data are provided as a Source Data file.



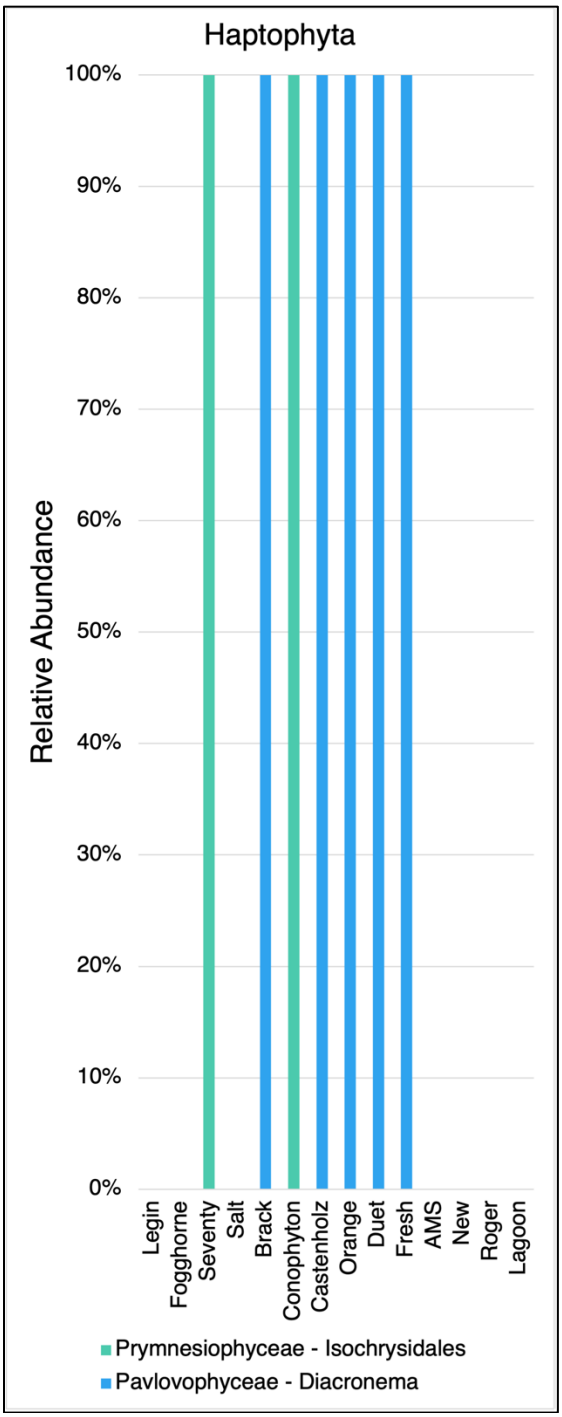
Supplementary Fig. 17. Relative abundances of microbial mat Centrohelida communities. The distributions of 18S rRNA gene sequences assigned to Centrohelida are shown and labelled for the ponds and the Bratina Lagoon. Source data are provided as a Source Data file.



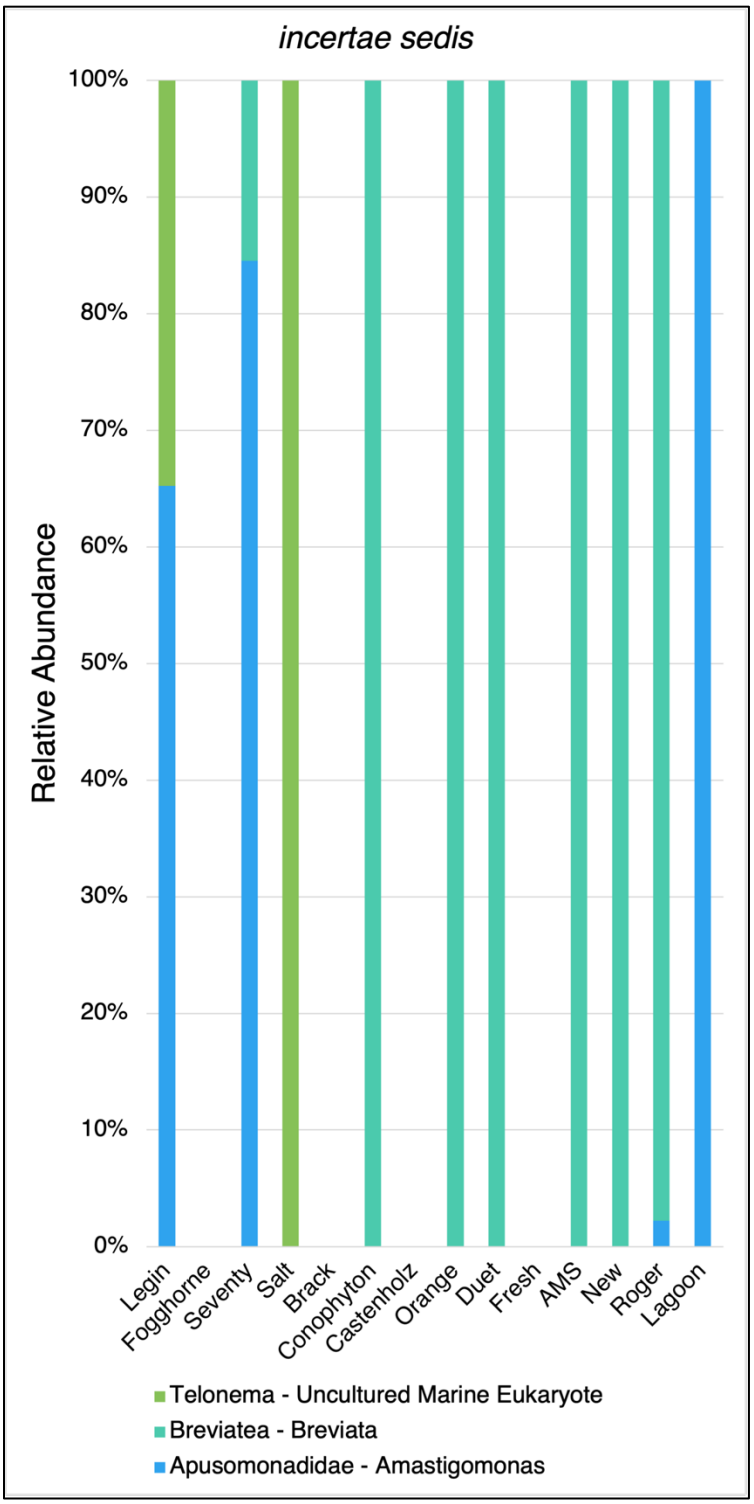
Supplementary Fig. 18. Relative abundances of microbial mat Cryptophyceae communities. The distributions of 18S rRNA gene sequences assigned to Cryptophyceae are shown and labelled for the ponds and the Bratina Lagoon. Source data are provided as a Source Data file.



Supplementary Fig. 19. Relative abundances of microbial mat Excavata communities. The distributions of 18S rRNA gene sequences assigned to Excavata are shown and labelled for the ponds and the Bratina Lagoon. Source data are provided as a Source Data file.



Supplementary Fig. 20. Relative abundances of microbial mat Haptophyta communities. The distributions of 18S rRNA gene sequences assigned to Haptophyta are shown and labelled for the ponds and the Bratina Lagoon. Source data are provided as a Source Data file.



Supplementary Fig. 21. Relative abundances of microbial mat *incertae sedis* communities. The distributions of 18S rRNA gene sequences assigned to *incertae sedis* are shown and labelled for the ponds and the Bratina Lagoon. Source data are provided as a Source Data file.

Supplementary References

1. Atwood, A. R., Volkman, J. K. & Sachs, J. P. Characterization of unusual sterols and long chain diols, triols, keto-ols and n-alkenols in El Junco Lake, Galápagos. *Org. Geochem.* **66**, 80–89 (2014).
2. Harvey, D. J. & Vouros, P. Mass Spectrometric Fragmentation of Trimethylsilyl and Related Alkylsilyl Derivatives. *Mass Spectrom. Rev.* **39**, 105–211 (2020).