Antarctic biodiversity predictions through substrate qualities and environmental DNA

Paul Czechowski, Michel de Lange, Marks Stevens, and Aleks Terauds

February 16, 2021

1 Supplemental information

1.1 Sequence data generation

- Czechowski, Clarke, et al. (2016) and Czechowski, White, et al. (2016) describe sequence data generation in high detail.
- Table 1 and Fig. 2 describe project sequencing effort.
- Table 3 lists of all analyzed libraries and their collection locations, excluding control reactions.

1.2 Data filtering and taxonomic assignments

- Our eDNA data reprocessing started with 16 524 031 sequences, which collectively represented 2 656 eukaryotes, as well as 5 441 non-eukaryotes (i.e., bacterial, archean or undetermined) ASVs. Unfiltered sample mean (min., med., max.) coverage was 107 299 reads (1 108 773 692 828), unfiltered ASV mean (min, median, max) coverage was 2 041 reads (2 70, 1 491 686). After application of package decontam some ASVs remained in control reactions, prior to their subtraction from field data: one ASV of indefinable origin in the extraction blanks, and 51 bacterial and eukaryote ASVs in the amplification blanks.
- Please refer to the external Megan (Huson et al., 2016) file for alignment qualities.
- Fig. 1 provides an overview of taxonomic composition before, during, and after data filtering and for sequencing controls after the application of package *decontam* (Davis et al., 2018).
- Example references for the five highest-covered species among our data are:
    - *Acanthothecis fontana*: Muscavitch and Lendemer (2016)
    - *Coccomyxa* sp.: Blanc et al. (2012)
    - *Mrakia frigida*: Xin and Zhou (2007)
    - *Pseudochilodonopsis quadrivacuolata*: Qu et al. (2015)
    - *Scottnema lindsayae*: Velasco-Castrillón et al. (2014)

- *Embata laticeps*: Myers (1931)
- *Mesobiotus furciger*: Velasco-Castrillón et al. (2014)

1.3 Environmental predictors

- Fig. 3 shows climatic predictor information, summarized per location, as explored in this work.
- Fig. 4 shows abiotic predictor information, summarized per location, after data filtering.
- Fig. 5 shows abiotic predictor information, summarized per location, after data filtering.
- Table 2 and Figs. 6, 7 show full results of regression analysis.

1.4 Analysis code

- Analysis code and other resources beyond the provided materials are available via https://github.com/OldMortality/eukaryotes.

References

Blanc, Guillaume et al. (2012). "The genome of the polar eukaryotic microalga Coccomyxa subellipsoidea reveals traits of cold adaptation." In: *Genome Biology* 13.5, R39. ISSN: 1465-6906. DOI: 10.1186/gb-2012-13-5-r39. URL: http://genomebiology.biomedcentral.com/articles/10.1186/gb-2012-13-5-r39.

Bolyen, Evan et al. (July 2019). "Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2." In: *Nature Biotechnology.* ISSN: 1087-0156. DOI: 10.1038/s41587-019-0209-9. URL: http://www.nature.com/articles/s41587-019-0209-9.

Callahan, Benjamin J., Paul J. McMurdie, and Susan P. Holmes (Dec. 2017). "Exact sequence variants should replace operational taxonomic units in marker-gene data analysis." In: *The ISME Journal* 11.12, p. 113597. ISSN: 1751-7362. DOI: 10.1038/ismej.2017.119. URL: http://dx.doi.org/10.1038/ismej.2017.119%20http://www.nature.com/doifinder/10.1038/ismej.2017.119%20http://biorxiv.org/content/early/2017/03/07/113597.

Czechowski, Paul, Laurence J Clarke, et al. (Apr. 2016). "Antarctic eukaryotic soil diversity of the Prince Charles Mountains revealed by high-throughput sequencing." In: *Soil Biology and Biochemistry* 95.4, pp. 112–121. ISSN: 00380717. DOI: 10.1016/j.soilbio.2015.12.013. URL: http://dx.doi.org/10.1016/j.soilbio.2015.12.013%20http://linkinghub.elsevier.com/retrieve/pii/S0038071715004459%20https://linkinghub.elsevier.com/retrieve/pii/S0038071715004459.

Czechowski, Paul, Duanne White, et al. (Dec. 2016). "Age-related environmental gradients influence invertebrate distribution in the Prince Charles Mountains, East Antarctica." In: *Royal Society Open Science* 3.12, p. 160296. ISSN: 2054-5703. DOI: 10.1098/rsos.160296. URL: https://royalsocietypublishing.org/doi/10.1098/rsos.160296.

Davis, Nicole M. et al. (Dec. 2018). "Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data." In: *Microbiome* 6.1, p. 226. ISSN: 2049-2618. DOI: 10.1186/s40168-018-0605-2. URL: https://microbiomejournal.biomedcentral.com/articles/10.1186/s40168-018-0605-2.

Huson, Daniel H. et al. (June 2016). “MEGAN Community Edition - Interactive Exploration and Analysis of Large-Scale Microbiome Sequencing Data.” In: *PLOS Computational Biology* 12.6. Ed. by Timothée Poisot, e1004957. ISSN: 1553-7358. DOI: 10.1371/journal.pcbi.1004957. URL: https://dx.plos.org/10.1371/journal.pcbi.1004957.

Muscavitch, Zachary M. and James C. Lendemer (2016). “A new species of Acanthothecis (Ostropales), highlights subtropical floristic elements of the southern Appalachian lichen biota in eastern North America.” In: *The Bryologist* 119.4, p. 350. ISSN: 0007-2745. DOI: 10.1639/0007-2745-119.4.350.

Myers, Frank J. (1931). “The distribution of rotifera on Mount Desert Island.” In: *American Museum Novitates* 494, pp. 1–12. URL: http://scholar.google.com/scholar?hl=en%7B%5C&%7DbtnG=Search%7B%5C&%7Dq=intitle:The+distribution+of+rotifera+on+mount+desert+island%7B%5C#%7D0.

Qu, Zhishuai et al. (2015). “Morphological and phylogenetic studies on three members of the genus pseudochilodonopsis (Ciliophora, cyrtophoria) isolated from brackish waters in china, including a novel species, pseudochilodonopsis quadrivacuolata sp. nov.” In: *International Journal of Systematic and Evolutionary Microbiology* 65.12, pp. 4323–4334. ISSN: 14665026. DOI: 10.1099/ijsem.0.000580.

Van Wessem, J. M. et al. (2014). “Improved representation of East Antarctic surface mass balance in a regional atmospheric climate model.” In: *Journal of Glaciology* 60.222, pp. 761–770. ISSN: 00221430. DOI: 10.3189/2014JoG14J051.

Velasco-Castrillón, Alejandro et al. (Jan. 2014). “Distribution and diversity of soil microfauna from East Antarctica: Assessing the link between biotic and abiotic factors.” In: *PLoS ONE* 9.1. Ed. by Xiujun Wang, e87529. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0087529. URL: http://dx.plos.org/10.1371/journal.pone.0087529.

Wessem, J. M. van et al. (Jan. 2014). “Updated cloud physics in a regional atmospheric climate model improves the modelled surface energy balance of Antarctica.” In: *The Cryosphere* 8.1, pp. 125–135. ISSN: 1994-0424. DOI: 10.5194/tc-8-125-2014. URL: https://tc.copernicus.org/articles/8/125/2014/.

Xin, Ming-xiu and Pei-jin Zhou (Apr. 2007). “Mrakia psychrophila sp. nov., a new species isolated from Antarctic soil.” In: *Journal of Zhejiang University SCIENCE B* 8.4, pp. 260–265. ISSN: 1673-1581. DOI: 10.1631/jzus.2007.B0260. URL: http://link.springer.com/10.1631/jzus.2007.B0260.

Table 2: Signs of coefficient estimates for each predictor resulting from lasso logistic regression. Signs indicate changes at Mawson Escarpment and Mount Menzies in comparison to the reference location (Lake Terrasovoje). X-ray spectral peaks of certain wavelength are compounds of mineral group as indicated by asterisk (*).

Phylym	Potassium	Sulphur	Conductivity	pH of CaCl2	ATP fluorescence	Feldspar	Titanite	Pyroxene / Amphibole / Garnet *	Micas	Dolomite	Kaolin / Chlorite*	Calcite	Chlorite	Slope
Arthropoda						+			-	-				
Ascomycota				-	+		+							
Bacillariophyta			-					-	-		+			-
Basidiomycota		-		-		+	+			+			-	
Cercozoa		-		-								-	-	
Chlorophyta	-			-							+	+		-
Chytridiomycota			-	-	-			-	-	+	+			
Ciliophora		-	-					-	-		+			
Nematoda		-	-	+										
Rotifera	-			-		+	+							
Streptophyta											-			

Table 1: Project sequencing effort and obtained read to yield unfiltered data (including controls), across all sequencing runs.

Run	Input	Filtered	Denoised	Non-chimeric	Libraries	Avg. Seq./Lib.
A	10 055 133	9 771 027	9 759 212	9 667 648	92	105 083
B	10 061 999	9 872 492	9 856 147	9 725 926	91	106 818
	20 117 132	19 643 519	19 615 359	19 393 574	183	105 975

Table 3: Sequence library names and locations of filtered data, as used to generate Figure 1 in main text.

Sample	Location	Longitude	Latitude
18S.1.1.C.PCM.MM	Mount Menzies	62.02444444	-73.38888889
18S.1.10.A.PCM.MM	Mount Menzies	61.88722222	-73.42555556
18S.1.10.B.PCM.MM	Mount Menzies	62.01055556	-73.39277778
18S.1.10.C.PCM.MM	Mount Menzies	62.12638889	-73.43972222
18S.1.11.A.PCM.MM	Mount Menzies	61.84666667	-73.42916667
18S.1.11.B.PCM.MM	Mount Menzies	62.15555556	-73.43222222
18S.1.11.C.PCM.MM	Mount Menzies	61.8775	-73.42027778
18S.1.2.B.PCM.MM	Mount Menzies	61.84388889	-73.41777778
18S.1.2.C.PCM.MM	Mount Menzies	62.14972222	-73.42333333

Table 3: Sequence library names and locations of filtered data, as used to generate Figure 1 in main text. (continued)

18S.1.3.B.PCM.MM	Mount Menzies	61.84305556	-73.42333333
18S.1.3.C.PCM.MM	Mount Menzies	62.04222222	-73.3875
18S.1.4.B.PCM.MM	Mount Menzies	61.85916667	-73.43944444
18S.1.5.B.PCM.MM	Mount Menzies	61.98027778	-73.42444444
18S.1.5.C.PCM.MM	Mount Menzies	62.09611111	-73.43722222
18S.1.6.B.PCM.MM	Mount Menzies	61.95666667	-73.42583333
18S.1.6.C.PCM.MM	Mount Menzies	62.09055556	-73.43666667
18S.1.7.A.PCM.MM	Mount Menzies	61.87111111	-73.42277778
18S.1.7.B.PCM.MM	Mount Menzies	62.01361111	-73.45111111
18S.1.7.C.PCM.MM	Mount Menzies	62.12305556	-73.44805556
18S.1.8.A.PCM.MM	Mount Menzies	62.02166667	-73.42888889
18S.1.8.B.PCM.MM	Mount Menzies	61.94027778	-73.42916667
18S.1.9.A.PCM.MM	Mount Menzies	62.05361111	-73.4
18S.1.9.B.PCM.MM	Mount Menzies	62.15305556	-73.44333333
18S.1.1.D.PCM.ME	Mawson Escarpment	68.31888889	-73.32027778
18S.1.1.F.PCM.ME	Mawson Escarpment	68.40666667	-73.30666667
18S.1.1.G.PCM.ME	Mawson Escarpment	68.38722222	-73.31722222
18S.1.1.H.PCM.ME	Mawson Escarpment	68.50194444	-73.30833333
18S.1.10.D.PCM.ME	Mawson Escarpment	68.45111111	-73.3075
18S.1.10.E.PCM.ME	Mawson Escarpment	68.39861111	-73.30416667
18S.1.10.F.PCM.ME	Mawson Escarpment	68.41111111	-73.31916667
18S.1.10.G.PCM.ME	Mawson Escarpment	68.50361111	-73.31055556
18S.1.11.D.PCM.ME	Mawson Escarpment	68.44666667	-73.30916667
18S.1.11.E.PCM.ME	Mawson Escarpment	68.43194444	-73.30777778
18S.1.11.G.PCM.ME	Mawson Escarpment	68.35361111	-73.3275
18S.1.2.D.PCM.ME	Mawson Escarpment	68.33555556	-73.31777778
18S.1.2.E.PCM.ME	Mawson Escarpment	68.44972222	-73.30444444
18S.1.2.F.PCM.ME	Mawson Escarpment	68.43194444	-73.30888889
18S.1.2.G.PCM.ME	Mawson Escarpment	68.38111111	-73.3175
18S.1.2.H.PCM.ME	Mawson Escarpment	68.34416667	-73.3275
18S.1.3.D.PCM.ME	Mawson Escarpment	68.32388889	-73.32
18S.1.3.E.PCM.ME	Mawson Escarpment	68.44666667	-73.3075
18S.1.3.H.PCM.ME	Mawson Escarpment	68.35416667	-73.32833333
18S.1.4.D.PCM.ME	Mawson Escarpment	68.33416667	-73.31888889
18S.1.4.E.PCM.ME	Mawson Escarpment	68.44166667	-73.30166667
18S.1.4.F.PCM.ME	Mawson Escarpment	68.46666667	-73.3025
18S.1.4.G.PCM.ME	Mawson Escarpment	68.4375	-73.32027778
18S.1.4.H.PCM.ME	Mawson Escarpment	68.37555556	-73.33
18S.1.5.D.PCM.ME	Mawson Escarpment	68.31638889	-73.32444444

Table 3: Sequence library names and locations of filtered data, as used to generate Figure 1 in main text. (continued)

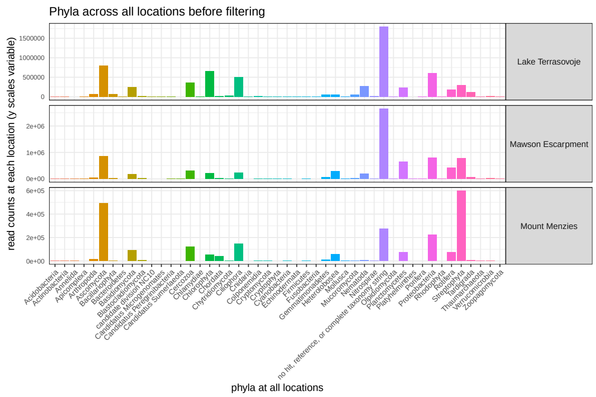
18S.1.5.E.PCM.ME	Mawson Escarpment	68.46055556	-73.30361111
18S.1.5.F.PCM.ME	Mawson Escarpment	68.42333333	-73.30916667
18S.1.5.H.PCM.ME	Mawson Escarpment	68.36805556	-73.32805556
18S.1.6.D.PCM.ME	Mawson Escarpment	68.14138889	-72.92361111
18S.1.6.E.PCM.ME	Mawson Escarpment	68.46555556	-73.30027778
18S.1.6.F.PCM.ME	Mawson Escarpment	68.40555556	-73.30805556
18S.1.6.G.PCM.ME	Mawson Escarpment	68.48722222	-73.30555556
18S.1.6.H.PCM.ME	Mawson Escarpment	68.37555556	-73.33
18S.1.7.D.PCM.ME	Mawson Escarpment	68.30138889	-73.32611111
18S.1.7.E.PCM.ME	Mawson Escarpment	68.35361111	-73.31583333
18S.1.7.F.PCM.ME	Mawson Escarpment	68.42666667	-73.31638889
18S.1.7.G.PCM.ME	Mawson Escarpment	68.36277778	-73.32333333
18S.1.7.H.PCM.ME	Mawson Escarpment	68.38444444	-73.32694444
18S.1.8.D.PCM.ME	Mawson Escarpment	68.34111111	-73.3225
18S.1.8.E.PCM.ME	Mawson Escarpment	68.38472222	-73.31055556
18S.1.8.F.PCM.ME	Mawson Escarpment	68.42166667	-73.31083333
18S.1.8.G.PCM.ME	Mawson Escarpment	68.36055556	-73.32333333
18S.1.8.H.PCM.ME	Mawson Escarpment	68.38055556	-73.31777778
18S.1.9.D.PCM.ME	Mawson Escarpment	68.42333333	-73.30555556
18S.1.9.E.PCM.ME	Mawson Escarpment	68.39638889	-73.2775
18S.1.9.F.PCM.ME	Mawson Escarpment	68.38944444	-73.31444444
18S.1.9.G.PCM.ME	Mawson Escarpment	68.5	-73.30972222
18S.1.9.H.PCM.ME	Mawson Escarpment	68.37222222	-73.31388889
18S.2.1.A.PCM.ME	Mawson Escarpment	68.35555556	-73.31583333
18S.2.1.B.PCM.ME	Mawson Escarpment	66.46972222	-73.01972222
18S.2.10.A.PCM.ME	Mawson Escarpment	66.445	-73.01194444
18S.2.2.A.PCM.ME	Mawson Escarpment	68.35833333	-73.31416667
18S.2.2.B.PCM.ME	Mawson Escarpment	68.41638889	-73.48416667
18S.2.2.H.PCM.ME	Mawson Escarpment	68.33944444	-73.32111111
18S.2.3.A.PCM.ME	Mawson Escarpment	68.49055556	-73.30888889
18S.2.4.A.PCM.ME	Mawson Escarpment	68.24444444	-73.29861111
18S.2.4.B.PCM.ME	Mawson Escarpment	66.79472222	-74.29027778
18S.2.5.A.PCM.ME	Mawson Escarpment	68.24777778	-73.30888889
18S.2.5.B.PCM.ME	Mawson Escarpment	65.60944444	-73.39638889
18S.2.6.A.PCM.ME	Mawson Escarpment	68.31777778	-73.31972222
18S.2.6.B.PCM.ME	Mawson Escarpment	68.15	-73.19611111
18S.2.7.A.PCM.ME	Mawson Escarpment	68.04166667	-72.81944444
18S.2.8.A.PCM.ME	Mawson Escarpment	68.47	-73.63777778
18S.2.9.A.PCM.ME	Mawson Escarpment	68.52944444	-73.64388889

Table 3: Sequence library names and locations of filtered data, as used to generate Figure 1 in main text. (continued)

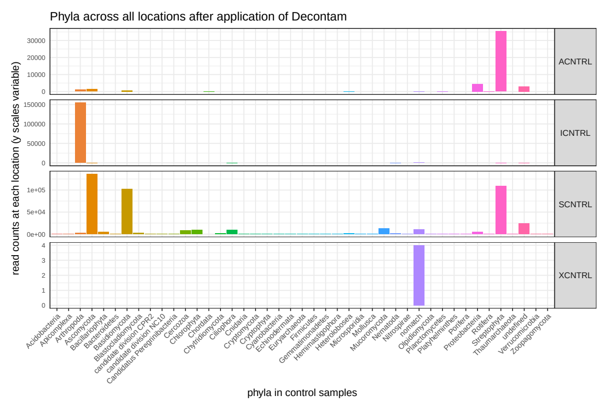
18S.2.1.C.PCM.LT	Lake Terrasovoje	67.88055556	-70.54194444
18S.2.1.D.PCM.LT	Lake Terrasovoje	67.89666667	-70.53083333
18S.2.1.E.PCM.LT	Lake Terrasovoje	67.94694444	-70.55916667
18S.2.1.F.PCM.LT	Lake Terrasovoje	67.79777778	-70.525
18S.2.1.G.PCM.LT	Lake Terrasovoje	68.2125	-70.82833333
18S.2.1.H.PCM.LT	Lake Terrasovoje	68.28472222	-73.41277778
18S.2.10.B.PCM.LT	Lake Terrasovoje	68.01638889	-70.51222222
18S.2.10.C.PCM.LT	Lake Terrasovoje	68.01388889	-70.53305556
18S.2.10.D.PCM.LT	Lake Terrasovoje	67.88972222	-70.52361111
18S.2.10.E.PCM.LT	Lake Terrasovoje	67.98361111	-70.54888889
18S.2.10.F.PCM.LT	Lake Terrasovoje	68.07111111	-70.84472222
18S.2.10.G.PCM.LT	Lake Terrasovoje	68.80805556	-70.42416667
18S.2.11.B.PCM.LT	Lake Terrasovoje	67.82555556	-70.52722222
18S.2.11.C.PCM.LT	Lake Terrasovoje	69.00805556	-70.53388889
18S.2.11.D.PCM.LT	Lake Terrasovoje	67.99305556	-70.54305556
18S.2.11.E.PCM.LT	Lake Terrasovoje	67.96916667	-70.54333333
18S.2.11.F.PCM.LT	Lake Terrasovoje	68.0425	-70.84472222
18S.2.2.C.PCM.LT	Lake Terrasovoje	67.83305556	-70.53027778
18S.2.2.D.PCM.LT	Lake Terrasovoje	68.00222222	-70.52888889
18S.2.2.E.PCM.LT	Lake Terrasovoje	67.93944444	-70.55694444
18S.2.2.F.PCM.LT	Lake Terrasovoje	67.81083333	-70.52666667
18S.2.2.G.PCM.LT	Lake Terrasovoje	68.02	-70.80861111
18S.2.3.C.PCM.LT	Lake Terrasovoje	68.05166667	-70.51722222
18S.2.3.D.PCM.LT	Lake Terrasovoje	67.92722222	-70.51694444
18S.2.3.E.PCM.LT	Lake Terrasovoje	67.91611111	-70.54805556
18S.2.3.F.PCM.LT	Lake Terrasovoje	67.82305556	-70.52722222
18S.2.3.G.PCM.LT	Lake Terrasovoje	68.00861111	-70.80916667
18S.2.4.C.PCM.LT	Lake Terrasovoje	68.04222222	-70.52138889
18S.2.4.D.PCM.LT	Lake Terrasovoje	67.89027778	-70.53138889
18S.2.4.E.PCM.LT	Lake Terrasovoje	67.77416667	-70.53972222
18S.2.4.F.PCM.LT	Lake Terrasovoje	67.86777778	-70.52277778
18S.2.4.G.PCM.LT	Lake Terrasovoje	68.17666667	-70.82916667
18S.2.5.C.PCM.LT	Lake Terrasovoje	67.87222222	-70.52666667
18S.2.5.D.PCM.LT	Lake Terrasovoje	68.00361111	-70.52638889
18S.2.5.F.PCM.LT	Lake Terrasovoje	68.03416667	-70.54027778
18S.2.5.G.PCM.LT	Lake Terrasovoje	64.60361111	-70.42944444
18S.2.6.C.PCM.LT	Lake Terrasovoje	67.87361111	-70.52388889
18S.2.6.D.PCM.LT	Lake Terrasovoje	67.99777778	-70.52027778
18S.2.6.E.PCM.LT	Lake Terrasovoje	67.85805556	-70.54583333

Table 3: Sequence library names and locations of filtered data, as used to generate Figure 1 in main text. (continued)

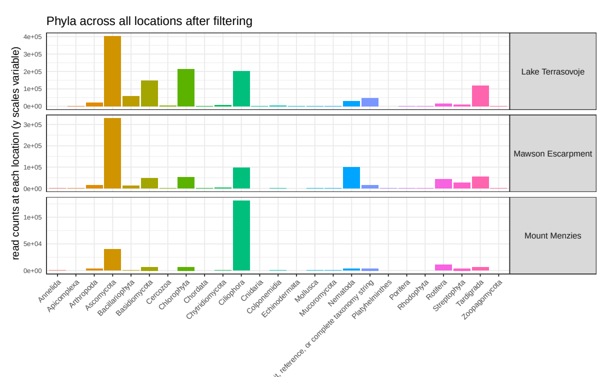
18S.2.6.F.PCM.LT	Lake Terrasovoje	68.04027778	-70.55027778
18S.2.7.C.PCM.LT	Lake Terrasovoje	68.00388889	-70.5175
18S.2.7.D.PCM.LT	Lake Terrasovoje	67.91583333	-70.51833333
18S.2.7.E.PCM.LT	Lake Terrasovoje	67.98805556	-70.58388889
18S.2.7.F.PCM.LT	Lake Terrasovoje	68.17722222	-70.80916667
18S.2.8.B.PCM.LT	Lake Terrasovoje	68.02944444	-70.51277778
18S.2.8.C.PCM.LT	Lake Terrasovoje	67.87861111	-70.52194444
18S.2.8.D.PCM.LT	Lake Terrasovoje	67.91861111	-70.53333333
18S.2.8.E.PCM.LT	Lake Terrasovoje	67.90722222	-70.54611111
18S.2.8.F.PCM.LT	Lake Terrasovoje	68.06777778	-70.81194444
18S.2.8.G.PCM.LT	Lake Terrasovoje	66.895	-70.56916667
18S.2.9.C.PCM.LT	Lake Terrasovoje	67.90611111	-70.54361111
18S.2.9.D.PCM.LT	Lake Terrasovoje	67.98694444	-70.54055556
18S.2.9.E.PCM.LT	Lake Terrasovoje	67.98138889	-70.57527778
18S.2.9.F.PCM.LT	Lake Terrasovoje	68.06472222	-70.83472222
18S.2.9.G.PCM.LT	Lake Terrasovoje	68.3825	-70.80944444



(a) Prior to filtering

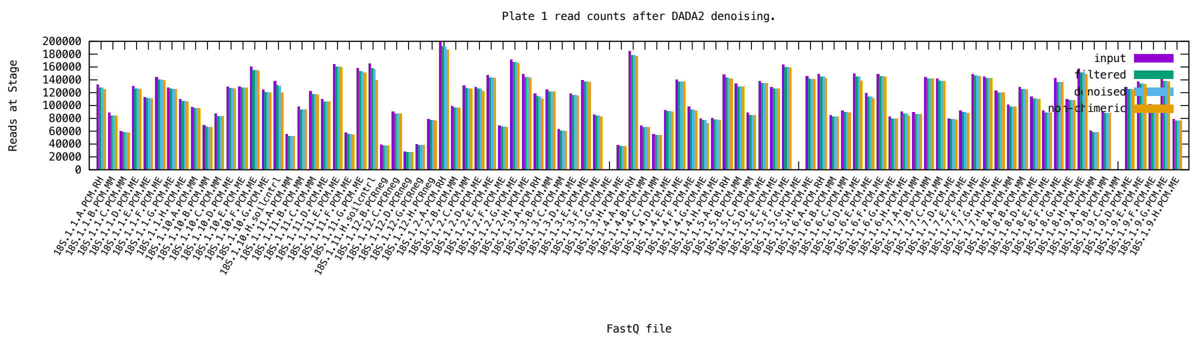


(b) Sequences remaining in control reaction after application of decontamination using R package *decontam* (Davis et al., 2018). ACNTROL: Amplification control; ICNTRL: non-Antarctic insect positive control; SCNTRL: Australian soil positive control; XCNTRL: Extraction control (Czechowski, White, et al., 2016).

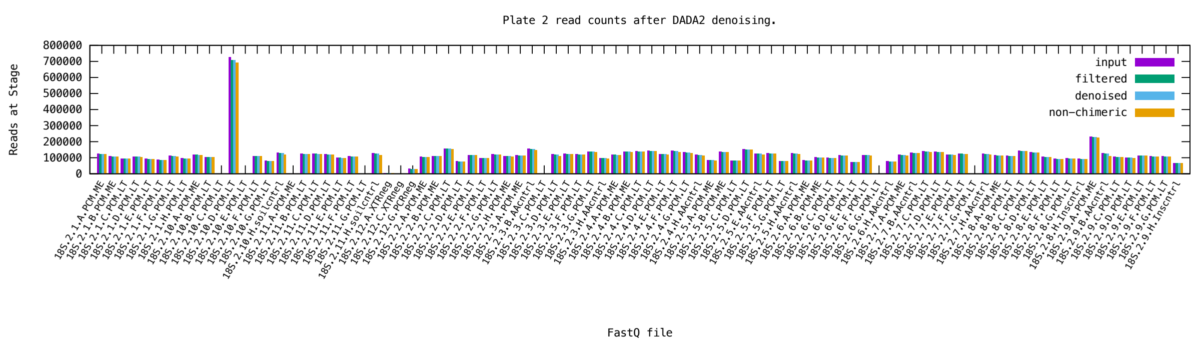


(c) Taxonomic assignments across 2 285 773 environmental 125 bp 18S Amplicon Sequence Variants (*sensu* Callahan, McMurdie, and Holmes, 2017) following NCBI hierarchy. Assignments derived from references sequences at least 50% identical to query sequences, assignment certainty e of 10-10, for matches of at least 90% query coverage, excluding environmental sequences. Collectively, ASVs were assigned to 495 species across 25 phyla.

Figure 1: Taxonomic composition overview of environmental Amplicon Sequence Variants before and after filtering stages. All species and reads contained in controls where removed from remaining data prior to further analysis.



(a) first plate



(b) second plate

Figure 2: Denoising and filtering statistics to obtain raw data as provided by Qiime (Bolyen et al., 2019).

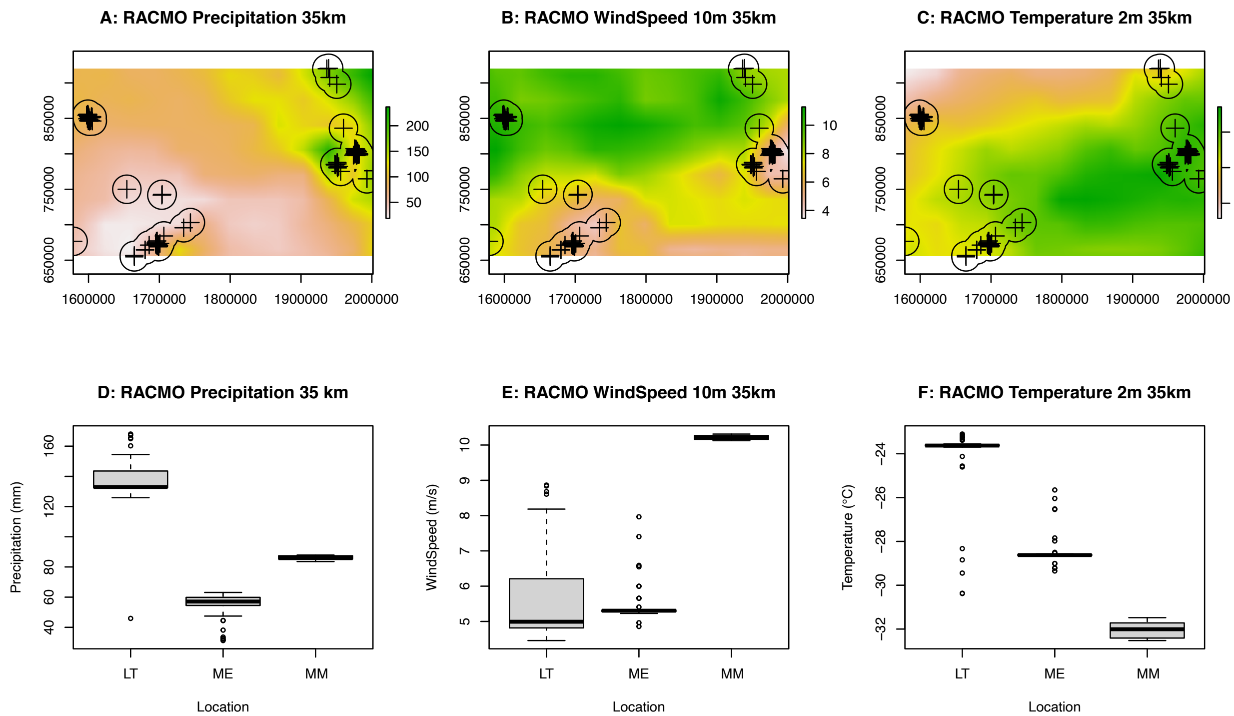


Figure 3: Upper: Climatic variables (Wessem et al., 2014; Van Wessem et al., 2014) in study extent after bilinear desegregation to 1 km pixel size. Indicated are sampling points and 20 km buffers for variable extraction. Lower: Extracted climatic variables per location, the median value in the buffered location was used.

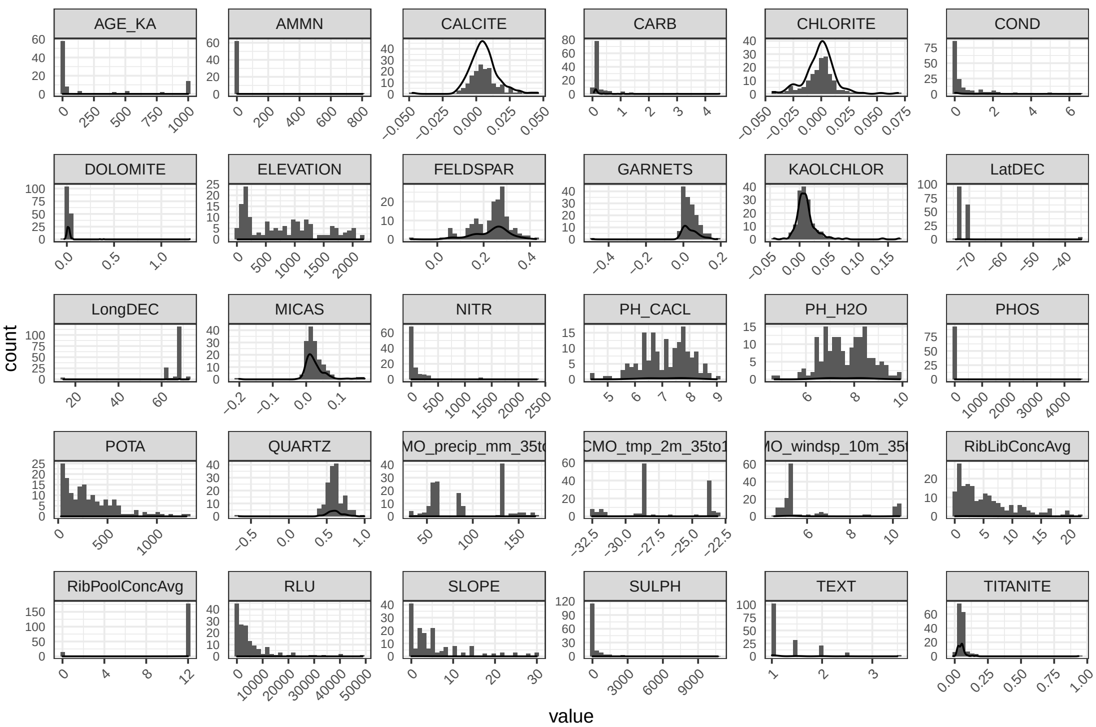
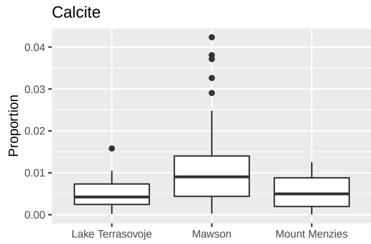
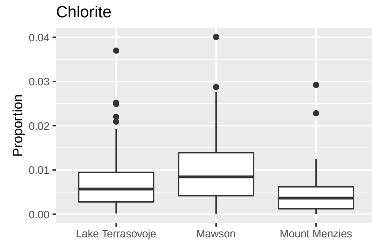


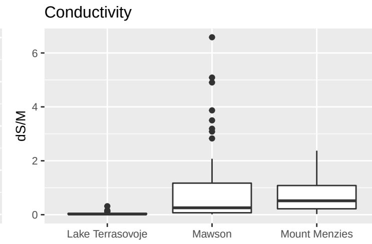
Figure 4: Distributions of raw predictors preceding data filtering. Also shown are predictors that were excluded due to missing data. “AGE_KA”: bedrock age in ka, estimated with geological maps; “AMMN” Ammonium N content (mg/Kg); “CALCITE”: calcite fraction % as integrated from X-ray diffraction spectra; “CARB”: organic carbon in %; “CHLORITE”: chlorite % as integrated from X-ray diffraction spectra; “COND”: conductivity in dS/m; “DOLOMITE”: Dolomite fraction % as integrated from X-ray diffraction spectra; “ELEVATION”: as measured with handheld GPS on site; “GARNET”: pyroxene, amphibole or garnet fraction % as integrated from X-ray diffraction spectra; “KAOCHLOR”: kaolin, chlorite and chlorite fraction % as integrated from X-ray diffraction spectra; “LatDEC and “LongDEC”, latitude and longitude in decimal degrees as measured with handheld GPS (WGS84) on site; “MICAS”: micas fraction % as integrated from X-ray diffraction spectra; “NITR”: Nitrate N (mg/Kg); “PH_CACL” and “PH_H2H”: pH values as measured for calcium dichloride and water; “PHOS” and “POTA”: phosphorus and potassium (mg/kg); “QUARTZ”: quartz % as integrated from X-ray diffraction spectra; “RACMO_precip_mm_35to1km”, “RACMO_windsp_10m_35to1km” and “RACMO_tmp_2m_35to1km”: precipitation, windspeed and temperature as extracted from climate rasters (also see Fig. 3); “RibLibConcAvg”: DNA concentration of pooled library replicates prior to equimolarisation; “RibPoolConcAvg”: DNA concentration of pooled library replicates after equimolarisation; “RLU”: relative light units across two replicates, as measured with an ATP meter during field work; “SLOPE”: Slope angle as measured at sampling site. “SULPH”: Sulphur content (mg/kg); “TEXT” coarseness of substrate. “TTITANITE”: titanite fraction % as integrated from X-ray diffraction spectra.



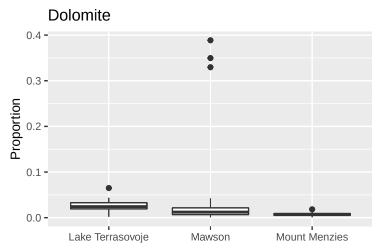
(a) calcite fraction



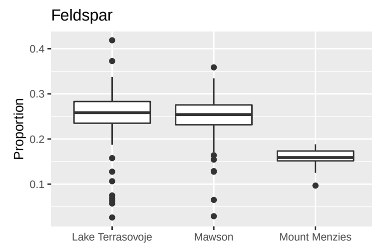
(b) chlorite fraction



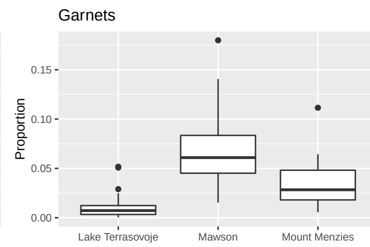
(c) salinity



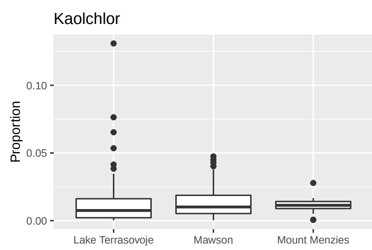
(d) dolomite fraction



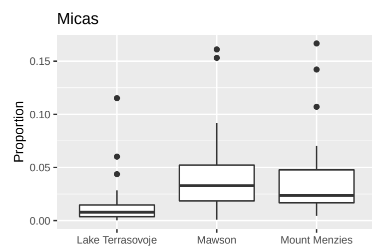
(e) feldspar fraction



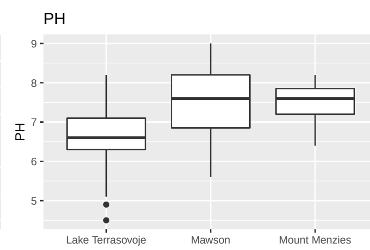
(f) pyroxene / amphibole / garnet fraction



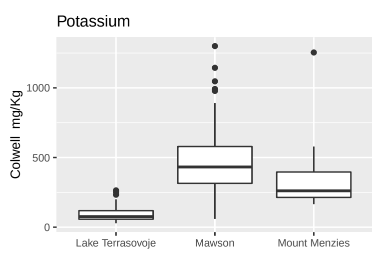
(g) kaolin / chlorite fraction



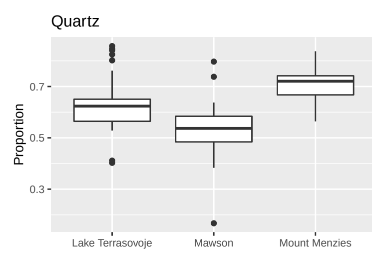
(h) micas fraction



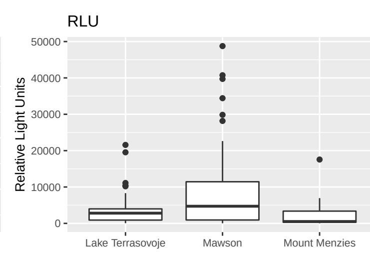
(i) pH of $CaCl_2$



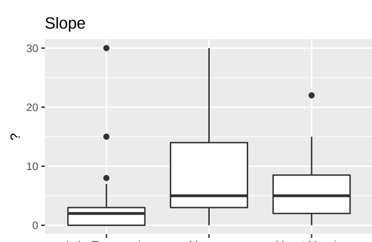
(j) potassium fraction



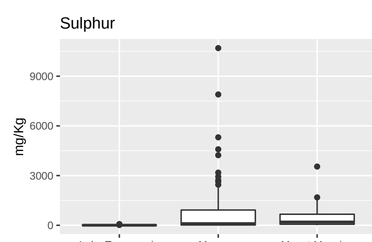
(k) quartz fraction



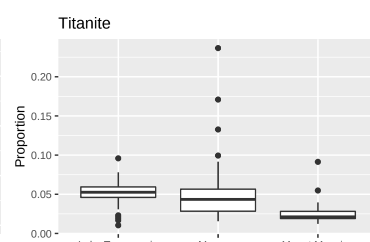
(l) ATP fluorescence



(m) slope angle

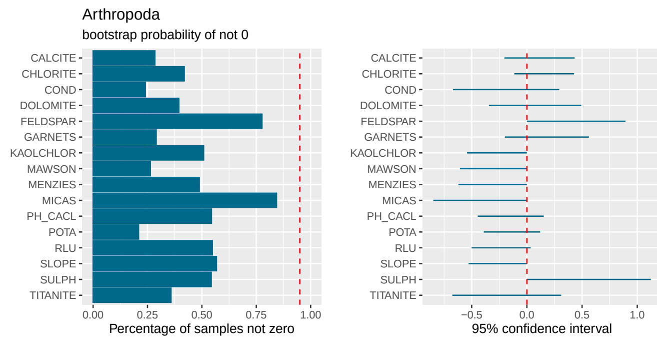


(n) sulphur

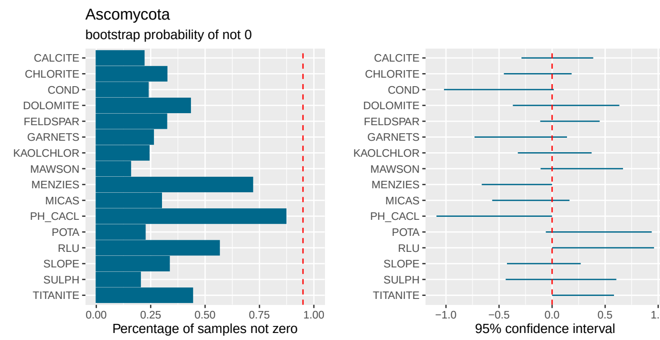


(o) titanite fraction

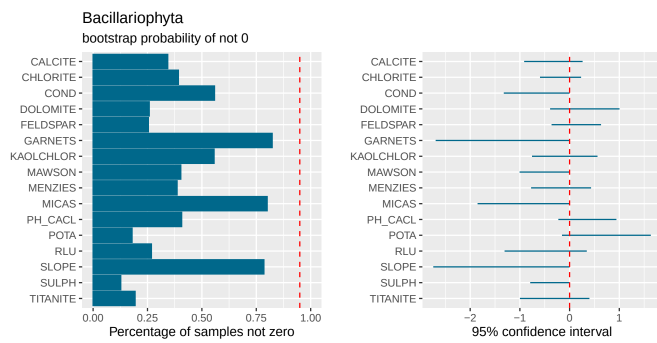
Figure 5: Predictor overview subsequent to data filtering, per location.



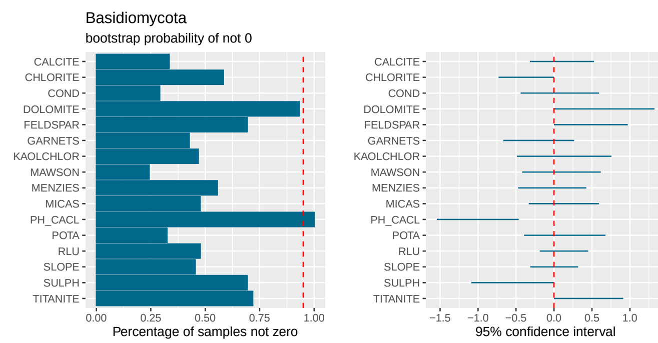
(a)



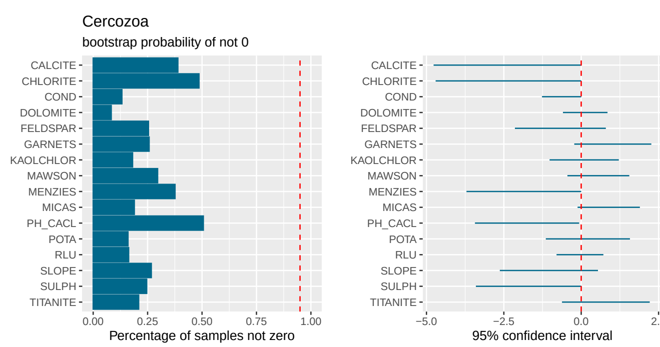
(b)



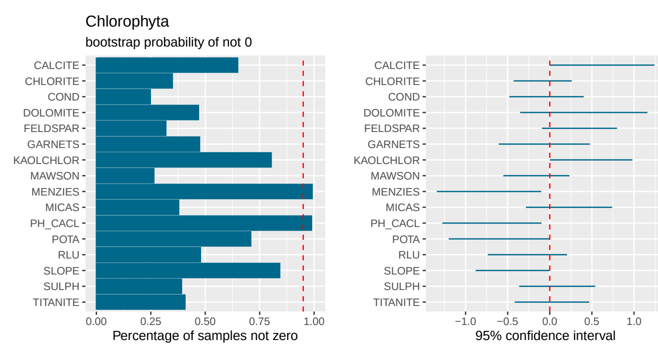
(c)



(d)



(e)



(f)

Figure 6: Full regression results, continued in Fig. 7

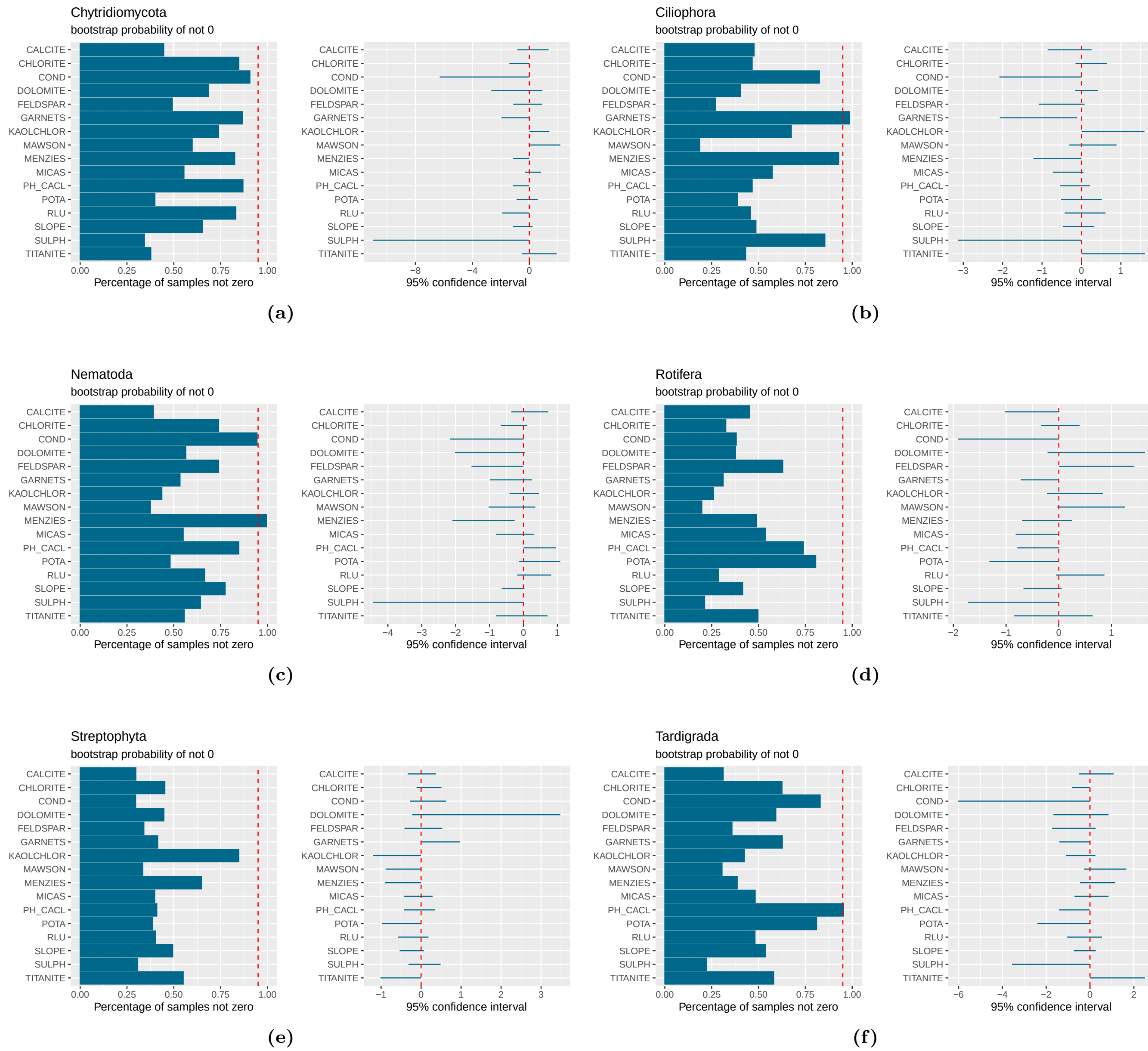


Figure 7: Full regression results, continued from Fig. 6