

Supplement of Biogeosciences, 17, 3203–3222, 2020
<https://doi.org/10.5194/bg-17-3203-2020-supplement>
© Author(s) 2020. This work is distributed under
the Creative Commons Attribution 4.0 License.



Supplement of

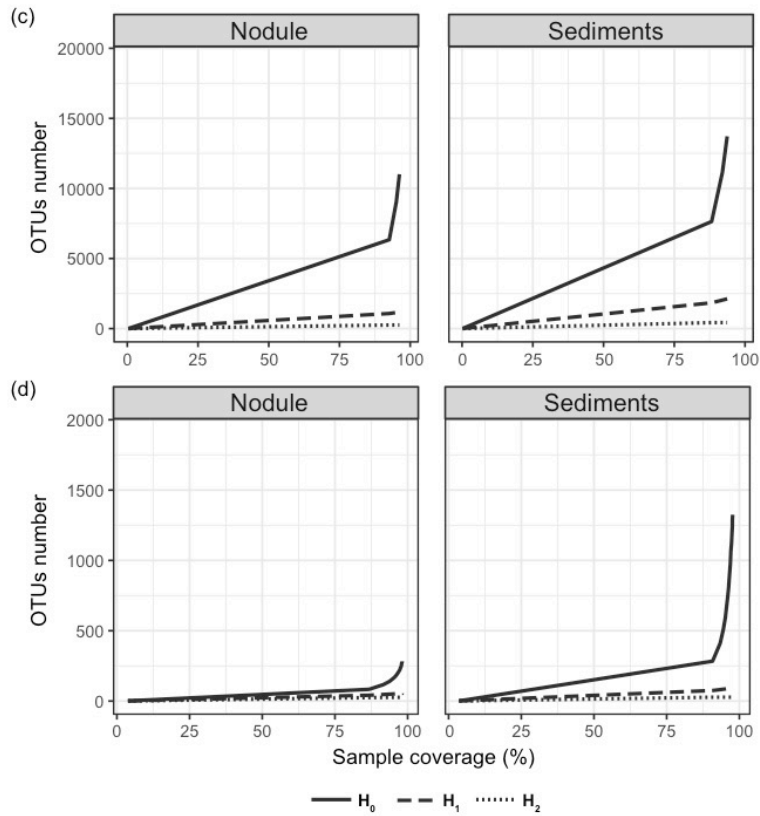
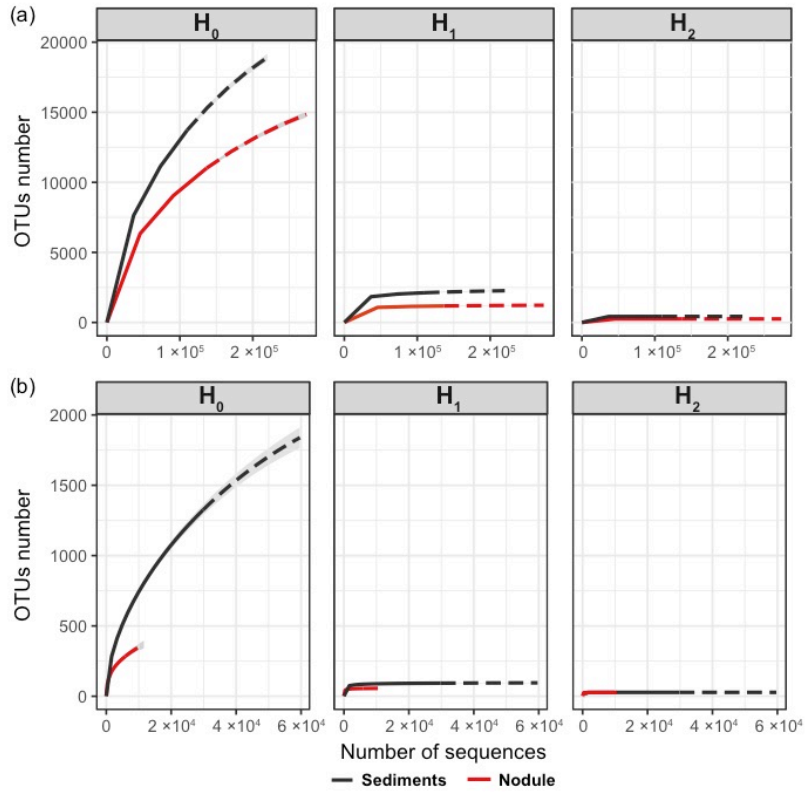
The contribution of microbial communities in polymetallic nodules to the diversity of the deep-sea microbiome of the Peru Basin (4130–4198 m depth)

Massimiliano Molari et al.

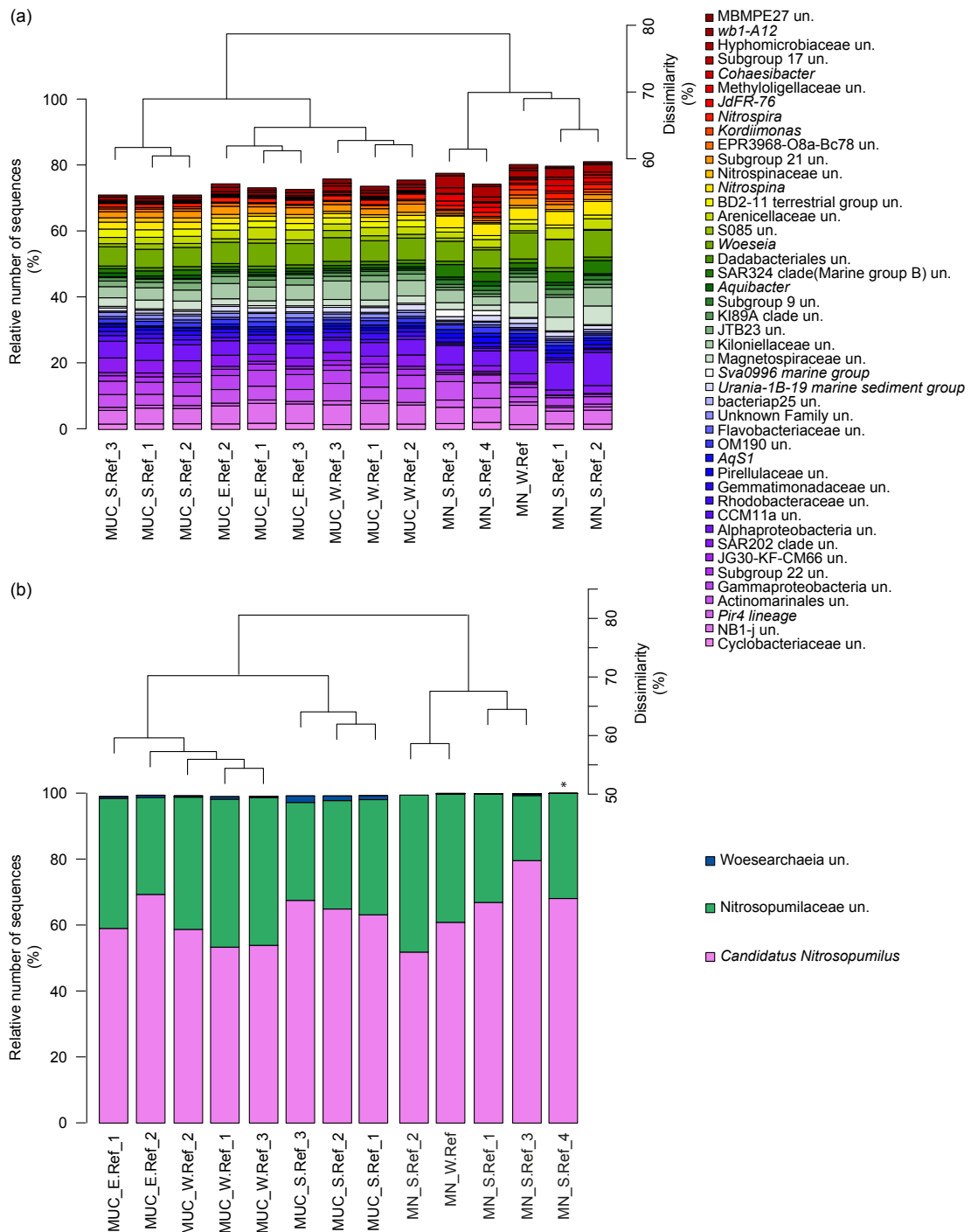
Correspondence to: Massimiliano Molari (mamolari@mpi-bremen.de)

The copyright of individual parts of the supplement might differ from the CC BY 4.0 License.

1 **Figure S1.** Rarefaction curves and diversity coverage in manganese nodules and sediments. a-b) show
2 sample-size-based rarefaction curve for Bacteria and Archaea, respectively; c-d) show coverage-based
3 curves for Bacteria and Archaea, respectively. The solid lines represent the observed accumulation
4 with the number of sequences sampled, and the dashed lines represent the extrapolated accumulation
5 up to double amount of sequences (only in a-b plots). Shaded area showed the 95 % confidence
6 intervals based on 100 bootstrap replications. Knots = 10 for Bacteria, and knots = 40 for Archaea.
7 H₀: number of OTUs (q=0); H₁: exponential Shannon (q=1); H₂: inverse Simpson (q=2).



9 **Figure S2.** Statistics of sequence and OTUs abundance, and proportion of absolute singletons,
 10 cosmopolitans and endemics for sediments (n=9) and nodule (n=5 for Bacteria, n=4 for Archaea)
 11 samples collected in Peru Basin. Absolute singletons: OTUs consisting of sequences occurring only
 12 once in the entire dataset; Cosmopolitan: OTUs present in 80 % of sediments and 80 % of nodule
 13 samples; Endemics: OTUs exclusively present only in 80 % sediments (and <20 % of nodule samples)
 14 or in 80 % nodule samples (and <20 % of sediments samples).



17 **Table S1.** Statistics of sequence and OTUs abundance, and proportion of absolute singletons,
 18 cosmopolitans and endemics for sediments (n=9) and nodule (n=5 for Bacteria, n=4 for Archaea)
 19 samples collected in Peru Basin. Absolute singletons: OTUs consisting of sequences occurring only
 20 once in the entire dataset; Cosmopolitan: OTUs present in 80 % of sediments and 80 % of nodule
 21 samples; Endemics: OTUs exclusively present only in 80 % sediments (and <20 % of nodule samples)
 22 or in 80 % nodule samples (and <20 % of sediments samples).

Bacteria	OTUs		Sequences	
	n.	%	n.	%
Entire dataset	557468		2271610	
Contaminants	20	0.0	15710	0.7
Absolute singletons	525169	94.2 (56 / 39) ^b	525169	23.1 (14 / 9) ^b
Working dataset ^a	32279	5.8	1730731	76.2
Sediments dataset ^a	28666	5.1 (88.8) ^c	1032246	45.4 (59.6) ^c
Nodule dataset ^a	19279	3.5 (59.7) ^c	698485	30.8 (40.3) ^c
Cosmopolitan OTUs ^a	1452	0.5 (8.9) ^c	1167668	58.4 (76.7) ^c
Endemics OTUs sediments ^a	1356	0.2 (4.2) ^c	39895	1.8 (2.3) ^c
Endemics OTUs nodules ^a	599	0.1 (1.9) ^c	52328	2.3 (3.0) ^c

23

Archaea	OTUs		Sequences	
	n.	%	n.	%
Entire dataset	51856		293098	
Contaminants	0	0.0	0	0.0
Absolute singletons	49482	95.4 (77 / 19) ^b	49482	16.9 (14 / 3) ^b
Working dataset ^a	2372	4.6	243616	83.1
Sediments dataset ^a	2356	4.5 (99.3) ^c	219460	74.9 (90.1) ^c
Nodule dataset ^a	591	1.1 (24.9) ^c	24156	8.2 (9.9) ^c
Cosmopolitan OTUs ^a	112	0.2 (4.7) ^c	194736	66.4 (79.9) ^c
Endemics OTUs sediments ^a	198	0.4 (8.3) ^c	10610	3.6 (4.4) ^c
Endemics OTUs nodules ^a	5	0.01 (0.2) ^c	121	0.04 (0.05) ^c

24

25 ^a after removal of contaminants (defined by negative control) and absolute singletons sequences (see Methods for details),
 26 percentage calculated on Entire dataset.

27 ^b contribution of sediments and nodules to absolute singletons, respectively.

28 ^c percentage calculated on working dataset.

29

30 **Table S2.** Output permutational multivariate analysis of variance on distance matrices
 31 (PERMANOVA).

OTUs CLR transformed and Euclidean distance															
Bacteria								Archaea							
Substrates	Df	SS	MS	F	R ²	P		Df	SS	MS	F	R ²	P		
Substrates	1	21547	21546.7	3.5043	0.22602	0.001		Substrates	1	3014.5	3014.53	4.9801	0.31164	0.001	
Residuals	12	73783	6148.6		0.77398			Residuals	11	6658.5	605.32		0.68836		
Total	13	95330			1			Total	12	9673			1		
Sites/Sediment	Df	SS	MS	F	R ²	P		Sites/Sediment	Df	SS	MS	F	R ²	P	
Sites	2	18360	9180.2	1.8699	0.38397	0.003		Sites	2	2542.1	1271.05	2.3048	0.47969	0.013	
Residuals	6	29458	4909.6		0.61603			Residuals	5	2757.3	551.47		0.52031		
Total	8	47818			1			Total	7	5299.4			1		
Reference.South/Substrate	Df	SS	MS	F	R ²	P		Reference.South/Substrate	Df	SS	MS	F	R ²	P	
Substrates	1	12771	12770.7	2.5848	0.34079	0.023		Substrates	1	3369.1	3369.1	7.5272	0.60087	0.029	
Residuals	5	24703	4940.7		0.65921			Residuals	5	2237.9	447.6		0.39913		
Total	6	37474			1			Total	6	5607			1		
Sites/Substrates (Strata=Site)	Df	SS	MS	F	R ²	P		Sites/Substrates (Strata=Site)	Df	SS	MS	F	R ²	P	
Sites	2	25842	12921	2.4356	0.27108	0.005		Sites	2	1837	918.5	1.8506	0.18991	0.021	
Sites:Substrates	2	21743	10871	2.0492	0.22808	0.005		Sites:Substrate	1	3369.1	3369.1	6.7881	0.3483	0.021	
Residuals	9	47745	5305		0.50084			Residuals	9	4466.9	496.3		0.46179		
Total	13	95330			1			Total	12	9673			1		

OTUs P/A table and Jaccard dissimilarity ^A															
Bacteria								Archaea							
Substrates	Df	SS	MS	F	R ²	P		Df	SS	MS	F	R ²	P		
Substrates	1	0.5986	0.59863	2.7963	0.18899	0.002		Substrates	1	0.45274	0.45274	2.2661	0.18474	0.003	
Residuals	12	2.5689	0.21408		0.81101			Residuals	10	1.9979	0.19979		0.81526		
Total	13	3.1676			1			Total	11	2.45064			1		
Sites/Sediment	Df	SS	MS	F	R ²	P		Sites/Sediment	Df	SS	MS	F	R ²	P	
Sites	2	0.50624	0.25312	1.3286	0.30693	0.002		Sites	2	0.52048	0.26024	1.4829	0.37231	0.003	
Residuals	6	1.14312	0.19052		0.69307			Residuals	5	0.87749	0.1755		0.62769		
Total	8	1.64936			1			Total	7	1.39798			1		
Reference.South/Substrate	Df	SS	MS	F	R ²	P		Reference.South/Substrate	Df	SS	MS	F	R ²	P	
Substrates	1	0.48253	0.48253	2.2875	0.31389	0.035		Substrates	1	0.41752	0.41752	2.0157	0.33507	0.1	
Residuals	5	1.0547	0.21094		0.68611			Residuals	4	0.82856	0.20714		0.66493		
Total	6	1.53722			1			Total	5	1.24609			1		
Sites/Substrates (Strata=Site)	Df	SS	MS	F	R ²	P		Sites/Substrates (Strata=Site)	Df	SS	MS	F	R ²	P	
Sites	2	0.5954	0.29772	1.4698	0.18798	0.006		Sites	2	0.46094	0.23047	1.2347	0.18809	0.027	
Sites:Substrates	2	0.7492	0.37458	1.8493	0.23651	0.006		Sites:Substrate	2	0.68307	0.34154	1.8297	0.27873	0.027	
Residuals	9	1.823	0.20255		0.57551			Residuals	7	1.30663	0.18666		0.53318		
Total	13	3.1676			1			Total	11	2.45064			1		

32 CLR: centered log-ratio; P/A: presence/absence; Df: degrees of freedom; SS: sum of the squares; F: statistic *F-ratio*; P:
 33 probability level.
 34

35 ^A based on 100 sequence re-samplings per sample to the smallest dataset (40613 sequences for Bacteria and 1835
 36 sequences for Archaea).
 37

38 **Table S3.** Genera differentially abundant in nodules and sediments (ALDEx2: glm adjusted $p < 0.01$;
39 KW adjusted $p < 0.05$). In bold the most abundant genera ($\geq 0.5\%$) at least two times more abundant in
40 nodule than in sediment; in italic the genera exclusively present (i.e. unique) in nodules. Base 2
41 logarithm of the ratios between geometric mean centred sequences number of nodule (Nod) and
42 sediment (Sed), and average of the sequences contribution of total number of sequences (%) retrieved
43 in nodules and in sediments are shown.

Enriched in Nodule	LOG2(Nod/Sed)	Nodule (%)	Sediment (%)	Enriched in Sediment	LOG2(Nod/Sed)	Nodule (%)	Sediment (%)
<i>Spingomonadaceae_unclassified</i>	-	0.04	0.00	Planctomycetales_unclassified	-0.02	0.44	0.49
<i>Filomicrobium</i>	-	0.01	0.00	Lutibacter	-1	0.00	0.02
Geminicoccaceae_unclassified	4	0.12	0.01	Chloroflexi_unclassified	-2	0.03	0.09
Methyloceanibacter	4	0.17	0.02	AT-s3-28_unclassified	-2	0.03	0.09
Robiginitomaculum	4	0.09	0.00	Chitinophagales_unclassified	-2	0.05	0.16
Mesorhizobium	3	0.25	0.01	Bacteriovoracaceae_unclassified	-2	0.04	0.14
Cohaesibacter	3	0.78	0.10	Nannocystaceae_unclassified	-2	0.02	0.07
OPB56_unclassified	3	0.03	0.00	Cellvibrionaceae_unclassified	-2	0.02	0.08
67-14_unclassified	3	0.31	0.06	OM182 clade_unclassified	-2	0.13	0.47
Syntrophaceae_unclassified	3	0.06	0.01	Candidatus Komeilibacteria_unclassified	-2	0.01	0.03
Maribacter	3	0.06	0.01	Roseobacter clade NAC11-7 lineage	-2	0.04	0.11
Methyloligellaceae_unclassified	2	1.46	0.31	Bacteroidia_unclassified	-2	0.03	0.11
Entotheonellaceae_unclassified	2	0.20	0.04	IS-44	-2	0.05	0.20
Blastocatella	2	0.18	0.04	Oligoflexaceae_unclassified	-2	0.01	0.08
Calorithrix	2	0.03	0.01	Lentimicrobiaceae_unclassified	-2	0.01	0.04
Hyphomicrobiaceae_unclassified	2	2.72	0.71	Marinoscillum	-3	0.02	0.08
Planctomicrobium	2	0.05	0.01	Anaerolineaceae_unclassified	-3	0.05	0.36
Simkaniaceae_unclassified	2	0.13	0.03	Colwelliaceae_unclassified	-3	0.01	0.13
Microtrichaceae_unclassified	2	0.13	0.03	Subgroup 7_unclassified	-3	0.00	0.03
LD1-PA32_unclassified	2	0.05	0.01	Peredibacter	-3	0.01	0.05
Subgroup 17_unclassified	2	1.03	0.27	Marinimicrobia (SAR406 clade)_unclassified	-4	0.01	0.07
JdFR-76	2	0.93	0.26	Total		1.00	2.91
Subgroup 9_unclassified	2	1.26	0.42				
Chlamydiales_unclassified	2	0.16	0.06				
SAR324 clade(Marine group B)_unclassified	2	3.12	1.10				
Vermiphilaceae_unclassified	2	0.14	0.05				
Acanthopleuribacter	1	0.04	0.01				
Bythopirellula	1	0.04	0.01				
Nitrospina	1	3.79	1.72				
Gemmataceae_unclassified	1	0.04	0.02				
Planctomycetacia_unclassified	1	0.05	0.03				
SM1A02	1	0.29	0.15				
Ekhidna	1	0.17	0.09				
Phycisphaeraceae_unclassified	1	0.49	0.27				
AqS1	1	1.10	0.66				
Microtrichales_unclassified	1	0.19	0.08				
Pirellulaceae_unclassified	1	1.31	0.75				
pltb-vmat-80_unclassified	1	0.05	0.00				
Pir4 lineage	1	1.60	0.91				
Alphaproteobacteria_unclassified	1	7.15	4.44				
Babeliales_unclassified	1	0.10	0.07				
Parvularculaceae_unclassified	1	0.06	0.04				
PAUC43f marine benthic group_unclassified	1	0.54	0.36				
Subgroup 10	0	0.75	0.67				
Aquibacter	0	0.83	0.73				
Cyclobacteriaceae_unclassified	0	1.76	1.70				
Gemmatimonadaceae_unclassified	0	1.17	1.15				
Rhodothermaceae_unclassified	0	0.38	0.39				
Total		35.37	17.82				