Supporting Information

Ghiglione et al. 10.1073/pnas.1208160109

SI Materials and Methods

Data Analysis Pipeline. The hypervariable V6 region of the SSU rRNA gene in bacteria was amplified using a combination of 967F (forward) and 1046R (reverse) primers under conditions described previously (1). Amplicons were sequenced using pyrosequencing technology (454 Life Sciences; GS20) at the Marine Biological Laboratory. Random sequencing errors were removed as described previously (2). To normalize the number of tags sequenced between samples and datasets and locations, tags were randomly resampled to the sample with the fewest tags using Daisychopper version 0.6 (3). Analyses of tag sequences were based on the frequency of resampled tags analyzed in Mothur (4) following a work flow that included multiple alignment to Silva database reference alignment (June 29, 2010), error-minimization using a "preclustering" step, and then a distance matrix was generated as input for "cluster," from which output files were created. Clusters based on a distance of 0.03 (≥97% identity) were used in all downstream analyses (these have up to 2 mismatches per 60-bp sequence). In addition, Mothur was used to define OTUs at similarity thresholds ranging from 87 to 100%. The number of remaining OTUs diminished with the decreasing SSU rRNA sequence similarity threshold used to define OTUs, and the percentage of remaining OTUs between one threshold and the next was plotted for each similarity value. Differences in the proportion of remaining OTUs between samples were used as a metric for comparison of the phylogenetic structure between communities (5). Taxonomic identification of the sequence reads (tags) followed the approach by Sogin et al. (6) and Huse et al. (7). Representative sequences for clustered tags originating from a global reference dataset at the 0.03 distance level are available at http://vamps.mbl.edu.

Similarity-Based Cluster and Statistical Analyses. Similarity-based cluster and statistical analyses were performed using the unweighted pair group method with arithmetic mean (UPGMA) (Primer version 6 software). A similarity profile test (SIMPROF) (Primer version 6) was performed on a null hypothesis that a specific subcluster can be recreated by permuting the entry species and samples. The significant branch (SIMPROF, $P < 0.05$) was used as a prerequisite for defining bacterial clusters. One-way analysis of similarity (ANOSIM) (Primer version 6) was performed on the same distance matrix to test the null hypothesis that there was no difference between bacterial communities of different clusters. Similarity percentage (SIMPER) (8) was used to determine which individual OTUs contributed most to the dissimilarity between groups of samples.

The resulting Bray–Curtis distance matrix was used to build a UPGMA tree of sample relationships. This resulting tree was imported in the web version of UniFrac (9), together with respective groups according to depth, distance from shore, and regions. UniFrac and P test significance were performed on all samples together and pairs of samples.

Relationships between bacterioplankton communities and environmental parameters were assessed using a direct gradient approach, canonical correspondence analysis (CCA) (CANOCO version 4.5) according to ref. 10. Spearman rank pairwise correlations between the environmental variables (depth, temperature, salinity, chlorophyll, latitude, and longitude; Table S1) helped to determine their significance. To statistically evaluate the significance of the first canonical axis and of all canonical axes together, we used Monte Carlo permutation full model test with 199 unrestricted permutations. Significant variables were chosen using a forward-selection procedure and 999 permutations. Furthermore, we specifically examined the relationship between latitude and OTU richness based on the Chao1 richness estimator calculated using the application in the Mothur package of executable programs. The analysis was limited to surface water samples at depths of 0–40 m (only two samples were collected at depths below 11 m) and deep samples collected at depths below 200 m. The Spearman Rank correlation (and significance) was calculated for latitude vs. Chao1 datasets in the "surface" and "deep."

1. Huber JA, et al. (2007) Microbial population structures in the deep marine biosphere. *Science* 318:97–100.
2. Huse SM, Huber JA, Morrison HG, Sogin ML, Welch DM (2007) Accuracy and quality of massively parallel DNA pyrosequencing. *Genome Biol* 8:R143.
3. Gilbert JA, et al. (2009) The seasonal structure of microbial communities in the Western English Channel. *Environ Microbiol* 11:3132–3139.
4. Schloss PD, et al. (2009) Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* 75:7537–7541.
5. Barberán A, Fernández-Guerra A, Auguet JC, Galand PE, Casamayor EO (2011) Phylogenetic ecology of widespread uncultured clades of the Kingdom Euryarchaeota. *Mol Ecol* 20:1988–1996.
6. Sogin ML, et al. (2006) Microbial diversity in the deep sea and the underexplored "rare biosphere". *Proc Natl Acad Sci USA* 103:12115–12120.
7. Huse SM, Welch DM, Morrison HG, Sogin ML (2010) Ironing out the wrinkles in the rare biosphere through improved OTU clustering. *Environ Microbiol* 12:1889–1898.
8. Clarke KR, Warwick RW (2001) *Change in Marine Communities: An Approach to Statistical Analysis and Interpretation* (PRIMER-E, Plymouth, UK), 2nd Ed, pp 1–72.
9. Lozupone C, Hamady M, Knight R (2006) UniFrac—an online tool for comparing microbial community diversity in a phylogenetic context. *BMC Bioinformatics* 7:371.
10. Ghiglione JF, Larcher M, Lebaron P (2005) Spatial and temporal scales of variation in bacterioplankton community structure in the NW Mediterranean Sea. *Aquat Microb Ecol* 40:229–240.

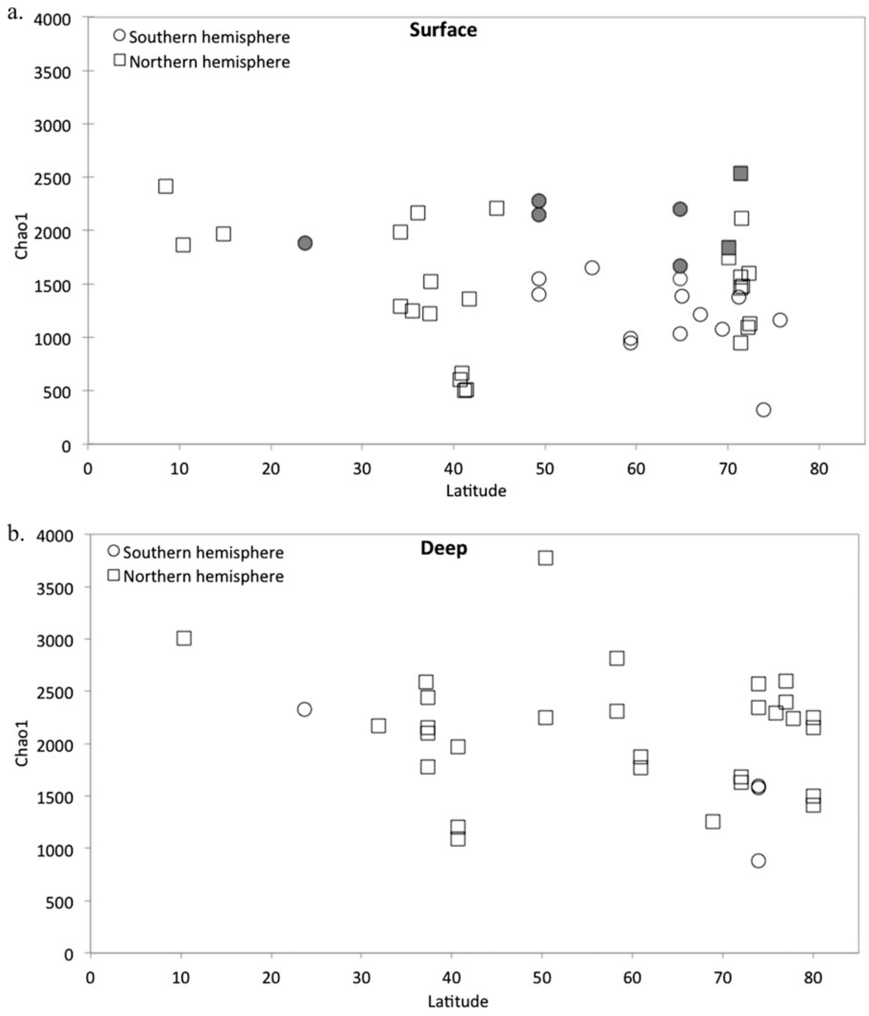

Fig. S1. Relationship between latitude and OTU richness as estimated by Chao1 for surface (0–40 m) (*A*) and deep (>200 m) (*B*) 16S rRNA gene tag datasets. The absolute value of latitude is plotted with the Southern and Northern hemisphere datasets indicated with circles (open circles for summer samples, and closed circles for winter samples) and squares, respectively.

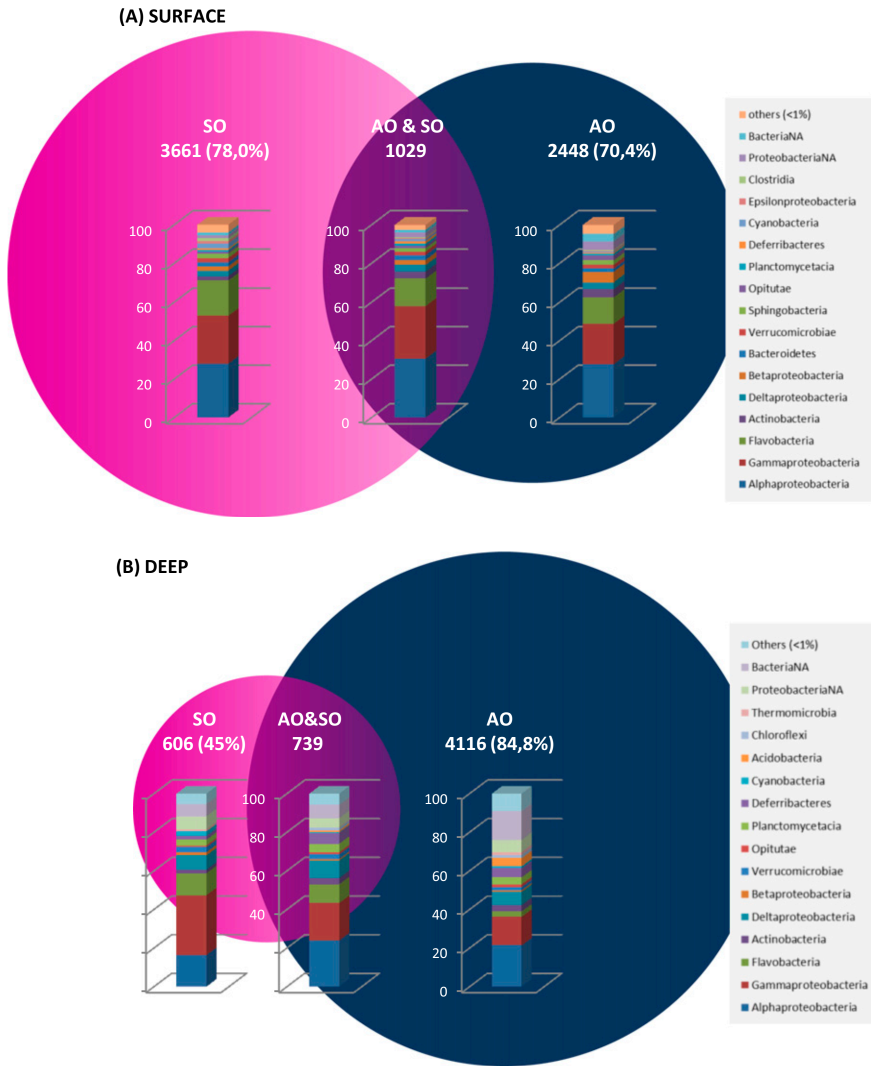

Fig. S2. Venn diagrams representing surface (*A*) and deep (*B*) OTUs found in Southern Ocean (SO) and Arctic Ocean (AO) bacterial datasets. Number of OTUs and their relative percentage to the total number of OTU (in brackets) are indicated in bold, and the corresponding number of tag sequences are indicated in brackets. Histograms of corresponding taxonomic composition are grouped at the class level for unique and shared OTUs between polar communities.

Surface ocean OTUs common to each pair of data sets

	SO	AO	ML
SO	3264	1029	817
AO	-	2002	666
ML	-		1104

Percents of surface ocean OTUs common to each pair of data sets

	SO	AO	ML
SO	69.6	14.8	5.1
AO	-	61.1	4.5
ML	-		91.3

Deep ocean OTUs common to each pair of data sets

	SO	AO	ML
SO	528	739	627
AO	-	2920	1744
ML	-	-	7371

Percents of deep ocean OTUs common to each pair of data sets

	SO	AO	ML
SO	39.2	11.9	5.9
AO	-	60.1	12.4
ML	-	-	80.2

Fig. S3. OTU groupings when all bacterial samples [Southern Ocean (SO), Arctic Ocean (AO), and midlatitude (ML)] were compared in surface and deep ocean datasets. Note that 420 OTUs (2.1%) were common to all three datasets in surface waters, and 548 OTUs (3.6%) were common in deep waters.

Table S1. Station and sample metadata, categories, and environmental parameters

Sample name	Station ID (ICoMM format) with sampling date	Depth (m)	Latitude	Longitude	Ocean region	Ocean subregion	Total sequences	Total OTUs (0.03)	Salinity (PSU)	Temperature (°C)	Chlorophyll (µg × L^{-1})
AO_BS1_C2	ACB_0001_2007_07_13	2	71.45	−156.06	Arctic	Beaufort Sea	15,600	786	20.0	2.5	0.33
AO_CS2_C2	ACB_0002_2007_07_11	2	71.43	−156.86	Arctic	Chukchi Sea	9,107	462	32.0	4.6	0.50
AO_CS3_C2	ACB_0003_2008_01_26	2	71.35	−156.68	Arctic	Chukchi Sea	17,516	387	35.0	−1.8	0.07
AO_CS4_C2	ACB_0004_2008_01_30	2	71.35	−156.68	Arctic	Chukchi Sea	19,118	666	35.0	−2.0	0.06
AO_CS5_O10	ACB_0005_2002_07_29	10	72.32	−151.98	Arctic	Chukchi Sea	19,737	517	29.4	−1.5	0.57
AO_CS6_O10	ACB_0006_2002_08_06	10	72.24	−159.34	Arctic	Chukchi Sea	14,575	472	29.5	−1.1	0.22
AO_BS7_O10	ACB_0007_2004_07_30	10	71.54	−150.89	Arctic	Beaufort Sea	17,551	465	29.9	5.2	0.64
AO_BS8_O40	ACB_0008_2004_08_21	40	72.42	−152.02	Arctic	Beaufort Sea	14,113	427	30.3	−1.2	0.27
AO_FB13_C3*	ACB_0013_2004_01_17	3	70.04	−126.30	Arctic	Franklin Bay	12,389	664	29.8	−1.6	0.04
AO_FB15_C3	ACB_0015_2004_07_16	3	70.05	−126.31	Arctic	Franklin Bay	50,884	527	29.0	4.0	0.61
AO_CS16_C2*	ACB_0016_2008_01_28	2	71.35	−156.68	Arctic	Chukchi Sea	19,084	799	35.0	−1.8	0.06
SO_AS1_C10	ASA_0001_2007_12_18	10	−73.94	−115.68	Antarctic	Amundsen Sea	11,411	182	33.9	−2.0	8.16
SO_AS10_C250	ASA_0010_2007_12_18	250	−73.94	−115.68	Antarctic	Amundsen Sea	15,636	364	34.0	−1.8	1.29
SO_AS13_C790	ASA_0013_2007_12_18	790	−73.96	−115.86	Antarctic	Amundsen Sea	23,704	802	34.5	0.4	NA
SO_AS14_C500	ASA_0014_2007_12_18	500	−73.96	−115.86	Antarctic	Amundsen Sea	27,983	717	34.1	−1.6	NA
SO_AP1_C6	CAM_0001_2002_01_17	6	−64.77	−64.05	Antarctic	Antarctic Peninsula	23,733	442	33.6	1.6	NA
SO_AP2_C6*	CAM_0002_2002_07_17	6	−64.77	−64.05	Antarctic	Antarctic Peninsula	10,798	826	33.8	−0.7	0.04
SO_AP3_C6*	CAM_0003_2002_08_20	6	−64.77	−64.05	Antarctic	Antarctic Peninsula	12,662	962	33.8	−0.37	0.07
SO_AP4_C6	CAM_0004_2006_02_28	6	−64.77	−64.07	Antarctic	Antarctic Peninsula	22,147	356	33.4	2.2	2.97
SO_KI5_C3*	CAM_0005_2007_07_23	3	−49.37	−70.20	Antarctic	Kerguelen Islands	38,796	743	32.8	1.9	0.16
SO_KI6_C3	CAM_0006_2007_10_09	3	−49.37	−70.20	Antarctic	Kerguelen Islands	32,147	563	33.3	3.2	5.04
SO_KI7_C3	CAM_0007_2007_02_07	3	−49.37	−70.20	Antarctic	Kerguelen Islands	18,711	562	33.8	6.9	0.35
SO_KI8_C3*	CAM_0008_2007_05_04	3	−49.37	−70.20	Antarctic	Kerguelen Islands	24,255	711	33.4	4.9	0.29
SO_WS9_O11	CAM_0009_2000_03_23	11	−55.13	−2.89	Antarctic	Weddell Sea	20,673	545	NA	0.9	0.27
SO_WS10_O11	CAM_0010_2000_03_25	11	−59.32	−0.40	Antarctic	Weddell Sea	29,214	364	34.2	0.3	0.04
SO_WS11_O11	CAM_0011_2000_03_26	11	−67.02	−5.47	Antarctic	Weddell Sea	19,583	412	NA	−0.8	0.05
SO_WS12_C11	CAM_0012_2000_03_29	11	−71.18	−12.46	Antarctic	Weddell Sea	21,290	496	34.2	−1.8	0.05
SO_RS13_O10	CAM_0013_2008_02_04	10	−59.40	175.56	Antarctic	Ross Sea	15,506	429	33.9	4.3	0.25
SO_RS14_O10	CAM_0014_2008_02_06	10	−65.01	179.97	Antarctic	Ross Sea	17,998	413	33.6	1.0	0.24
SO_RS15_O10	CAM_0015_2008_02_29	10	−69.43	−179.09	Antarctic	Ross Sea	23,149	441	33.5	−1.8	0.38
SO_RS16_C10	CAM_0016_2008_02_13	10	−75.66	169.77	Antarctic	Ross Sea	19,471	461	NA	−0.8	0.69
AO_BB1_O1000	DAO_0001_2007_07_12	1,000	68.83	−61.76	Arctic	Baffin Bay	27,165	720	34.5	0.9	0.00
AO_BS3_O1000	DAO_0003_2007_08_04	1,000	71.96	−150.23	Arctic	Beaufort Sea	18,873	916	34.9	0.0	NA
AO_BS4_O400	DAO_0004_2007_08_04	400	71.96	−150.23	Arctic	Beaufort Sea	24,725	806	34.8	0.8	0.01
AO_BS5_O1000	DAO_0005_2007_08_12	1,000	79.99	−149.99	Arctic	Beaufort Sea	17,654	871	34.9	0.0	0.00
AO_BS7_O400	DAO_0007_2007_08_12	400	79.99	−149.99	Arctic	Beaufort Sea	31,400	821	34.8	0.1	0.00
AO_BS9_O1000	DAO_0009_2007_08_15	1,000	77.00	−140.19	Arctic	Beaufort Sea	14,217	1,190	34.9	0.0	NA
AO_BS10_O400	DAO_0010_2007_08_15	400	77.00	−140.19	Arctic	Beaufort Sea	20,328	1,026	34.8	0.9	0.00
AO_BS11_O900	DAO_0011_2007_08_20	900	75.84	−128.64	Arctic	Beaufort Sea	21,373	1,062	34.9	0.1	NA
AO_BS12_O1000	DAO_0012_2007_08_23	1,000	73.97	−140.09	Arctic	Beaufort Sea	22,612	1,109	34.9	0.0	0.00
AO_BS13_O400	DAO_0013_2007_08_23	400	73.97	−140.09	Arctic	Beaufort Sea	16,487	1,000	32.8	0.7	0.00
AO_ES14_O1000	DAO_0014_2007_09_18	1,000	77.75	126.00	Arctic	East Siberian Sea	36,728	894	34.9	0.0	0.00
AO_ES15_O1200	DAO_0015_2007_09_21	1,200	79.94	142.39	Arctic	East Siberian Sea	26,671	882	34.9	−0.3	0.00
AO_ES16_O250	DAO_0016_2007_09_21	250	79.94	142.39	Arctic	East Siberian Sea	15,704	967	34.8	1.4	0.00

Table S1. Cont.

Sample name	Station ID (ICoMM format) with sampling date	Depth (m)	Latitude	Longitude	Ocean region	Ocean subregion	Total sequences	Total OTUs (0.03)	Salinity (PSU)	Temperature (°C)	Chlorophyll (µg × L^{-1})
NA112_O4000	KCK_NADP_Bv6;112R_4121m	4,000	50.40	−25.00	North Atlantic	North Atlantic Deep Water	15,497	1,356	34.9	2.3	NA
NA115_O550	KCK_NADP_Bv6;115R_550m	550	50.40	−25.00	North Atlantic	North Atlantic Deep Water	16,334	1,021	35.1	7.0	NA
NA137_O1700	KCK_NADP_Bv6;137_1710m	1,700	60.90	38.52	North Atlantic	North Atlantic Deep Water	13,759	946	34.9	3.0	NA
NA138_O700	KCK_NADP_Bv6;138–710m	700	60.90	38.52	North Atlantic	North Atlantic Deep Water	12,980	969	34.9	3.5	NA
NA53_O1400	KCK_NADP_Bv6;53R_1400m	1,400	58.30	29.13	North Atlantic	North Atlantic Deep Water	12,763	1,078	NA	NA	NA
NA55_O500	KCK_NADP_Bv6;55R_500m	500	58.30	29.13	North Atlantic	North Atlantic Deep Water	9,903	1,219	35.1	7.1	NA
MS2_C5	BMO_0002_2007_09_20	5	41.65	2.80	Mediterranean	NW Mediterranean Sea	18,361	580	37.9	23.3	0.09
MS3_C5	BMO_0003_2007_09_21	5	41.40	2.80	Mediterranean	NW Mediterranean Sea	22,253	326	37.8	23.8	0.07
MS4_O5	BMO_0004_2007_09_21	5	41.15	2.82	Mediterranean	NW Mediterranean Sea	17,311	360	37.9	23.8	0.08
MS5_O5	BMO_0005_2007_09_22	5	40.91	2.85	Mediterranean	NW Mediterranean Sea	22,951	456	37.9	21.0	0.10
MS6_O5	BMO_0006_2007_09_23	5	40.65	2.85	Mediterranean	NW Mediterranean Sea	21,884	346	37.3	24.3	0.08
MS9_O500	BMO_0009_2007_09_22	500	40.66	2.86	Mediterranean	NW Mediterranean Sea	23,347	711	38.516	13.185	NA
MS10_O2000	BMO_0010_2007_09_22	2,000	40.66	2.86	Mediterranean	NW Mediterranean Sea	14,527	987	38.479	13.213	NA
MS11_C500	BMO_0011_2007_09_22	500	40.65	2.85	Mediterranean	NW Mediterranean Sea	12,585	764	38.516	13.185	NA
NA_1_O47	AOT_0001_2008_11_07	47	42.85	−11.64	North Atlantic	North Atlantic	26,972	890	35.92	16	NA
NA_2_O1100	AOT_0002_2008_11_09	1,100	37.13	−13.36	North Atlantic	North Atlantic	18,090	1,332	36.017	10.9265	NA
NA_3_O73	AOT_0003_2008_11_09	73	37.13	−13.36	North Atlantic	North Atlantic	24,795	1,140	36.0697	15.7954	NA
NA_4_O89	AOT_0004_2008_11_14	89	22.51	−20.50	North Atlantic	North Atlantic	25,787	951	36.8478	22.7556	NA
NA_5_O7	AOT_0005_2008_11_16	7	14.76	−30.00	North Atlantic	North Atlantic	21,824	708	35.65	***26.473***	***0.17***
SA_6_O28*	AOT_0006_2008_11_29	28	−23.71	8.52	South Atlantic	South Atlantic	18,532	702	35.5673	18.0667	**0.18**
NA_7_O7	AOT_0007_2008_11_17	7	44.70	−20.36	North Atlantic	North Atlantic	27,567	847	35.44	***15.6***	***0.32***
NA_8_O100	AOT_0008_2008_11_17	100	10.63	−20.13	North Atlantic	North Atlantic	11,722	1,167	35.5739	15.2887	NA
NA_9_O48	AOT_0009_2008_11_17	48	10.63	−20.13	North Atlantic	North Atlantic	46,751	945	35.9724	20.4348	NA
NA_10_O11	AOT_0010_2008_11_17	11	10.37	−20.06	North Atlantic	North Atlantic	20,884	723	35.09	***28.214***	***0.39***
NA_12_O1300	AOT_0012_2008_11_17	1,300	10.37	−20.06	North Atlantic	North Atlantic	16,117	1,150	***34.955***	***4.608***	NA
NA_13_O7	AOT_0013_2008_11_18	7	8.50	−19.50	North Atlantic	North Atlantic	23,799	714	34.48	***28.608***	***0.21***
SA_14_O4600*	AOT_0014_2008_11_29	4,600	−23.71	8.52	South Atlantic	South Atlantic	10,151	949	35.6	***3.447***	NA
SA_16_O48*	AOT_0016_2008_11_27	48	−17.74	3.13	South Atlantic	South Atlantic	17,126	545	35.88	***18.701***	NA
AZ_1_O0	AWP_0001_2007_08_23	0	36.05	−28.35	Azores	Azores	17,185	727	36.73	23.6	***0.08***
AZ_3_O0	AWP_0003_2007_08_25	0	34.21	−32.12	Azores	Azores	22,064	746	36.932	25.3	***0.07***
AZ_6_O0	AWP_0006_2007_08_25	0	34.21	−32.12	Azores	Azores	20,528	623	36.932	25.3	***0.07***
AZ_7_O0	AWP_0007_2007_08_25	0	35.48	−32.43	Azores	Azores	17,419	614	36.863	24.6	***0.08***
AZ_8_O0	AWP_0008_2007_08_26	0	37.48	−30.57	Azores	Azores	25,041	596	36.172	23.1	***0.09***
AZ_9_O3660	AWP_0009_2007_06_11	3,660	37.34	−18.88	Azores	Azores	25,840	1,044	34.92	2.56	NA
AZ_10_O2100	AWP_0010_2007_06_11	2,100	37.34	−18.88	Azores	Azores	24,343	972	35	3.69	NA
AZ_11_O1200	AWP_0011_2007_06_11	1,200	37.34	−18.88	Azores	Azores	20,825	1,238	35.84	9.11	NA
AZ_12_O800	AWP_0012_2007_06_11	800	37.34	−18.88	Azores	Azores	23,238	1,301	35.67	10.33	NA
AZ_13_O100	AWP_0013_2007_06_11	100	37.34	−18.88	Azores	Azores	24,435	842	36.19	16.37	NA
AZ_14_O0	AWP_0014_2007_06_11	0	37.34	−18.88	Azores	Azores	25,370	585	36.4	18.54	***0.12***
AZ_15_O800	AWP_0015_2007_06_02	800	31.86	−27.88	Azores	Azores	21,776	1,131	36.19	10	NA
AZ_16_O100	AWP_0016_2007_06_02	100	31.86	−27.88	Azores	Azores	27,904	1,398	37	18.54	NA
SP_2_O110	KNX_0002_2006_12_24	110	−26.05	−156.89	South Pacific	South Pacific	26,953	1,154	35.57	18.13	0.39
SP3_O120	KNX_0003_2006_12_27	120	−27.94	−148.59	South Pacific	South Pacific	23,273	1,118	35.5	18.5	0.36
SP_4_O150	KNX_0004_2006_12_30	150	−26.48	−137.94	South Pacific	South Pacific	18,093	1,198	35.5	19.37	0.22

Table S1. Cont.

Sample name	Station ID (ICoMM format) with sampling date	Depth (m)	Latitude	Longitude	Ocean region	Ocean subregion	Total sequences	Total OTUs (0.03)	Salinity (PSU)	Temperature (°C)	Chlorophyll ($\mu g \times L^{-1}$)
SP_5_O145	KNX_0005_2007_01_01	145	−28.45	−131.39	South Pacific	South Pacific	24,661	1,217	35.39	18.1	0.35
SP_6_O200	KNX_0006_2007_01_07	200	−27.74	−117.62	South Pacific	South Pacific	30,918	1,277	35.5	18.76	0.27
SP_7_O110	KNX_0007_2007_01_11	110	−38.06	−133.09	South Pacific	South Pacific	21,932	891	34.3	11.5	0.30
SP_8_O100	KNX_0008_2007_01_13	100	−39.31	−139.80	South Pacific	South Pacific	16,099	792	34.36	11.2	0.40

Terms for sample name were designated by abbreviations derived from the ocean region, subregion, and depth. Environmental parameters derived from remotely sensed data are indicated in boldface italics (select temperature and chlorophyll values). NA, not available.

Asterisks (*) indicate samples collected in the winter.