

## **Supplementary Figures**

### **Diversity and dynamics of rare and of resident bacterial populations in coastal sands**

Angélique Gobet, Simone Böer, Susan M. Huse, Justus E.E. van Beusekom, Christopher Quince, Mitchell L. Sogin, Antje Boetius, Alban Ramette

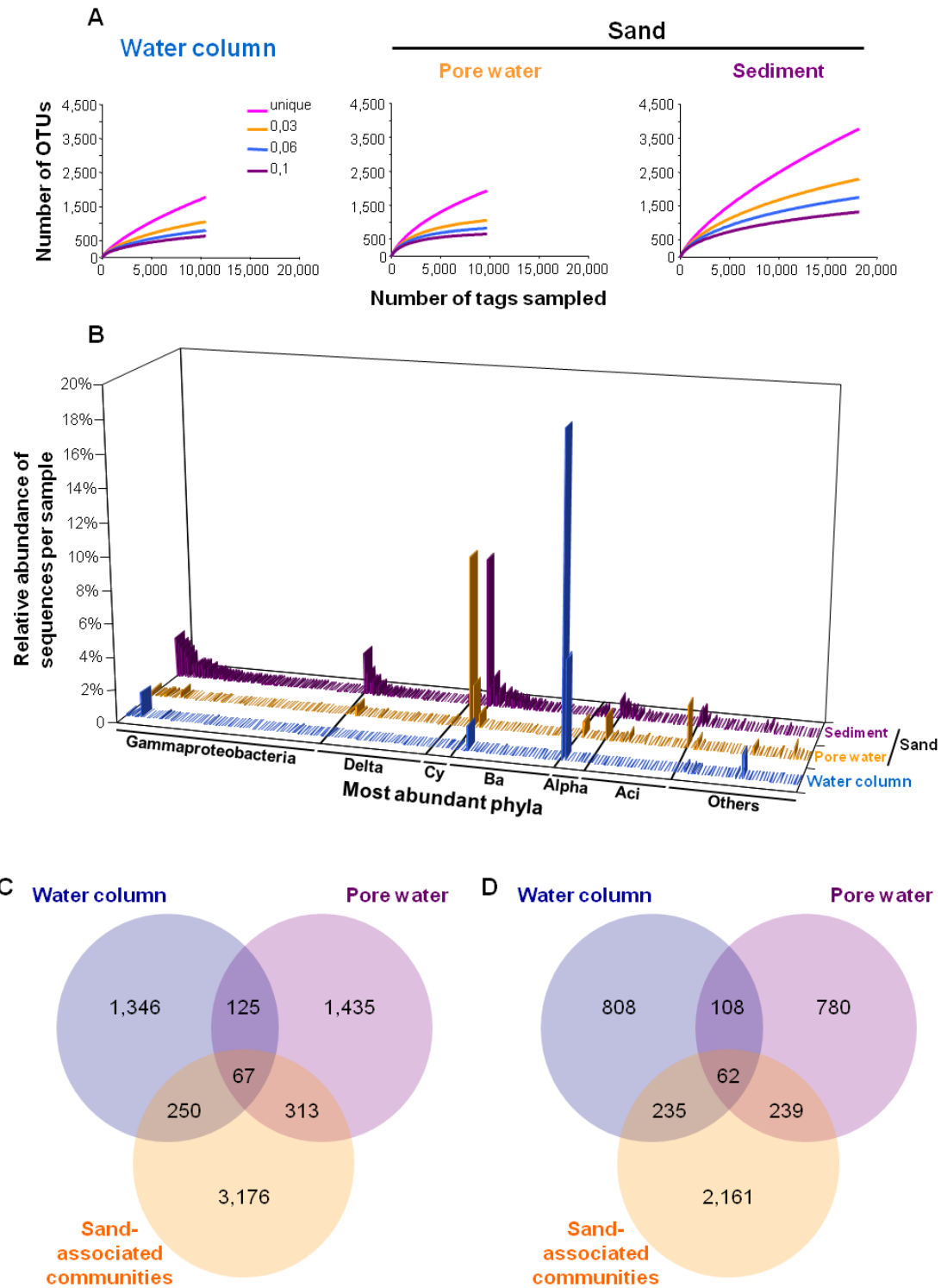
**Fig. S1. Bacterial composition, distribution and turnover in different sand-associated compartments and in the water column.**

**Fig. S2. Relative sequence abundance of phyla in the sand through depth and time.**

**Fig. S3. Turnover of the bacterial community between sediment depth layers or sampling dates.**

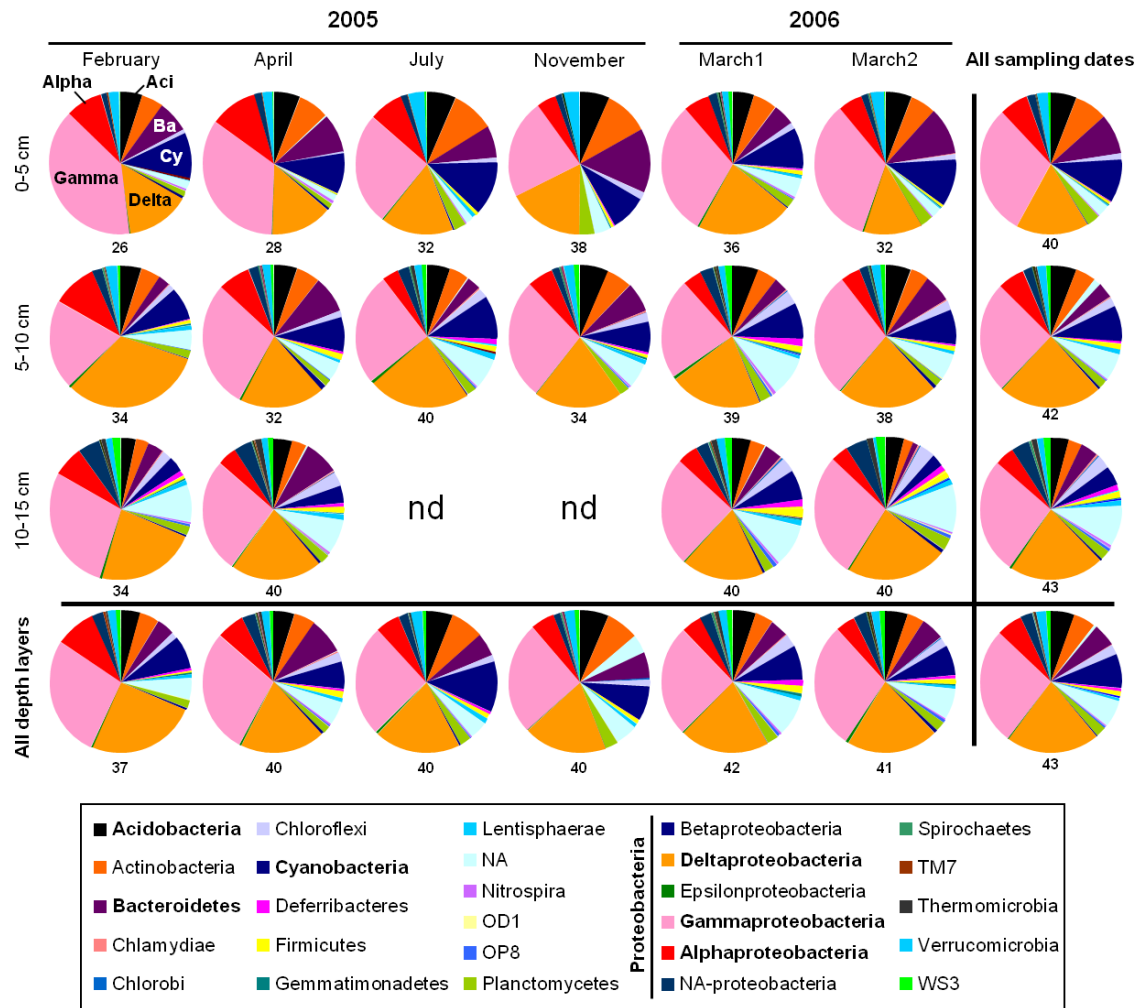
**Fig. S4. Partitioning of the biological variation in the bacterial community structure as a function of explanatory variables.**

**Fig. S5. Effects of environmental conditions on the proportion of SSO<sub>rel</sub> per sample.**

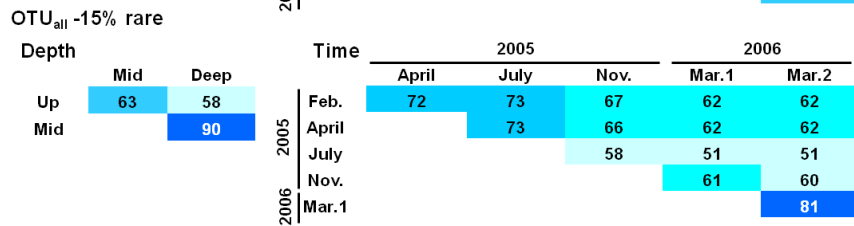
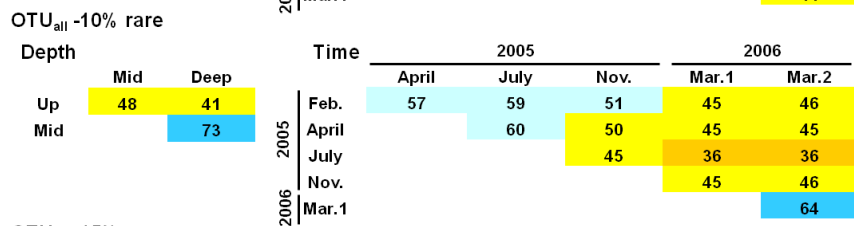
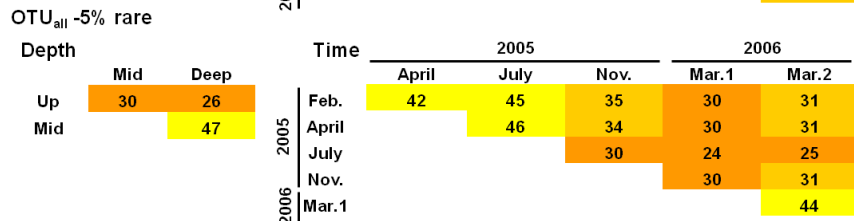
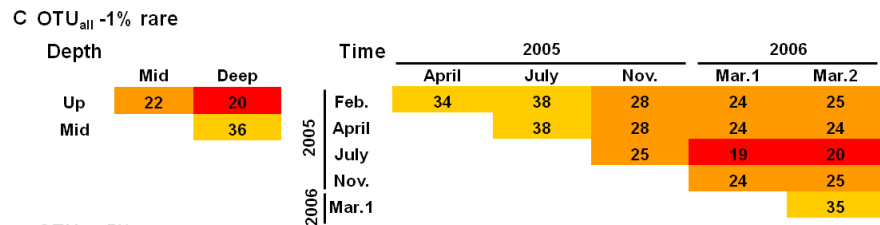
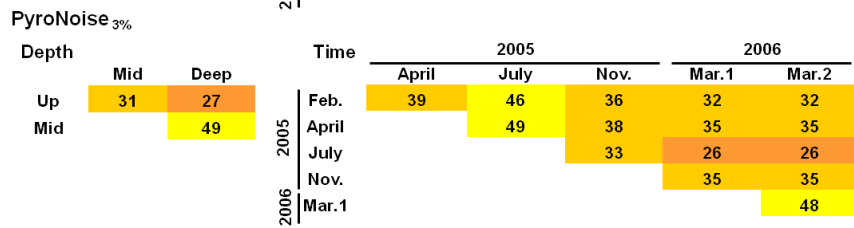
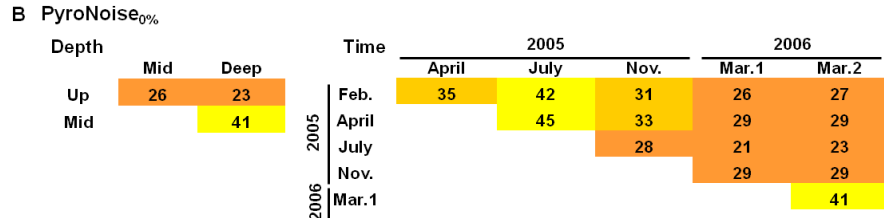
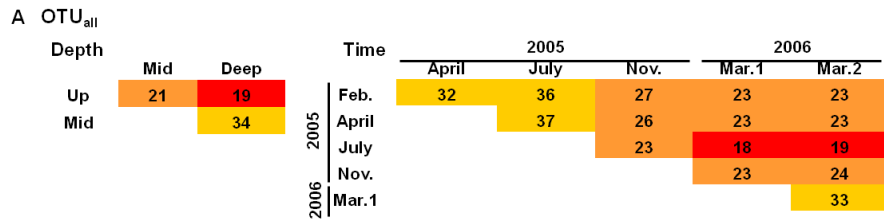


**Fig. S1. Bacterial composition, distribution and turnover in different sand-associated compartments and in the water column. (A)** Rarefaction curves in different compartments of the sand and in the water column at the unique, 3%, 6% and 10% dissimilarities to define OTU. **(B)** Sequence distribution in

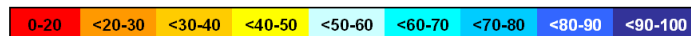
different compartments of the sand and in the water column in April 2008. Each bar represents an OTU<sub>unique</sub> (only OTU<sub>unique</sub> occurring more than 10 times in the OTU<sub>all</sub> data set are shown on the skyline plot). The *Proteobacteria* phylum was separated into its corresponding classes for higher resolution. Gamma, *Gammaproteobacteria*; Delta, *Deltaproteobacteria*; Cy, *Cyanobacteria*; Ba, *Bacteroidetes*; Alpha, *Alphaproteobacteria*; Aci, *Acidobacteria*; Others: *Actinobacteria*, NA-*Proteobacteria* (*Proteobacteria* with class annotation missing), *Planctomycetes*, *Chloroflexi*, *Verrucomicrobia*, WS3, *Firmicutes*, *Lentisphaerae*, *Deferribacteres*, *Epsilonproteobacteria*, *Gemmatimonadetes*. (C, D) Total number of shared OTU between the three compartments: Sand grain-associated biofilm, sand porewater, and overlying water column. (C) OTU<sub>unique</sub>, (D) PyroNoise-corrected OTU<sub>3%</sub>. Each entire circle represents the total number of OTU in a given compartment (here, one sample).



**Fig. S2. Relative sequence abundance of phyla in the sand through depth and time.** Some of the most abundant phyla were indicated on the first pie chart and in bold in the legend: Gamma, *Gammaproteobacteria*; Delta, *Deltaproteobacteria*; Cy, *Cyanobacteria*; Ba, *Bacteroidetes*; Alpha, *Alphaproteobacteria*; Aci, *Acidobacteria*. The *Proteobacteria* phylum was separated into its corresponding classes for higher resolution. Unassigned OTU<sub>unique</sub> were grouped as one phylum in this figure (NA). nd, missing samples. The total number of phyla (including the *Proteobacteria* divided as classes) is indicated under each pie chart. The legend indicates the phyla color code of the pie charts. Phyla with too few sequences to be visible on the pie charts were removed from the legend: *Aquificae*, *Deinococcus-Thermus*, *Fibrobacteres*, *Fusobacteria*, *Tenericutes*, *Thermodesulfobacteria* and, the candidate divisions: BRC1, JS1, OP10, OP11, OP3, OP5, OPB7, TG1, TM6 and, WS1.



% shared OTUs



C OTU<sub>all</sub> -20% rare

Depth	Mid	Deep	Time	2005			2006		
				April	July	Nov.	Mar.1	Mar.2	
Up	76	74	2005	Feb.	83	84	80	78	78
				April		84	80	78	77
Mid		97	2006	July		72	67	67	
				Nov.				77	76
				Mar.1				92	

OTU<sub>all</sub> -25% rare

Depth	Mid	Deep	Time	2005			2006		
				April	July	Nov.	Mar.1	Mar.2	
Up	87	86	2005	Feb.	92	92	92	90	90
				April		92	91	90	90
Mid		100	2006	July		85	82	82	
				Nov.				87	87
				Mar.1				98	

OTU<sub>all</sub> -30% rare

Depth	Mid	Deep	Time	2005			2006		
				April	July	Nov.	Mar.1	Mar.2	
Up	94	94	2005	Feb.	98	97	97	97	97
				April		97	97	96	96
Mid		100	2006	July		93	92	92	
				Nov.				93	93
				Mar.1				100	

OTU<sub>all</sub> -35% rare

Depth	Mid	Deep	Time	2005			2006		
				April	July	Nov.	Mar.1	Mar.2	
Up	97	97	2005	Feb.	99	99	99	99	99
				April		99	99	99	99
Mid		100	2006	July		97	96	97	
				Nov.				97	96
				Mar.1				100	

OTU<sub>all</sub> -40% rare

Depth	Mid	Deep	Time	2005			2006		
				April	July	Nov.	Mar.1	Mar.2	
Up	99	99	2005	Feb.	100	100	100	100	100
				April		99	99	99	99
Mid		100	2006	July		99	98	99	
				Nov.				98	98
				Mar.1				100	

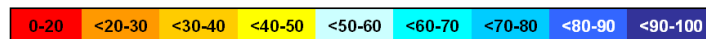
OTU<sub>all</sub> -45% rare

Depth	Mid	Deep	Time	2005			2006		
				April	July	Nov.	Mar.1	Mar.2	
Up	100	100	2005	Feb.	100	100	100	100	100
				April		100	100	100	100
Mid		100	2006	July		100	100	100	
				Nov.				99	99
				Mar.1				100	

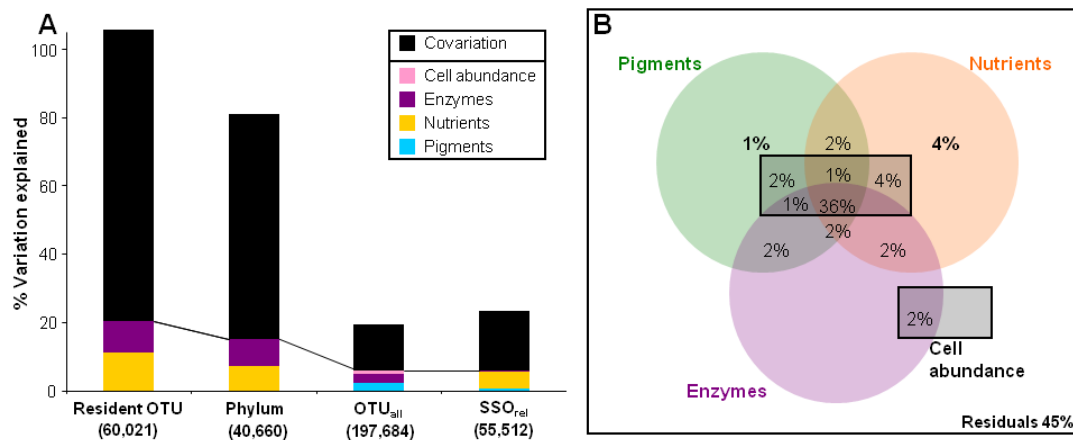
OTU<sub>all</sub> -50% rare

Depth	Mid	Deep	Time	2005			2006		
				April	July	Nov.	Mar.1	Mar.2	
Up	100	100	2005	Feb.	100	100	100	100	100
				April		100	100	100	100
Mid		100	2006	July		100	100	100	
				Nov.				100	100
				Mar.1				100	

% shared OTUs

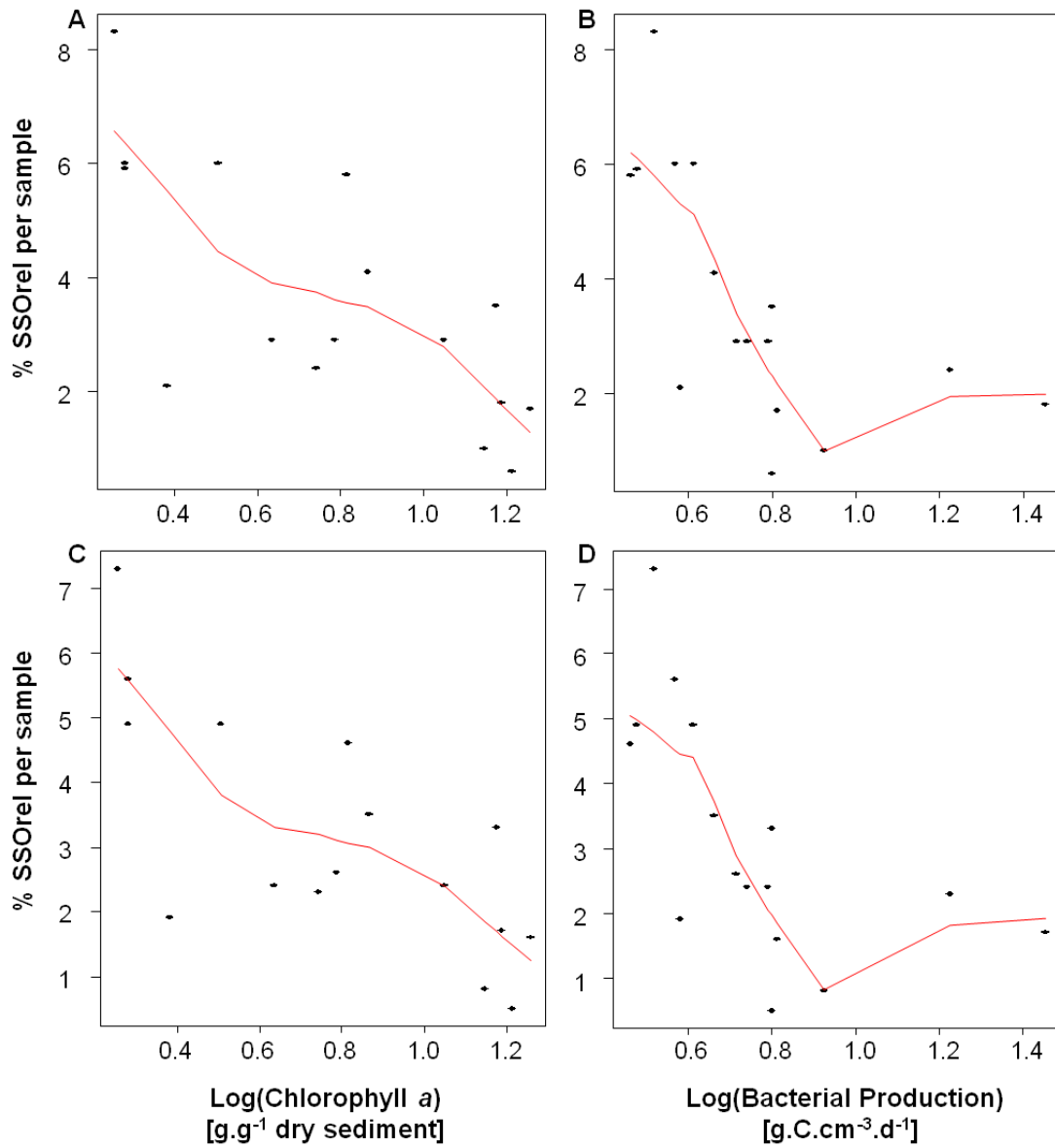


**Fig. S3. Turnover of the bacterial community between sediment depth layers or sampling dates (A)** on the original OTU<sub>all</sub> data set, **(B)** after PyroNoise correction of the 454 MPTS data set, at 0% and 3% OTU clustering (PyroNoise<sub>0%</sub> and PyroNoise<sub>3%</sub>, respectively), **(C)** after removing successive percentages of rare OTU<sub>unique</sub> (1%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45% and 50%) from the OTU<sub>all</sub> data set by the application of MultiCoLA. The percentage of OTU shared between a sampling depth (or date) and the previous one was calculated and values were represented according to heatmap matrices. OTU<sub>all</sub> represents the original data set with all OTU<sub>unique</sub>, used here as a reference to test for the effects of correction and clustering on the resolution of bacterial community dynamics.



**Fig. S4. Partitioning of the biological variation in the bacterial community structure as a function of explanatory parameters** based on (A) the Phylum level, OTU<sub>all</sub>, resident OTU and SSO<sub>rel</sub> data sets, and on (B) potential pathogens (including *Parachlamydia*, *Arcobacter*, *Francisella*, *Acinetobacter*, *Rickettsiella*, *Pseudomonas*, *Ralstonia*). Environmental parameters accounted for included pigments (chlorophyll *a* and pheophytin), nutrients (silicate, phosphate, nitrite, nitrate, ammonium), extra-cellular enzyme activities (chitinase,  $\alpha$ -glucosidase,  $\beta$ -glucosidase, lipase, aminopeptidase, phosphatase), cell abundance and their combined effects. Unexplained variation is not shown. In (A), The black line in each panel separates the pure factor effects from their covariations. Covariation of any of the 4 environmental factors is represented under one category “covariation”. Here, the OTU<sub>all</sub> level includes also sequences without complete annotation. The total number of sequences in each data set is indicated in parentheses.





**Fig. S5. Effects of environmental conditions on the proportion of SSO<sub>rel</sub> per sample.** PyroNoise-corrected data were clustered to define OTU at 0% (A, B) and 3% (C, D) sequence dissimilarity levels. The red line in each plot represents the best local fitting regression line between the variables.