Supplementary Information for

Structural heterogeneity yet high similarity of the microbial community on reverse osmosis membrane-driven biofilms during seawater desalination

Supporting Figure Legends

Fig. S1. Morphological features of unused RO membrane surface. Structural analysis of surface (a) and cross-sectional thickness measurements (b) by scanning electron microscopy (SEM).

Fig. S2. Diatom observation on fouled RO membrane surface of (a) M1 (first located RO membrane element) and (b) M2 (last located RO membrane element).

Fig. S3. Confocal laser scanning microscopy of the bacteria and archaea on the RO membrane (a) M1 and (b) M2. The bacteria were hybridized FAM labeled bacteria specific probes (green) and the archaea were hybridized CY5 labeled archaea specific probes (red) (bar, 20 μm).

Fig. S4. Rarefaction curve of (a) bacterial and (b) archaeal communities. The number of OTUs was defined at 95% similarity.

Fig. S5. Neighbor-joining tree of unclassified dominant bacteria (above 1%) with proportion and the closest type strains. RO feed (R), the front membrane (M1), and the last membrane (M2). The phylogeny was analyzed using 500 bootstrap replicates and was rooted using the archaeal strain *Nitrosoarchaeum limnia* SFB1.

Fig. S1

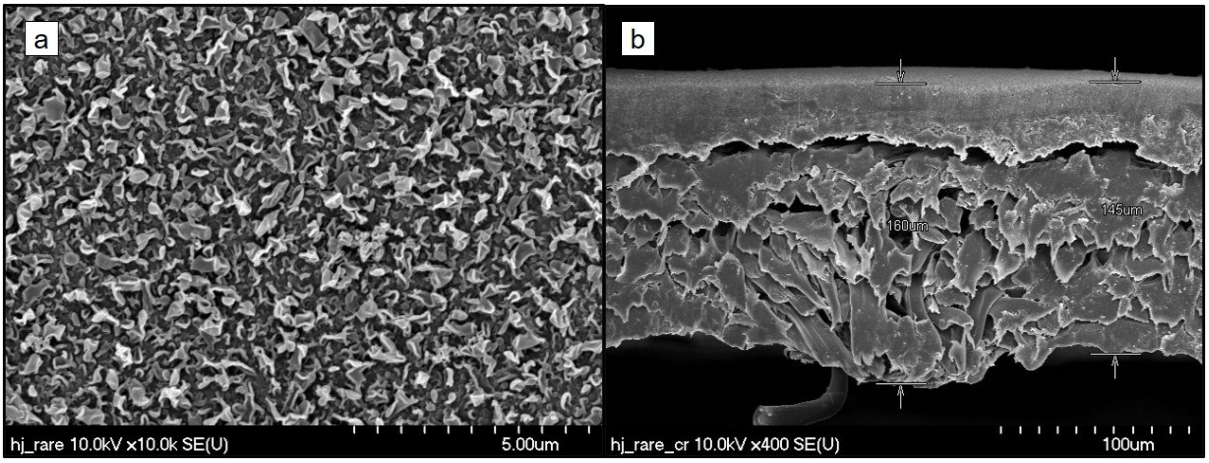
a
hj_rare 10.0kV x10.0k SE(U)
5.00um
b
160um
145um
hj_rare_cr 10.0kV x400 SE(U)
100um

Fig. S2

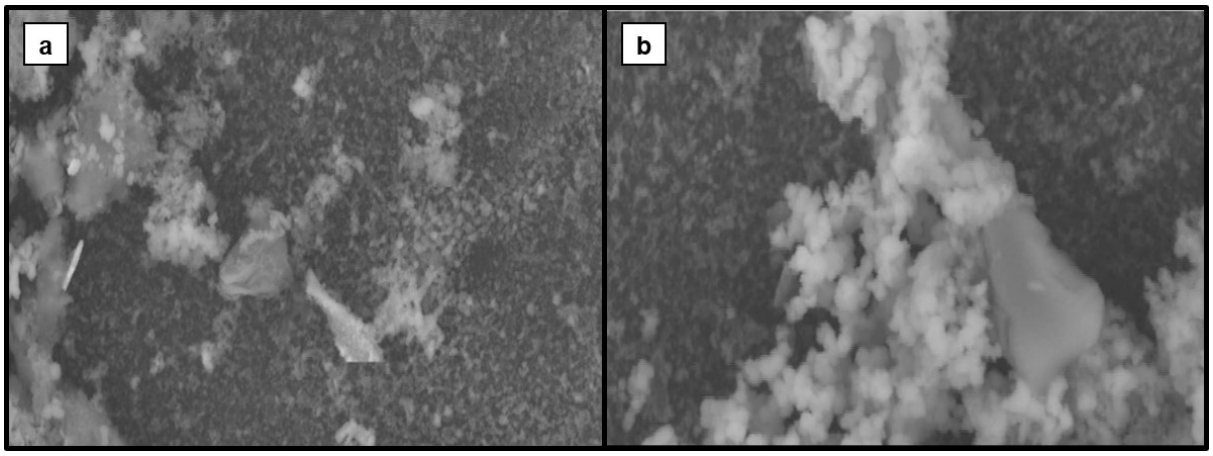

Figure S2. Diatom observation on fouled RO membrane surface of (a) M1 (first located RO membrane element) and (b) M2 (last located RO membrane element).

Fig. S3

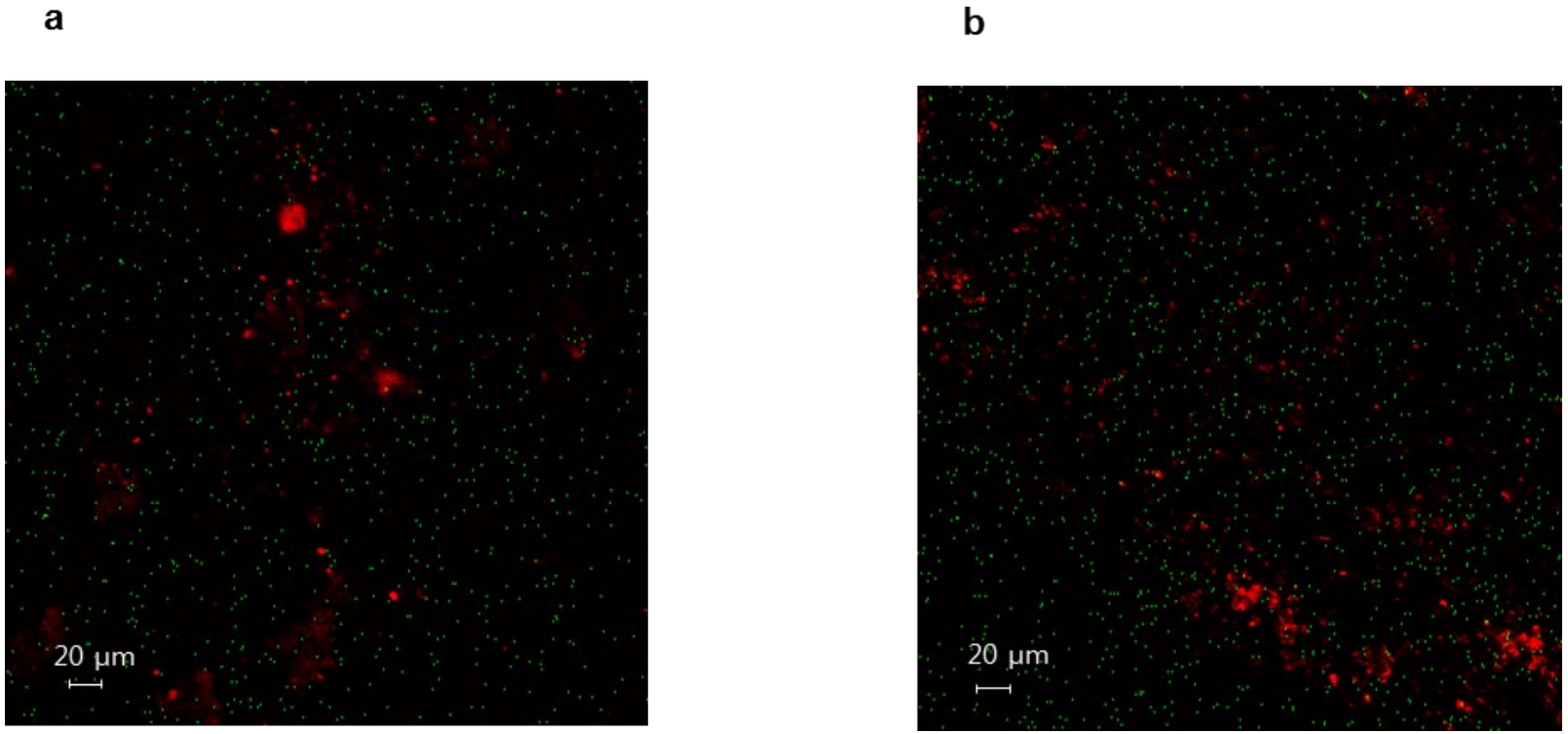

Figure S3. Confocal laser scanning laser microscopy of the bacteria and archaea on the RO membrane (a) M1 and (b) M2. The bacteria were hybridized FAM labeled bacteria specific probes (green) and the archaea were hybridized CY5 labeled archaea specific probes (red) (bar, 20 µm).

Additional Explanation: Materials and methods for fluorescence in situ hybridization and electron microscope.

Nine randomly collected membrane samples (approximately 2 mm × 2 mm) were immersed into 30 ml buffer solution (130 mM NaCl, 10 mM Na_2HPO_4/NaH_2PO_4, pH 7.4). The microbial cell detachment was performed using an ultrasonic cleaner (5510E-DTH, Bransonic Ultrasonics Corporation, CT, USA) at a 42 kHz output, and fixation and hybridization were performed as described by (Llobet-Brossa et al., 1998) and (Manz et al., 1992) respectively. The oligonucleotide probes used for FISH were specific for archaea (Arc344; 5'-TCGCGCCTGCTGCICCCCGT-3' and Arc915; 5'-GTGCTCCCCCGCCAATTCCT-3') and bacteria (Eub338, 5'-GCTGCCTCCCGTAGGAGT-3'). Here, bacterial and archaeal probes were labeled with 5-carboxyfluorescein (5' FAM, green) and 5-N-Nitdiethyltetramethylindodicarbocyanine (5' Cy5, red), respectively. All microbial cells were obtained from DAPI-stained cells; the total numbers of DAPI-stained and fluorophore dye labeled microbial cells were determined using a confocal laser scanning microscope (Zeiss LSM 5 Pascal, Zeiss, Jena, Germany) from 30 randomly chosen fields of view and then calculated as cells per area.

Fig. S4

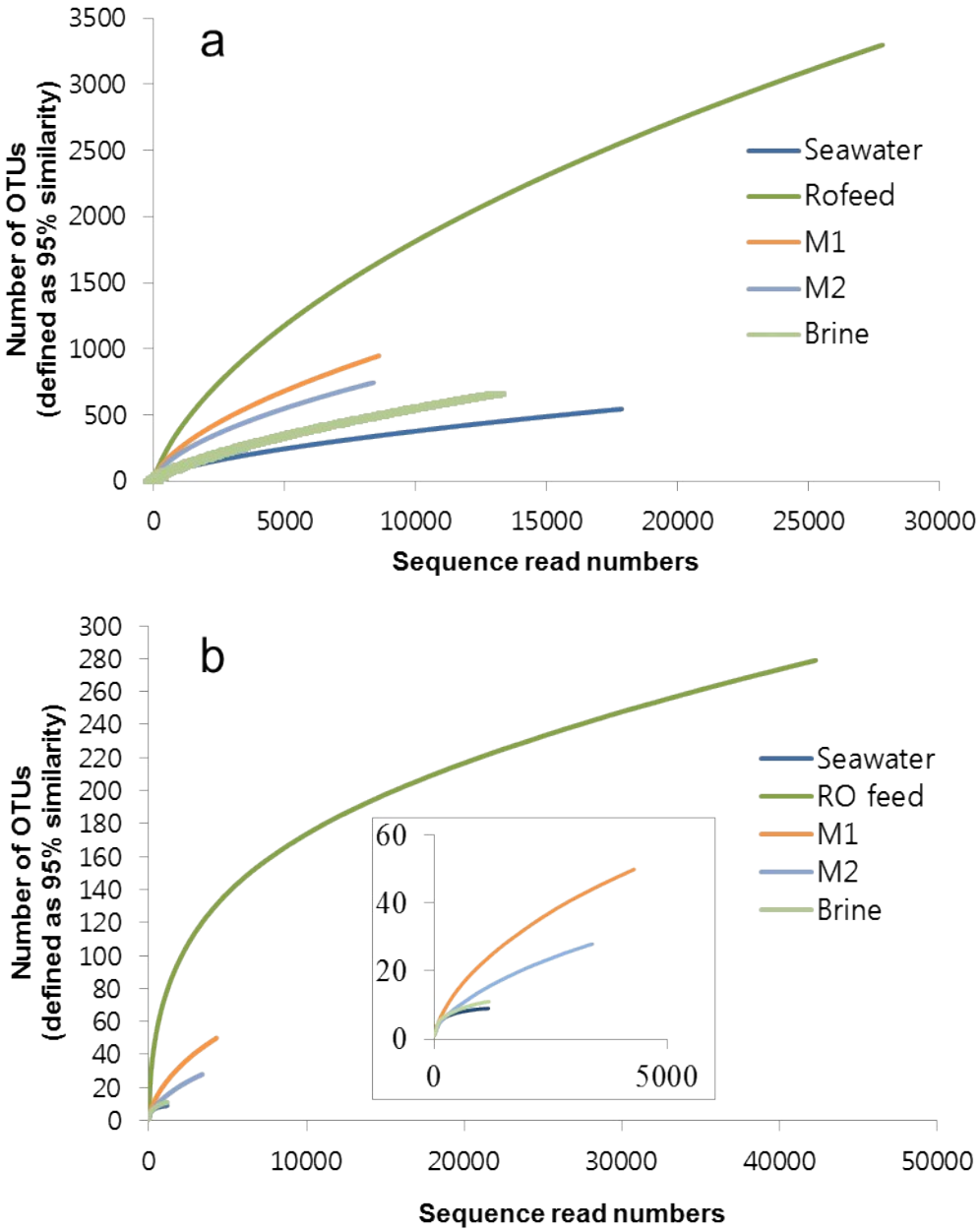

Figure S4. Rarefaction curve of (a) bacterial and (b) archaeal communities. The number of OTUs was defined at 95% similarity.

Fig. S5

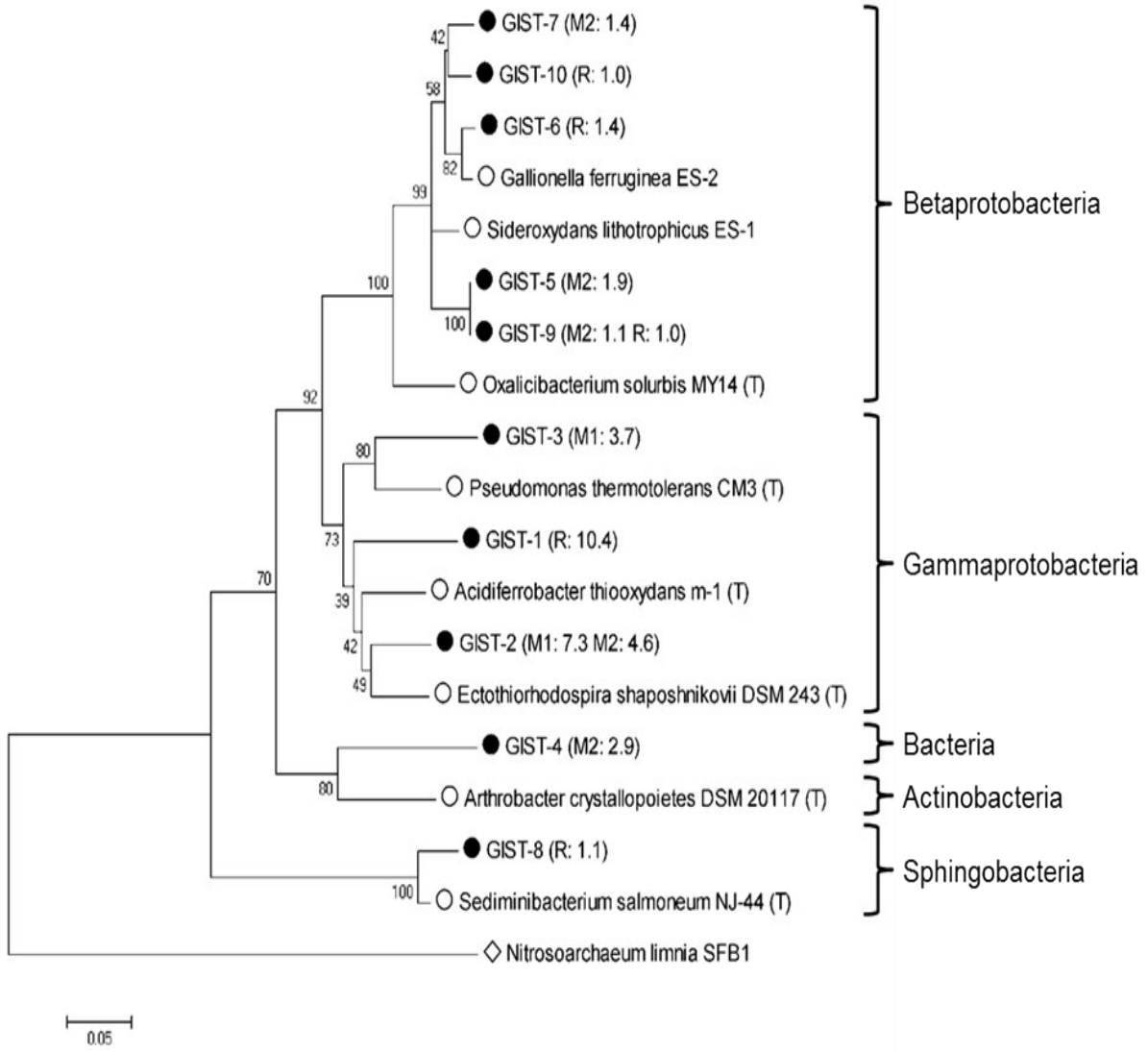

Figure S5. Neighbor-joining tree of unclassified dominant bacteria (above 1%) with proportion and the closest type strains. RO feed (R), the front membrane (M1), and the last membrane (M2). The phylogeny was analyzed using 500 bootstrap replicates, and was rooted using the archaeal strain *Nitrosoarchaeum limnia* SFB1.

Additional Explanation:

To classify unknown dominant bacteria, a neighbor-joining tree was created to investigate dominant bacteria clusters having the most similar type stains. Unknown dominant bacteria were taxonomically classified by comparison with the most similar type stains. In Figure S3, most unclassified sequences are classified as either Betaproteobacteria or Gammaproteobacteria. The most dominant unclassified sequences (GIST-1 and GIST-2), were defined in Gammaproteobacteria. Unknown dominant bacterial genera, more than 1% of the total community, were subsequently selected for further examination and classification; they were then compared to reported uncultured bacteria and type strains based on a pairwise similarity comparison (Table S2). The RO feed, M1, and M2 had relatively high proportions of total unclassified bacterial genus (74.9%, 75.6%, and 72.6%, data not shown) compared to seawater and the brine (5.4%, 8.2%). The unidentified reads were determined to be important bacteria related to the biofouling community, because of the high portions of unidentified sequences. In the case of M1 and M2, several common unclassified sequences existed; e.g., GIST-2, GIST-3, GIST-4, GIST-5, GIST-7, and GIST-9. On the other hand, in the seawater and brine most dominant sequences were classified genera.

Table S1. Obtained qualified sequence reads and ecological indices (Observed species, Chao1, PD, and Equitability) determined by 97% similarity. Ecological indices were obtained after all samples were rarefied by 8,911 and 1,166 sequence reads for bacteria and archaea, respectively.

Sample		Qualified sequences	Observed species	Chao1	PD	Equitability
Bacteria	Seawater	18,804	198	289	20.8	0.40
	RO feed	30,571	799	1137	78.1	0.77
	M1†	8,911	405	504	41.0	0.73
	M2‡	9,196	383	498	38.1	0.70
	Brine	13,208	244	331	24.0	0.37
Archaea	Seawater	1,166	6.0	6.0	0.5	0.09
	RO feed	42,458	21.8	31.6	1.0	0.33
	M1†	4,287	4.9	5.4	0.1	0.20
	M2‡	3,395	4.5	4.7	0.2	0.10
	Brine	1,178	10.9	11.1	0.7	0.12

† First located RO membrane in RO unit.
‡ Last located RO membrane in RO unit.

Additional Explanation:

Pyrosequencing produced a total of about 80,700 and 52,500 qualified reads of bacteria and archaea, respectively. The species richness and diversity were then compared among sites and the numbers of operating taxonomic units (OTUs) were determined at 95% similarity. The bacterial and archaeal ecological indices revealed the same patterns within the samples; overall, the bacterial indices showed more OTUs and more diverse communities than archaea. In detail, the seawater and brine displayed the lowest number of species, diversity, and Shannon evenness index. There were, however, noticeable bacterial community changes from the seawater to RO feed after pretreatment. The most diverse and the highest number of species counted were found in the RO feed within the desalination process. The biofouling communities of the RO membranes displayed the lowest raw sequence numbers, with M1 having similar community indices to M2, though these bacterial communities were more varied and more evenly distributed than in the seawater.

Table S2. Identification of dominant unclassified bacteria (≥ 1 %) in the RO feed (R), the first located RO membrane (M1), the last located RO membrane (M2), with 97% similarity clustering. The pairwise similarity comparison was performed using an EzTaxon-e server.

Unclassified ID (relative abundance)	Closest relative (Pairwise Similarity/Accession#)	Closet cultured (Pairwise Similarity/Accession#)
GIST-1 (R: 10.4%)	Uncultured gammaproteobacterium from uranium mining wastes /94.30%/FM877664	*Acidiferrobacter thiooxydans* strain m-1(T) /88.12%/AF387301
GIST-2 (M1: 7.3%, M2: 4.6%)	Uncultured soil bacterium from oil polluted soil/99.17%/DQ378269	*Ectothiorhodospira shaposhnikovii* DSM 243(T) /90.64%/M59151
GIST-3 (M1: 3.7%)	Uncultured gamma proteobacterium from a contaminated coastal sediment/89.94%/FM242357	*Pseudomonas thermotolerans* CM3(T) /86.2%/AJ311980
GIST-4 (M2: 2.9%)	Uncultured bacterium from sediments of eutrophic reservoir/93.63%/AY955088	*Arthrobacter crystallopoietes* DSM 20117(T) / 84.21%/X80738
GIST-5 (M2: 1.9%)	Uncultured bacterium clone from uranium-contaminated aquifer/93.15 %/ AY532538	*Sideroxydans lithotrophicus* ES-1/90.74%/ACVF01000013
GIST-6 (R: 1.4%)	Uncultured *Gallionella* sp. clone TrefC4 16S ribosomal RNA gene from streamer in acidic, iron-rich spa water/98.04%/AY766002	*Gallionella ferruginea* ES-2/97.70%/ CP002159
GIST-7 (M2: 1.4%)	Uncultured bacterium clone BG.f2 16S ribosomal RNA gene from bench glacier/96.64%/ DQ228376	*Sideroxydans lithotrophicus* ES-1/94.40%/ACVF01000013
GIST-8 (R: 1.1%)	Flavobacteria bacterium KF030 from freshwater/99.36%/AB269814	*Sediminibacterium salmoneum* NJ-44(T)/92.89%/EF407879
GIST-9 (M2: 1.1%, R: 1.0%)	Uncultured bacterium clone 1013-1-CG10 16S ribosomal RNA gene from uranium-contaminated aquifer/93.14%/ AY532538	*Oxalicibacterium solurbis* MY14(T)/90.72%/ AB008503
GIST-10 (R: 1.0%)	Uncultured bacterium clone BG.f2 16S ribosomal RNA gene from bench glacier/97.99%/ DQ228376	*Sideroxydans lithotrophicus* ES-1/94.42%/ACVF01000013

Additional Explanation:

The unidentified reads were determined to be important bacteria related to the biofouling community, because of the high portions of unidentified sequences. Unknown dominant bacteria were taxonomically classified by comparison with the most similar type stains. In Figure S3, most unclassified sequences are classified as either Betaproteobacteria or Gammaproteobacteria. The most dominant unclassified sequences (GIST-1 and GIST-2), were defined in Gammaproteobacteria. Unknown dominant bacterial genera, more than 1% of the total community, were subsequently selected for further examination and classification; they were then compared to reported uncultured bacteria and type strains based on a pairwise similarity comparison (Table S2). The unidentified reads were determined to be important bacteria related to the biofouling community, because of the high portions of unidentified sequences. In the case of M1 and M2, several common unclassified sequences existed; e.g., GIST-2, GIST-3, GIST-4, GIST-5, GIST-7, and GIST-9. On the other hand, in the seawater and brine most dominant sequences were classified genera.

References

Llobet-Brossa, E., Rosselló-Mora, R., Amann, R. 1998. Microbial Community Composition of Wadden Sea Sediments as Revealed by Fluorescence In Situ Hybridization. *Applied and Environmental Microbiology*, **64**(7), 2691-2696.

Manz, W., Amann, R., Ludwig, W., Wagner, M., Schleifer, K.-H. 1992. Phylogenetic Oligodeoxynucleotide Probes for the Major Subclasses of Proteobacteria: Problems and Solutions. *Systematic and Applied Microbiology*, **15**(4), 593-600.