

TABLE S1. List of chitinase gene sequences extracted from the GOS data set and corresponding 16S rRNA gene sequences used for UniFrac analyses. Values for the average genome equivalents for sampling site GS001-GS047 were adopted from previous publications (Howard et al. 2008 (14); Biers et al. 2009 (3)). Values for the remaining sampling sites were estimated analogously by using the number of reads containing sequences for single copy genes from the respective samples. The fraction of bacteria containing chitinase genes was estimated by assuming 5 chitinase genes per cell (Cottrell et al. 1999 (7)).

sampling site	chitinase-hits (after removing duplicate reads)	groupA	groupC	included 16S reads	average genome equivalents	% chitin degrader
GS000a	60 (58)	56	4	10	257	4.5%
GS000b	6	6	0	10	245	0.5%
GS000c	13	13	0	10	261	1.0%
GS000d	9	9	0	10	229	0.8%
GS001a*	2	2	0	0	27	1.5%
GS001b*	1	1	0	0	25	0.8%
GS001c	2	2	0	4	82	0.5%
GS002	0	0	0	0	53	nd
GS003	3	3	0	8	28	2.1%
GS004	0	0	0	0	28	nd
GS005	3	3	0	8	21	2.8%
GS006	1	1	0	4	30	0.7%
GS007	0	0	0	0	29	nd
GS008	2	1	1	10	75	0.5%
GS009	2	2	0	9	40	1.0%
GS010	0	0	0	0	53	nd
GS011	3	3	0	10	72	0.8%
GS012	6	6	0	10	67	1.8%
GS013	7	6	1	10	47	3.0%

GS014	1	1	0	10	81	0.2%
GS015	3	3	0	10	75	0.8%
GS016	2	2	0	10	73	0.5%
GS017	10	10	0	10	132	1.5%
GS018	4	4	0	10	86	0.9%
GS019	1	1	0	10	80	0.3%
GS020	21	20	1	30	135	3.1%
GS021	5 (4)	5	0	10	69	1.2%
GS022	3 (2)	3	0	10	74	0.5%
GS023	0	0	0	0	81	nd
GS025*	4	4	0	0	20	4.1%
GS026	3	3	0	8	59	1.0%
GS027	4	4	0	10	135	0.6%
GS028	3	3	0	10	114	0.5%
GS029	6	6	0	10	79	1.5%
GS030	29 (21)	29	0	10	265	1.6%
GS031	10	10	0	10	281	0.7%
GS032	5	5	0	10	67	1.5%
GS033	19 (18)	19	0	30	286	1.3%
GS034	1	1	0	10	78	0.3%
GS035	3	3	0	10	84	0.7%
GS036	2	2	0	10	45	0.9%
GS037	0	0	0	0	39	nd
GS038	0	0	0	0	<1	nd
GS039	0	0	0	0	<1	nd
GS040	0	0	0	0	<1	nd
GS041	0	0	0	0	<1	nd
GS042	0	0	0	0	<1	nd
GS043	0	0	0	0	<1	nd
GS044	0	0	0	0	<1	nd
GS045	0	0	0	0	<1	nd

GS046	0	0	0	0	<1	nd
GS047	2	2	0	8	41	1.0%
GS048a	1	1	0	7	42	0.5%
GS049	0	0	0	0	56	nd
GS050	1	1	0	1	<1	nd
GS051	0	0	0	0	63	nd
GS108a	0	0	0	0	23	nd
GS108b*	2	2	0	0	16	2.5%
GS109	0	0	0	0	36	nd
GS110a	2	2	0	10	46	0.9%
GS110b*	1	1	0	0	12	1.7%
GS111	0	0	0	0	32	nd
GS112a	1	1	0	10	48	0.4%
GS112b*	1	1	0	0	12	1.6%
GS113	0	0	0	0	63	nd
GS114	2	2	0	10	171	0.2%
GS115	0	0	0	0	33	nd
GS116	1	1	0	6	33	0.6%
GS117a	5	5	0	10	175	0.6%
GS117b*	0	0	0	0	16	nd
GS119	0	0	0	0	35	nd
GS120	0	0	0	0	24	nd
GS121	2	2	0	10	65	0.6%
GS122a	2	2	0	10	52	0.8%
GS122b*	3	3	0	0	10	5.8%
GS123	1	1	0	10	61	0.3%
GS148	7	7	0	9	58	2.4%
GS149	4	4	0	10	63	1.3%
total	297 (284)	290	7	468	4277	1.3%

* derived from particle size fractions >0.8 and therefore not included into UniFrac analyses. nd: not detected

TABLE S2. Sampling sites categories and environmental data for individual sampling sites from the GOS metagenome (Rusch et al. 2007 (32)).

Latitude classification was defined by allocating sampling sites with a latitude $< 23.5^\circ$ to the tropic climate zone. The remaining sampling sites were grouped into a non tropic climate zone.

Sampling site	Habitat Type	Salinity habitat	Latitude habitat	Geographic Location habitat	Latitude	Wat. Dep. [m]	Chlorophyll [μgL^{-1}]	Salinity [psu]	Temp [$^\circ\text{C}$]	Coll. Date
GS000	Open Ocean	marine	non Tropics	Sargasso Sea	31°10'30N"	4200	0.17	36.7	20.5	2/25/2003
GS001	Open Ocean	marine	non Tropics	Sargasso Sea	32°10'00N"	4200	0.1	36.7	22.9	5/15/2003
GS003	Coastal	marine	non Tropics	North American East Coast	42°51'10N"	119	1.4	29.9	11.7	8/21/2003
GS005	Embayment	marine	non Tropics	North American East Coast	44°41'25N"	64	6	30.2	15	8/23/2003
GS006	Estuary	estuarine	non Tropics	North American East Coast	45°6'42N"	11	2.8	NA	11.2	8/23/2003
GS008	Coastal	marine	non Tropics	North American East Coast	41°29'9N"	12	2.2	26.5	9.4	11/17/2003
GS009	Coastal	marine	non Tropics	North American East Coast	41°5'28N"	32	4	31	11	11/17/2003
GS011	Estuary	estuarine	non Tropics	North American East Coast	38°56'24N"	8	4.8	NA	11	11/18/2003
GS012	Estuary	estuarine	non Tropics	North American East Coast	39°25'4N"	25	21	3.5	1	12/18/2003
GS013	Coastal	marine	non Tropics	North American East Coast	38°56'49N"	20	3	NA	9.3	12/19/2003
GS014	Coastal	marine	non Tropics	North American East Coast	36°0'14N"	31	1.7	NA	18.6	12/21/2003
GS015	Coastal	marine	non Tropics	Caribbean Sea	32°30'25N"	47	0.2	36	25	1/8/2004
GS016	Coastal Sea	marine	non Tropics	Caribbean Sea	24°29'18N"	3333	0.16	35.8	26.4	1/8/2004
GS017	Open Ocean	marine	Tropics	Caribbean Sea	24°10'29N"	4513	0.13	35.8	27	1/9/2004
GS018	Open Ocean	marine	Tropics	Caribbean Sea	20°31'21N"	4470	0.14	35.4	27.4	1/10/2004
GS019	Coastal	marine	Tropics	Caribbean Sea	18°2'12N"	3336	0.23	35.4	27.7	1/12/2004
GS020	Fresh Water	freshwater	Tropics	Panama Canal	10°42'59N"	4.2	NA	0.1	28.6	1/15/2004
GS021	Coastal	marine	Tropics	Eastern Tropical Pacific	9°9'52N"	76	0.5	30.7	27.6	1/20/2004
GS022	Open Ocean	marine	Tropics	Eastern Tropical Pacific	8°7'45N"	2431	0.33	32.3	29.3	1/21/2004
GS025	Fringing Reef	marine	Tropics	Eastern Tropical Pacific	6°29'34N"	30	0.11	31.4	28.3	1/28/2004
GS026	Open Ocean	marine	Tropics	Galapagos Islands	5°33'10N"	2386	0.22	32.6	27.8	2/2/2004
GS027	Coastal	marine	Tropics	Galapagos Islands	1°15'51N"	2.3	0.4	34.9	25.5	2/4/2004
GS028	Coastal	marine	Tropics	Galapagos Islands	1°12'58S"	156	0.35	NA	NA	2/5/2004
GS029	Coastal	marine	Tropics	Galapagos Islands	1°13'1S"	12	0.4	34.5	26.2	2/9/2004

GS030	Warm Seep	marine	Tropics	Galapagos Islands	0°12'0S"	19	NA	NA	26.9	2/9/2004
GS031	Coastal upwelling	marine	Tropics	Galapagos Islands	0°16'20N"	19.6	0.35	NA	18.6	2/10/2004
GS032	Mangrove	estuarine	Tropics	Galapagos Islands	0°18'4S"	1.6	NA	NA	25.4	2/11/2004
GS033	Hypersaline	hypersaline	Tropics	Galapagos Islands	0°35'38S"	0.3	NA	63.4	37.6	2/19/2004
GS034	Coastal	marine	Tropics	Galapagos Islands	1°13'42S"	35	0.36	NA	27.5	2/20/2004
GS035	Coastal	marine	Tropics	Galapagos Islands	0°22'59S"	71	0.28	34.5	21.8	3/2/2004
GS036	Coastal	marine	Tropics	Galapagos Islands	1°23'21N"	67	0.65	34.6	25.8	3/2/2004
GS047	Open Ocean	marine	Tropics	Tropical South Pacific	0°1'15S"	2400	NA	37.3	28.6	3/29/2004
GS048	Coral Reef	marine	Tropics	Polynesia Archipelagos	10°7'53S"	34	0.095	35.1	28.9	5/18/2004
GS050	Coral Atoll	marine	Tropics	Polynesia Archipelagos	17°28'33S"	24	NA	NA	27.8	5/19/2004
GS108	Lagoon Reef	marine	Tropics	Indian Ocean	15°16'40S"	7	0.108	32.4	25.8	8/4/2005
GS110	Open Ocean	marine	Tropics	Indian Ocean	12°5'33S"	1220	0.128	32.7	27	8/6/2005
GS112a	Open Ocean	marine	Tropics	Indian Ocean	10°26'46S"	4573	0.134	32.5	26.6	8/8/2005
GS114	Open Ocean	marine	Tropics	Indian Ocean	8°30'18S"	3649	0.14	33.1	28.2	8/15/2005
GS116	Open Ocean	marine	Tropics	Indian Ocean	8°30'18S"	2150	0.287	33.1	26.2	8/17/2005
GS117a	Coastal	marine	Tropics	Indian Ocean	4°59'25S"	14	0.206	35.5	26.4	9/9/2005
GS121	Open Ocean	marine	non Tropics	Indian Ocean	4°38'6S"	4309	0.14	35.4	23.1	9/29/2005
GS122	Open Ocean	marine	non Tropics	Indian Ocean	4°36'49S"	4921	0.151	35.8	20.2	9/30/2005
GS123	Open Ocean	marine	non Tropics	Indian Ocean	29°20'56S"	1860	0.228	35.8	20.4	10/1/2005
GS148	Fringing Reef	marine	Tropics	Indian Ocean	30°53'54S"	1	NA	NA	NA	9/11/2005
GS149	Harbor	marine	Tropics	Indian Ocean	32°23'57S"	5	NA	NA	NA	9/12/2005

NA: not available

TABLE S3. RDP Classification of 16S rRNA genes extracted from the GOS database and used for UniFrac analyses.

bacterial class	number of sequences	percent
Spirochaetes	1	0.2%
Mollicutes	1	0.2%
Nitrospira	1	0.2%
Deferribacteres	1	0.2%
Ktedonobacteria	1	0.2%
WS3	1	0.2%
Epsilonproteobacteria	2	0.4%
Deinococci	2	0.4%
Clostridia	3	0.6%
Holophaga	5	1.1%
Opitute	5	1.1%
Betaproteobacteria	12	2.6%
Deltaproteobacteria	15	3.2%
Gammaproteobacteria	19	4.1%
Sphingobacteria	28	6.0%
Actinobacteria	54	11.5%
Cyanobacteria	69	14.7%
Flavobacteria	93	19.9%
Alphaproteobacteria	149	31.8%

References

- Biers, E. J., S. L. Sun, et al. (2009). "Prokaryotic genomes and diversity in surface ocean waters:interrogating the Global Ocean Sampling metagenome." Applied and Environmental Microbiology **75**(7): 2221-2229.
- Cottrell, M. T., J. A. Moore, et al. (1999). "Chitinases from uncultured marine microorganisms." Applied and Environmental Microbiology **65**(6): 2553-2557.
- Howard, E. C., S. L. Sun, et al. (2008). "Abundant and diverse bacteria involved in DMSP degradation in marine surface waters." Environmental Microbiology **10**(9): 2397-2410.
- Rusch, D. B., A. L. Halpern, et al. (2007). "The Sorcerer II Global Ocean Sampling Expedition: Northwest Atlantic through Eastern Tropical Pacific." PLoS Biol **5**(3): e77.