Ecological Genomics of the Uncultivated Marine Roseobacter Lineage CHAB-I-5

Yao Zhang, Ying Sun, Nianzhi Jiao, Ramunas Stepanauskas, Haiwei Luo

Supplemental Results

Organic carbon uptake and catabolism

The CHAB-I-5 and DC5-80-3 lineages are likely to be more versatile in the uptake of various carbohydrate compounds compared to the other two uncultivated lineages SAG-O19 and NAC11-7, and most of the six cultured pelagic strains (Table 4). The alpha-glucoside transport system (*aglEFGK*), the rhamnose transport system (*rhaPSQT*), the fructose transport system (*frcABC*), the ribose transport system (*rbsABC*), the sorbitol/mannitol transport system (*smoEFG*), and the D-xylose transport system (*xylFGH*), for instance, are consistently present in the CHAB-I-5 and DC5-80-3 lineages and differentially present in the SAG-O19 and NAC11-7 lineages, but are completely or largely missing in the cultured pelagic strains Y4I, DSS-3, HTCC2597, and GAI101. Most of these carbohydrates are taken up by the type strain *Planktomarina temperate* RCA23 of the DC5-80-3 lineage (1), and most of these transporters are functional in *Phaeobacter gallaeciensis* BS107 (2), a scallop-associated strain (3) closely related to *Phaeobacter* sp Y4I analyzed here. These mechanisms are therefore likely to be important for the ecological success of lineages CHAB-I-5 and DC5-80-3, and perhaps lineages SAG-O19 and NAC11-7 as well. Several other sugar uptake pathways, however, are completely or largely missing in the four dominant lineages and instead are present in a few cultured strains (Table 4). These include the raffinose/stachyose/melibiose

transport system (*msmEFG*), the maltose/maltodextrin transport system (*malEFG*), the glucose/mannose transport system (*gtsABC*), and the glycerol transport system (*glpVPQST*). In addition to the ABC transporters, bacteria often rely on another type of energy-intensive phosphotransferase system (PTS) for the uptake of various monosaccharides, disaccharides, amino sugars, polyols, and other sugar derivatives (4), and this system is universal among Roseobacters.

All analyzed genomes encode similar potential for organic acid uptake through secondary transporters (Table 4). For instance, the tripartite ATP-independent periplasmic (TRAP) transporter (*dctPQM*) (5) for the uptake of the C_4-dicarboxylate compounds (e.g., malate, fumarate, succinate) and the tripartite tricarboxylate transporter (TTT) system (*tctABC*) for the uptake of tricarboxylic acids, as well as their regulatory genes (*dctBD* and *tctDE*, respectively) are prevalent among all these pelagic Roseobacters. Utilization of aromatic compounds derived from vascular plant as a carbon source is another important feature of Roseobacters (6). The key gene encoding an aromatic-ring-cleaving enzyme, protocatechuate 3,4-dioxygenase (*pcaH*) (7), is present in the SAG-O19 and DC5-80-3 lineages and in four of the six cultured genomes (Table 4). Likewise, The feruloyl-CoA synthetase (*ferA*) crucial for the degradation of ferulic acid in lignin (8) is similarly preserved among these Roseobacters.

Nitrogen, phosphorus, and iron acquisition, and vitamin synthesis

The ocean surface waters in general are a nitrogen (N)-limiting environment (9). Extracellular proteins are an important N source for marine microbes. This is evident by using the activity of leucine aminopeptidase (*pepL*) as a proxy for the potential of extracellular peptide use (10). This gene, along with those encoding other

aminopeptidases such as proline aminopeptidase (*pepP*), alanine aminopeptidase (*pepN*), and methionine aminopeptidase (*pepM*), are consistently present in all Roseobacters analyzed here, suggesting that extracellular proteins may be a common substrate for pelagic Roseobacters. Other genes universal to these Roseobacters for the uptake of nitrogenous compounds include the high-affinity ammonium transporter (*amtB*), the L-amino acid transport system (*aapJQMP*), the branched amino acid transport system (*livGFHMK*), assimilation of amine into glutamine (*glnA*) and glutamate (*gltBD*), and their regulatory system P-II (*glnBD*, *ntrBC*) and the *ntrXY* two-component system.

The polyamine compounds (e.g., spermidine, putrescine, cadaverine, norspermidine, and spermine) comprise an important N source for marine bacteria (11). These compounds are taken by the spermidine-preferential (*potABCD*) and the putrescine-specific (*potFGHI*) uptake systems (12, 13). At least one of these transporters was identified in each of the Roseobacters analyzed here. Another important nitrogenous compound is urea, making up to over 50% of total N uptake by planktonic community in some ocean regions (14, 15). The genes encoding urease (*ureABCDEFG*) are missing only in one dominant lineage (NAC11-7) and one cultured strain (Y4I), though not all urease-containing Roseobacters carry genes for urea transporters (*urtABCDE*). Opines are derivatives of amino acids and thus likely to be a relevant N source for some ocean bacteria such as SAR11 (16), but it may not be used by pelagic Roseobacters, since the relevant transporter systems (*occTPMQ*) are missing in all of them. Utilization of nitrate/nitrite as N sources involves genes encoding nitrate/nitrite transport system (*nasDEF*) and assimilatory nitrate reductase (*nasA*) and nitrite reductase (*nirB*). These are completely missing in lineages SAG-O19, DC5-80-3, and NAC11-7 and four out of the

six cultured pelagic strains (Table 4). The presence of the transporter but absence of the reductases in CHAB-I-5 is likely due to the incomplete genome recovery, and further sequencing is needed to validate this hypothesis.

Inorganic phosphate is the preferred phosphorus (P) source for marine microbes, but it is often depleted in surface water and even below the detection limit in some regions, such as Sargasso Sea and Mediterranean Sea (17, 18). Under phosphate starving condition, bacteria often use the high-affinity phosphate transport system (*pstABCS*) for uptake of phosphate. This system was identified in all the Roseobacters except lineage NAC11-7. In addition, bacteria explore phosphoesters through alkaline phosphatases (APases) (19, 20) and phosphonates via C-P lyase (*phnGHIJM*) (21), the latter mechanism is completely missing in the four largely uncultivated lineages. Among the three gene families (*phoA*, *phoX*, *phoD*) encoding APases, the *phoX* is present in CHAB-I-5 but missing in the other three dominant lineages, and the *phoA* and *phoD* are completely missing in all four major lineages. Notably, *phoX* allows Roseobacters to utilize phosphoesters at low concentration (22), and thus may confer an ecological advantage for CHAB-I-5 to thrive in P depleted waters. These genes are better represented in the cultured pelagic strains. While *phoA* is completely missing, *phoX*, *phoD*, and C-P lyase were found in at least half of the six cultured pelagic strains (Table 4).

Iron and vitamin often becomes limiting to ocean microbes (9, 23). The iron (III) transport system (*afuAB*) and the ferrous iron transporter (*feoB*) were identified in lineage CHAB-I-5. These genes are largely missing in the other three dominant lineages, except the *afuA* found in SAG-O19 and the *afuB* in NAC11-7. They are similarly infrequent in

the six cultured pelagic Roseobacters (Table 4). For example, the *afuAB* system is only present in HTCC2516 but incomplete in GAI101 and CCS1, and the *feoB* system is missing in all the six genomes. Some other iron transporters were exclusively found in the cultured strains, though they are all rare. The iron (III) dicitrate transporter (*fecA*) and the iron permease (*ftr1*), for instance, are present only in one and two strains, respectively. Siderophore and heme are important iron-binding ligands in a few Roseobacter strains (24, 25). The gene clusters encoding the heme uptake system, the siderophore biosynthesis, and the siderophore transporter are completely missing in the four largely uncultivated lineages. While the former two clusters are similarly rare in the six cultured pelagic strains, the siderophore transport system is present in four of them. All the Roseobacters analyzed here are likely to make vitamin B_{12} (cobalamin) and B_6 (pyridoxine), since the key genes for synthesizing the former (*cobCK*) and the latter (*pdxHJ*) are universal to them. By contrast, the key genes responsible for vitamin B_1 (thiamine; *thiCL*) and B_7 (biotin; *bioB*) biosynthesis are missing in all the dominant lineages, and present in a subset of the six cultured pelagic strains (Table 4).

References

1. **Giebel H-A, Kalhoefer D, Gahl-Janssen R, Choo Y-J, Lee K, Cho J-C, Tindall BJ, Rhiel E, Beardsley C, Aydogmus ÖO, Voget S, Daniel R, Simon M, Brinkhoff T.** 2013. *Planktomarina temperata* gen. nov., sp. nov., belonging to the globally distributed RCA cluster of the marine *Roseobacter* clade, isolated from the German Wadden Sea. International Journal of Systematic and Evolutionary Microbiology **63:**4207-4217.
2. **Wiegmann K, Hensler M, Wöhlbrand L, Ulbrich M, Schomburg D, Rabus R.** 2014. Carbohydrate catabolism in *Phaeobacter inhibens* DSM 17395, a member of the marine *Roseobacter* clade. Applied and Environmental Microbiology doi:10.1128/aem.00719-14.
3. **Ruiz-Ponte C, Cilia V, Lambert C, Nicolas JL.** 1998. *Roseobacter gallaeciensis* sp. nov., a new marine bacterium isolated from rearings and collectors of the scallop *Pecten maximus*. Int J Syst Bacteriol **48:**537-542.
4. **Kotrba P, Inui M, Yukawa H.** 2001. Bacterial phosphotransferase system (PTS) in carbohydrate uptake and control of carbon metabolism. Journal of Bioscience and Bioengineering **92:**502-517.
5. **Mulligan C, Fischer M, Thomas GH.** 2011. Tripartite ATP-independent periplasmic (TRAP) transporters in bacteria and archaea. FEMS Microbiol Rev **35:**68-86.
6. **Buchan A, Gonzalez JM, Moran MA.** 2005. Overview of the marine *Roseobacter* lineage. Applied and Environmental Microbiology **71:**5665-5677.
7. **Buchan A, Collier LS, Neidle EL, Moran MA.** 2000. Key aromatic-ring-cleaving enzyme, protocatechuate 3,4-dioxygenase, in the ecologically important marine Roseobacter lineage. Applied and Environmental Microbiology **66:**4662-4672.
8. **Masai E, Harada K, Peng X, Kitayama H, Katayama Y, Fukuda M.** 2002. Cloning and characterization of the ferulic acid catabolic genes of *Sphingomonas paucimobilis* SYK-6. Applied and Environmental Microbiology **68:**4416-4424.
9. **Moore CM, Mills MM, Arrigo KR, Berman-Frank I, Bopp L, Boyd PW, Galbraith ED, Geider RJ, Guieu C, Jaccard SL, Jickells TD, La Roche J, Lenton TM, Mahowald NM, Maranon E, Marinov I, Moore JK, Nakatsuka T, Oschlies A, Saito MA, Thingstad TF, Tsuda A, Ulloa O.** 2013. Processes and patterns of oceanic nutrient limitation. Nature Geosci **6:**701-710.
10. **Nagata T.** 2008. Organic matter-bacteria interactions in seawater. *In* Kirchman DL (ed), Microbial ecology of the oceans. John Wiley & Sons, New Jersey.
11. **Lee C, Jørgensen NG.** 1995. Seasonal cycling of putrescine and amino acids in relation to biological production in a stratified coastal salt pond. Biogeochemistry **29:**131-157.
12. **Terui Y, Saroj S, Sakamoto A, Yoshida T, Higashi K, Kurihara S, Suzuki H, Toida T, Kashiwagi K, Igarashi K.** 2014. Properties of putrescine uptake by PotFGHI and PuuP and their physiological significance in *Escherichia coli*. Amino Acids **46:**661-670.

13. **Hanfrey CC, Pearson BM, Hazeldine S, Lee J, Gaskin DJ, Woster PM, Phillips MA, Michael AJ.** 2011. Alternative spermidine biosynthetic route is critical for growth of Campylobacter jejuni and is the dominant polyamine pathway in human gut microbiota. The Journal of Biological Chemistry **286:**43301-43312.
14. **Kudela RM, Cochlan WP.** 2000. Nitrogen and carbon uptake kinetics and the influence of irradiance for a red tide bloom off southern California. Aquatic Microbial Ecology **21:**31-47.
15. **Solomon CM, Collier JL, Berg GM, Glibert PM.** 2010. Role of urea in microbial metabolism in aquatic systems: a biochemical and molecular review. Aquatic Microbial Ecology **59:**67-88.
16. **Smith DP, Thrash JC, Nicora CD, Lipton MS, Burnum-Johnson KE, Carini P, Smith RD, Giovannoni SJ.** 2013. Proteomic and transcriptomic analyses of "*Candidatus* Pelagibacter ubique" describe the first PII-independent response to nitrogen limitation in a free-living Alphaproteobacterium. mBio **4:**e00133-00112.
17. **Wu J, Sunda W, Boyle EA, Karl DM.** 2000. Phosphate Depletion in the Western North Atlantic Ocean. Science **289:**759-762.
18. **Thingstad TF, Krom MD, Mantoura RFC, Flaten GAF, Groom S, Herut B, Kress N, Law CS, Pasternak A, Pitta P, Psarra S, Rassoulzadegan F, Tanaka T, Tselepides A, Wassmann P, Woodward EMS, Riser CW, Zodiatis G, Zohary T.** 2005. Nature of phosphorus limitation in the ultraoligotrophic eastern Mediterranean. Science **309:**1068-1071.
19. **Duhamel S, Bjorkman KM, Wambeke FV, Moutin T, Karl DM.** 2011. Characterization of alkaline phosphatase activity in the North and South Pacific Subtropical Gyres: Implications for phosphorus cycling. Limnology and Oceanography **56:**1244-1254.
20. **Martinez J, Smith DC, Steward GF, Azam F.** 1996. Variability in ectohydrolytic enzyme activities of pelagic marine bacteria and its significance for substrate processing in the sea. Aquatic Microbial Ecology **10:**223-230.
21. **Dyhrman ST, Chappell PD, Haley ST, Moffett JW, Orchard ED, Waterbury JB, Webb EA.** 2006. Phosphonate utilization by the globally important marine diazotroph Trichodesmium. Nature **439:**68-71.
22. **Sebastian M, Ammerman JW.** 2011. Role of the phosphatase PhoX in the phosphorus metabolism of the marine bacterium *Ruegeria pomeroyi* DSS-3. Environmental Microbiology Reports **3:**535-542.
23. **Sañudo-Wilhelmy SA, Cutter LS, Durazo R, Smail EA, Gómez-Consarnau L, Webb EA, Prokopenko MG, Berelson WM, Karl DM.** 2012. Multiple B-vitamin depletion in large areas of the coastal ocean. Proceedings of the National Academy of Sciences **109:**14041-14045.
24. **Thole S, Kalhoefer D, Voget S, Berger M, Engelhardt T, Liesegang H, Wollherr A, Kjelleberg S, Daniel R, Simon M, Thomas T, Brinkhoff T.** 2012. *Phaeobacter gallaeciensis* genomes from globally opposite locations reveal high similarity of adaptation to surface life. ISME J **6:**2229-2244.

25. **Roe KL, Hogle SL, Barbeau KA.** 2013. Utilization of heme as an iron source by marine *Alphaproteobacteria* in the *Roseobacter* Clade. Applied and Environmental Microbiology **79:**5753-5762.

Supplemental Figure Legends

Figure S1. The maximum likelihood16S rRNA gene tree of cultured Roseobacters and the uncultivated lineages (CHAB-I-5, SAG-O19, DC5-80-3, NAC11-7). Numbers near the uncultivated lineages are bootstrap values of these branches. The five single cells sequenced in the present study are marked with star.

Figure S2. Estimate of the complete genome size of the five new single-cell amplified genomes (red) based on a polynomial regression model (R^2=0.95). The 95% prediction interval (PI) is also shown. The x axis shows the ratio of the number of conserved single-copy genes universally present in fully sequenced Roseobacter genomes to the number of predicted protein-coding genes in a genome; the y axis is the number of nucleotides sequenced.

Figure S3. Estimate of the relative abundance of the Roseobacter clade in bacterioplankton communities in global oceans. The equator and boundaries of the main climate zones (tropic, temperate, polar) are marked in dashed lines. Each solid dot represents at least one metagenomic samples observed within the ±0.5 ° latitude/longitude range represented by the center of the dot, and is colored according to the relative abundance of a particular lineage in the Roseobacter community.

Figure S4. Geographic distribution of samples. The genomic samples of the four largely uncultivated Roseobacter lineages are colored, and the metagenomic samples from are symbolized using various shapes. Round circles represent samples from GOS, GOS Baltic Sea and GOS Banyoles projects; Xs represent GOM project; Top-up triangles represent CalCOFI project; Asterisks represent Tara Ocean Expedition project; Crosses represent metagenome data from Monterey Bay; Top-down triangles represent data from Western Channel; Left-pointing triangles represent microbial community studies of North Atlantic Spring Bloom; Right-pointing triangle represent metagenome data for marine bacteriaplankton; Squares represent studies for microbial community in a shallow-sea hydrothermal system near Kueishantao; Rhombuses represent samples from the Hawaii Ocean Time-series (HOT) station ALOHA, as well as data from Mediterranean deep chlorophyll maximum studies; Stars represent the Antarctica aquatic microbial metagenome project (Table S2).

Figure S5. Photosynthetic gene cluster (PGC) structure and arrangement in Roseobacters. Green: *bch* genes; orange: *crt* genes; red: *puf* and regulators genes; pink: *puh* genes; cyan: *pucC* genes; black: *acsf* genes; blue: *hem* and *cyc* genes; yellow: *LhaA* genes; white: uncertain or unrelated genes. The horizontal arrows represent putative transcripts. According to the orientation of the two conserved regions, the

PGCs are divided into three different groups. type I, forward *bchFNBHLM-LhaA-puh* plus forward *crtF-bchCXYZ-puf*; type II, forward *bchFNBHLM-LhaA-puh* plus reverse *crtF-bchCXYZ-puf*; type III, forward *crtF-bchCXYZ-puf* plus forward *bchFNBHLM-LhaA-puh*. Double slash indicates the genes are located outside the PGC.

Figure S6. RAxML Maximum likelihood phylogenetic analysis of the four types of NRPS/PKS proposed by Martens et al., Microb. Ecol. 54:31-42, 2007. Among the four largely uncultivated Roseobacter lineages and the six cultured pelagic Roseobacters, only two genes (red) homologous to these NRPS/PKS were identified. Other Roseobacter sequences from Martens et al. 2007 are in blue.

Fig. S1

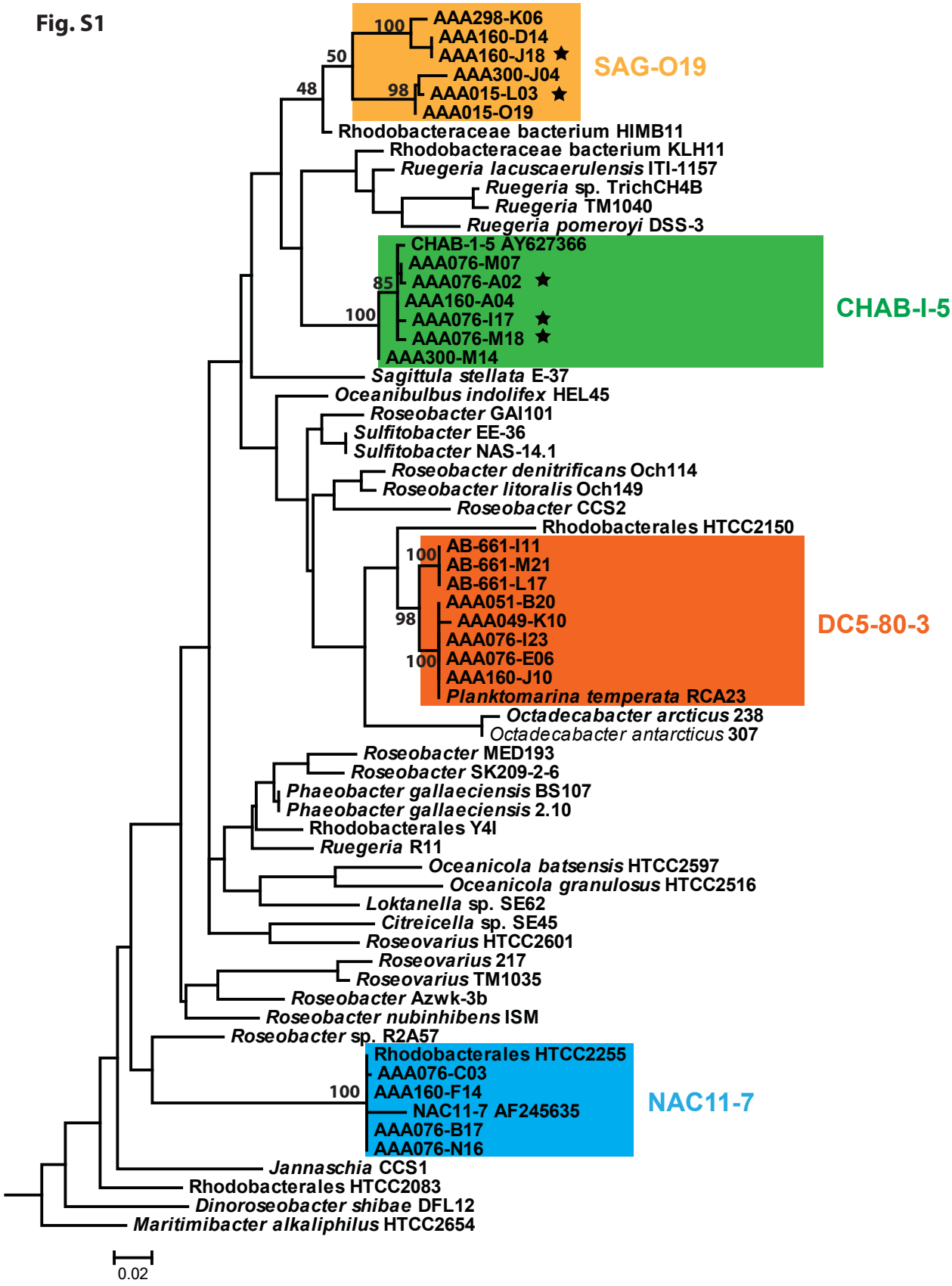
AAA298-K06
AAA160-D14
AAA160-J18
AAA300-J04
AAA015-L03
AAA015-O19
SAG-O19
Rhodobacteraceae bacterium HIMB11
Rhodobacteraceae bacterium KLH11
Ruegeria lacuscaerulensis ITI-1157
Ruegeria sp. TrichCH4B
Ruegeria TM1040
Ruegeria pomeroyi DSS-3
CHAB-1-5 AY627366
AAA076-M07
AAA076-A02
AAA160-A04
AAA076-I17
AAA076-M18
AAA300-M14
CHAB-I-5
Sagittula stellata E-37
Oceanibulbus indolifex HEL45
Roseobacter GAI101
Sulfitobacter EE-36
Sulfitobacter NAS-14.1
Roseobacter denitrificans Och114
Roseobacter litoralis Och149
Roseobacter CCS2
Rhodobacterales HTCC2150
AB-661-I11
AB-661-M21
AB-661-L17
AAA051-B20
AAA049-K10
AAA076-I23
AAA076-E06
AAA160-J10
Planktomarina temperata RCA23
DC5-80-3
Octadecabacter arcticus 238
Octadecabacter antarcticus 307
Roseobacter MED193
Roseobacter SK209-2-6
Phaeobacter gallaeciensis BS107
Phaeobacter gallaeciensis 2.10
Rhodobacterales Y4I
Ruegeria R11
Oceanicola batsensis HTCC2597
Oceanicola granulosus HTCC2516
Loktanella sp. SE62
Citreicella sp. SE45
Roseovarius HTCC2601
Roseovarius 217
Roseovarius TM1035
Roseobacter Azwk-3b
Roseobacter nubinhibens ISM
Roseobacter sp. R2A57
Rhodobacterales HTCC2255
AAA076-C03
AAA160-F14
NAC11-7 AF245635
AAA076-B17
AAA076-N16
NAC11-7
Jannaschia CCS1
Rhodobacterales HTCC2083
Dinoroseobacter shibae DFL12
Maritimibacter alkaliphilus HTCC2654
100
50
48
98
85
100
100
98
100
100
0.02

Fig. S2

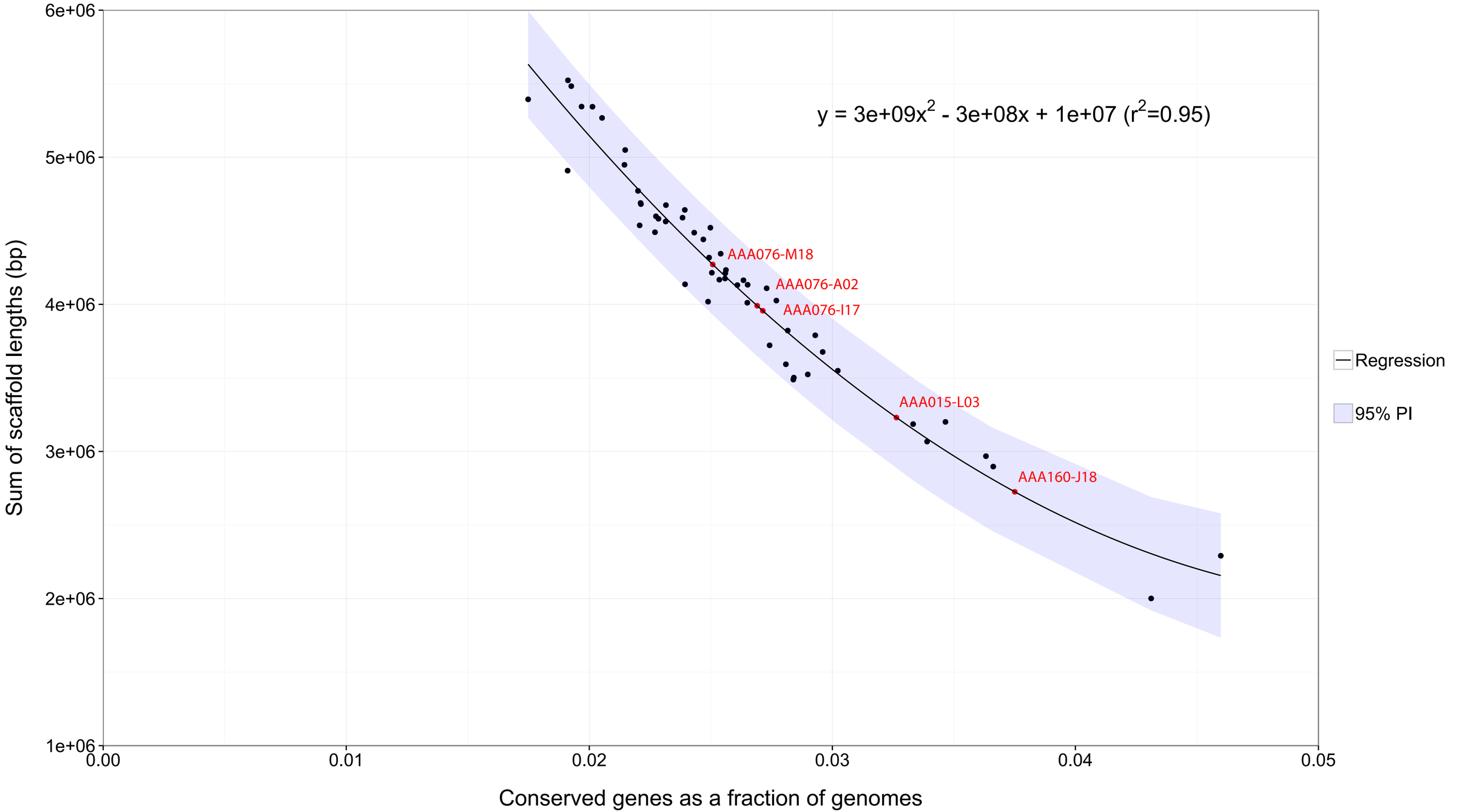
y = 3e+09x^2 - 3e+08x + 1e+07 (r^2=0.95)
AAA076-M18
AAA076-A02
AAA076-I17
AAA015-L03
AAA160-J18
Regression
95% PI
Sum of scaffold lengths (bp)
6e+06
5e+06
4e+06
3e+06
2e+06
1e+06
0.00
0.01
0.02
0.03
0.04
0.05
Conserved genes as a fraction of genomes

Fig. S3

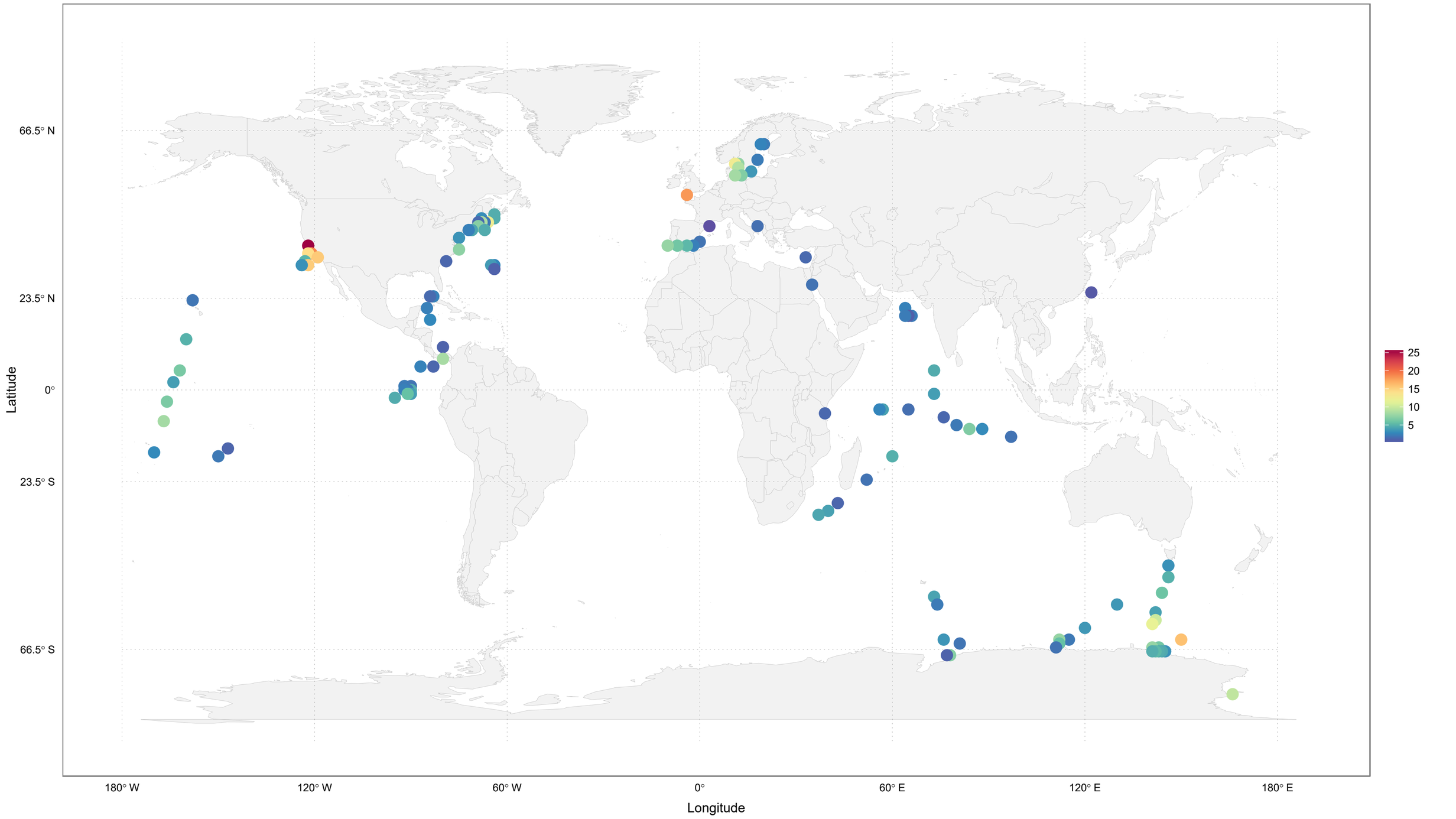
66.5° N
23.5° N
0°
23.5° S
66.5° S
Latitude
180° W
120° W
60° W
0°
60° E
120° E
180° E
Longitude
25
20
15
10
5

Fig. S4

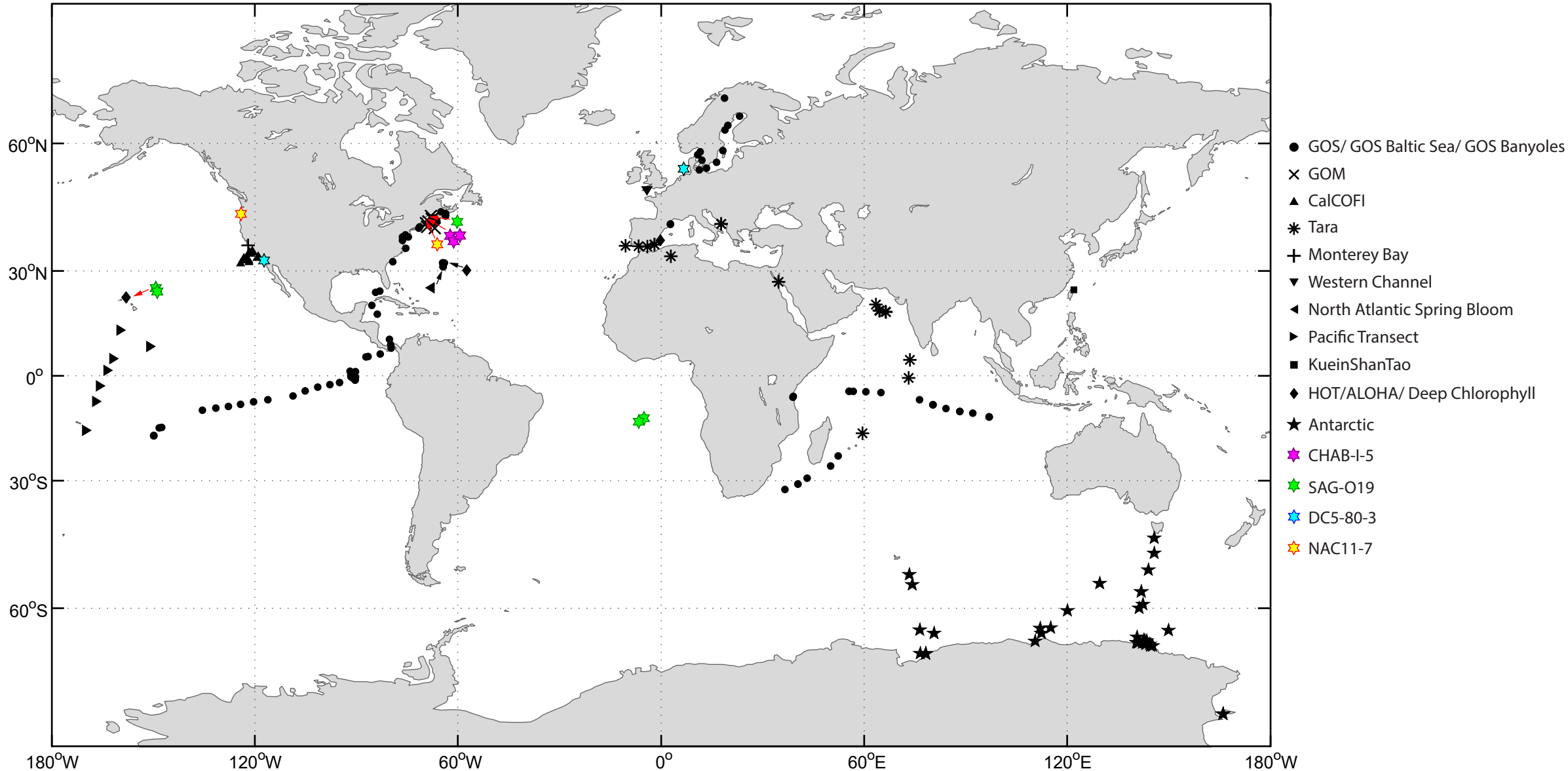
60°N
30°N
0°
30°S
60°S
180°W
120°W
60°W
0°
60°E
120°E
180°W
GOS/ GOS Baltic Sea/ GOS Banyoles
GOM
CalCOFI
Tara
Monterey Bay
Western Channel
North Atlantic Spring Bloom
Pacific Transect
KueinShanTao
HOT/ALOHA/ Deep Chlorophyll
Antarctic
CHAB-I-5
SAG-O19
DC5-80-3
NAC11-7

Fig. S5

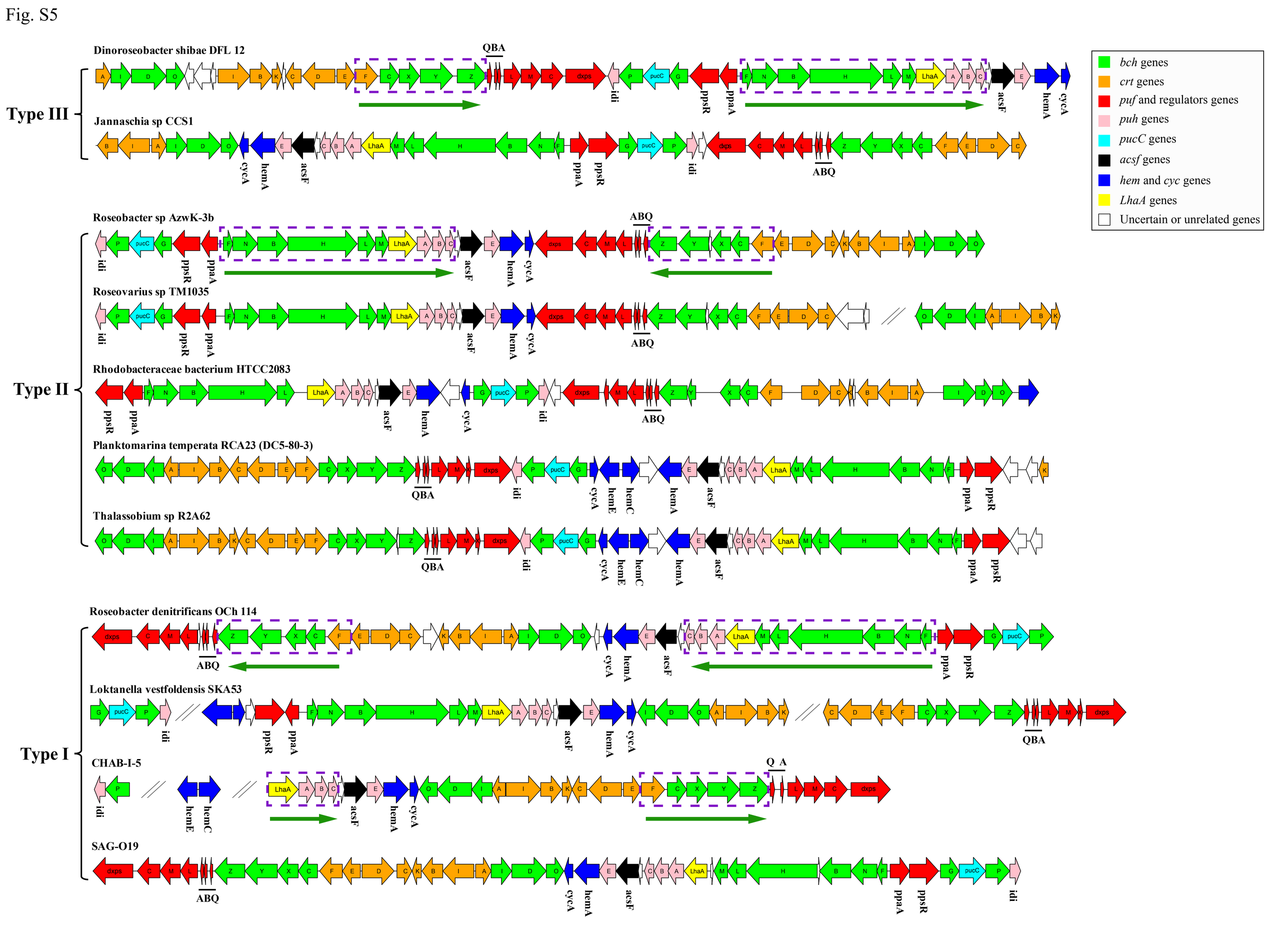
Type III
Dinoroseobacter shibae DFL 12
Jannaschia sp CCS1
Type II
Roseobacter sp AzwK-3b
Roseovarius sp TM1035
Rhodobacteraceae bacterium HTCC2083
Planktomarina temperata RCA23 (DC5-80-3)
Thalassobium sp R2A62
Type I
Roseobacter denitrificans OCh 114
Loktanella vestfoldensis SKA53
CHAB-I-5
SAG-O19
QBA
ABQ
idi
ppsR
ppaA
acsF
hemA
cycA
hemE
hemC
bch genes
crt genes
puf and regulators genes
puh genes
pucC genes
acsf genes
hem and cyc genes
LhaA genes
Uncertain or unrelated genes

Fig. S6

Rattus norvegicus|AAA57219.1
fatty acid synthase
Nitrosomonas europaea ATCC 19718|NP_841435.1
type I polyketide synthase WcbR
Bordetella parapertussis 12822|NP_885141.1
type I polyketide synthase
Bordetella bronchiseptica RB50|NP_889454.1
type I polyketide synthase
type I PKS Proteobacteria
Streptomyces fradiae|AAB66506.1
tylactone synthase modules 4 & 5
Streptomyces hygroscopicus subsp. ascomyceticus|AAF86396.1
FkbA
Saccharopolyspora erythraea NRRL 2338|CAA44448.1
6-deoxyerythronolide B synthase II
Streptomyces rochei|BAA87896.1
lankamycin synthase, partial (plasmid)
Saccharopolyspora erythraea NRRL 2338|CAA44449.1
6-deoxyerythronolide B synthase III
Streptomyces antibioticus|AAF82408.1
8,8a-deoxyoleandolide synthase 1
Streptomyces fradiae|AAB66507.1
tylactone synthase module 6
Micromonospora griseorubida|BAA76543.1
polyketide synthase
Streptomyces natalensis|CAC20930.1
PimS0 protein
Streptomyces venezuelae|AAC69329.1
type I polyketide synthase PikAI
Streptomyces caelestis|AAC46028.1
polyketide synthase module 7
Streptomyces sp. MA6548|AAC68815.1
FK506 polyketide synthase
Streptomyces hygroscopicus subsp. ascomyceticus|AAF86393.1
FkbB
Streptomyces fradiae|AAB66505.1
tylactone synthase module 3
type I PKS Actinobacteria
Chrysosporum ovalisporum|AAM33470.1
polyketide synthase, partial
Microcystis aeruginosa PCC 7806|AAF00959.1
McyD
Anabaena sp. 90|AAO62584.1
polyketide synthase type 1
Nodularia spumigena|AAO64405.1
NdaD
Melittangium lichenicola|CAD89777.1
MelF protein
Stigmatella aurantiaca DW4/3-1|AAF19814.1
MtaF
Nodularia spumigena|AAO64407.1
NdaF
Microcystis aeruginosa PCC 7806|AAF00958.1
McyE
type I PKS Cyanobacteria and Myxobacteria
Myxococcus xanthus|CAB38084.1
Ta1, partial
Pseudomonas syringae pv. tomato str. DC3000|NP_792409.1
yersiniabactin polyketide/non-ribosomal peptide synthetase
Nostoc punctiforme PCC 73102|ZP_00108701.2
COG3321: Polyketide synthase modules and related proteins
Nostoc sp. GSV224|AAF15892.2
NosB
Thalassobius gelatinovorus|AAX51697.1
putative polyketide synthase, partial
Rhodobacteraceae bacterium TY|AAX51695.1
putative polyketide synthase, partial
Paracoccus denitrificans PD1222|EAN68445.1
Beta-ketoacyl synthase:Thioesterase:Acyl transferase reon:Short-chain dehydrogenase/reductase SDR:Phosphopantetheine-binding
Ruegeria pomeroyi DSS-3|SPO_RS04280
Phaeobacter inhibens DSM 16374|AAX51694.1
putative polyketide synthase, partial
Phaeobacter gallaeciensis DSM 26640|AAX51696.1
putative polyketide synthase, partial
Silicibacter sp. TM1040|EAN55851.1
Beta-ketoacyl synthase:Thioesterase:Acyl transferase reon:Phosphopantetheine-binding
marine bacterium HP12|AAX51698.1
putative polyketide synthase, partial
Rhodobacterales bacterium Y4I|RBY4I_RS02870
Roseovarius nubinhibens ISM|ZP_00960961.1
non-ribosomal peptide synthase
Bacillus cereus|AAR87760.2
ZmaK
Streptomyces verticillus|AAG02357.1
polyketide synthase
Chrysosporum ovalisporum|AAM33468.1
peptide synthetase, partial
Nodularia spumigena|AAO64404.1
NdaC
Microcystis aeruginosa PCC 7806|AAF00957.1
McyG
Microcystis aeruginosa|BAB12213.1
peptide synthetase and polyketide synthase
Planktothrix agardhii NIVA-CYA 126/8|CAD29795.1
peptide synthetase
Lyngbya majuscula|AAN32979.1
BarE
Stigmatella aurantiaca DW4/3-1|AAF19812.1
MtaD
Sorangium cellulosum|AAF26920.1
polyketide synthase
Hybrid PKS/NRPS complexes
Bacillus subtilis subsp. spizizenii ATCC 6633|AAF08795.1
MycA
Stigmatella aurantiaca DW4/3-1|AAF19813.1
MtaE
type I PKS Cyanobacteria and Myxobacteria
0.50

Table S1. Summary statistics of the five assembled single-cell amplified genomes

	CHAB-I-5			SAG-O19	
	AAA076-A02	AAA076-M18	AAA076-I17	AAA015-L03	AAA160-J18
Assembly Statistics					
Assembly Size (Mbp)	1.89	2.32	3.14	1.18	2.21
Estimated Genome Size (Mbp)	4.10	4.36	4.07	3.40	2.97
Estimated Genome Recovery (%)	46.02	53.15	77.31	34.71	74.41
Number of Contigs	62	87	98	78	47
Smallest Contig Size (Mbp)	0.002	0.002	0.002	0.002	0.002
Largest Contig Size (Mbp)	0.32	0.31	0.33	0.07	0.33
GC Content (%)	49.70	49.61	49.52	39.61	40.51
N50 (Mbp)	0.09	0.08	0.11	0.03	0.13
Non-Coding RNA					
Number of tRNA	18	25	38	17	34
Number of rRNA	0	3	3	3	3

Table S2 Metadata for the metagenomic datasets used to estimate the relative abundance of the CHAB-I-5, DC5-80-3, NAC11-7, and SAG-O19 lineages in the pelagic Roseobacter communities.

Part I. Project Information and Statistics

Full Name of The Project	PubID	iMicrobeID	# Raw Reads	# Qualified Reads	Length (bp)
Antarctica Aquatic Microbial Metagenome	21124488	4	64,626,265	56,769,086	21,885,405,020
CalCOFI Project	22278668	NA	4,247,911	3,587,839	1,008,676,775
Deep Chlorophyll Maximum Metagenome	20393571	116	1,204,321	1,072,794	301,092,101
Kueishantao Shallow-Sea Hydrothermal System	23940820	NA	565,611	557,579	211,344,022
Gulf of Maine	22050608	NA	2,810,805	2,807,924	2,832,208,962
Global Ocean Sampling Expedition (GOS)	17504484	26	12,672,518	12,453,518	11,555,226,255
Global Ocean Sampling Baltic Sea	NA	114	23,909,459	20,325,101	7,775,555,569
Global Ocean Sampling Banyoles	NA	125	3,990,132	3,658,922	1,462,065,025
Microbial Community at the HOT/ALOHA	16439655	31	5,234,409	4,425,531	1,519,175,594
Marine Bacterioplankton Metagenomes	19444203	7	1,314,590	1,111,670	269,165,248
Monterey Bay Microbial Study	17472632	48	1,200,014	1,073,521	265,391,390
North Atlantic Spring Bloom	22323811	6	6,784,781	5,496,974	1,308,920,188
Tara Oceans Expeditions	23575371	NA	8,162,564	7,511,565	2,734,688,433
Western Channel Observatory	21124740	72	5,752,026	5,001,406	1,854,605,771
Total	-	-	142,475,406	125,853,430	54,983,520,353

Part II. Metadata for Metagenome Data in Detail

SAMPLE	FILTER MIN (um)	FILTER MAX (um)	Latitude	Longitude	Sample Depth(m)	# Qualified Reads
Station_368	0.1	0.8	-44.72	145.76	2.00	549,926
Station_368	0.8	3	-44.72	145.76	2.00	205,939
Station_368	3	200	-44.72	145.76	2.00	205,939
Station_367	0.1	0.8	-48.25	145.81	2.00	553,511
Station_367	0.8	3	-48.25	145.81	2.00	414,686
Station_367	3	200	-48.25	145.81	2.00	366,697
Station_366	0.1	0.8	-52.02	144.07	2.00	604,892
Station_366	0.8	3	-52.02	144.07	2.00	180,763
Station_366	3	200	-52.02	144.07	2.00	248,417
Station_394	0.1	0.8	-53.02	73.38	1.00	438,254
Station_394	NA	NA	-53.02	73.38	1.00	424,739
Station_386	0.1	0.8	-54.95	129.62	1.00	458,352
Station_386	0.8	3	-54.95	129.62	1.00	138,789
Station_386	3	200	-54.95	129.62	1.00	298,965
Station_393	0.1	0.8	-55.26	74.26	1.00	737,147
Station_393	NA	NA	-55.26	74.26	1.00	400,756
Station_364	0.1	0.8	-56.70	141.87	2.00	569,669
Station_364	0.8	3	-56.70	141.87	2.00	190,720
Station_364	3	200	-56.70	141.87	2.00	240,416
Station_365	0.1	0.8	-56.70	141.88	3,693.00	633,520
Station_365	0.8	3	-56.70	141.88	3,693.00	232,482

Station_365	3	200	-56.70	141.88	3,693.00	91,423
Station_346	0.1	0.8	-59.31	142.46	2.00	339,550
Station_346	0.8	3	-59.31	142.46	2.00	418,282
Station_346	3	200	-59.31	142.46	2.00	246,757
Station_363	0.1	0.8	-60.00	141.23	2.00	477,204
Station_363	0.8	3	-60.00	141.23	2.00	291,961
Station_363	3	200	-60.00	141.23	2.00	286,552
Station_387	0.1	0.8	-60.50	120.05	1.00	451,225
Station_387	NA	NA	-60.50	120.05	1.00	340,595
Station_388	0.1	0.8	-63.82	115.17	1.00	465,963
Station_388	NA	NA	-63.82	115.17	1.00	367,159
Station_236	0.1	0.8	-63.89	112.07	2.00	1,019,859
Station_236	3	200	-63.89	112.07	2.00	258,250
Station_392	0.1	0.8	-64.20	76.46	1.00	664,003
Station_392	0.8	3	-64.20	76.46	1.00	270,670
Station_392	3	200	-64.20	76.46	1.00	176,910
Station_358	0.1	0.8	-64.30	150.01	2.00	523,368
Station_358	0.8	3	-64.30	150.01	2.00	218,627
Station_358	3	200	-64.30	150.01	2.00	186,381
Station_389	0.1	0.8	-64.80	112.38	1.00	514,996
Station_389	NA	NA	-64.80	112.38	1.00	422,070
Station_390	0.1	0.8	-64.83	80.72	1.00	493,689
Station_390	NA	NA	-64.83	80.72	1.00	335,502
Station_362	0.1	0.8	-65.54	140.72	2.00	537,017
Station_362	0.8	3	-65.54	140.72	2.00	223,381
Station_362	3	200	-65.54	140.72	2.00	261,356
Station_347	0.1	0.8	-66.02	142.67	2.00	558,649
Station_347	0.8	3	-66.02	142.67	2.00	236,989
Station_347	3	200	-66.02	142.67	2.00	233,739
Station_357	0.1	0.8	-66.17	142.94	2.00	639,130
Station_357	0.8	3	-66.17	142.94	2.00	335,608
Station_357	3	200	-66.17	142.94	2.00	375,542
Station_359	0.1	0.8	-66.19	143.49	2.00	480,202
Station_359	0.8	3	-66.19	143.49	2.00	706,529
Station_359	3	200	-66.19	143.49	2.00	403,615
Station_235	0.1	0.8	-66.27	110.53	1.83	173,213
Station_235	0.8	3	-66.27	110.53	1.83	571,550
Station_235	3	200	-66.27	110.53	1.83	158,492
Station_348	0.1	0.8	-66.34	142.99	1.50	429,571
Station_348	0.8	3	-66.34	142.99	1.50	281,648
Station_348	3	200	-66.34	142.99	1.50	195,752
Station_361	0.1	0.8	-66.47	140.54	1,170.00	531,595
Station_361	0.8	3	-66.47	140.54	1,170.00	206,386
Station_361	3	200	-66.47	140.54	1,170.00	243,771
Station_351	0.1	0.8	-66.56	143.34	1.50	494,126
Station_351	0.8	3	-66.56	143.34	1.50	474,757
Station_351	3	200	-66.56	143.34	1.50	700,917
Station_349	0.1	0.8	-66.57	142.32	1.50	548,378
Station_349	0.8	3	-66.57	142.32	1.50	231,214
Station_349	3	200	-66.57	142.32	1.50	217,790
Station_350	0.1	0.8	-66.57	142.32	330.00	554,483
Station_350	0.8	3	-66.57	142.32	330.00	280,323
Station_350	3	200	-66.57	142.32	330.00	240,873
Station_360	0.1	0.8	-66.58	140.88	2.00	496,486
Station_360	0.8	3	-66.58	140.88	2.00	255,924
Station_360	3	200	-66.58	140.88	2.00	178,368

Station_355	0.1	0.8	-66.76	144.33	2.00	473,774
Station_355	0.8	3	-66.76	144.33	2.00	456,349
Station_355	3	200	-66.76	144.33	2.00	338,571
Station_356	0.1	0.8	-66.76	144.33	890.00	494,195
Station_356	0.8	3	-66.76	144.33	890.00	402,615
Station_356	3	200	-66.76	144.33	890.00	448,121
Station_352	0.1	0.8	-66.77	143.29	2.50	556,684
Station_352	0.8	3	-66.77	143.29	2.50	467,204
Station_352	3	200	-66.77	143.29	2.50	387,214
Station_353	0.1	0.8	-67.05	144.67	2.00	533,045
Station_353	0.8	3	-67.05	144.67	2.00	246,833
Station_353	3	200	-67.05	144.67	2.00	270,886
Station_354	0.1	0.8	-67.07	145.20	1,320.00	601,668
Station_354	0.8	3	-67.07	145.20	1,320.00	462,720
Station_354	3	200	-67.07	145.20	1,320.00	395,663
Station_391	0.1	0.8	-68.40	76.68	1.00	397,515
Station_391	0.8	3	-68.40	76.68	1.00	194,952
Station_391	3	200	-68.40	76.68	1.00	2
Site227	0.1	0.8	-68.40	78.18	23.00	516,553
Site227	0.8	3	-68.40	78.18	23.00	597,573
Site227	3	200	-68.40	78.18	23.00	281,508
Site228	0.1	0.8	-68.40	78.18	18.00	377,213
Site228	0.8	3	-68.40	78.18	18.00	576,338
Site228	3	200	-68.40	78.18	18.00	297,908
Site229	0.1	0.8	-68.40	78.18	14.00	431,973
Site229	0.8	3	-68.40	78.18	14.00	477,621
Site229	3	200	-68.40	78.18	14.00	313,907
Site230	0.1	0.8	-68.40	78.18	12.70	457,150
Site230	0.8	3	-68.40	78.18	12.70	559,951
Site230	3	200	-68.40	78.18	12.70	220,480
Site231	0.1	0.8	-68.40	78.18	11.50	561,283
Site231	0.8	3	-68.40	78.18	11.50	508,260
Site231	3	200	-68.40	78.18	11.50	384,538
Site232	0.1	0.8	-68.40	78.18	5.00	575,374
Site232	0.8	3	-68.40	78.18	5.00	496,417
Site232	3	200	-68.40	78.18	5.00	166,811
Station_379	0.1	0.8	-68.46	78.21	1.00	458,509
Station_379	0.8	3	-68.46	78.21	1.00	474,948
Station_379	3	200	-68.46	78.21	1.00	256,680
Station_233	0.1	0.8	-68.46	78.20	0.00	411,766
Station_233	0.8	3	-68.46	78.20	0.00	470,453
Station_233	3	200	-68.46	78.20	0.00	496,829
Station_374	0.1	0.8	-68.46	78.20	1.80	517,512
Station_374	0.8	3	-68.46	78.20	1.80	491,795
Station_374	3	200	-68.46	78.20	1.80	323,274
Station_375	0.1	0.8	-68.46	78.20	4.30	563,211
Station_375	0.8	3	-68.46	78.20	4.30	489,468
Station_375	3	200	-68.46	78.20	4.30	158,918
Station_376	0.1	0.8	-68.46	78.20	5.60	376,866
Station_376	0.8	3	-68.46	78.20	5.60	506,884
Station_376	3	200	-68.46	78.20	5.60	295,559
Station_377	0.1	0.8	-68.46	78.20	6.50	531,265
Station_377	0.8	3	-68.46	78.20	6.50	498,427
Station_377	3	200	-68.46	78.20	6.50	196,239
Station_378	0.1	0.8	-68.46	78.20	6.80	547,538
Station_378	0.8	3	-68.46	78.20	6.80	585,018

Station_378	3	200	-68.46	78.20	6.80	208,895
Station_369	0.1	0.8	-77.68	166.01	2.70	316,473
Station_369	0.8	3	-77.68	166.01	2.70	323,286
Station_369	3	200	-77.68	166.01	2.70	428,944
Station_372	0.1	0.8	-77.68	166.00	NA	204,104
Station_372	0.8	3	-77.68	166.00	NA	275,194
Station_372	3	200	-77.68	166.00	NA	321,696
Station_370	0.1	0.8	-77.71	166.08	1.00	558,917
Station_370	0.8	3	-77.71	166.08	1.00	155,507
Station_370	3	200	-77.71	166.08	1.00	197,431
Station_371	0.1	0.8	-77.71	166.08	NA	294,025
Station_371	0.8	3	-77.71	166.08	NA	280,902
Station_371	3	200	-77.71	166.08	NA	210,735
GS264	3	200	35.09	-120.78	2.00	139,065
GS264	0.1	0.8	35.09	-120.78	2.00	227,280
GS264	0.8	3	35.09	-120.78	2.00	126,369
GS263	0.1	0.8	34.72	-121.55	2.00	193,836
GS263	0.8	3	34.72	-121.55	2.00	154,698
GS263	3	200	34.72	-121.55	2.00	145,591
GS257	3	200	33.66	-118.97	2.00	168,177
GS257	0.8	3	33.66	-118.97	2.00	210,057
GS262	0.1	0.8	33.15	-123.22	2.00	127,780
GS262	0.8	3	33.15	-123.22	2.00	193,116
GS262	3	200	33.15	-123.22	2.00	174,478
GS260	3	200	32.91	-122.14	2.00	276,363
GS260	0.8	3	32.91	-122.14	2.00	296,068
GS258	0.1	0.8	32.33	-121.71	2.00	227,565
GS258	0.8	3	32.33	-121.71	2.00	207,779
GS258	3	200	32.33	-121.71	2.00	166,429
GS259	3	200	31.91	-124.17	2.00	179,854
GS259	0.1	0.8	31.91	-124.17	2.00	220,947
GS259	0.8	3	31.91	-124.17	2.00	152,387
Mediterranean_DCM	0.22	0.8	38.07	-0.23	50.00	1,072,794
GOM04	0.1	0.8	44.12	-67.97	1.50	938,897
GOM13	0.1	0.8	43.38	-67.68	1.50	925,178
GOM03	0.1	0.8	42.77	-68.67	1.50	453,620
GOM14	0.1	0.8	42.35	-69.38	1.50	9,728
GOM06	0.1	0.8	41.47	-69.10	1.50	10,037
GOM12	0.1	0.8	41.13	-66.88	1.50	470,464
GS006	0.1	0.8	45.11	-64.95	1.00	59,660
GS005	0.1	0.8	44.69	-63.64	1.00	61,028
GS004	0.1	0.8	44.14	-63.64	2.00	52,934
GS007	0.1	0.8	43.63	-66.85	1.00	50,947
GS003	0.1	0.8	42.85	-66.22	1.00	61,578
GS002	0.1	0.8	42.50	-67.24	1.00	121,481
GS008	0.1	0.8	41.49	-71.35	1.00	129,582
GS009	0.1	0.8	41.09	-71.60	1.00	79,205
GS011	0.1	0.8	39.42	-75.50	1.00	124,342
GS012	0.1	0.8	38.95	-76.42	13.20	126,035
GS010	0.1	0.8	38.94	-74.69	1.00	78,253
MOVE858	0.002	0.22	38.13	-76.38	2.07	11,493
GS013	0.1	0.8	36.00	-75.39	2.10	137,735
GS014	0.1	0.8	32.51	-79.26	1.00	128,823
GS000c	0.22	0.8	32.17	-64.01	5.00	367,572
GS001a	3	20	32.17	-64.50	5.00	135,457
GS001b	0.8	3	32.17	-64.50	5.00	90,014

GS001c	0.1	0.8	32.17	-64.50	5.00	92,259
GS000a	0.1	0.8	31.18	-64.32	5.00	642,806
GS000b	0.22	0.8	31.18	-64.32	5.00	315,716
GS000d	0.22	0.8	31.18	-64.32	5.00	330,486
GS015	0.1	0.8	24.49	-83.07	1.70	127,291
GS016	0.1	0.8	24.17	-84.34	2.00	126,919
GS017	0.1	0.8	20.52	-85.41	2.00	257,415
GS018	0.1	0.8	18.04	-83.78	1.70	142,685
GS019	0.1	0.8	10.72	-80.25	1.70	135,238
GS020	0.1	0.8	9.16	-79.84	2.00	296,264
GS021	0.1	0.8	8.13	-79.69	1.60	131,748
GS022	0.1	0.8	6.49	-82.90	2.00	121,621
GS023	0.1	0.8	5.64	-86.57	2.00	133,010
GS025	0.8	3	5.55	-87.09	1.10	120,309
GS035	0.1	0.8	1.39	-91.82	1.70	140,746
GS026	0.1	0.8	1.26	-90.30	2.00	102,669
GS030	0.1	0.8	0.27	-91.63	19.00	359,013
GS036	0.1	0.8	-0.02	-91.20	2.10	77,506
GS029	0.1	0.8	-0.20	-90.84	2.10	131,468
GS031	0.1	0.8	-0.30	-91.65	12.00	436,224
GS034	0.1	0.8	-0.38	-90.28	2.10	134,315
GS032	0.1	0.8	-0.59	-91.07	0.10	147,468
GS027	0.1	0.8	-1.22	-90.42	2.20	221,993
GS028	0.1	0.8	-1.22	-90.32	2.00	188,997
GS033	0.1	0.8	-1.23	-90.43	0.20	691,906
GS037	0.1	0.8	-1.97	-95.01	1.80	65,641
GS038	0.1	0.8	-2.58	-97.85	1.80	741
GS039	0.1	0.8	-3.34	-101.37	2.00	759
GS040	0.1	0.8	-4.50	-105.07	2.20	736
GS117a	0.1	0.8	-4.61	55.51	1.80	346,836
GS117b	0.8	3	-4.61	55.51	1.80	50,561
GS116	0.1	0.8	-4.64	56.84	1.50	60,917
GS115	0.1	0.8	-4.66	60.52	1.50	60,986
GS114	0.1	0.8	-4.99	64.98	1.50	348,692
GS041	0.1	0.8	-5.93	-108.69	2.00	678
GS149	0.1	0.8	-6.12	39.12	1.50	110,959
GS148	0.1	0.8	-6.32	39.01	0.30	107,576
GS113	0.1	0.8	-7.01	76.33	1.80	109,641
GS042	0.1	0.8	-7.11	-116.12	1.70	699
GS043	0.1	0.8	-7.66	-120.40	1.90	711
GS044	0.1	0.8	-8.42	-124.24	2.00	676
GS112a	0.1	0.8	-8.51	80.38	1.80	99,756
GS112b	0.8	3	-8.51	80.38	1.80	52,018
GS112	0.1	0.8	-8.51	80.38	1.80	1,043,098
GS045	0.1	0.8	-9.02	-127.77	1.70	730
GS046	0.1	0.8	-9.57	-131.49	1.90	624
GS111	0.1	0.8	-9.60	84.20	1.80	59,058
GS047	0.1	0.8	-10.13	-135.45	30.00	66,004
GS110a	0.1	0.8	-10.45	88.30	1.50	99,248
GS110b	0.8	3	-10.45	88.30	1.50	49,517
GS109	0.1	0.8	-10.94	92.06	1.50	59,786
GS108a	0.1	0.8	-12.09	96.88	1.80	51,756
GS108b	0.8	3	-12.09	96.88	1.80	49,560
GS108	0.1	0.8	-12.09	96.88	1.80	1,296,088
GS051	0.1	0.8	-15.14	-147.44	1.00	128,958
GS050	0.1	0.8	-15.28	-148.22	1.20	713

GS049	0.1	0.8	-17.45	-149.80	1.40	92,438
GS048a	0.1	0.8	-17.48	-149.81	1.40	90,456
GS048b	0.8	3	-17.48	-149.81	1.40	47,639
GS119	0.1	0.8	-23.22	52.31	2.00	60,961
GS120	0.1	0.8	-26.04	50.12	2.80	45,992
GS121	0.1	0.8	-29.35	43.22	1.50	110,699
GS122a	0.1	0.8	-30.90	40.42	1.90	101,506
GS122b	0.8	3	-30.90	40.42	1.90	49,973
GS123	0.1	0.8	-32.40	36.59	2.20	107,920
GS667	0.1	0.8	68.36	18.82	1.00	213,819
GS667	0.8	3	68.36	18.82	1.00	166,458
GS667	3	200	68.36	18.82	1.00	213,130
GS665	0.1	0.8	65.17	23.23	1.00	315,270
GS665	0.8	3	65.17	23.23	1.00	236,358
GS665	3	200	65.17	23.23	1.00	221,727
GS666	0.1	0.8	65.17	23.23	10.00	207,420
GS666	0.8	3	65.17	23.23	10.00	412,840
GS666	3	200	65.17	23.23	10.00	415,837
GS673	0.1	0.8	63.50	19.81	1.00	468,586
GS673	0.8	3	63.50	19.81	1.00	619,209
GS673	3	200	63.50	19.81	1.00	138,350
GS674	0.1	0.8	63.50	19.81	16.00	261,402
GS674	0.8	3	63.50	19.81	16.00	214,022
GS674	3	200	63.50	19.81	16.00	223,350
GS659	0.1	0.8	62.65	18.95	1.00	298,393
GS659	0.8	3	62.65	18.95	1.00	624,016
GS659	3	200	62.65	18.95	1.00	536,378
GS660	0.1	0.8	62.65	18.95	12.00	512,331
GS660	0.8	3	62.65	18.95	12.00	462,008
GS660	3	200	62.65	18.95	12.00	262,144
GS677	0.1	0.8	58.58	18.23	9.00	279,205
GS677	0.8	3	58.58	18.23	9.00	366,024
GS677	3	200	58.58	18.23	9.00	363,345
GS678	0.1	0.8	58.58	18.23	74.00	314,794
GS678	0.8	3	58.58	18.23	74.00	499,410
GS678	3	200	58.58	18.23	74.00	448,400
GS687	0.1	0.8	58.32	11.54	1.00	280,654
GS687	0.8	3	58.32	11.54	1.00	266,481
GS687	3	200	58.32	11.54	1.00	175,602
GS688	0.1	0.8	58.32	11.54	30.00	275,823
GS688	0.8	3	58.32	11.54	30.00	259,768
GS688	3	200	58.32	11.54	30.00	440,911
GS689	0.1	0.8	58.32	11.54	72.00	278,925
GS689	0.8	3	58.32	11.54	72.00	287,725
GS689	3	200	58.32	11.54	72.00	186,404
GS694	0.1	0.8	57.80	10.87	1.00	157,019
GS694	0.8	3	57.80	10.87	1.00	196,555
GS694	3	200	57.80	10.87	1.00	198,236
GS695	0.1	0.8	57.80	10.87	19.00	338,026
GS695	0.8	3	57.80	10.87	19.00	329,091
GS695	3	200	57.80	10.87	19.00	294,173
GS685	0.1	0.8	56.66	12.12	1.00	316,197
GS685	0.8	3	56.66	12.12	1.00	711,352
GS685	3	200	56.66	12.12	1.00	129,464
GS686	0.1	0.8	56.66	12.12	20.00	136,894
GS686	0.8	3	56.66	12.12	20.00	276,335

GS686	3	200	56.66	12.12	20.00	483,078
GS679	0.1	0.8	56.17	16.38	1.00	214,330
GS679	0.8	3	56.17	16.38	1.00	1,074,321
GS679	3	200	56.17	16.38	1.00	206,659
GS681	0.1	0.8	54.93	13.48	1.00	249,059
GS681	0.8	3	54.93	13.48	1.00	361,349
GS681	3	200	54.93	13.48	1.00	355,466
GS682	0.1	0.8	54.93	13.48	24.00	390,373
GS682	0.8	3	54.93	13.48	24.00	354,283
GS682	3	200	54.93	13.48	24.00	365
GS683	0.1	0.8	54.57	11.33	1.00	315,025
GS683	0.8	3	54.57	11.33	1.00	182,830
GS683	3	200	54.57	11.33	1.00	443,411
GS684	0.1	0.8	54.57	11.33	15.00	211,945
GS684	0.8	3	54.57	11.33	15.00	487,677
GS684	3	200	54.57	11.33	15.00	165,069
GS728	0.1	0.8	42.13	2.77	1.75	283,037
GS728	0.8	3	42.13	2.77	1.75	275,249
GS728	3	200	42.13	2.77	1.75	249,861
GS729	0.1	0.8	42.13	2.77	4.50	243,126
GS729	0.8	3	42.13	2.77	4.50	407,779
GS729	3	200	42.13	2.77	4.50	275,434
GS730	0.1	0.8	42.12	2.75	22.25	397,015
GS730	0.8	3	42.12	2.75	22.25	280,423
GS730	3	200	42.12	2.75	22.25	281,092
GS731	0.1	0.8	42.12	2.75	24.00	283,007
GS731	0.8	3	42.12	2.75	24.00	445,694
GS731	3	200	42.12	2.75	24.00	237,205
BATS216_20M_SG	0.22	1.6	31.67	-64.17	20.00	140,565
BATS216_50M_SG	0.22	1.6	31.67	-64.17	50.00	132,100
HF10_10-07-02	1/0/1900	1.6	22.75	-158.00	10.00	6,953
HF130_10-06-02	1/0/1900	1.6	22.75	-158.00	130.00	4,195
HF200_10-06-02	1/0/1900	1.6	22.75	-158.00	200.00	6,499
HF4000_12-21-03	1/0/1900	1.6	22.75	-158.00	4,000.00	10,228
HF500_10-06-02	1/0/1900	1.6	22.75	-158.00	500.00	8,169
HF70_10-07-02	1/0/1900	1.6	22.75	-158.00	70.00	10,120
HF770_12-21-03	1/0/1900	1.6	22.75	-158.00	770.00	11,107
HOT179_125M_GDNA	NA	NA	22.75	-158.00	125.00	248,409
HOT179_25M_GDNA	NA	NA	22.75	-158.00	25.00	271,426
HOT179_500M_GDNA	NA	NA	22.75	-158.00	500.00	264,793
HOT179_75M_GDNA	NA	NA	22.75	-158.00	75.00	306,043
HOT179_125M_SG	0.22	1.6	22.73	-158.03	125.00	151,719
HOT179_25M_SG	0.22	1.6	22.73	-158.03	25.00	129,094
HOT179_500M_SG	0.22	1.6	22.73	-158.03	500.00	132,282
HOT179_75M_SG	0.22	1.6	22.73	-158.03	75.00	148,059
HOT186_110M_GDNA	0.22	1.6	22.73	-158.03	110.00	425,140
HOT186_25M_GDNA	0.22	1.6	22.73	-158.03	25.00	535,032
HOT186_500M_GDNA	0.22	1.6	22.73	-158.03	500.00	895,630
HOT186_75M_GDNA	0.22	1.6	22.73	-158.03	75.00	587,968
Kueishantao_Surface	0.22	0.8	24.83	121.95	0.00	297,277
Kueishantao_Vent	0.22	0.8	24.83	121.95	17.20	260,302
S_35179	0.2	5	13.48	-159.88	5.00	199,041
S_35171	0.2	5	8.63	-151.08	5.00	54,660

S_35162	5	5	5.08	-161.97	5.00	49,617
S_35163	0.2	5	5.08	-161.97	5.00	238,303
S_35155	0.2	5	1.67	-163.67	5.00	243,049
S_35147	0.2	5	-3.03	-166.00	5.00	82,924
S_35139	0.2	5	-7.52	-167.00	5.00	64,729
S_35131	0.2	5	-16.00	-170.00	5.00	179,347
MB2000JD298_1	0.2	NA	36.75	-122.02	0.00	197,581
MB2000JD298_2	0.2	NA	36.75	-122.02	0.00	182,512
MB2001JD115_1	0.2	NA	36.75	-122.02	0.00	169,978
MB2001JD115_2	0.2	NA	36.75	-122.02	0.00	159,644
MB2001JD135_1	0.2	NA	36.75	-122.02	0.00	200,520
MB2001JD135_2	0.2	NA	36.75	-122.02	0.00	163,286
BATS-174-1	NA	NA	31.71	-64.21	120.00	217,918
BATS-174-2	NA	NA	31.71	-64.21	NA	234,036
BATS-173-120	0.2	NA	31.68	-64.17	40.00	469,168
BATS-173-40	0.2	NA	31.68	-64.17	NA	485,710
BATS-174a-80	0.2	NA	31.66	-64.12	80.00	457,334
BATS-179-1	NA	NA	31.66	-64.16	NA	178,547
BATS-179-2	NA	NA	31.66	-64.16	NA	207,731
BATS-167-0	0.2	NA	31.65	-64.26	1.00	531,905
BATS-167-120	0.2	NA	31.65	-64.26	120.00	421,731
BATS-167-160	0.2	NA	31.65	-64.26	160.00	287,951
BATS-167-200	0.2	NA	31.65	-64.26	200.00	568,776
BATS-167-250	0.2	NA	31.65	-64.26	250.00	484,444
BATS-167-40	0.2	NA	31.65	-64.26	40.00	457,926
BATS-167-80	0.2	NA	31.65	-64.26	80.00	493,797
st23_D	0.1	0.8	42.17	17.72	DCMNA(56 m)	223,698
st23_S	0.1	0.8	42.17	17.72	SRF	182,856
st7_D	0.1	0.8	37.04	-1.95	DCMNA(42 m)	856,705
st7_S	0.1	0.8	37.04	-1.95	SRF	307,202
st3_S	0.1	0.8	36.73	-10.47	SRF	59,195
st4_S	0.1	0.8	36.55	-6.57	SRF	132,612
st6_S	0.1	0.8	36.52	-4.01	SRF	89,804
st30_S	0.1	0.8	33.92	32.89	SRF	244,969
st31_S	0.1	0.8	27.14	34.81	SRF	587,699
st36_S	0.1	0.8	20.82	63.51	SRF	687,836
st38_S	0.1	0.8	19.04	64.49	SRF	573,463
st38_Z	0.1	0.8	19.04	64.56	OMZNA(350 m)	602,254
st39_Z	0.1	0.8	18.73	66.39	OMZNA(270 m)	640,380
st39_S	0.1	0.8	18.57	66.49	SRF	561,923
st43_S	0.1	0.8	4.66	73.49	SRF	504,252
st46_S	0.1	0.8	-0.66	73.16	SRF	616,110
st49_S	0.1	0.8	-16.81	59.50	SRF	640,607
SMPL_AUG_10AM_PF	0.22	NA	50.27	-4.15	2.00	294,987
SMPL_AUG_4AM_PF	0.22	NA	50.27	-4.15	2.00	458,942
SMPL_JAN_NIGHT_PF	0.22	NA	50.26	-4.24	2.00	707,309
SMPL_AUG_10PM_PF	0.22	NA	50.25	-4.20	2.00	479,204
SMPL_AUG_4PM_PF	0.22	NA	50.25	-4.20	2.00	561,201
SMPL_APR_NIGHT_PF	0.22	NA	50.25	-4.19	2.00	435,993
SMPL_APR_DAY	0.22	NA	50.25	-4.21	2.00	328,169

SMPL_APR_DAY_PF	0.22	NA	50.25	-4.21	2.00	579,226
SMPL_JAN_DAY	0.22	NA	50.25	-4.21	2.00	667,498
SMPL_JAN_DAY_PF	0.22	NA	50.25	-4.21	2.00	488,877

Table S3. Count of metagenomic reads hitting to Roseobacters in general and percentage of the four largely uncultivated lineages in Roseobacters, respectively. DCM, deep chlorophyll maximum; FL, free-living; PA, particle-associated. The Tara Ocean metagenomes were obtained from Hingamp et al. 2013 (ISME J). The new Tara data (from Sunagawa et al. 2015, Science) were not available at the time of this analysis.

PROJ	Living Style and Temperate Zone	# Roseobacter Reads	# CHAB-I-5 reads	%	# DC5-80-3 reads	%	# SAG-O19 reads	%	# NAC11-7 reads	%
Southern Ocean	FL_Tropical	0	0	0.00	0	0.00	0	0.00	0	0.00
Southern Ocean	FL_Temperate	318,264	11,146	3.50	130,351	40.96	25,730	8.08	49,124	15.43
Southern Ocean	FL_Polar	184,216	6,034	3.28	64,630	35.08	6,623	3.60	26,972	14.64
Southern Ocean	PA_Tropical	0	0	0.00	0	0.00	0	0.00	0	0.00
Southern Ocean	PA_Temperate	206,254	8,075	3.92	90,763	44.01	11,060	5.36	22,416	10.87
Southern Ocean	PA_Polar	275,983	6,627	2.40	73,086	26.48	5,154	1.87	22,787	8.26
calcofi	FL_Tropical	0	0	0.00	0	0.00	0	0.00	0	0.00
calcofi	FL_Temperate	47,963	5,245	10.94	11,416	23.80	7,424	15.48	16,030	33.42
calcofi	FL_Polar	0	0	0.00	0	0.00	0	0.00	0	0.00
calcofi	PA_Tropical	0	0	0.00	0	0.00	0	0.00	0	0.00
calcofi	PA_Temperate	28,879	5,668	19.63	3,594	12.45	2,464	8.53	10,506	36.38
calcofi	PA_Polar	0	0	0.00	0	0.00	0	0.00	0	0.00
Mediterranean DCM	FL_Tropical	0	0	0.00	0	0.00	0	0.00	0	0.00
Mediterranean DCM	FL_Temperate	12,801	888	6.94	579	4.52	5,436	42.47	871	6.80
Mediterranean DCM	FL_Polar	0	0	0.00	0	0.00	0	0.00	0	0.00
Mediterranean DCM	PA_Tropical	0	0	0.00	0	0.00	0	0.00	0	0.00
Mediterranean DCM	PA_Temperate	0	0	0.00	0	0.00	0	0.00	0	0.00
Mediterranean DCM	PA_Polar	0	0	0.00	0	0.00	0	0.00	0	0.00
Gulf of Maine	FL_Tropical	0	0	0.00	0	0.00	0	0.00	0	0.00
Gulf of Maine	FL_Temperate	79,525	17,397	21.88	16,949	21.31	12,161	15.29	10,808	13.59
Gulf of Maine	FL_Polar	0	0	0.00	0	0.00	0	0.00	0	0.00
Gulf of Maine	PA_Tropical	0	0	0.00	0	0.00	0	0.00	0	0.00
Gulf of Maine	PA_Temperate	0	0	0.00	0	0.00	0	0.00	0	0.00
Gulf of Maine	PA_Polar	0	0	0.00	0	0.00	0	0.00	0	0.00
GOS	FL_Tropical	238,264	15,599	6.55	7,068	2.97	110,268	46.28	12,327	5.17
GOS	FL_Temperate	61,770	6,922	11.21	10,099	16.35	19,819	32.09	6,304	10.21
GOS	FL_Polar	0	0	0.00	0	0.00	0	0.00	0	0.00
GOS	PA_Tropical	3,331	360	10.81	95	2.85	957	28.73	140	4.20
GOS	PA_Temperate	1,798	956	53.17	52	2.89	150	8.34	82	4.56
GOS	PA_Polar	0	0	0.00	0	0.00	0	0.00	0	0.00
GOS_Baltic_sea	FL_Tropical	0	0	0.00	0	0.00	0	0.00	0	0.00
GOS_Baltic_sea	FL_Temperate	43,522	4,252	9.77	9,617	22.10	2,314	5.32	14,670	33.71
GOS_Baltic_sea	FL_Polar	282	14	4.96	17	6.03	9	3.19	14	4.96
GOS_Baltic_sea	PA_Tropical	0	0	0.00	0	0.00	0	0.00	0	0.00
GOS_Baltic_sea	PA_Temperate	128,515	24,588	19.13	27,558	21.44	3,188	2.48	12,690	9.87
GOS_Baltic_sea	PA_Polar	203	9	4.43	11	5.42	10	4.93	8	3.94
GOS_banyoles	FL_Tropical	0	0	0.00	0	0.00	0	0.00	0	0.00
GOS_banyoles	FL_Temperate	898	31	3.45	40	4.45	168	18.71	48	5.35
GOS_banyoles	FL_Polar	0	0	0.00	0	0.00	0	0.00	0	0.00
GOS_banyoles	PA_Tropical	0	0	0.00	0	0.00	0	0.00	0	0.00
GOS_banyoles	PA_Temperate	699	24	3.43	28	4.01	173	24.75	18	2.58
GOS_banyoles	PA_Polar	0	0	0.00	0	0.00	0	0.00	0	0.00
HOT_ALOHA	FL_Tropical	25,674	1,890	7.36	1,368	5.33	7,557	29.43	2,201	8.57
HOT_ALOHA	FL_Temperate	3,303	471	14.26	175	5.30	706	21.37	396	11.99
HOT_ALOHA	FL_Polar	0	0	0.00	0	0.00	0	0.00	0	0.00
HOT_ALOHA	PA_Tropical	0	0	0.00	0	0.00	0	0.00	0	0.00
HOT_ALOHA	PA_Temperate	0	0	0.00	0	0.00	0	0.00	0	0.00
HOT_ALOHA	PA_Polar	0	0	0.00	0	0.00	0	0.00	0	0.00
Kueishantao Island	FL_Tropical	0	0	0.00	0	0.00	0	0.00	0	0.00
Kueishantao Island	FL_Temperate	1,806	92	5.09	43	2.38	1,117	61.85	55	3.05
Kueishantao Island	FL_Polar	0	0	0.00	0	0.00	0	0.00	0	0.00
Kueishantao Island	PA_Tropical	0	0	0.00	0	0.00	0	0.00	0	0.00
Kueishantao Island	PA_Temperate	0	0	0.00	0	0.00	0	0.00	0	0.00
Kueishantao Island	PA_Polar	0	0	0.00	0	0.00	0	0.00	0	0.00
Pacifc Ocean transect	FL_Tropical	14,932	801	5.36	509	3.41	8,214	55.01	671	4.49
Pacifc Ocean transect	FL_Temperate	0	0	0.00	0	0.00	0	0.00	0	0.00
Pacifc Ocean transect	FL_Polar	0	0	0.00	0	0.00	0	0.00	0	0.00
Pacifc Ocean transect	PA_Tropical	0	0	0.00	0	0.00	0	0.00	0	0.00
Pacifc Ocean transect	PA_Temperate	0	0	0.00	0	0.00	0	0.00	0	0.00
Pacifc Ocean transect	PA_Polar	0	0	0.00	0	0.00	0	0.00	0	0.00
Monterey_bay	FL_Tropical	0	0	0.00	0	0.00	0	0.00	0	0.00
Monterey_bay	FL_Temperate	109,490	15,541	14.19	13,796	12.60	2,315	2.11	66,701	60.92
Monterey_bay	FL_Polar	0	0	0.00	0	0.00	0	0.00	0	0.00
Monterey_bay	PA_Tropical	0	0	0.00	0	0.00	0	0.00	0	0.00
Monterey_bay	PA_Temperate	0	0	0.00	0	0.00	0	0.00	0	0.00
Monterey_bay	PA_Polar	0	0	0.00	0	0.00	0	0.00	0	0.00
north Atlantic	FL_Tropical	0	0	0.00	0	0.00	0	0.00	0	0.00
north Atlantic	FL_Temperate	24,397	4,538	18.60	1,131	4.64	5,035	20.64	1,579	6.47
north Atlantic	FL_Polar	0	0	0.00	0	0.00	0	0.00	0	0.00
north Atlantic	PA_Tropical	0	0	0.00	0	0.00	0	0.00	0	0.00
north Atlantic	PA_Temperate	0	0	0.00	0	0.00	0	0.00	0	0.00
north Atlantic	PA_Polar	0	0	0.00	0	0.00	0	0.00	0	0.00
Tara	FL_Tropical	69,375	5,039	7.26	2,316	3.34	32,825	47.32	2,954	4.26
Tara	FL_Temperate	71,894	9,325	12.97	2,612	3.63	30,351	42.22	5,456	7.59
Tara	FL_Polar	0	0	0.00	0	0.00	0	0.00	0	0.00
Tara	PA_Tropical	0	0	0.00	0	0.00	0	0.00	0	0.00
Tara	PA_Temperate	0	0	0.00	0	0.00	0	0.00	0	0.00
Tara	PA_Polar	0	0	0.00	0	0.00	0	0.00	0	0.00
Western Channel	FL_Tropical	0	0	0.00	0	0.00	0	0.00	0	0.00
Western Channel	FL_Temperate	226,737	16,098	7.10	34,942	15.41	17,110	7.55	132,097	58.26
Western Channel	FL_Polar	0	0	0.00	0	0.00	0	0.00	0	0.00
Western Channel	PA_Tropical	0	0	0.00	0	0.00	0	0.00	0	0.00
Western Channel	PA_Temperate	0	0	0.00	0	0.00	0	0.00	0	0.00
Western Channel	PA_Polar	0	0	0.00	0	0.00	0	0.00	0	0.00

Table S4. Survey of select genes and metabolic pathways in pelagic Roseobacter representatives (Continued from Table 4).

Functional Category	Pathway	CHAB-I-5	SAG-O19	DC5-80-3	NAC11-7	CCS1	HTCC2597	HTCC2516	Y4I	GAI101	DSS-3
carbohydrate uptake	*aglEFGK*, alpha-glucoside uptake	•	•	•		•		•			
	rhaPSQT, rhamnose uptake	•		•		•		•			
	frcABC, fructose uptake	•	•	•	•	•				•	
	rbsABC, ribose uptake	•		•		•	•	•		•	
	smoEFG, sorbitol/mannitol uptake	•	•	•	•	•	•	•		•	
	xylFGH, D-xylose uptake	•	•	•	•		•	•		•	•
	msmEFG, raffinose/stachyose/melibiose uptake					•		•			•
	malEFG, maltose/maltodextrin uptake							•			
	gtsABC, glucose/mannose uptake	•				•		•	•		
	glpVPGST, glycerol uptake	•				•	•		•	•	•
	PTS, phosphotransferase system	•	•	•	•	•	•	•	•	•	•
organic acid uptake	TRAP, tripartite ATP-independent periplasmic transporter	•	•	•	•	•	•	•	•	•	•
	TTT, tripartite tricarboxylate transporter	•	•	•	•	•		•	•	•	•
N-acetyltaurine catabolism	*naaTSRABC*			•						•	•
aromatics catabolism	*pcaH*, protocatechuate 3,4-dioxygenase		•	•		•	•		•	•	
	ferA, feruloyl-CoA synthetase	•	•	•	•	•	•		•	•	•
nitrogen acquisition	*nasDEF*, nitrate/nitrite uptake	•				•				•	
	nasA/nirB, assimilatory nitrate reductase					•				•	
	pepL, leucine aminopeptidase	•	•	•	•	•	•	•	•	•	•
	pepP, proline aminopeptidase	•	•	•	•	•	•	•	•	•	•
	pepN, alanine aminopeptidase	•	•	•	•	•	•	•	•	•	•
	pepM, methionine aminopeptidase	•	•	•	•	•	•	•	•	•	•
	amtB, ammonium uptake	•	•	•	•	•	•	•	•	•	•
	aapJQMP, L-amino acid uptake	•	•	•	•	•	•	•	•	•	•
	livGFHMK, branched amino acid uptake	•	•	•	•	•	•	•	•	•	•
	glnA, glutamine synthetase	•	•	•	•	•	•	•	•	•	•
	gltBD, glutamate synthase	•	•	•	•	•	•	•	•	•	•
	glnBD, *ntrBC*, regulatory	•	•	•	•	•	•	•	•	•	•
	ntrXY, regulatory	•	•	•	•	•	•	•	•	•	•
	potABCD, polyamine uptake		•	•	•	•	•	•	•	•	•
	potFGHI, polyamine uptake	•	•		•		•		•	•	•
	urtABCDE, urea uptake	•	•				•	•		•	•
	ureABCDEFG, urease	•	•	•		•	•	•		•	•
	occTPMQ, opine uptake										
phosphorus uptake	*pstABCS*, phosphate uptake	•	•	•		•	•	•	•	•	•
	phoX, alkaline phosphatase	•				•			•	•	•
	phoA, alkaline phosphatase										
	phoD, alkaline phosphatase								•	•	•
	phnGHIJM, C-P lyase						•		•	•	•
iron uptake	*afuAB*, iron (III) uptake	•						•			
	feoB, iron (II) uptake	•									
	fecA, iron (III) dicitrate uptake									•	
	ftr1, iron permease									•	•
	heme uptake										
	siderophore synthesis										
	siderophore uptake						•	•	•	•	
vitamin synthesis	B_{12}, cobalamin	•	•	•	•	•	•	•	•	•	•
	B_6, pyridoxine	•	•	•	•	•	•	•	•	•	•
	B_1, thiamine						•		•		•
	B_7, biotin								•		•