Genome sequence of *Synechococcus* CC9311: Insights into adaptation to a coastal environment

Brian Palenik*, Qinghu Ren†, Chris L. Dupont*, Garry S. Myers†, John F. Heidelberg†, Jonathan H. Badger†, Ramana Madupu†, William C. Nelson†, Lauren M. Brinkac†, Robert J. Dodson†, A. Scott Durkin†, Sean C. Daugherty†, Stephen A. Sullivan†, Hoda Khouri†, Yasmin Mohamoud†, Rebecca Halpin†, and Ian T. Paulsen†‡

*Scripps Institution of Oceanography, University of California at San Diego, La Jolla, CA 92093; and †The Institute for Genomic Research, Rockville, MD 20850

Edited by Robert Haselkorn, University of Chicago, Chicago, IL, and approved July 10, 2006 (received for review April 11, 2006)

Coastal aquatic environments are typically more highly productive and dynamic than open ocean ones. Despite these differences, cyanobacteria from the genus *Synechococcus* are important primary producers in both types of ecosystems. We have found that the genome of a coastal cyanobacterium, *Synechococcus* sp. strain CC9311, has significant differences from an open ocean strain, *Synechococcus* sp. strain WH8102, and these are consistent with the differences between their respective environments. CC9311 has a greater capacity to sense and respond to changes in its (coastal) environment. It has a much larger capacity to transport, store, use, or export metals, especially iron and copper. In contrast, phosphate acquisition seems less important, consistent with the higher concentration of phosphate in coastal environments. CC9311 is predicted to have differences in its outer membrane lipopolysaccharide, and this may be characteristic of the speciation of some cyanobacterial groups. In addition, the types of potentially horizontally transferred genes are markedly different between the coastal and open ocean genomes and suggest a more prominent role for phages in horizontal gene transfer in oligotrophic environments.

cyanobacteria | genomics | marine

Coastal waters typically have higher nutrient concentrations than open ocean waters because of wind-driven upwelling of nutrients from deeper depths and inputs from land and sediments. The higher nutrient concentrations lead to higher primary productivity. The spectral quality of light is typically different because of the presence of terrestrial material and algal biomass. These conditions contrast strongly with the low-nutrient blue-light-dominated ecosystems of the open ocean. Although each coastal environment has unique elements, these generalizations help us understand the adaptations likely to be found in coastal compared to open ocean microorganisms.

Some adaptations of photosynthetic microorganisms to the open ocean vs. coastal environment have included adaptations to nutrient levels and light. Differences in the pigments of coastal vs. open ocean *Synechococcus* have been well documented (1–4). In terms of nutrients, Carpenter (5) noted that coastal phytoplankton (diatom) species had a higher K_s (half-saturation constant for transport) for nitrate, whereas related open ocean diatom species had a lower K_s. The minimum amount of iron and other metals for growth of open ocean phytoplankton is less than that needed for coastal species, suggesting that adaptation to *in situ* metal levels is a significant factor in phytoplankton speciation (6–8). Recently, it has been shown, again in diatoms, that adaptation to low iron in the open ocean involves changes in the cellular concentration of the iron-rich photosynthetic reaction center proteins of photosystem I (9) and the use of plastocyanin, a copper containing protein, instead of iron (10).

We report here the genome sequence of *Synechococcus* sp. strain CC9311. This organism was isolated from the edge of California Current after nitrate enrichment and low light incubation (11). Strains related to CC9311 have been isolated from coastal environments such as Vineyard Sound (12, 13) and have been highly represented in *rpoC* gene sequence libraries of Southern California coastal waters and in the water column of the California Current when it displayed a coastal type chlorophyll profile (ref. 14; B.P., unpublished work). CC9311 possesses an ability to adapt to light quality (blue to green light ratios) not seen in open ocean *Synechococcus* strains such as WH8102, further indicating a coastal ecosystem niche for this strain (12). The availability of the genome sequence of CC9311 (Fig. 1) allows us to compare it to the genome sequence of *Synechococcus* sp. strain WH8102 (15), an open ocean strain, and to begin to understand the adaptation of bacterial genomes to the coastal vs. open ocean environments.

Results and Discussion

Gene Regulation and Two-Component Regulatory Systems. One of the insights from the genome of the open ocean *Synechococcus* WH8102 was that it and other open ocean cyanobacteria have minimal regulatory systems, particularly two-component regulatory systems consisting of a sensor and response regulator pair (15–17). There are only five histidine kinase sensors and nine response regulators in WH8102, and it was suggested that this was due to adaptation to a relatively constant ecosystem. As one would predict from adaptation to the more variable coastal environment, CC9311 has nearly double this number, with 11 histidine kinase sensors and 17 response regulators (Fig. 2). Interestingly, these additional systems occur in pairs in the genome, which is not always the case in WH8102. The function of these sensors is not predictable from their sequences at this time but may regulate the more complex metal metabolism in CC9311.

Despite the presence of additional sensor kinases, based on BLAST and phylogenetic analyses, CC9311 apparently lacks a phosphate sensor-response regulator system seen in other cyanobacteria and bacteria in general (18). Consistent with this, several alkaline phosphatases present in WH8102 are absent, and CC9311 has fewer periplasmic phosphate-binding proteins used in ABC transporter systems. These differences between the open ocean and coastal *Synechococcus* types likely reflect the higher phosphate concentrations in coastal environments compared to some surface ocean environments where phosphate can become limiting.

Metals and CC9311. CC9311 has a number of metal enzymes or cofactors not found in WH8102, suggesting that it has a greater

Conflict of interest statement: No conflicts declared.

This paper was submitted directly (Track II) to the PNAS office.

Data deposition: The sequence reported in this paper has been deposited in the GenBank database (accession no. CP000435). The *Synechococcus* CC9311 strain has been deposited in the Provasoli–Guillard National Center for Culture of Marine Phytoplankton (http://ccmp.bigelow.org) under catalog no. CCMP2515.

‡To whom correspondence should be addressed. E-mail: ipaulsen@tigr.org.

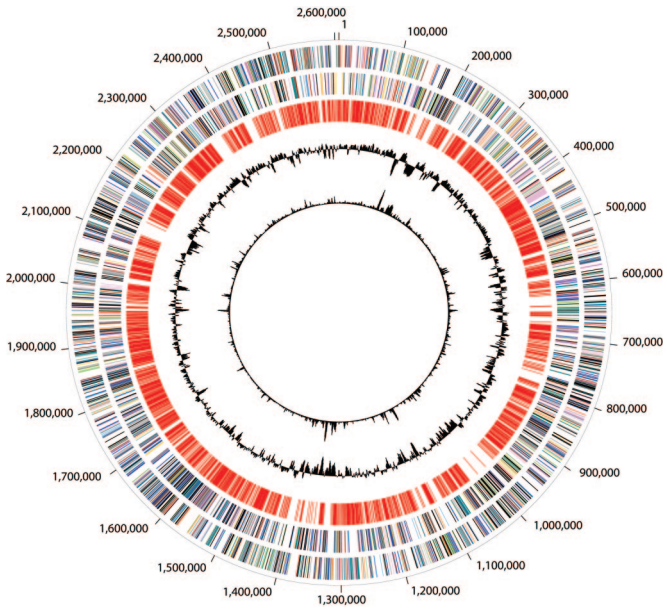

Fig. 1. Circular representation of the *Synechococcus* CC9311 overall genome structure. The outer scale designates coordinates in base pairs. The first circle shows predicted coding regions on the plus strand, color coded by role categories: violet, amino acid biosynthesis; light blue, biosynthesis of cofactors, prosthetic groups, and carriers; light green, cell envelope; red, cellular processes; brown, central intermediary metabolism; yellow, DNA metabolism; light gray, energy metabolism; magenta, fatty acid, and phospholipid metabolism; pink, protein synthesis and fate; orange, purines, pyrimidines, nucleosides, and nucleotides; olive, regulatory functions and signal transduction; dark green, transcription; teal, transport, and binding proteins; gray, unknown function; salmon, other categories; and blue, hypothetical proteins. The second circle shows predicted coding regions on the minus strand color coded by role categories. The third circle shows in red the set of 1,730 genes conserved between *Synechococcus* CC9311 and WH8102, the fourth circle shows percentage G+C in relation to the mean G+C in a 2,000-bp window in black, and the fifth circle shows the trinucleotide composition in black.

use for iron (Fig. 3). This is consistent with higher metal quotas for iron of coastal vs. open ocean phytoplankton, such as diatoms (6), and adds a mechanistic basis to these previous studies. Iron-dependent metalloenzymes unique to CC9311 include a cytochrome P450-like encoding ORF (sync_2424), two additional cytochrome *c* molecules (sync_1753 and sync_1742), and one or two additional ferrodoxins (sync_1953 and sync_0980, the latter truncated). It also has a putative iron-dependent alcohol dehydrogenase (sync_2669).

CC9311 appears to have a greater use for copper than WH8102, because it has a copper zinc superoxide dismutase not seen in marine cyanobacteria (sync_1771) until this work and the recent availability of two marine cyanobacterial genomes (CC9902 and CC9605; http://genome.jgi-psf.org/mic_home.html). It has a putative multicopper oxidase (sync_1489), which could be involved in oxidation of organic compounds or detoxifying high levels of reduced copper (19). Interestingly, it has been shown that *Synechococcus* sp. strain WH8016, a strain in the same clade as CC9311, was more resistant to copper than oligotrophic strains (20) .

For other metal usage, there appears to be a putative vanadium-dependent bromoperoxidase (sync_2681). The latter gene is very interesting, because it is highly similar to one in marine red algae. In red algae, this enzyme generates brominated compounds using hydrogen peroxide (21, 22). Cyanobacteria have been shown to produce brominated compounds such as bromodiphenyl ethers through an unknown mechanism, with the best-studied case being a filamentous cyanobacterial symbiont of a sponge (23). These brominated compounds have been found recently to cause leakage of fungal cell membranes (24), but the role of brominated compounds, if any, in CC9311 is open to speculation.

Possibly because of its more intensive use of metals, CC9311 has some metal transporters not seen in WH8102, including an FeoA/B transporter for iron(II) (sync_0681-0682). Total iron concentrations are higher in coastal environments, and reduced iron(II) may be more abundant as well, because it is likely produced from photochemical reactions of iron and organic matter (25, 26). CC9311 also has three cation-dependent efflux transporters (sync_0686, sync_1861, and sync_1510) compared to two in WH8102, suggesting that it may have an increased capacity to export toxic metal levels if needed.

In contrast, the oligotrophic ocean strain WH8102 has systems predicted for the efflux of arsenite (preceded by its reduction) and chromate (15) that are not found in CC9311. It has been suggested that high arsenate to phosphate ratios in oligotrophic regions result in the need of microorganisms to deal with excess arsenate (27).

Coastal *Synechococcus* strain CC9311 has a greatly enhanced capacity for metal storage. This is seen in the four copies of *smtA*, a gene for bacterial metallothionein (sync_1081, sync_2426, sync_0853, and sync_2379) compared to one in *Synechococcus* WH8102 and none in some *Prochlorococcus* strains. Gene amplification of *smtA* has been found in freshwater *Synechococcus* PCC6301 in response to higher trace metal levels such as cadmium (28). However, in this case, *smtA* copies occur in tandem, not disbursed throughout the genome as seen in CC9311.

CC9311 also has a greatly enhanced capacity specifically for iron storage. It has five copies of bacterial ferritin (sync_0854, sync_0687, sync_1077, sync_1539, and sync_0680) compared to one in most cyanobacterial genomes including *Synechococcus* WH8102. It also has a ferritin-related protein DpsA (DNA-protecting protein under starved conditions) that binds iron. The later is not found in WH8102 but is found in some *Prochlorococcus* strains (PMT2218 in MIT9313).

It is unclear whether the greatly enhanced transport and metal storage capacity for iron and other metals in CC9311 is due to a greater need for metals, the need to respond to excess metal levels, or the possibility that the cells see episodic metal concentrations. Iron concentrations in California coastal environments can vary from limiting to replete with rapid fluctuations (29), thus the ability to store iron may be advantageous. Taken together, these results suggest a much more metal-dependent ecological strategy for CC9311 (Fig. 3 and Table 1, which is published as supporting information on the PNAS web site).

Organic Nitrogen and Other Transporters. CC9311 and WH8102 also differ in other aspects of their membrane transporter complement that may reflect differences in the nutrients they are exposed to in their different environments. Interestingly, CC9311 has multiple AMT family ammonia transporters and based on this, ammonia is arguably its most important nitrogen source, but determining this will require *in situ* gene expression studies. CC9311 encodes a TRAP family dicarboxylate transporter as well as a DASS family transporter that may also be specific for carboxylates and a formate/nitrite transporter that is not present in WH8102. CC9311 also encodes a second type of predicted urea transporter and two APC-type amino acid transporters that are not present in WH8102. These capabilities are consistent with the coastal isolate CC9311 being exposed to more organic matter than its oligotrophic ocean relative WH8102. There is a significant expansion of mechanosensitive ion channels in CC9311, which has five MscS and two MscL members compared with only two MScS channels in WH8102. Mechanosensitive ion channels can function as "emergency

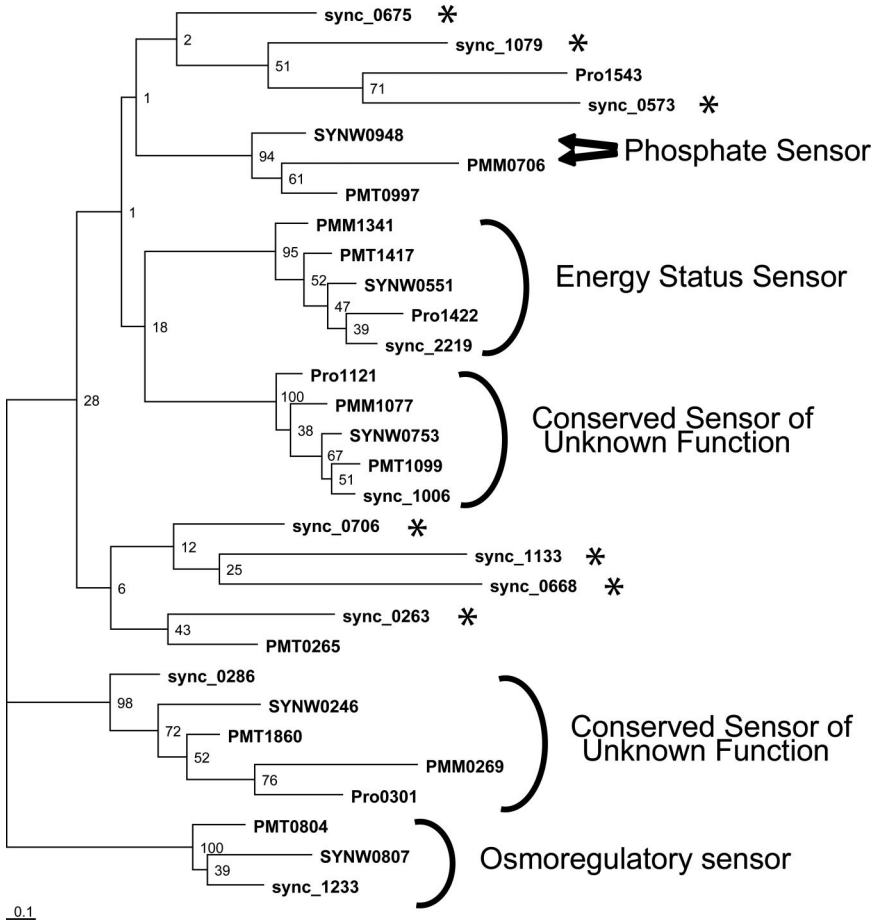

Fig. 2. Phylogenetic tree of sensor kinases from *Synechococcus* CC9311 (sync_xxxx), and WH8102 (SYNWxxxx), *Prochlorococcus marinus* MI9313 (PMTxxxx), MED4 (PMMxxxx), and SS120 (Proxxxx). This maximum-likelihood phylogenetic tree was generated by using PHYLIP, and bootstrap values are indicated next to the branch nodes. Orthologous clusters conserved in all of the cyanobacteria shown are highlighted by lines on the side, the phosphate sensor is labeled, and the divergent sensors unique to CC9311 are highlighted with asterisks.

relief valves" during conditions of osmotic shock, implying that the coastal isolate CC9311 may be subject to a more osmotically challenging environment (30).

Light and CC9311. The predicted ORFs associated with photosynthesis and light harvesting are relatively similar to WH8102. One exception is the much greater number of high light-inducible protein (HLIP) gene family members in CC9311 (with 14) compared to WH8102 (with eight). Increased HLIP content has been associated with cyanobacteria found in high light environments (16), thus these results predict that CC9311 would have the capacity to live in high light surface waters or under changing light conditions found during mixing of the water column.

Some differences in the ORFs clustered in the phycobilisome-encoding region were found between WH8102 and CC9311, and these may play a role in the type IV chromatic light adaptation discovered in CC9311 (12). The genome sequence identifies two ORFs (sync_0485 and sync_0486) as phycobiliprotein lyases not found in WH8102; such proteins were predicted to be involved in chromatic adaptation in a recent biochemical study (31). These ORFs clearly merit further attention.

Horizontal Gene Transfer. Strains WH8102 and CC9311 share 1,730 ORFs. Mapping these on the CC9311 genome indicated they were unevenly distributed, with a number of intervening regions that essentially lacked any genes conserved with WH8102 (Fig. 1). Analysis of these regions indicated that some ($\approx$116 ORFs with 19 regions of $>$3 kb) displayed an atypical trinucleotide composition and GC percentage, suggesting they may be novel genomic "islands" relatively recently acquired by CC9311 (Table 2, which is published as supporting information on the PNAS web site). Previous analysis of the WH8102 genome (15) had also identified putative similar islands based on their atypical nucleotide content. These WH8102 putative islands also essentially lacked any of the 1,730 conserved *Synechococcus* genes.

The putative genomic islands with atypical nucleotide content from CC9311 and WH8102 appear to differ significantly in terms of gene function. The majority of the WH8102 islands consist largely of hypothetical genes, often flanked by phage integrase genes, suggesting they may be of phage origin. In contrast, none of the CC9311 islands contain phage integrase genes or other identifiable phage genes. It has been hypothesized that lysogenic phages would be more common in nutrient-poor environments such as the open ocean (discussed in ref. 32). The residual phage-related genes in open ocean WH8102 but not CC9311 are some of the first data consistent with this hypothesis.

Both genomes have unique islands consisting of different polysaccharide biosynthesis genes that may be important in changing cell surface characteristics, perhaps in response to phage or grazing selection pressure. Other islands unique to CC9311 encode an ABC secretion system and an RTX family toxin homologue, a predicted secreted nuclease and protease, and some two-component regulatory system genes. Some of the

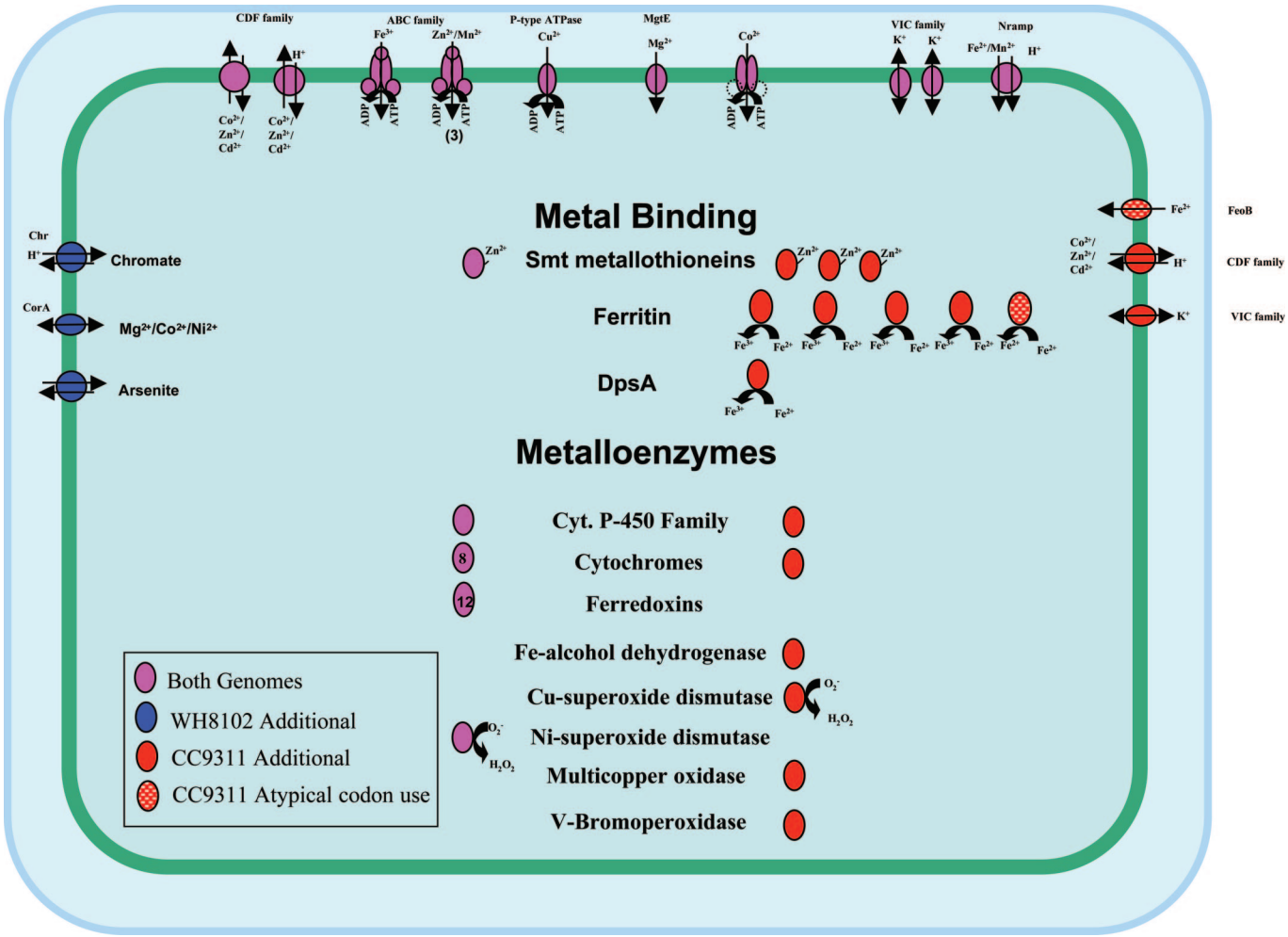

Fig. 3. Overview of metal transport and metabolism in *Synechococcus* CC9311 and WH8102. Metal ion transporters are shown in the membrane, with the arrows indicating the direction of transport. Metal-binding proteins and metalloenzymes are shown inside the cell, and the number of copies of each system is shown in parentheses or within the protein. The color shading of the proteins indicates their distribution: magenta, present in both WH8102 and CC9311; red, present only in CC9311; and blue, present only in WH8102. Hatching indicates that the gene is located in a region with atypical trinucleotide content.

previously mentioned metal metabolism genes, including a ferritin and ferrous iron transport genes, are also found in these islands. The presence of metal-related (especially iron) genes in these islands with atypical codon usage is interesting, because it suggests that metal usage may also be under strong selection. Genes with new physiological capabilities for metal use may be highly favored and maintained in CC9311, if acquired through horizontal gene transfer.

Cell Surfaces: LPS and Pili. CC9311 (relative to WH8102) is missing the genes for the synthesis of KDO, a molecule necessary for the biosynthesis of a typical LPS, and is missing genes for one pathway for the biosynthesis of the sugar rhamnose, a potential component of LPS. The genes for the synthesis of lipid A, the lipid part of LPS, were found. At its simplest level, this suggests that CC9311 has differences in its LPS compared to WH8102. Preliminary LPS analyses suggest this to be the case (B.P., B. Brahamsha, P. Azadi, and S. Snyder, unpublished work). A greatly altered LPS could drastically change the sensitivity of CC9311 to particular phages; because of its abundance at the cell surface, LPS is often a phage receptor (33).

CC9311 has seven putative genes for pili and pilin biosynthesis. Thus it may have pili that would be available for twitching motility or DNA uptake. Both of these could be potentially useful in coastal ecosystems where CC9311 is more likely to encounter surfaces or DNA than in the open ocean. In contrast, CC9311 is missing two major cell surface proteins (*SwmA* and *SwmB*) involved in swimming motility in WH8102 (34, 35). The use of the CC9311 genome and other nonmotile *Synechococcus* genomes will help determine genes unique to WH8102 and thus other genes that could be involved in its unique form of swimming motility. However, our examination of these WH8102 "unique" genes so far has not yielded clues, because many of these genes are annotated only as hypothetical or conserved hypothetical.

Summary. The coastal strain CC9311 has dramatic differences in gene complement compared to the open ocean strain WH8102. Many of these differences are consistent with adaptation to a coastal environment. Because the genus marine *Synechococcus* contains multiple clades (potential species), it will be interesting to see which of these coastal/open ocean differences will be conserved across all clades or whether, even within coastal clades, different strategies exist for adapting to this complex environment.

Methods

Genome Sequencing, Annotation, and Characteristics. The complete genome sequence of *Synechococcus* CC9311 was determined by

using the whole-genome shotgun method (36). Physical and sequencing gaps were closed by using a combination of primer walking, generation and sequencing of transposon-tagged libraries of large-insert clones, and multiplex PCR (37). Identification of putative protein-encoding genes and annotation of the genome were performed as described (38). An initial set of ORFs predicted to encode proteins was initially identified by using GLIMMER (39). ORFs consisting of <30 codons and those containing overlaps were eliminated. Frame shifts and point mutations were corrected or designated "authentic." Functional assignment, identification of membrane-spanning domains, and determination of paralogous gene families were performed as described (38). Sequence alignments and phylogenetic trees were generated by using the methods described (38). The CC9311 genome was found to be composed of one circular chromosome of 2,606,748 bp (Fig. 1), with an average GC content of 52.5%. A total of 3,065 ORFs, 2 rRNA operons, and 44 tRNAs were identified within the CC9311 genome.

Trinucleotide Composition. Distribution of all 64 trinucleotides (3 mers) was determined, and the 3-mer distribution in 2,000-bp windows that overlapped by half their length (1,000 bp) across the genome was computed. For each window, we computed the χ^2 statistic on the difference between its 3-mer content and that of the whole chromosome. A large value for χ^2 indicates the 3-mer composition in this window is different from the rest of the chromosome. Probability values for this analysis are based on assumptions that the DNA composition is relatively uniform throughout the genome, and that 3-mer composition is independent. Because these assumptions may be incorrect, we prefer to interpret high χ^2 values as indicators of regions on the chromosome that appear unusual and demand further scrutiny.

Comparative Genomics. The *Synechococcus* CC9311 and WH8102 genomes were compared at the nucleotide level by suffix tree analysis by using MUMmer (40), and their ORFs were compared by a reciprocal best BLAST match analysis by using an E-value cutoff of 10^{-5}.

We thank The Institute for Genomic Research faculty, sequencing facility, and informatics group for expert advice and assistance. This work was supported by National Science Foundation Grant EF0333162.

1. Wood, A. M., Phinney, D. A. & Yentsch, C. S. (1998) *Mar. Ecol. Prog. Ser.* **162,** 25–31.
2. Olson, R. J., Chisholm, S. W., Zettler, E. R. & Armbrust, E. V. (1988) *Deep-Sea Res.* **35,** 425–440.
3. Olson, R. J., Chisholm, S. W., Zettler, E. R. & Armbrust, E. V. (1990) *Limnol. Oceanogr.* **35,** 45–58.
4. Wood, A. M., Lipsen, M. & Coble, P. (1999) *Deep-Sea Res. II* **46,** 1769–1790.
5. Carpenter, E. J. & Guillard, R. R. L. (1971) *Ecology* **52,** 183–185.
6. Sunda, W. G., Swift, D. G. & Huntsman, S. A. (1991) *Nature* **351,** 55–57.
7. Brand, L. E., Sunda, W. G. & Guillard, R. R. L. (1983) *Limnol. Oceanogr.* **28,** 1182–1198.
8. Ryther, J. H. & Kramer, D. D. (1961) *Ecology* **42,** 444–446.
9. Strzepek, R. F. & Harrison, P. J. (2004) *Nature* **431,** 689–692.
10. Peers, G. & Price, N. M. (2006) *Nature* **441,** 341–344.
11. Toledo, G. & Palenik, B. (1997) *Appl. Environ. Microbiol.* **63,** 4298–4303.
12. Palenik, B. (2001) *Appl. Environ. Microbiol.* **67,** 991–994.
13. Waterbury, J. B. & Rippka, R. (1989) in *Bergey's Manual of Systematic Bacteriology*, eds. Staley, J. T., Bryant, M. P., Pfennig, N. & Holt, J. B. (Williams & Wilkins, Baltimore), Vol. 3, pp. 1728–1746.
14. Ferris, M. J. & Palenik, B. (1998) *Nature* **396,** 226–228.
15. Palenik, B., Brahamsha, B., Larimer, F. W., Land, M., Hauser, L., Chain, P., Lamerdin, J., Regala, W., Allen, E. A., McCarren, J., *et al.* (2003) *Nature* **424,** 1037–1042.
16. Rocap, G., Larimer, F. W., Lamerdin, J., Malfatti, S., Chain, P., Ahlgren, N. A., Arellano, A., Coleman, M., Hauser, L., Hess, W. R., *et al.* (2003) *Nature* **424,** 1042–1047.
17. Dufresne, A., Salanoubat, M., Partensky, F., Artiguenave, F., Axmann, I. M., Barbe, V., Duprat, S., Galperin, M. Y., Koonin, E. V., Le Gall, F., *et al.* (2003) *Proc. Natl. Acad. Sci. USA* **100,** 10020–10025.
18. Hirani, T., Suzuki, I., Murata, N., Hayashi, H. & Eaton-Rye, J. (2001) *Plant Mol. Biol.* **45,** 133–144.
19. Grass, G., Thakali, K., Klebba, P., Thieme, D., Muller, A., Wildner, G. & Rensing, C. (2004) *J. Bacteriol.* **186,** 5826–5833.
20. Brand, L. E., Sunda, W. G. & Guillard, R. R. L. (1986) *J. Exp. Mar. Biol. Ecol.* **96,** 225–250.
21. Pedersén, M. (1976) *Physiol. Plant* **37,** 6–11.
22. Carter, J. N., Beatty, K. E., Simpson, M. T. & Butler, A. (2002) *J. Inorg. Biochem.* **91,** 59–69.
23. Unson, M. D., Holland, N. D. & Faulkner, D. J. (1994) *Mar. Biol.* **119,** 1–11.
24. Sionov, E., Roth, D., Sandovsky-Losica, H., Kashman, Y., Rudi, A., Chill, L., Berdicevsky, I., Segal, E., *et al.* (2005) *J. Infect.* **50,** 453–460.
25. Kuma, K., Nakabayashi, S., Suzuki, Y., Kudo, I. & Matsunaga, K. (1992) *Marine Chemistry* **37,** 15–27.
26. Barbeau, K., Rue, E. L., Trick, C. G., Bruland, K. W. & Butler, A. (2003) *Limnol. Oceanogr.* **48,** 1069–1078.
27. Cutter, G., Cutter, L., Featherstone, A. & Lohrenz, S. E. (2001) *Deep-Sea Res. II* **48,** 2895–2915.
28. Gupta, A., Whitton, B. A., Morby, A. P., Huckle, J. W. & Robinson, N. J. (1992) *Proc. R. Soc. London Ser. B* **248,** 273–281.
29. Bruland, K. W., Rue, E. L. & Smith, G. J. (2001) *Limnol. Oceanogr.* **46,** 1661–1674.
30. Booth, I. R. & Louis, P. (1999) *Curr. Opin. Microbiol.* **2,** 166–169.
31. Everroad, C., Six, C., Partensky, F., Thomas, J. C., Holtzendorff, J. & Wood, A. M. (2006) *J. Bacteriol.* **188,** 3345–3356.
32. Ortmann, A. C., Lawrence, J. E. & Suttle, C. A. (2002) *Microb. Ecol.* **43,** 225–231.
33. Traurig, M. & Misra, R. (1999) *FEMS Microbiol. Lett.* **181,** 101–108.
34. McCarren, J., Heuser, J., Roth, R., Yamada, N., Martone, M. & Brahamsha, B. (2005) *J. Bacteriol.* **187,** 224–230.
35. McCarren, J. & Brahamsha, B. (2005) *J. Bacteriol.* **187,** 4457–4462.
36. Fraser, C. M., Casjens, S., Huang, W. M., Sutton, G. G., Clayton, R., Lathigra, R., White, O., Ketchum, K. A., Dodson, R., Hickey, E. K., *et al.* (1997) *Nature* **390,** 580–586.
37. Tettelin, H., Radune, D., Kasif, S., Khouri, H. & Salzberg, S. L. (1999) *Genomics* **62,** 500–507.
38. Paulsen, I. T., Seshadri, R., Nelson, K. E., Eisen, J. A., Heidelberg, J. F., Read, T. D., Dodson, R. J., Umayam, L., Brinkac, L. M., Beanan, M. J., *et al.* (2002) *Proc. Natl. Acad. Sci. USA* **99,** 13148–13153.
39. Salzberg, S. L., Delcher, A. L., Kasif, S. & White, O. (1998) *Nucleic Acids Res.* **26,** 544–548.
40. Delcher, A. L., Phillippy, A., Carlton, J. & Salzberg, S. L. (2002) *Nucleic Acids Res.* **30,** 2478–2483.