Plant Cell Physiol. 38(11): 1171-1176 (1997)

Mini Review

Complete Genome Structure of the Unicellular Cyanobacterium *Synechocystis* sp. PCC6803

Takakazu Kaneko and Satoshi Tabata

Kazusa DNA Research Institute, 1532-3 Yana, Kisarazu, Chiba, 292 Japan

Cyanobacteria are photoautotrophic organisms capable of oxygen-producing photosynthesis similar to that in eukaryotic algae and plants, and because of this, they have been used as model organisms for the study of the mechanism and regulation of oxygen-producing photosynthesis. To understand the entire genetic system in cyanobacteria, the nucleotide sequence of the entire genome of the unicellular cyanobacterium *Synechocystis* sp. PCC6803 has been determined. The total length of the circular genome is 3,573,470 bp, with a GC content of 47.7%. A total of 3,168 potential protein coding genes were assigned. Of these, 145 (4.6%) were identical to reported genes, and 1,259 (39.6%) and 342 (10.8%) showed similarity to reported and hypothetical genes, respectively. The remaining 1,422 (45.0%) showed no apparent similarity to any genes registered in the databases. Classification of the genes by their biological function and comparison of the gene complement with those of other organisms have revealed a variety of features of the genetic information characteristic of a photoautotrophic organism. The sequence data, as well as other information on the *Synechocystis* genome, is presented in CyanoBase on WWW [http://www.kazusa.or.jp/cyano/].

Key words: Cyanobacteria — Genome sequencing — Photosynthesis — Plastid genome — *Synechocystis* sp. PCC6803.

Cyanobacteria, also called "blue-green algae", are aquatic prokaryotes capable of oxygenic photosynthesis; over 1,500 species with various morphologies have been reported so far. Cyanobacteria are clearly separated from other photosynthetic bacteria, such as purple and green bacteria, because they utilize H_2O as an electron donor; others do not. It has generally been accepted that the ancestors of cyanobacteria which acquired the ability to conduct oxygen-producing photosynthesis in the early stage of evolution gave rise to plant plastids by endosymbiotic events, thereby conferring the ability for photosynthesis to the extant algae and plants. There are many properties, both in the structure and mechanism of photosynthesis, common to cyanobacteria, algae, and plants. Various genetic engineering techniques which facilitate studies of gene function and regulation are applicable to cyanobacteria. Bacause of these, cyanobacteria have long been used as model organisms for studying photosynthesis in higher plants, where a more complex genetic system regulates the whole photosynthetic process.

Abbreviations: IS, insertion sequence; ORF, open reading frame.

Although cyanobacteria comprise one of the largest constituents of the gram-negative bacteria, only a limited number of strains have been used for physiological and genetic studies. Among these, *Synechocystis* sp. PCC6803, a unicellular cyanobacterium with the advantages of naturally-transformable characteristics, is able to grow heterotrophically in the absence of light, allowing characterization of mutants of both PSI and II. In 1996, the nucleotide sequence of the entire genome of *Synechocystis* sp. PCC6803 was determined. This was the first report of the whole genome in a photoautotrophic organism. By taking a close look at the structure of each gene and the organization of the entire genome, genetic information characteristic of photoautotrophic bacteria is becoming better understood.

Genome structures of cyanobacteria—Although the structure and function of individual genes in cyanobacteria have been intensively studied, information on genome structure and on organization of the genes throughout the entire genome is still inadequate. Physical and gene maps of four cyanobacterial strains, *Anabaena* sp. PCC7120 (Bancroft et al. 1989), *Synechococcus* sp. PCC6301 (Kaneko et al. 1996a), *Synechocystis* sp. PCC6803 (Kotani et al. 1994) and *Synechococcus* sp. PCC7002 (Chen and Widger 1993) have been reported previously. The sizes and average GC contents of these four genomes are summarized in Table 1. PCC7120 is a filamentous cyanobacterium capable of nitrogen fixation; *Synechococcus* sp. PCC6301, *Synechococcus* sp. PCC7002, and *Synechocystis* sp. PCC6803 are unicellular cyanobacteria. The genome size and GC content differ from strain to strain, although two *Synechococcus* strains have genomes of similar sizes. No similarity is observed not only in the physical maps but in the relative locations of the genes which have been mapped. Because a variety of bacteria with only the ability for oxygenic photo-

Table 1 Sizes and avarage GC contents of four cyanobacterial genomes

	Genome size (Mb)	GC-content (%)
Anabaena sp. PCC7120	6.4[a]	42.5[e]
Synechococcus sp. PCC6301	2.7[b]	55.1[e]
Synechococcus sp. PCC7002	2.7[c]	49.1[e]
Synechocystis sp. PCC6803	3.5[d]	47.7[d]

[a] Bancroft et al. 1996.
[b] Kaneko et al. 1996a.
[c] Chen et al. 1993.
[d] Kaneko et al. 1996b.
[e] Herdman et al. 1979.

synthesis in common are classified as cyanobacteria, it may be natural that there is little apparent similarity among their genomic structures. Even though extant cyanobacterial strains may originate from a common ancestor, rearrangement of the genomes, gene shuffling, gene multiplication, and base substitution should repeatedly occur during evolution to give rise to a variety of genomes with differing sizes, gene locations, and GC contents.

Genomic structure of Synechocystis sp. PCC6803—The nucleotide sequence of the entire genome of *Synechocystis* sp. PCC6803 was determined in 1996 as the first example of a photoautotrophic organism. The circular genome was covered by 107 cosmid clones, 15 lambda clones, and 17 long-PCR products, whose locations were confirmed on the physical map. Then the nucleotide sequence of each of the clones and the long-PCR products was determined, and connected based on the physical map. The nucleotide sequence of the entire genome of *Synechocystis* sp. PCC6803 thus deduced was 3,573,470 bp long, and the average GC content was 47.7% (Kaneko et al. 1996b).

One of the notable features of the *Synechocystis* sp. PCC6803 genome is its high content of two types of repetitive sequences, IS-like elements and a HIP1 (highly iterated palindromic) sequence. There are over 70 IS-like elements containing ORFs which show similarity to bacterial transposases spread all over the genome (Kaneko, personal communication). They were classified into 9 groups on the basis of similarity and/or structure of 17–36 bp inverted repeats on both termini. Interestingly, only 26 of them seemed to hold the coding capacity of full-length transposase proteins. The remaining ORFs were disrupted by mutations such as frame-shift or deletion or by insertion of other IS-like elements. In addition, traces of inversion of the large portions of the genome due to homologous recombination between the identical IS-like sequences at different positions in the genome were observed. These observations strongly suggest that the IS-like elements have played a significant role in dynamic alteration of the genome structure.

HIP1 is a 8 base sequence, GCGATCGC, first reported in the genomes of *Synechococcus* species and some other cyanobacterial strains (Gupta et al. 1993). By surveying the entire genome sequence of *Synechocystis* sp. PCC6803, 3,160 copies of HIP1 were found to be fairly evenly distributed in the genome (Tanaka, personal communication). The average frequency of occurrence was 1 copy/1,131 bp. Approximately 90% of the HIP1 were located in the potential protein coding regions, indicating that these sequences are translated. Interestingly, more than 70% of them are translated as Ala-Ile-Ala, while the rest are either as Gly-Asp-Arg or Arg-Ser. The origin and functional significance of these HIP1 in cyanobacterial genomes remain to be studied.

Structural RNA genes in the Synechocystis sp. PCC6803 genome—To assign the genes coding for structural RNAs, the nucleotide sequence of the entire genome was subjected to computer assisted analyses, including a similarity search and prediction by computer programs. As a result, 2 copies of an rRNA gene cluster, 42 tRNA genes, and a gene for a RNA subunit of RNase P were identified.

The two rRNA gene clusters, each containing genes for 16S rRNA, Ile-tRNA, 23S rRNA, and 5S rRNA in this order, were identified on the basis of sequence similarity to registered RNA sequences. The nucleotide sequences of these two clusters, consisting of the 5,028 bp, are completely identical and are located 870 kb apart from each other in reverse orientation in the genome. Other cyanobacteria and the most of the plant plastids also have two copies of the rRNA gene cluster in their genomes (Sugiura 1993, Kaneko et al. 1996a), whereas other eubacteria have various numbers of copies (1–7) of this cluster (Fleischmann et al. 1995, Fraser et al. 1995, Himmelreich et al. 1996, Bult et al. 1996). These data may reflect the phylogenetic relationship among the genomes of cyanobacteria and plant plastids.

Forty-two tRNA genes, which were identified in the *Synechocystis* sp. PCC6803 genome based on their sequence similarity to the reported tRNAs and computer prediction by the tRNAscan computer program, represent 41 tRNA species (Nakamura, personal communication). Most of the genes are present in a single copy; the only exception is 2 copies of trnI-GAU genes located in the duplicated rRNA gene clusters. The number of the tRNA genes in the *Synechocystis* sp. PCC6803 genome is limited compared with that in the *Escherichia coli* genome (86 tRNA genes) (Ikemura, personal communication), but is sufficient for dealing with all the codons used by this species. The genes are scattered all over the genome and exist as single units; only the genes for trnY-GUA and trnT-GGU are tandemly aligned with a 8 bp interval. This is quite different from the *E. coli* genome in which many of the tRNA genes form clusters consisting of 5 to 9 genes

(Komine et al. 1990). It is also notable that most of the tRNA genes in *Synechocystis* sp. PCC6803 showed high degrees of sequence similarity to those in the plant plastids (Nakamura, personal communication).

Protein coding genes in the Synechocystis sp. PCC6803 genome—Assignment of potential protein coding regions was performed by combination of two different methods, similarity search and computer prediction. All the ORFs consisting of longer than 50 sense codons in the six translation frames were translated into amino acid sequences and subjected to a similarity search against the public protein database. The ORFs which showed similarity to known genes were preferentially taken into account, and ORFs showing no apparent similarity were assigned to the unoccupied spaces. If two ORFs overlapped, the longer one was chosen, unless the function of the shorter one could be anticipated. Then the short DNA stretches between the assigned ORFs were subjected to search for the known short genes. As a consequence, 3,001 potential protein coding regions were assigned in the genome.

In parallel with the similarity search described above, the nucleotide sequence of the entire genome was subjected to prediction of coding regions using a computer program, GeneMark, which looks for differences in the frequency of regional appearance of oligonucleotides between coding and non-coding regions (Hirosawa et al. 1995). The GeneMark program predicted 167 additional coding regions; 282 ORFs assigned by similarity search escaped from the prediction. By combining the results of these two independent analyses together, a total of 3,168 potential protein coding genes were assigned in the *Synechocystis* sp. PCC6803 genome. The gene density is 1 gene per 1.1 kb, a typical value for bacterial genomes. The average length of the translated gene products was 326 amino acid residues, and the potential protein coding regions, as a whole, occupied 87.0% of the genome.

Of the 3,168 potential protein coding genes, 145 (4.6%) were identical to the reported *Synechocystis* sp. PCC6803 genes, 935 (29.4%) showed high degree of similarity to reported genes, and 324 (10.2%) and 342 (10.8%) showed similarity to known and hypothetical genes, respectively. The remaining 1,424 (45.0%) did not show any apparent similarity to sequences in the public DNA and protein databases. One thousand four hundred sixteen genes whose functions were deduced were classified into 15 categories according to their biological function, as shown in Table 2.

Characteristic features of the Synechocystis sp. PCC6803 genes—The nucleotide sequences of entire genomes have been reported for several bacteria, including gram-negative bacteria and an archaebacterium (Fleischmann et al. 1995, Fraser et al. 1995, Himmelreich et al. 1996, Bult et al. 1996). By comparing the entire gene complement of the *Synechocystis* sp. PCC6803 genome with those in other bacterial genomes, characteristic traits of cyanobacteria can be defined.

(1) *Photosynthetic genes*—One of the most characteristic categories for the cyanobacterial genome listed in Table

Table 2 Functional categories of putative protein gene products in *Synechocystis* sp. PCC6803

Category	Gene number
Amino acid biosynthesis	84
Biosynthesis of cofactors, prosthetic groups, and carriers	108
Cell envelope	64
Cellular processes	62
Central intermediary metabolism	31
Energy metabolism	86
Fatty acid, phospholipid, and sterol metabolism	35
Photosynthesis and respiration	131
Nucleic acid metabolism	38
General regulatory functions	147
DNA replication, recombination, and repair	49
Transcription	24
Translation	144
Transport and binding proteins	158
Other categories	255
Function unknown	1,752
Total	3,168

2 is "Photosynthesis and respiration". One hundred twenty eight genes involved in various stages of photosynthesis were identified on the basis of sequence similarity to the photosynthetic genes reported in *Synechocystis* sp. PCC6803 and other organisms. The relative locations of the genes in the genome are shown in Fig. 1. Gene multiplication was observed for the following genes: *cpcC, cpcG, ctaC, ctaD, ctaE, ndhD, ndhF, petC, petF, psaK, psbA, psbC*(*isiA*) and *psbD*. The genes for small subunits of PSI and PSII, including *psaG, psaH, psaN, psbP, psbQ, psbR, psbS, psbT,* and *psbW*, which are commonly present in the nuclear genomes of higher plants, were not found in the *Synechocystis* sp. PCC6803 genome.

(2) *Relationship to plant plastids*—It has been generally accepted that unicellular cyanobacteria are candidates for the progenitors of plastids in algae and plants. If this is the case, traces of evolution might be found in the genomes of extant cyanobacteria and plastids. The genomes of algal and plant plastids are 120 to 160 kb long, much shorter than those of cyanobacteria, and contain only 100 to 200 genes. Up to the present, the entire sequences of 10 plastid genomes have been reported (Reardon and Price 1995), and complete lists of potential gene complements have been compiled. Comparison of 3,168 genes in the entire

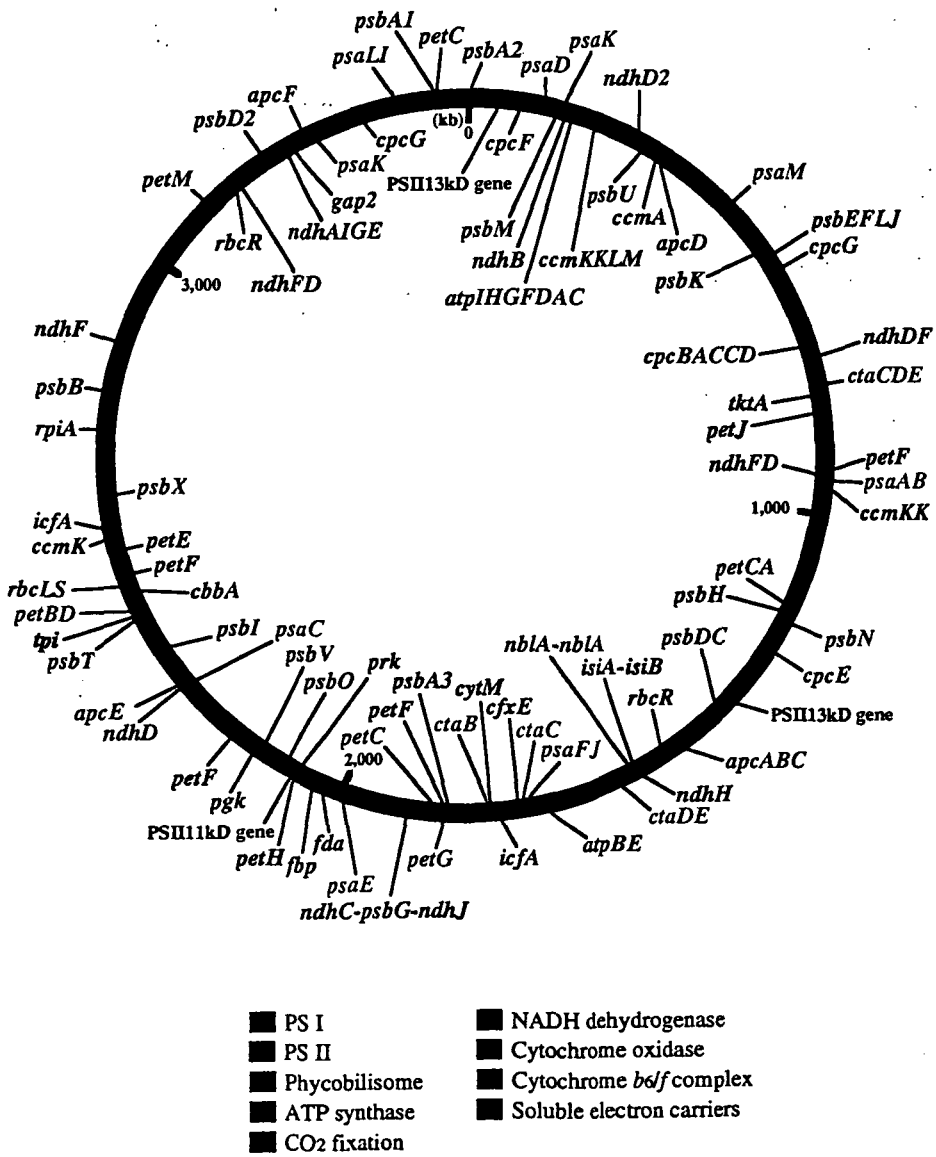

Fig. 1 Locations of 89 photosynthetic genes and gene clusters in the *Synechocystis* sp. PCC6803 genome. The genes on the direct DNA strand are indicated outside of the gray circle based the nucleotide sequence and ones on the complementary strand are on the inside. The genes are color-coded according to their functional categories.

Synechocystis sp. PCC6803 genome with the genes in 10 plastid genomes indicates that 224 of *Synechocystis* sp. PCC6803 genes have homologues in the plastid genomes. These include those for a variety of biological pathways, PSI and PSII, the electron transfer system, CO_2 fixation, transcription, and translation. The most significant similarity as a whole genome was observed between *Synechocystis* sp. PCC6803 and the plastid of *Porphyra purpurea* (red alga), in which 94% of the plastid genes showed homology to the genes in *Synechocystis* sp. PCC6803 genome (Kaneko, personal communication). These observations strongly suggest that the progenitor(s) of the extant cyanobacteria has given rise to plastids and that most of the genes were lost and/or transferred to nuclear genomes during the process of evolution.

(3) *Genes for signal transduction systems*—Bacteria have a signal transduction system for adaptive responses to various changes in their environment. This system is referred to as a "two-component system", because it typically consists of two signal transducers, a sensory kinase and a response regulator. The sensory kinases monitor the environmental conditions, and the response regulators regulate gene expression according to the signals from the sensory kinases. In *Synechocystis* sp. PCC6803, a total of 80 genes for two component signal transducers, including 26 genes for sensory kinases, 38 genes for response regulators, and 16 genes for hybrid sensory kinases containing both transmitter and receiver domains, have been identified (Mizuno et al. 1996). The genes for the two component system were also found in the plastid genomes of red alga *P. purpurea* (Reith and Munholland 1995) and Glaucocystophyte *Cyanophora paradoxa* (Stirewalt et al. 1995). Some of these genes show sequence similarity to those in the *Synechocystis* sp. PCC6803 genome. The gene products of *ycf26, ycf27* and *ycf29* are similar to those of sll0698 (drug sensory protein A) (Bartsevich and Shestakov 1995), slr0947 coding for a regulator belonging to OmpR subfamily, and slr1783 and slr1909 which are regulators of the NarL subfamily, respectively. The two component system was also reported in a higher plant, *Arabidopsis thaliana* (Chang et al. 1993, Kakimoto 1996).

A similarity search indicated that several putative gene products of sensory kinase genes in *Synechocystis* sp. PCC6803 possessed chromophore-binding domains commonly found in phytochromes of green plants (Kaneko et al. 1996). The most remarkable example of such genes is an ORF denoted as slr0473 (Mizuno et al. 1996). The N-terminus of the putative gene product of slr0473 showed significant sequence similarity to that of phytochrome C of higher plants. It was reported that the slr0473 gene product produced in *E. coli* cells bound efficiently to phycocyanobilin, a linear tetrapyrrole chromophore, and that the gene product showed the spectra characteristic of the plant phytochrome after irradiation with red and far-red light (Hughes et. al. 1997). The results indicate that the gene product of slr0473 is a functional phytochrome. Another gene for a sensory kinase, sll1124, also contained the chromophore-binding domain (Wilde et al. 1997). Introduction of mutation into sll1124 had no effect on photoautotrophic growth under white or far red light, but a drastic effect under blue light, suggesting that the gene may act as one of the light receptors involved in photoautotrophic growth.

Database for the genomic information from Synechocystis sp. PCC6803—The nucleotide sequence and other information on the assigned protein genes in the *Synechocystis* sp. PCC6803 genome are available in both the public DNA databases and the web database named CyanoBase [http://www.kazusa.or.jp/cyano/]. In CyanoBase, the graphical images of the circular physical map and the linear gene map are presented. By clicking on each gene box on the linear map, gene information such as the translated amino acid, the nucleotide sequences of the gene, and the results of similarity search can be obtained. The entire nucleotide sequence of the genome and the amino acid sequences of the putative gene products are available from the ftp site, which makes it possible to re-analyze the sequence information in one's own system. Using the BLAST search engine equipped in the database, the question "Is there a homologue of gene X in the cyanobacterial genome?" can be easily addressed. CyanoBase, therefore, is not only a site where information on the cyanobacterial genome is provided, but also a tool for a variety of studies on gene structure and function, gene organization, and genomic evolution.

Future perspective—Though the structures of all the gene constituents in *Synechocystis* sp. PCC6803 were deduced by analyzing all of the genomic information with the aid of computers, 55% of the assigned genes did not show any sequence similarity to those of known function. One of the most immediate and urgent goals is to determine the biological function of these genes by experimental methods. Systematic disruption of the genes whose functions are unknown is currently underway to examine the possible role of the gene product. Another promising strategy is to purify the proteins when they are produced in large quantities in *E. coli* cells and characterize them in vitro. This has been proved effective by the following examples. The product of slr0473, which exhibited sequence similarity to the plant phytochromes, showed several characteristics of plant phytochromes according to the biochemical analysis using proteins overproduced in *E. coli* as described in the previous section. The gene products of slr0088 and slr0611 were shown to be β-carotene ketolase which produces echinenone in the carotenoid biosynthesis pathway (Fernandez-Gonzalez et al. 1997), and solanesyl diphosphate synthase which is related to biosynthesis of ubiquinone 9 (Okada et al. 1997), respectively.

Since information on the primary structures of genes

and their putative products is known, it is possible to monitor the expression of many or even all of the genes in the genome utilizing sequence information of each gene as a tag. With microarray technology using nucleotide sequence information, the transcriptional level of each of all the gene constituents can be measured on a panel at once. Changes in translation can also be monitored by proteome analysis. Soluble proteins in the cell extract of *Synechocystis* sp. PCC6803 were resolved on two-dimensional gels, and the partial amino acid sequences of the N-termini of 96 major proteins were determined (Sazuka and Ohara 1996). Each of the sequences obtained was back-translated and correlated to the corresponding gene in the genome [http://www.kazusa.or.jp/tech/sazuka/cyano/proteome.html]. Using this system, one can easily identify the genes responsible for the proteins whose expression has changed under different experimental conditions.

In this mini-review, we described the several examples of characteristic features of the cyanobacterial genome which have become clear by analysis of the entire genome sequence. To fully utilize the sequence information, however, careful re-examination of the data should be carried out from various points of view. This will undoubtedly extract more information from the genome sequences and lead to better understanding of the genetic systems in cyanobacteria.

Referances

Bancroft, I., Wolk, C.P. and Oren, E.V. (1989) *J. Bacteriol.* 171: 5940–5948.

Bartsevich, V.V. and Shestakov, S.V. (1995) *Microbiology* 141: 2915–2920.

Bult, C.J., White, O., Olsen, G.J., Zhou, L., Fleischmann, R.D., Sutton, G.G., Blake, J.A., FitzGerald, L.M., Clayton, R.A., Gocayne, J.D., Kerlavage, A.R., Dougherty, B.A., Tomb, J.F., Adams, M.D., Reich, C.I., Overbeek, R., Kirkness, E.F., Weinstock, K.G., Merrick, J.M., Glodek, A., Scott, J.L., Geoghagen, N.S.M., Weidman, J.F., Fuhrmann, J.L., Nguyen, D., Utterback, T.R., Kelley, J.M., Peterson, J.D., Sadow, P.W., Hanna, M.C., Cotton, M.D., Roberts, K.M., Hurst, M.A., Kaine, B.P., Borodovsky, M., Klenk, H.P., Fraser, C.M., Smith, H.O., Woese, C.R. and Venter, J.C. (1996) *Science* 273: 1058–1073.

Chang, C., Kwok, S.F., Bleecker, A.B. and Meyerowitz, E.M. (1993) *Science* 262: 539–544.

Chen, X. and Widger, W.R. (1993) *J. Bacteriol.* 175: 5106–5116.

Fernandez-Gonzalez, B., Sandmann, G. and Vioque, A. (1997) *J. Biol. Chem.* 272: 9728–9733.

Fleischmann, R.D., Adams, M.D., White, O., Clayton, R.A., Kirkness, E.F., Kerlavage, A.R., Bult, C.J., Tomb, J.F., Dougherty, B.A., Merrick, J.M., McKenney, K., Sutton, G., FitzHugh, W., Fields, C., Gocayne, J.D., Scott, J., Shirley, R., Liu, L., Glodek, A., Kelly, J.M., Weidman, J.F., Phillips, C.A., Spriggs, T., Hedblom, E., Cotton, M.D., Utterback, T.R., Hanna, M.C., Nguyen, D.T., Saudek, D.M., Brandon, R.C., Fine, L.D., Fuhrmann, J.L., Geoghagen, N.S.M., Gnehm, C.L., McDonald, L.A., Small, K.V., Frazer, C.M., Smith, H.O. and Venter, J.C. (1995) *Science* 269: 496–512.

Fraser, C.M., Gocayne, J.D., White, O., Adams, M.D., Clayton, R.A., Fleischmann, R.D., Bult, C.J., Kerlavage, A.R., Sutton, G., Kelley, J.M., Fritchman, J.L., Weidman, J.F., Small, K.V., Sandusky, M., Fuhrmann, J., Nguyen, D., Utterback, T.R., Saudek, D.M., Phillips, C.A., Merrick, J.M., Tomb, J.F., Dougherty, B.A., Bott, K.F., Hu, P.C., Lucier, T.S., Peterson, S.N., Smith, H.O., Hutchison III, C.A. and Venter, J.C. (1995) *Science* 270: 397–403.

Gupta, A., Morby, A.P., Turner, J.S., Whitton, B.A. and Robinson, N.J. (1993) *Mol. Microbiol.* 7: 189–195.

Herdman, M., Janvier, M., Waterbury, J.B., Rippka, R. and Stanier, R.Y. (1979) *J. Gen. Microbiol.* 111: 63–71.

Himmelreich, R., Hilbert, H., Plagens, H., Pirkl, E., Li, B.C. and Herrmann, R. (1996) *Nucl. Acids Res.* 24: 4420–4449.

Hirosawa, M., Kaneko, T., Tabata, S., McIninch, J.D., Hayes, W.S., Borodovsky, M. and Isono, K. (1995) *DNA Res.* 2: 239–246.

Hughes, J., Lamparter, T., Mittmann, F., Hartmann, E., Gärtner, W., Wilde, A. and Börner, T. (1997) *Nature* 386: 663.

Kakimoto, T. (1996) *Science* 274: 982–985.

Kaneko, T., Matsubayashi, T., Sugita, M. and Sugiura, M. (1996a) *Plant Mol. Biol.* 31: 193–201.

Kaneko, T., Sato, S., Kotani, H., Tanaka, A., Asamizu, E., Nakamura, Y., Miyajima, N., Hirosawa, M., Sugiura, M., Sasamoto, S., Kimura, T., Hosouchi, T., Matsuno, A., Muraki, A., Nakazaki, N., Naruo, K., Okumura, S., Shimpo, S., Takeuchi, C., Wada, T., Watanabe, A., Yamada, M., Yasuda, M. and Tabata, S. (1996b) *DNA Res.* 3: 109–136.

Komine, Y., Adachi, T., Inokuchi, H. and Ozeki, H. (1990) *J. Mol. Biol.* 212: 579–598.

Kotani, H., Kaneko, T., Matsubayashi, T., Sato, S., Sugiura, M. and Tabata, S. (1994) *DNA Res.* 1: 303–307.

Mizuno, T., Kaneko, T. and Tabata, S. (1996) *DNA Res.* 3: 407–414.

Okada, K., Minehira, M., Zhu, X., Suzuki, K., Nakagawa, T., Matsuda, H. and Kawamukai, M. (1997) *J. Bacteriol.* 179: 3058–3060.

Reardon, E.M. and Price, C.A. (1995) *Plant Mol. Biol. Rep.* 13: 320–326.

Reith, M. and Munholland, J. (1995) *Plant Mol. Biol. Rep.* 13: 333–335.

Sazuka, T. and Ohara, O. (1996) *DNA Res.* 3: 225–232.

Stirewalt, V.L., Michalowski, C.B., Löffelhardt, W., Bohnert, H.J. and Bryant, D.A. (1995) *Plant Mol. Biol. Rep.* 13: 327–332.

Sugiura, M. (1993) *Plant Mol. Biol.* 19: 149–168.

Wilde, A., Churin, Y., Schubert, H. and Börner, T. (1997) *FEBS Lett.* 406: 89–92.

(Received July 24, 1997; Accepted September 17, 1997)