J Mol Evol (2001) 52:333–341
DOI: 10.1007/s002390010163

JOURNAL OF MOLECULAR EVOLUTION

© Springer-Verlag New York Inc. 2001

Horizontal Transfer of the Photosynthesis Gene Cluster and Operon Rearrangement in Purple Bacteria

Naoki Igarashi, Jiro Harada, Sakiko Nagashima, Katsumi Matsuura, Keizo Shimada, Kenji V.P. Nagashima

Department of Biology, Tokyo Metropolitan University, Minamiohsawa 1-1, Hachioji, Tokyo 192-0397, Japan

Received: 23 August 2000 / Accepted: 13 December 2000

Abstract. A 37-kb photosynthesis gene cluster was sequenced in a photosynthetic bacterium belonging to the β subclass of purple bacteria (*Proteobacteria*), *Rubrivivax gelatinosus.* The cluster contained 12 bacteriochlorophyll biosynthesis genes (*bch*), 7 carotenoid biosynthesis genes (*crt*), structural genes for photosynthetic apparatuses (*puf* and *puh*), and some other related genes. The gene arrangement was markedly different from those of other purple photosynthetic bacteria, while two superoperonal structures, *crtEF–bchCXYZ–puf* and *bchFNBHLM–lhaA–puhA,* were conserved. Molecular phylogenetic analyses of these photosynthesis genes showed that the photosynthesis gene cluster of *Rvi. gelatinosus* was originated from those of the species belonging to the α subclass of purple bacteria. It was concluded that a horizontal transfer of the photosynthesis gene cluster from an ancestral species belonging to the α subclass to that of the β subclass of purple bacteria had occurred and was followed by rearrangements of the operons in this cluster.

Key words: Horizontal gene transfer — Photosynthesis gene cluster — Purple bacteria — *Rubrivivax gelatinosus* — Superoperon — Phylogenetic tree

Introduction

It has been well established that a large cluster of genes required for photosynthesis exists in purple bacteria (*Proteobacteria*), although the significance of this cluster has not been fully clarified. In the purple bacterium, *Rhodobacter capsulatus,* the photosynthesis gene cluster consists of several overlaps of transcriptional units (Beatty 1995) and shares more than 1% of the total genome. Photosynthesis gene clustering has not been reported in any other classes of eubacteria with the exception of heliobacteria (Xiong et al. 1998).

The photosynthetic apparatus of purple bacteria consists of a reaction-center complex and light-harvesting complexes. These complexes contain bacteriochlorophylls and carotenoids to absorb light energy and convert it to electrochemical energy. The genes coding for the two core proteins of the reaction-center complex of purple bacteria, the L and M subunits, form an operon called *puf* with the genes for the two small hydrophobic proteins of the core light-harvesting (LH1) complex, the α and β subunits. The H subunit of the reaction-center complex is coded by the *puhA* gene and is included in another operon, *puh* (Bauer et al. 1991). We have previously described the horizontal transfer of the *puf* operon in purple bacteria. The phylogenetic tree based on the genes coding for the L and M subunits was inconsistent with those based on the sequences of 16S rRNA and soluble cytochrome *c* (Nagashima et al. 1993, 1994, 1997). It was suggested that the *puf* genes of *Rvi. gelatinosus* and the related species were obtained from other purple bacteria distantly related to these species by horizontal gene transfer (Nagashima et al. 1997).

Genes required for the syntheses of bacteriochlorophyll *a* and carotenoids have been mapped on the DNAs of some purple nonsulfur bacteria using the techniques of gene disruption and complementation (Taylor et al.

Correspondence to: Kenji V.P. Nagashima; *e-mail:* nagashima-kenji@c.metro-u.ac.jp

1983; Coomber et al. 1990; Yildiz et al. 1992). In two closely related purple bacteria, *Rba. capsulatus* and *Rhodobacter sphaeroides,* most of the genes required for the biosyntheses of bacteriochlorophylls and carotenoids are flanked by *puh* and *puf* operons (Youvan et al. 1984; Naylor et al. 1999). Pigment biosynthesis genes also form operons, which are further organized into transcriptional units called "superoperons" (Young et al. 1989; Wellington et al. 1991; Bauer et al. 1991). The total size of this photosynthesis gene cluster is approximately 46 kb in *Rba. capsulatus* (Youvan et al. 1984). *Rhodocista centenaria* also seems to have a similar arrangement of photosynthesis genes (Yildiz et al. 1992). In addition, several photosynthesis genes that have been cloned and sequenced in two other purple bacteria, *Rhodospirillum rubrum* and *Acidiphilium rubrum,* showed arrangements consistent with those in the *Rhodobacter* species (Berard and Gingras 1991; Masuda et al. 1999). A genome-sequencing project for *Rhodopseudomonas palustris* is now in progress, and most of the photosynthesis genes appear to be clustered in this species (http://spider.jgi-psf.org/JGI_microbial/html/). These species all belong to the α subclass of purple bacteria, which is one of three subclasses, α, β, and γ, containing photosynthetic species (Woese 1987).

In *Rvi. gelatinosus,* a carotenoid biosynthesis gene, *crtD,* was recently shown to be located downstream of the *puf* operon (Nagashima et al. 1995), indicating that the gene arrangement in the photosynthesis gene cluster of *Rvi. gelatinosus* differs from those of other purple bacteria (Ouchane et al. 1997a). In this study, the whole nucleotide sequence of the 37-kb photosynthesis gene cluster of *Rvi. gelatinosus* IL144 was determined. The photosynthesis genes in *Rvi. gelatinosus* showed significantly high sequence identities to those in the species of the α subclass, although the gene arrangement in *Rvi. gelatinosus* was markedly different from those in previously characterized photosynthesis gene clusters. A portion of these results has already been published in the form of a preliminary report (Igarashi et al. 1998).

Materials and Methods

Cloning and Sequencing. The *Rvi. gelatinosus* strain IL144 was grown anaerobically in light at 30°C in a PYS medium (Nagashima et al. 1996). The genomic DNA was purified, partially digested with *Sau*3AI, and ligated to a unique *Bam*HI site of the SuperCos1 cosmid vector (Stratagene, La Jolla, CA). The cosmid library was screened by colony hybridization using a ^{32}P-labeled 0.9-kb DNA fragment containing the 5′ region of the *crtD* gene of *Rvi. gelatinosus* cloned previously (Nagashima et al. 1996). Three positive colonies were picked up and named pgc#6, pgc#11, and pgc#12. The nucleotide sequences of two of these clones, pgc#6 and pgc#12, have been determined. The cosmids were partially and randomly digested with DNaseI or *Hae*III and subcloned into pUC118 or pHSG396 plasmids. More than 400 subclones were sequenced using a Dye Terminator Cycle Sequencing Kit and PRISM 377 or PRISM 310 DNA sequencer (PE Applied Biosystems, Foster City, CA). A part of the sequencing was performed using synthetic oligonucleotides as primers. The nucleotide sequence reported in this study was determined completely on both strands. The nucleotide sequence data were analyzed using a DNASIS software package (Hitachi Soft, Yokohama, Japan) and a FASTA online DNA homology search system provided by the DNA Data Bank of Japan (DDBJ).

Phylogenetic Tree Construction. Phylogenetic trees were drawn using the programs ClustalX (Thompson et al. 1997) and MEGA (Kumar et al. 1993). All gaps in the sequence alignment were omitted in a pairwise manner. Construction of the trees was performed by the neighbor-joining method, applying the Tamura and Nei distance as a distance estimator. Only transversional replacement were taken into account. However, the distance estimator for the construction of the phylogenetic tree of 16S rRNA was the Kimura two-parameter distance, in which both transitional and transversional replacements were taken into account. The sequence data obtained in this study have been submitted to the DDBJ/EMBL/GenBank databases under accession number AB034704. The accession numbers of the sequences used for sequence comparison and construction of phylogenetic trees are as follows: *Rba. capsulatus,* Z11165; *Rba. sphaeroides,* AJ010302; *Rsp. rubrum,* AF018954; *A. rubrum,* AB017351 and AB005218; and *H. mobilis,* AF080002. Other gene sequences used for the tree construction were also obtained from the DDBJ/EMBL/GenBank databases. Preliminary sequence data for *Rps. palustris* were obtained from the DOE Joint Genome Institute (JGI) at http://spider.jgi-psf.org/JGI_microbial/html/. DNA sequences of the presumed *ppx* and *carA* genes of *Rba. sphaeroides* were obtained from the *Rhodobacter sphaeroides* genome project site at http://www-mmg.med.uth.tmc.edu/sphaeroides/.

Results

Genes Contained in the Photosynthesis Gene Cluster. The insert DNA fragments of the two cosmid clones containing photosynthesis genes of *Rvi. gelatinosus,* pgc#6 and pgc#12, were sequenced. We found 55 open reading frames (ORFs) on this sequence. Based on comparisons between the predicted amino acid sequences and the known sequences of other purple bacteria (Table 1), 12 and 7 of these ORFs were respectively identified as genes coding for biosynthesis enzymes for bacteriochlorophyll *a* and carotenoids.

The five *puf* genes coding for apoproteins of the light-harvesting and the reaction-center complexes and the putative *puhA* gene coding for the H subunit of the reaction center were also detected among these genes. Figure 1 shows the predicted gene arrangement on the approximately 59-kb nucleotide sequence determined in this study. The 17- to 21-kb region from the left side contained the *puf* operon, which has five genes coding for the subunits of the reaction center and the light-harvesting complexes, in addition to two ORFs with unknown functions. The nucleotide sequence of the *puf* operon obtained in this study was consistent with the sequence published previously (Nagashima et al. 1994) except for three errors (a one-base insertion in the ORF48 and two mismatches in the *pufC* gene, respectively).

In the downstream region of the *puf* operon, the carotenoid biosynthesis genes, *crtD, crtC,* and *crtB,* were

Table 1 Amino acid sequence identities (%) of the gene products predicted from *Rvi. gelatinosus* photosynthesis gene cluster to the photosynthesis proteins known in other purple bacteria

Gene product	*Rba. capsulatus*	*Rba. sphaeroides*	*Rsp. rubrum*	*A. rubrum*	*Rps. palustris*
CrtE	50	53 (64)[a]			57 (52)
CrtF	37	40 (54)			49 (41)
BchC	59	57 (73)			64 (59)
BchX	74	78 (80)		76 (72)	82 (74)
BchY	69	70 (79)	71 (70)	65 (63)	75 (70)
BchZ	70	70 (82)	75 (68)	69 (63)	76 (69)
PufB	47	45 (80)	47 (44)	59 (36)	60 (40)
PufA	52	52 (76)	49 (48)	43 (55)	56 (52)
PufL	66	69 (78)	72 (70)	76 (65)	75 (67)
PufM	64	66 (77)	72 (61)	72 (61)	73 (59)
CrtA	31	31 (47)			
CrtD	41	45 (55)			52 (42)
CrtC	39	40 (59)			48 (44)
CrtB	47	47 (65)			49 (49)
ORF276	38				44 (37)
ORF154	25		31 (27)		36 (27)
ORF227	41		47 (36)		39 (36)
PuhA	44	41 (66)	45 (38)		50 (41)
LhaA	51	57 (67)	63 (57)		59 (52)
BchM	53	54 (66)			67 (47)
BchL	68	70 (80)			75 (70)
BchH	58	58 (75)		60 (53)	70 (56)
BchB	65	69 (72)		70 (62)	75 (63)
BchN	65	68 (74)		69 (62)	69 (63)
BchF	69	65 (81)		61 (62)	64 (63)
Ppa	28	26 (43)			48 (26)
PpsR	31	33 (53)			52 (32)
CrtI	49	51 (69)			58 (48)
BchG	56	59 (71)			67 (54)
ORF440	53	53 (67)			59 (53)
BchP	61	60 (74)			66 (59)

[a] Values in parentheses are amino acid sequence identities (%) to the corresponding gene products of *Rba. capsulatus.*

detected, a finding which was consistent with the results for another strain of *Rvi. gelatinosus,* strain S1 (Ouchane et al. 1997b). The amino acid sequences of the products of these three genes were highly conserved (90, 95, and 89% identical, respectively) between the two strains. Extreme overlapping of *crtD* and *crtC* locations (373 bp) was also common to these two strains. However, an ORF that has not been reported in the strain S1 was found here between *pufC* and *crtD.* This ORF was identified as *crtA* coding for a spheroidene monooxygenase, since a mutant lacking this gene did not produce spheroidenone but spheroidene (unpublished data).

The upstream region of the *Rvi. gelatinosus puf* operon contained possible carotenoid biosynthesis genes, *crtE* and *crtF,* and possible bacteriochlorophyll biosynthesis genes, *bchC, bchX, bchY,* and *bchZ,* as in the *Rhodobacter* species. Unlike what was observed in the *Rhodobacter* species, however, the region farther upstream, which was about 10 kb or more from the *crtE,* contained no homologues of photosynthesis genes.

The putative *puhA* gene coding for the H subunit of the reaction center in *Rvi. gelatinosus* was located about 8 kb downstream from *pufC* on the complementary strand to the *puf* genes and was followed by four ORFs. Three of these, ORF276, ORF154, and ORF227, showed considerably high sequence identities to the ORF274, ORF162b, and ORF214 of *Rba. capsulatus,* respectively, at the corresponding positions. However, the genes homologous to the ORF274 and ORF55 found in *Rba. capsulatus* were not detected in *Rvi. gelatinosus* at the corresponding positions. Instead, ORF358, whose product showed 42.7% amino acid sequence identity to PNZIP, a Leu zipper protein of a higher plant (*Pharbitis nil*) chloroplast (Zheng et al. 1998), was located in *Rvi. gelatinosus.* The region upstream of the putative *puhA* contained six possible bacteriochlorophyll biosynthesis genes, *bchF, bchN, bchB, bchH, bchL,* and *bchM,* whose products showed more than 53% amino acid sequence identity to the products of *Rba. capsulatus.* Between *puhA* and *bchM* was located an ORF that was similar to *lhaA* of *Rba. capsulatus* and that has been suggested previously to function in the assembly process of LH1 (Young and Beatty 1998).

A possible repressor gene homologous to *crtJ* of *Rba. capsulatus* and *ppsR* of *Rba. sphaeroides* was found in *Rvi. gelatinosus* in the region upstream of the presumed *bchF* and was tentatively designated *ppsR.* This possible repressor was not labeled *crtJ* to avoid confusion with carotenoid biosynthesis genes. In the *Rhodobacter* species, *crtJ* and *ppsR* have been shown to code for transcriptional repressors of pigment biosynthesis gene expression (Ponnampalam et al. 1995; Penfold and Pemberton 1994; Gomelsky and Kaplan 1995). The possible *ppsR* gene in *Rvi. gelatinosus* was preceded by an ORF coding for 238 amino acid residues showing significant sequence similarities to the products of ORF192 of *Rba. capsulatus* and *ppa* of *Rba. sphaeroides* located at the corresponding position. A gene located in proximity to the possible *ppsR* of *Rvi. gelatinosus* was identified as a carotenoid biosynthesis gene, *crtI.* Mutant cells of *Rvi. gelatinosus* lacking this gene accumulated phytoene molecules, confirming that the product of this gene is phytoene desaturase (unpublished data). In the region upstream of the *crtI,* genes presumed to be the bacteriochlorophyll biosynthesis genes *bchG* and *bchP* were found. An ORF located between these genes and coding 440 amino acids showed considerably high sequence identity to the ORF located between the corresponding *bch* genes in *Rhodobacter* species. No other possible photosynthesis genes were found in the 11-kb region downstream from the presumed *bchP* in *Rvi. gelatinosus.*

Genes Located Outside the Photosynthesis Gene Cluster. Twenty ORFs were found outside the 37-kb photosynthesis gene cluster in the 59-kb DNA fragment. The proteins showing the highest sequence identities to these ORF products are summarized in Table 2. Four ORFs in the region upstream of *crtE* (Fig. 1, left)

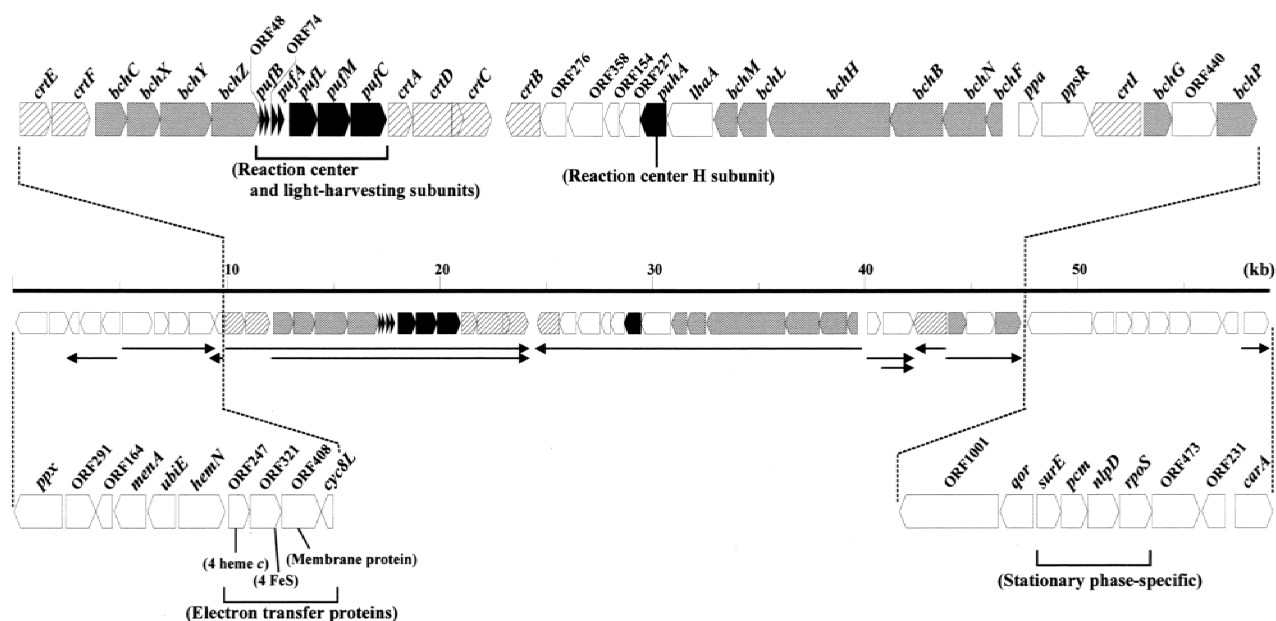

Fig. 1. Arrangement of genes predicted from the nucleotide sequence of the 59-kb DNA fragment of *Rvi. gelatinosus* cloned in this study. The genes are presented as arrows pointing in the direction of their transcriptions. Genes coding for the light-harvesting and the reaction—center apoproteins are indicated by *solid arrows.* Genes assigned to bacteriochlorophyll and carotenoid biosynthesis genes are shown by *gray* and *hatched arrows,* respectively. The presumed operon structures (transcription units) are given by *thin arrows,* based on the locations of the possible promoter sequences shown in Fig. 2.

Table 2 Proteins showing the highest amino acid sequence identities to the deduced products of *Rvi. gelatinosus* genes located outside the photosynthesis gene cluster

Gene/ORF	Best-matched protein found in databases	% Identity
ppx	*Pseudomonas aeruginosa* exopolyphosphatase protein	40.1
ORF291	*Methylobacterium extorquens* hypothetical protein	46.7
ORF164	*Pseudomonas aeruginosa* hypothetical protein	35.0
menA	*Mycobacterium tuberculosis* MenA protein (hypothetical)	30.2
ubiE	*Escherichia coli* ubiquinone/menaquinone methyltransferase	44.0
hemN	*Pseudomonas aeruginosa* oxygen-independent coproporphyrinogen III	31.7
ORF247	None	
ORF321	*Archaeoglobus fulgidus* molybdopterin oxidoreductase	40.9
ORF408	*Escherichia coli* NrfD protein	23.1
cyc8L	*Rubrivivax gelatinosus* ATCC17011 small cytochrome *c*	94.2
ORF1001	None	
qor	*Mycobacterium tuberculosis* fadB4 protein (hypothetical Qor)	38.6
surE	*Legionella pneumophila* survival protein homologue	49.2
pcm	*Escherichia coli* L-isoaspartyl protein carboxyl methyltransferase	41.8
nlpD	*Ralstonia solanacearum* novel lipoprotein NlpD	56.6
rpoS	*Ralstonia solanacearum* alternative RNA σ factor RpoS	52.6
ORF473	*Vibrio* sp. hypothetical protein	40.7
ORF231	*Escherichia coli* ygcA protein (hypothetical)	35.5
carA	*Pseudomonas aeruginsa* carbamoylphosphate synthetase small subunit	66.2

were suggested to be genes coding for exopolyphosphatase (*ppx*), the menaquinone biosynthesis enzyme (*menA*), the ubiquinone/menaquinone biosynthesis enzyme (*ubiE*), and coproporphyrinogen III oxidase (*hemN*), respectively. The region directly upstream of *crtE* contained a gene coding for a soluble cytochrome *c,* designated *cyc8L.* This cytochrome is preferably produced under anaerobic conditions (Menin et al. 1999) and functions as an electron donor to the reaction center (Osyczka et al. 1997). The region downstream of the presumed *hemN* contained three ORFs, ORF247, ORF321, and ORF408, predicted to code for a soluble cytochrome *c,* an FeS-binding protein, and a membrane protein, respectively. The deduced amino acid sequence of the ORF247 product contained four binding motifs to *c*-type hemes, –C–X–X–C–H. No proteins homologous to this cytochrome were found in the databases.

The region downstream of *bchP* contained nine

```
E. coli (σ70)                                  TTGACA      N15-19      TATAAT
Rba. capsulatus    bchC    aaaaagTGTctaatcaaaTTGACAgtcgggcgtgtaagttc-AATGaTAcacacagg (45)
Rba. sphpaeroides  bchC    tggcacTGTccaataaagTTGACActtcacgatgtcccgtt-AATGTTAcacctgaa (235)
Rba. capsulatus    bchF    cctgagTGTaagttttcaTTGACActttctcgtgacaagaccAgTtTTAcggcagag (30)
Rba. capsulatus    crtI    gacagtTGTaaatcggaaTTGACgacctatcatcccccc---AATGcaAcctgaaac (13)
Rba. capsulatus    crtE    cttgggTGTaagtttcagTTtACAcaggtaggtgcgaatgccAATGTgcgtcgtgac (18)
Consensus                                    aaaTTGACA                   AATGttA
                                               -35                         -10

Rvi. gelatinosus   menA    ttaaagTGTcgagtcagccTGACActctcctcgacgaacgcctgcccagaccgatga (831)
                   hemN    ggagagTGTcaggctgacTcGACActttaactggctcgccgaaagcgcgtcgcatgc (99)
                   cyc8L   ccaaagTGTcaactgattTTGACAcgcgacagacctctgtcaagaaagcggaactga (45)
                   crtE    gtcgcgTGTcaaaatcagTTGACActttggggtgacatgacgaccttacagcggtgc (32)
                   crtE    cgcatcaGTtccgctttcTTGACAgaggtctgtcgcgtgtcaaaatcagttgacact (63)
                   bchC    aaggcaTGTccagccaccTTGACAcactgaactgtcagcttaagatgacgtcgtgat (81)
                   bchF    gcatgacGTcaacccctgTTGACAtcaggggtgtccccgtctagattctgtgccatg (-3)
                   ppa     ccctgaTGTcaacaggggTTGACgtcatgcgaacctgacactacccttgagtcgtgc (270)
                   ppa     ggttgacGTcatgcgaaccTGACActacccttgagtcgtgcgggtaccgttcgcccg (254)
                   ppsR    gggaggTGgtctgcgaacTTGACAgccgatcgatagatcccgtgcgccgggtgcgcc (349)
                   crtI    acagagcGTcaattgtcgTTGACAcaatccctgccttaggacaagctcgacggatgc (2)
                   crtI    tggggcTGaatcttgcgcTTGACAaaacagagcgtcaattgtcgttgacacaatccc (28)
                   bchG    ggattgTGTcaacgacaaTTGACgctctgttttgtcaagcgcaagattcagccccat (-2)
                   bchG    gacaatTGacgctctgttTTGtCAagcgcaagattcagccccatggcaaggcccgcg (-15)
                   carA    ccgggtTGTggctcgcggTTGACAtggcgctgcggcacttcatcgccgcattggcgt (162)
Consensus                        TGTcaa       TTGACA
```

Fig. 2. Comparison of possible promoter sequences found in this study and those reported in other purple bacterial photosynthesis genes. All 15 sequences found in the *Rvi. gelatinosus* 59-kb DNA fragment and showing a high similarity (allowing up to two mismatches) to the –35 region of the consensus promoter sequences for the photosynthesis genes reported in the *Rhodobacter* species, TGT–N_9–TTGACA, are shown. The number in parentheses indicates the number of nucleotides between the start codon, ATG, and the right end of the sequence.

ORFs. The region directly downstream of *bchP* was shared by an unknown gene coding for 1001 amino acids and a possible *qor* gene coding for quinone oxidoreductase. The region farther downstream contained four ORFs suggested to be *surE, pcm, nlpD,* and *rpoS* (also known as *katF*). In *E. coli,* the *surE* and *pcm* genes are cotranscribed and have been known to be expressed under conditions adverse to survival (Visick et al. 1998; Li et al. 1997). In *E. coli,* the product of *rpoS* (or *katF*) is known to be a σ factor, σ^s (or σ^{38}), which controls gene expression in response to starvation and when the cell growth reaches a stationary phase (Tanaka et al. 1993). The downstream region of the presumed *rpoS* contained two ORFs and a presumed gene coding for the carbamoylphosphate synthetase small subunit.

Possible Promoter Sequences. In *Rba. capsulatus* and *Rba. sphaeroides,* it has been shown that the expression of photosynthesis genes is highly regulated by environmental factors, i.e., oxygen tension and light intensity (Bauer 1995; Zeilstra-Ryalls et al. 1998). In these species, the regions directly upstream of the bacteriochlorophyll and carotenoid biosynthesis genes have a consensus promoter with a palindromic sequence recognized by a repressor protein, *ppsR* or *crtJ,* working mainly under aerobic conditions (Young et al. 1989; Ma et al. 1993; Gomelsky and Kaplan 1995). *Rvi. gelatinosus* also has a homologue of these genes, suggesting that the photosynthesis gene expression in *Rvi. gelatinosus* is controlled by mechanisms similar to those shown in the two *Rhodobacter* species. As shown in Fig. 2, 15 possible *ppsR* recognition motifs with a –35 σ^{70} recognition motif, TGT–N_9–TTGACA, which was well matched with the consensus sequence shown in the *Rhodobacter* species, were detected within the 59-kb nucleotide sequence of *Rvi. gelatinosus.* The –10 σ^{70} recognition sequence was not apparent in *Rvi. gelatinosus.* Eight of the possible recognition sequences located within a 100-bp region upstream of *crtE, bchC, bchF, crtI,* and *bchG* of *Rvi. gelatinosus,* respectively, were consistent with the locations reported in the *Rhodobacter* species, with the exception of that of *bchG.* A possible consensus promoter sequence for the *puf* operon reported in other purple bacteria, however, was not detected in the *Rvi. gelatinosus* photosynthesis gene cluster, although the accumulation of a large amount of the *puf* mRNA was apparent (Nagashima et al. 1994).

Discussion

Conservation of Photosynthesis Superoperons. As shown in Fig. 3, the gene arrangement in the *Rvi. gelatinosus* photosynthesis gene cluster was markedly different from those of the *Rhodobacter* species. The *puf* and *puh* structural genes for the reaction center of *Rvi. gelatinosus* are located among pigment biosynthesis genes. On the other hand, the local arrangements of some photosynthesis genes are conserved. The *puf* operon of *Rba capsulatus* is cotranscribed with the upstream *crtEF* and *bchCXYZ* genes, although a large amount of *puf* mRNA is transcribed by its own promoter (Young et al. 1989; Wellington et al. 1991). *Rvi. gelatinosus* also has an arrangement of *crtEF–bchCXYZ–puf* that can be transcribed under the same control mechanism as that reported in *Rba. capsulatus.* The detection of the consensus promoter-like sequences for the pigment biosynthesis genes at the regions directly upstream of *crtE* and *bchC* in *Rvi. gelatinosus* suggests that superoperonal structures containing the *puf* operon are conserved. However, the *Rvi. gelatinosus puf*-containing superoperon may include the *crtA, crtD,* and *crtC* genes, since carotenoid biosynthesis was abolished when the transcriptional terminator was inserted into the *puf* operon

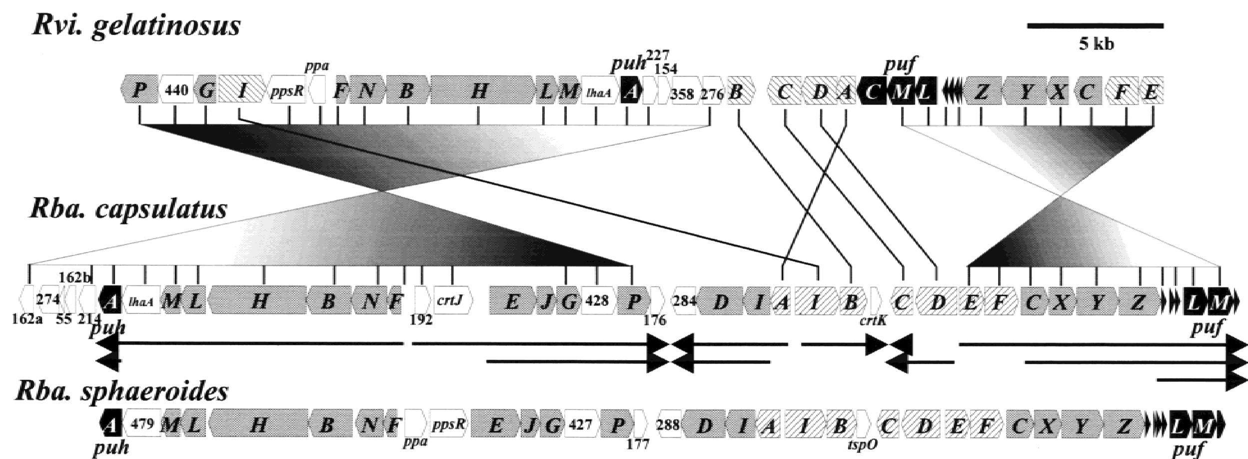

Fig. 3. Comparison of photosynthesis gene arrangements between *Rvi. gelatinosus* and the *Rhodobacter* species. *Arrows* show the directions and extents of the transcripts reported in *Rba. capsulatus.* The symbols showing the photosynthetic apparatus apoprotein genes and bacteriochlorophyll and carotenoid biosynthesis genes are as described in the legend to Fig. 1.

(Ouchane et al. 1997a). The absence of any possible promoter sequences in the upstream regions of these *crt* genes supports this idea.

The superoperonal gene arrangement, *bchFNBHLM–lhaA–puhA,* reported in *Rba. capsulatus* (Bauer et al. 1991; Burke et al. 1993a) was also found in *Rvi. gelatinosus.* The consensus promoter-like sequence for the pigment biosynthesis genes was also found in the region directly upstream of *bchF* in *Rvi. gelatinosus,* as in *Rba. capsulatus* (Alberti et al. 1995) and *Rba. sphaeroides* (Gomelsky and Kaplan 1995). In addition to this superoperonal structure, three of the four ORFs located directly downstream of *puhA* in *Rvi. gelatinosus* were highly homologous to the ORFs at the corresponding positions in *Rba. capsulatus.* These ORFs may be transcribed as a part of the *puh* superoperon and are likely to work in the process of assembly of the photosynthetic apparatus, as suggested in *Rba. capsulatus* (Wong et al. 1996). A subcluster containing *puhA, bchP*–ORF440–*bchG–ppsR–ppa–bchFNBHLM–lhaA–puhA*–ORFs, was also conserved in *Rvi. gelatinosus,* although the locations corresponding to the *bchE* and *bchJ* genes in *Rhodobacter* species were shared by the *crtI* gene in *Rvi. gelatinosus* (Fig. 3).

One explanation for the conservation of these two superoperonal structures containing *puf* and *puh* genes, respectively, is the presence of strong selection pressures. It has been suggested that one of the advantages gained by the clustering of photosynthesis genes is improved control of the synthesis of the photosynthetic apparatus in response to changes in environmental conditions, especially changes in light-intensity and redox conditions (Yildiz et al. 1992; Wellington et al. 1992). At the very least, this hypothesis seems to explain the *puf*- and *puh*-containing superoperonal structures, since *Rvi. gelatinosus* quickly and radically alters the expression levels of the photosynthetic apparatus in response to environmental changes as well.

Horizontal Transfer of the Photosynthesis Gene Cluster. Rvi. gelatinosus is a member of the β subclass of purple bacteria and is phylogenetically distinct from the species belonging to the α subclass (Woese 1987). However, most of the photosynthesis gene products of *Rvi. gelatinosus* showed very high sequence identities to the gene products of *Rps. palustris,* a species belonging to the α2 subgroup of purple bacteria, and these values were considerably higher than the corresponding identities among the α-purple bacteria (Table 1). This phylogenetical inconsistency, as well as the strong conservation of the photosynthesis superoperonal structures, suggests a horizontal transfer of the whole photosynthesis gene cluster, as expected from our previous studies showing the horizontal transfer of the *puf* operon between α-purple bacteria and β- and γ-purple bacteria (Nagashima et al. 1993, 1997).

Figures 4B and C show the phylogenetic trees of the bacteriochlorophyll biosynthesis genes shown in Table 1, *bchNBH* (approx. 6.6 kb) and *bchYZ* (approx. 3.0 kb), based on the available nucleotide sequence data. It has been known that the products of *bchN* and *bchB* form an enzyme complex catalyzing tetrapyrrole reduction, which converts a protochlorophyllide into a chlorophyllide *a.* This reaction is followed by a second reduction to produce a bacteriochlorophyllide *a,* which is catalyzed by an enzyme complex of the products of *bchY* and *bchZ* in *Rba. capsulatus* (Burke et al. 1993a). In addition, the amino acid sequences of the products of *bchNB* show significant similarity to those of *bchYZ,* suggesting that these gene sets have a common origin (Burke et al. 1993b). The sequence data of *bchYZ* have not been reported in species other than purple bacteria; therefore, the *bchNB* sequence of *Rba. capsulatus* was used as an outgroup in the phylogenetic tree of *bchYZ.* The phylogenetic trees of both *bchNBH* and *bchYZ* consistently showed that *Rvi. gelatinosus* is located within the α subclass of purple bacteria, which was consistent with the

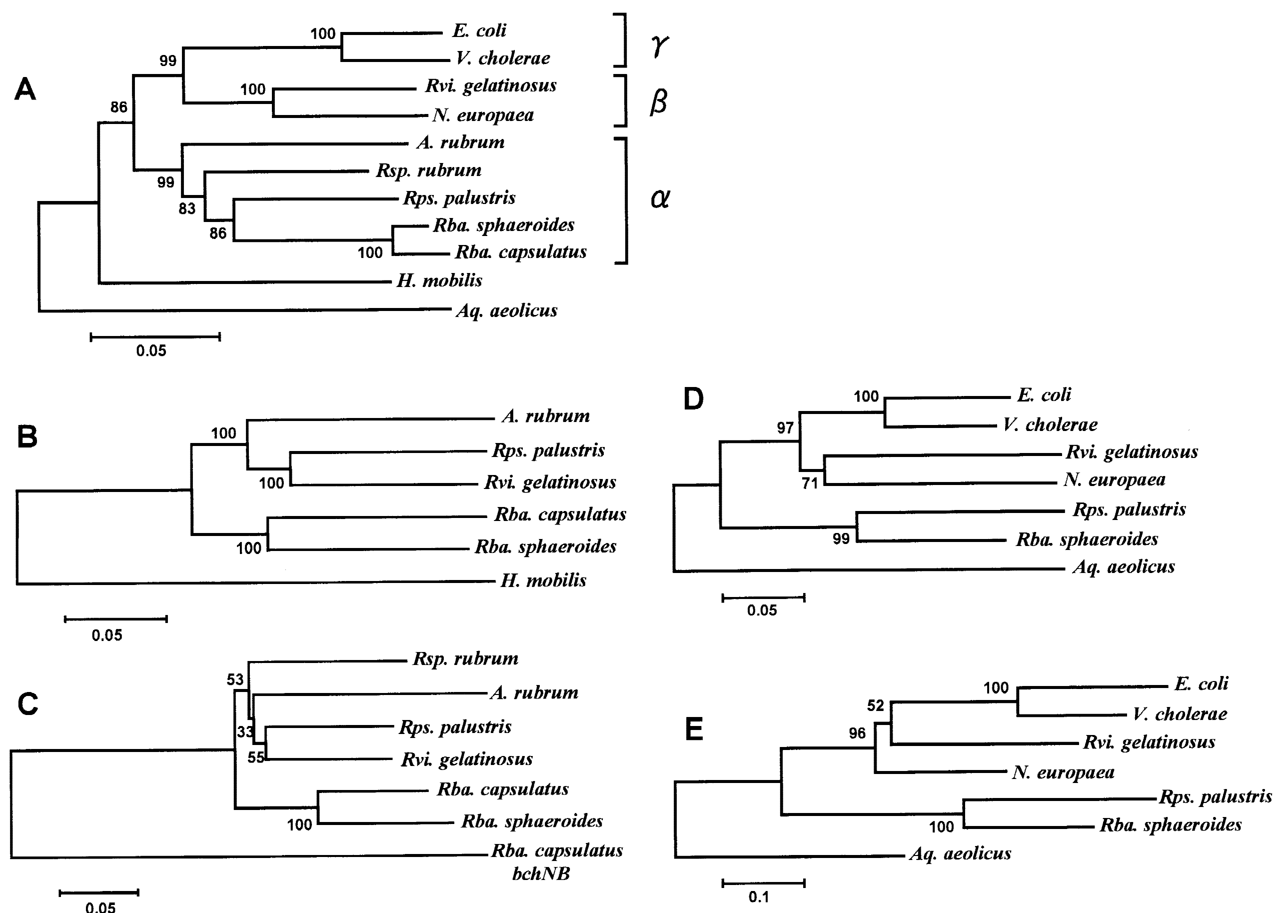

Fig. 4. Phylogenetic trees based on the nucleotide sequences of 16S rRNA (**A**), *bchNBH* (**B**), *bchYZ* (**C**), *carA* (**D**), and *ppx* (**E**). A gram-positive bacterium, *Heliobacillus mobilis,* was used as an outgroup in the phylogenetic tree of *bchNBH.* In the *bchYZ* tree, the *bchN* and *bchB* sequences of *Rba. capsulatus* were used as an outgroup, since these genes have slight sequence similarity to the *bchY* and *bchZ* genes, respectively. In other trees a nonphotosynthetic species, *Aquifex aeolicus,* was used as an outgroup. Other nonphotosynthetic species, *Escherichia coli, Vibrio cholerae,* and *Nitrosomonas europaea,* were also used in the trees of 16S rRNA, *carA,* and *ppx.* The obtained bootstrap values are presented at the corresponding nodes. The trees were constructed based on the neighbor-joining method applying the Kimura two-parameter distance for 16S rRNA and the Tamura and Nei distance for the others as distance estimators. Other necessary information for the tree construction is given under Materials and Methods.

phylogenetic relationships based on the nucleotide sequences of *puf* genes. Phylogenetic trees based on the deduced amino acid sequences of most bacteriochlorophyll and carotenoid biosynthesis enzymes and other related photosynthesis gene products listed in Table 1 consistently support this tree topology (data not shown). However, phylogenetic trees based on the nucleotide sequences of genes outside the photosynthesis gene cluster, *ppx* and *carA,* showed that *Rvi. gelatinosus* is distinct from the α-purple bacteria and is closely related to other β- and γ-purple bacteria, which is consistent with the relationships shown in the phylogenetic tree of 16S rRNA (Figs. 4A, D, and E). These data indicate that horizontal transfer of the photosynthesis gene cluster occurred between ancestral species of *Rps. palustris* and *Rvi. gelatinosus.*

As shown in Fig. 3, the remarkable difference between the structures of the photosynthesis gene clusters of *Rvi. gelatinosus* and those of other species belonging to the α subclass can be attributed mainly to the locations of two subclusters, one containing the *puf* genes and the other containing the *puh* gene. It is possible that a donor species of the horizontal gene transfer already had the two proximal superoperonal structures, *crtEF–bchCXYZ–puf* and *bchFNBHLM–lhaA–puhA,* as well as some of the regulatory and structural genes, i.e., *ppa, ppsR,* and some other *bch* genes. One possible explanation for the evolutionary development of the photosynthesis gene cluster is that the donor species had the *Rhodobacter*-like gene cluster and that the horizontal gene transfer from this donor, possibly an ancestral species of the α2 subgroup, to the ancestral species of the β subclass was followed by rearrangement of the two superoperonal structures, i.e., by two inversions and/or by translocations in *Rvi. gelatinosus,* as shown in Fig. 3. If this is true, the other remarkable difference—i.e., that the four essential bacteriochlorophyll biosynthesis genes, *bchE, bchJ, bchI,* and *bchD,* were absent in the photosynthesis gene cluster of *Rvi. gelatinosus*—might have been due to translocations of these genes away from the

cluster in *Rvi. gelatinosus* after the horizontal transfer of the whole photosynthesis gene cluster.

The Hypothetical Gene Cluster for Photosynthetic Electron Transfer. Recently, it was shown that a photosynthetic bacterium of gram-positive lineage, *Heliobacillus mobilis,* also has a photosynthesis gene cluster, and that genes coding for the subunits of the cytochrome *bc* complex, the *pet* genes, are placed in proximity to this cluster (Xiong et al. 1998). This gene arrangement seems to be reasonable since the photosynthetic reaction center and the cytochrome *bc* complex form a close linkage of electron transfer. *Rvi. gelatinosus* also has ORFs for some presumed electron transfer proteins, such as ORF247, ORF321, and ORF408, and for possible biosynthesis enzymes for hemes and quinones at the flanking region of the photosynthesis gene cluster. The presence of a consensus promoter-like sequence for pigment biosynthesis gene expression in the region directly upstream of these genes (Fig. 2) suggests that the hypothetical electron transfer proteins of *Rvi. gelatinosus* are synthesized with photosynthetic apparatuses under anaerobic conditions and are involved in the light-driven electron transfer pathways. This may be supported by the facts that the consensus promoter-like sequence for expression of the pigment biosynthesis gene is located in the region directly upstream of *cyc8L* and that the synthesis of the product, a soluble cytochrome c_8 with a low midpoint potential, has been shown to be enhanced under photosynthetic conditions (Menin et al. 1999).

In summary, we have provided evidence here of the horizontal transfer of a large gene cluster for photosynthesis. Sequence analyses of the whole genome of *E. coli* suggested that 755 of 4288 ORFs in *E. coli* have been obtained through at least 234 horizontal gene transfer events after the divergence from the *Salmonella* lineage (Lawrence and Ochman 1998). Some of these gene transfer events may have contributed new phenotypes to the host species. In the present study, it was shown that even a phenotype as advanced as that of photosynthesis can be transferred between distantly related species.

Acknowledgments. Preliminary sequence data for *Rps. palustris* were obtained from the DOE Joint Genome Institute (JGI) at http://spider.jgi-psf.org/JGI_microbial/html/. This work was supported in part by grants from the Ministry of Education, Science, and Culture of Japan and funds for a Special Research Project (1999) at Tokyo Metropolitan University.

References

Alberti M, Burke DH, Hearst JE (1995) Structure and sequence of the photosynthesis gene cluster. In: Blankenship RE, Madigan MT, Bauer CE (eds) Anoxygenic photosynthetic bacteria. Kluwer Academic, Dordrecht, The Netherlands, pp 1083–1106

Bauer CE (1995) Regulation of photosynthesis gene expression. In: Blankenship RE, Madigan MT, Bauer CE (eds) Anoxygenic photosynthetic bacteria. Kluwer Academic, Dordrecht, The Netherlands, pp 1221–1234

Bauer CE, Buggy JJ, Yang Z, Marrs BL (1991) The superoperonal organization of genes for pigment biosynthesis and reaction center proteins is a conserved feature in *Rhodobacter capsulatus:* Analysis of overlapping *bchB* and *puhA* transcripts. Mol Gen Genet 228: 433–444

Beatty JT (1995) Organization of photosynthesis gene transcripts. In: Blankenship RE, Madigan MT, Bauer CE (eds) Anoxygenic photosynthetic bacteria. Kluwer Academic, Dordrecht, The Netherlands, pp 1209–1219

Berard J, Gingras G (1991) The *puh* structural gene coding for the H subunit of the *Rhodospirillum rubrum* photoreaction center. Biochem Cell Biol 69:122–131

Burke DH, Alberti M, Hearst JE (1993a) *bchFNBH* bacteriochlorophyll synthesis genes of *Rhodobacter capsulatus* and identification of the third subunit of light-independent protochlorophyllide reductase in bacteria and plants. J Bacteriol 175:2414–2422

Burke DH, Hearst JE, Sidow A (1993b) Early evolution of photosynthesis: Clues from nitrogenase and chlorophyll iron proteins. Proc Natl Acad Sci USA 90:7134–7138

Coomber SA, Chaudhri M, Connor A, Britton G, Hunter CN (1990) Localized transposon Tn5 mutagenesis of the photosynthetic gene cluster of *Rhodobacter sphaeroides.* Mol Microbiol 4:977–989

Gomelsky M, Kaplan S (1995) Genetic evidence that PpsR from *Rhodobacter sphaeroides* 2.4.1 functions as a repressor of *puc* and *bchF* expression. J Bacteriol 177:1634–1637

Igarashi N, Shimada K, Matsuura K, Nagashima KVP (1998) Photosynthetic gene cluster in purple bacterium, *Rubrivivax gelatinosus.* In: Garab G (ed) Photosynthesis: Mechanisms and effects, Vol IV. Kluwer Academic, Dordrecht, The Netherlands, pp 2889–2892

Kumar S, Tamura K, Nei M (1993) MEGA: Molecular evolutionary genetics analysis, version 1.0. The Pennsylvania State University, University Park 16802

Lawrence JG, Ochman H (1998) Molecular archaeology of the *Escherichia coli* genome. Proc Natl Acad Sci USA 95:9413–9417

Li C, Wu P-Y, Hsieh M (1997) Growth-phase-dependent transcriptional regulation of the *pcm* and *surE* genes required for stationary-phase survival of *Escherichia coli.* Microbiology 143:3513–3520

Ma D, Cook DN, O'Brien CA, Hearst JE (1993) Analysis of the promoter and regulatory sequences of an oxygen-regulated *bch* operon in *Rhodobacter capsulatus* by site-directed mutagenesis. J Bacteriol 175:2037–2045

Masuda T, Inoue K, Masuda M, Nagayama M, Tamaki A, Ohta H, Shimada H, Takamiya K (1999) Magnesium insertion by magnesium cheletase in the biosynthesis of zinc bacteriochlorophyll *a* in an aerobic acidophilic bacterium *Acidiphilium rubrum.* J Biol Chem 274:33594–33600

Menin L, Yoshida M, Jaquinod M, Nagashima KVP, Matsuura K, Parot P, Verméglio A (1999) Dark aerobic growth conditions induce the synthesis of a high potential cytochrome c_8 in the photosynthetic bacterium *Rubrivivax gelatinosus.* Biochemistry 38:15238–15244

Nagashima KVP, Shimada K, Matsuura K (1993) Phylogenetic analysis of photosynthetic genes of *Rhodocyclus gelatinosus:* Possibility of horizontal gene transfer in purple bacteria. Photosynth Res 36: 185–191

Nagashima KVP, Matsuura K, Ohyama S, Shimada K (1994) Primary structure and transcription of genes encoding B870 and photosynthetic reaction center apoproteins from *Rubrivivax gelatinosus.* J Biol Chem 269:2477–2484

Nagashima KVP, Shimada K, Matsuura K (1995) Effects of inactivation of genes coding for the reaction center-bound cytochrome subunit on growth and electron transfer in purple photosynthetic bacterium, *Rubrivivax gelatinosus.* In: Mathis P (ed) Photosynthesis: From light to biosphere, Vol I. Kluwer Academic, Dordrecht, The Netherlands, pp 599–602

Nagashima KVP, Shimada K, Matsuura K (1996) Shortcut of the photosynthetic electron transfer in a mutant lacking the reaction center-

bound cytochrome subunit by gene disruption in a purple bacterium, *Rubrivivax gelatinosus.* FEBS Lett 385:209–213

Nagashima KVP, Hiraishi A, Shimada K, Matsuura K (1997) Horizontal transfer of genes coding for the photosynthetic reaction centers of purple bacteria. J Mol Evol 45:131–136

Naylor GW, Addlesee HA, Gibson LCD, Hunter CN (1999) The photosynthesis gene cluster of *Rhodobacter sphaeroides.* Photosynth Res 62:121–139

Osyczka A, Yoshida M, Nagashima KVP, Shimada K, Matsuura K (1997) Electron transfer from high-potential iron-sulfur protein and low-potential cytochrome *c*-551 to the primary donor of *Rubrivivax gelatinosus* reaction center mutationally devoid of the bound cytochrome subunit. Biochim Biophys Acta 1321:93–97

Ouchane S, Picaud M, Vernotte C, Reiss-Husson F, Astier C (1997a) Pleiotropic effects of *puf* interposon mutagenesis on carotenoid biosynthesis in *Rubrivivax gelatinosus.* J Biol Chem 272:1670–1676

Ouchane S, Picaud M, Vernotte C, Astier C (1997b) Photooxidative stress stimulates illegitimate recombination and mutability in carotenoid-less mutants of *Rubrivivax gelatinosus.* EMBO J 16:4777–4787

Penfold RJ, Pemberton JM (1994) Sequencing, chromosomal inactivation, and functional expression in *Escherichia coli* of *ppsR,* a gene which represses carotenoid and bacteriochlorophyll synthesis in *Rhodobacter sphaeroides.* J Bacteriol 176:2869–2876

Ponnampalam SN, Buggy JJ, Bauer CE (1995) Characterization of an aerobic repressor that coordinately regulates bacteriochlorophyll, carotenoid, and light harvesting-II expression in *Rhodobacter capsulatus.* J Bacteriol 177:2990–2997

Tanaka K, Takayanagi Y, Fujita N, Ishihama A, Takahashi H (1993) Heterogeneity of the principal σ factor in *Escherichia coli:* The *rpoS* gene product, σ^{38}, is a second principal σ factor of RNA polymerase in stationary-phase *Escherichia coli.* Proc Natl Acad Sci USA 90:3511–3515

Taylor DP, Cohen SN, Clark WG, Marrs BL (1983) Alignment of genetic and restriction maps of the photosynthesis region of the *Rhodopseudomonas capsulata* chromosome by a conjugation-mediated marker rescue technique. J Bacteriol 154:580–590

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25:4876–4882

Visick JE, Cai H, Clarke S (1998) The L-isoaspartyl protein repair methyltransferase enhances survival of aging *Escherichia coli* subjected to secondary environmental stresses. J Bacteriol 180:2623–2629

Wellington CL, Taggart AKP, Beatty JT (1991) Functional significance of overlapping transcripts of *crtEF, bchCA,* and *puf* photosynthesis gene operons in *Rhodobacter capsulatus.* J Bacteriol 173:2954–2961

Wellington CL, Bauer CE, Beatty JT (1992) Photosynthesis gene superoperons in purple nonsulfur bacteria: The tip of the iceberg? Can J Microbiol 38:20–27

Woese CR (1987) Bacterial evolution. Microbiol Rev 51:221–271

Wong DK-H, Collins WJ, Harmer A, Lilburn TG, Beatty JT (1996) Directed mutagenesis of the *Rhodobacter capsulatus puhA* gene and Orf 214: Pleiotropic effects on photosynthetic reaction center and light-harvesting 1 complexes. J Bacteriol 178:2334–2342

Xiong J, Inoue K, Bauer CE (1998) Tracking molecular evolution of photosynthesis by characterization of a major photosynthesis gene cluster from *Heliobacillus mobilis.* Proc Natl Acad Sci USA 95:14851–14856

Yildiz FH, Gest H, Bauer CE (1992) Conservation of the photosynthesis gene cluster in *Rhodospirillum centenum.* Mol Microbiol 6:2683–2691

Young C, Beatty JT (1998) Topological model of the *Rhodobacter capsulatus* light-harvesting complex I assembly protein LhaA (previously known as ORF1696). J Bacteriol 180:4742–4745

Young DA, Bauer CE, Williams JC, Marrs BL (1989) Genetic evidence for superoperonal organization of genes for photosynthetic pigments and pigments binding proteins in *Rhodobacter capsulatus.* Mol Gen Genet 218:1–12

Youvan DC, Bylina EJ, Alberti M, Begusch H, Hearst JE (1984) Nucleotide and deduced polypeptide sequences of the photosynthetic reaction-center, B870 antenna, and flanking polypeptides from *R. capsulata.* Cell 37:949–957

Zeilstra-Ryalls J, Gomelsky M, Eraso JM, Yeliseev A, O'Gara J, Kaplan S (1998) Control of photosystem formation in *Rhodobacter sphaeroides.* J Bacteriol 180:2801–2809

Zheng CC, Porat R, Lu P, O'Neill SD (1998) *PNZIP* is a novel mesophyll-specific cDNA that is regulated by phytochrome and the circadian rhythm and encodes a protein with a leucine zipper motif. Plant Physiol 116:27–35