

RESEARCH ARTICLE

Open Access

# Complete genome sequence of the filamentous anoxygenic phototrophic bacterium *Chloroflexus aurantiacus*

Kuo-Hsiang Tang<sup>1</sup>, Kerrie Barry<sup>2</sup>, Olga Chertkov<sup>3</sup>, Eileen Dalin<sup>2</sup>, Cliff S Han<sup>3</sup>, Loren J Hauser<sup>4</sup>, Barbara M Honchak<sup>1</sup>, Lauren E Karbach<sup>1,7</sup>, Miriam L Land<sup>4</sup>, Alla Lapidus<sup>5</sup>, Frank W Larimer<sup>4</sup>, Natalia Mikhailova<sup>5</sup>, Samuel Pitluck<sup>2</sup>, Beverly K Pierson<sup>6</sup> and Robert E Blankenship<sup>1\*</sup>

## Abstract

**Background:** *Chloroflexus aurantiacus* is a thermophilic filamentous anoxygenic phototrophic (FAP) bacterium, and can grow phototrophically under anaerobic conditions or chemotrophically under aerobic and dark conditions. According to 16S rRNA analysis, *Chloroflexi* species are the earliest branching bacteria capable of photosynthesis, and *Cfl. aurantiacus* has been long regarded as a key organism to resolve the obscurity of the origin and early evolution of photosynthesis. *Cfl. aurantiacus* contains a chimeric photosystem that comprises some characters of green sulfur bacteria and purple photosynthetic bacteria, and also has some unique electron transport proteins compared to other photosynthetic bacteria.

**Methods:** The complete genomic sequence of *Cfl. aurantiacus* has been determined, analyzed and compared to the genomes of other photosynthetic bacteria.

**Results:** Abundant genomic evidence suggests that there have been numerous gene adaptations/replacements in *Cfl. aurantiacus* to facilitate life under both anaerobic and aerobic conditions, including duplicate genes and gene clusters for the alternative complex III (ACIII), auracyanin and NADH:quinone oxidoreductase; and several aerobic/anaerobic enzyme pairs in central carbon metabolism and tetrapyrroles and nucleic acids biosynthesis. Overall, genomic information is consistent with a high tolerance for oxygen that has been reported in the growth of *Cfl. aurantiacus*. Genes for the chimeric photosystem, photosynthetic electron transport chain, the 3-hydroxypropionate autotrophic carbon fixation cycle, CO<sub>2</sub>-anaplerotic pathways, glyoxylate cycle, and sulfur reduction pathway are present. The central carbon metabolism and sulfur assimilation pathways in *Cfl. aurantiacus* are discussed. Some features of the *Cfl. aurantiacus* genome are compared with those of the *Roseiflexus castenholzii* genome. *Roseiflexus castenholzii* is a recently characterized FAP bacterium and phylogenetically closely related to *Cfl. aurantiacus*. According to previous reports and the genomic information, perspectives of *Cfl. aurantiacus* in the evolution of photosynthesis are also discussed.

**Conclusions:** The genomic analyses presented in this report, along with previous physiological, ecological and biochemical studies, indicate that the anoxygenic phototroph *Cfl. aurantiacus* has many interesting and certain unique features in its metabolic pathways. The complete genome may also shed light on possible evolutionary connections of photosynthesis.

\* Correspondence: Blankenship@wustl.edu

<sup>1</sup>Department of Biology and Department of Chemistry, Campus Box 1137, Washington University in St. Louis, St. Louis, MO 63130, USA  
Full list of author information is available at the end of the article

## Background

The thermophilic bacterium *Chloroflexus aurantiacus* was the first filamentous anoxygenic phototrophic (FAP) bacterium (also known as the green non-sulfur bacterium or green gliding bacterium) to be discovered [1]. The type strain *Cfl. aurantiacus* J-10-fl was found in a microbial mat together with cyanobacteria when isolated from a hot spring near Sokokura, Hakone district, Japan. *Cfl. aurantiacus* can grow phototrophically under anaerobic conditions or chemotrophically under aerobic and dark conditions.

The photosystem of *Cfl. aurantiacus* includes the peripheral antenna complex known as a chlorosome, the B808-866 light-harvesting core complex, and a quinone-type (or type-II) reaction center [2,3]. While *Cfl. aurantiacus* primarily consumes organic carbon sources (i.e. acetate, lactate, propionate, and butyrate) that are released by the associated cyanobacteria in the *Chloroflexus*/cyanobacterial mats of its natural habitat, it can also assimilate CO<sub>2</sub> with the 3-hydroxypropionate (3HOP) autotrophic carbon fixation cycle [4,5]. Further, studies have reported carbon, nitrogen and sulfur metabolisms of *Cfl. aurantiacus* [1].

According to 16S rRNA analysis, *Chloroflexi* species are the earliest branching bacteria capable of photosynthesis [6-8] (Figure 1) and have long been considered to be critical to understanding the evolution of photosynthesis [9-16]. However, there are also indications that there has been widespread horizontal gene transfer of photosynthesis genes, so the evolutionary history of photosynthesis is still poorly understood [17].

During the transition from an anaerobic to an aerobic world, organisms needed to adapt to the aerobic environment and to become more oxygen-tolerant. Most of the gene products can function with or without oxygen, whereas several proteins and enzymes are known to be exclusively functional in either aerobic or anaerobic environments. Thus, gene replacements have been found in the evolution of many metabolic processes [18-20]. Some aspects of the genome annotation of *Chloroflexi* species have been discussed by Bryant, Ward and coworkers [5,21].

Several genes encoding aerobic and anaerobic enzyme pairs, as well as a number of duplicated gene clusters, have been identified in the *Cfl. aurantiacus* genome. In this report, we use genomic annotation, together with previous physiological and biochemical studies, to illustrate how *Cfl. aurantiacus* may be a good model system for understanding the evolution of metabolism during the transition from anaerobic to aerobic conditions. Some of the genomic information is compared with that of the genome of *Roseiflexus castenholzii*, a recently characterized FAP bacterium that lacks chlorosomes [22].

## Results and Discussion

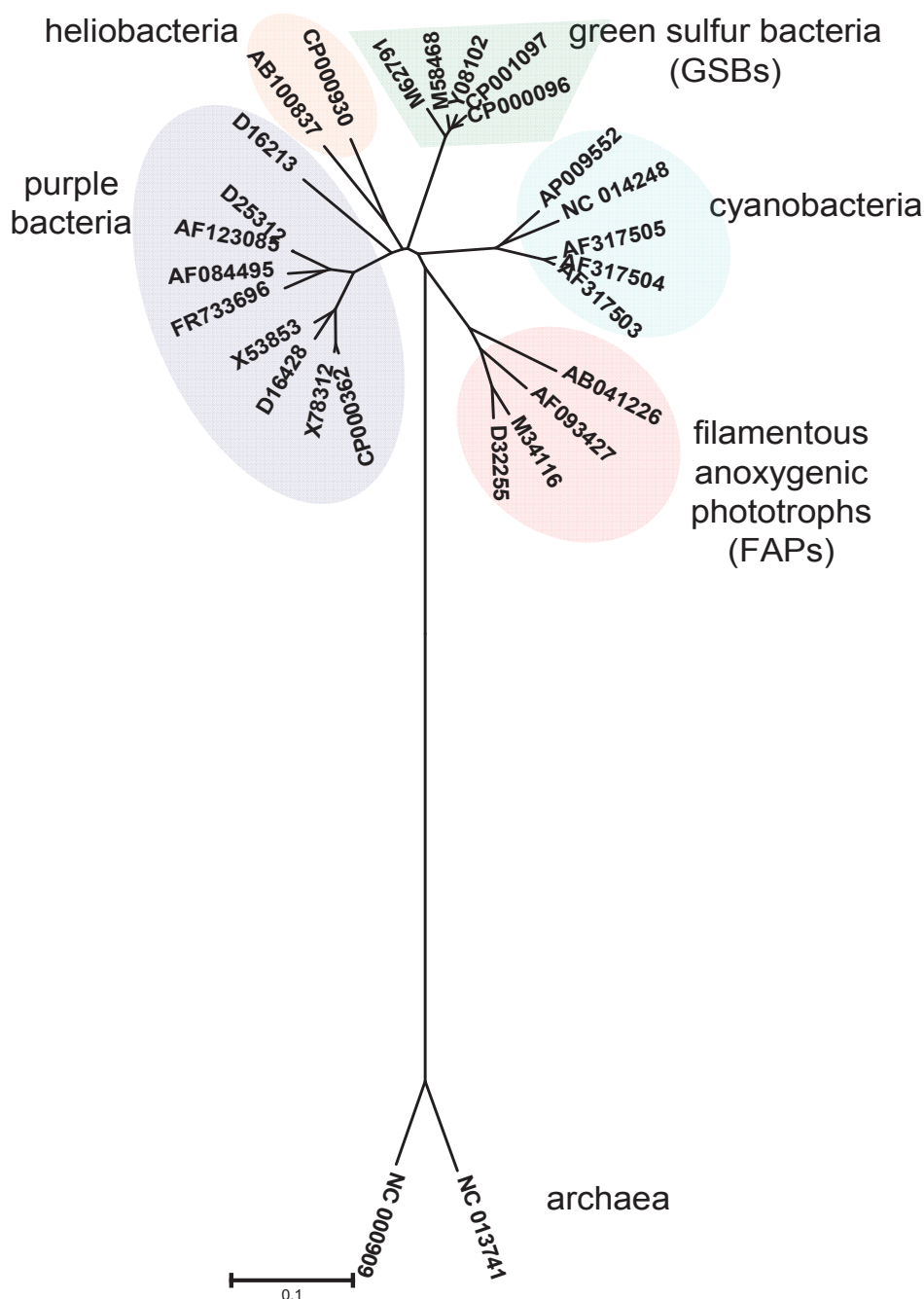
### Genome properties

The genome size of *Cfl. aurantiacus* J-10-fl (5.3-Mb) (Table 1 and Figure 2) is comparable to that of other phototrophic *Chloroflexi* species: *Chloroflexus* sp. Y-400-fl (5.3-Mb), *Chloroflexus aggregans* (4.7-Mb), *Roseiflexus* sp. RS-1 (5.8-Mb), and *Roseiflexus castenholzii* DSM 13941 (5.7-Mb). Here, we summarize several unique features in the *Cfl. aurantiacus* genome, and compare some of the features with other *Chloroflexi* species and various photosynthetic and non-photosynthetic microorganisms. The complete genome has been deposited in GenBank with accession number CP000909 (RefSeq entry NC\_010175). Further information is available at the Integrated Microbial Genome database ([http://img.jgi.doe.gov/cgi-bin/pub/main.cgi?section=TaxonDetail&page=taxonDetail&taxon\\_oid=641228485](http://img.jgi.doe.gov/cgi-bin/pub/main.cgi?section=TaxonDetail&page=taxonDetail&taxon_oid=641228485)). The *oriC* origin is at twelve o'clock of the circular genome map (Figure 2). Like in many prokaryotes, AT-rich repeated sequence can be found in the origin of replication.

No special replication patterns can be found in the genome, and genes responsible for DNA replication and repair do not form a cluster (e.g., *dnaA* (Caur\_0001), *dnaB* (Caur\_0951), genes encoding DNA polymerase III (Caur\_0523, Caur\_0987, Caur\_1016, Caur\_1621, Caur\_1930, Caur\_2419, Caur\_2639, Caur\_2725 and Caur\_3069), *polA* (Caur\_0341), genes encoding other DNA polymerases (Caur\_1575, Caur\_1899, Caur\_2099, Caur\_2101, Caur\_2868, Caur\_3077, Caur\_3225, Caur\_3495, Caur\_3509 and Caur\_3510), *recD* (Caur\_0824), *recF* (Caur\_3876), *recG* (Caur\_0261), *recQ* (Caur\_2049), *rho* (Caur\_0274), *gyrA* (Caur\_1241), *gyrB* (Caur\_3009), *uvrD* (Caur\_0724) and others).

### A. Photosynthetic antenna and reaction center genes

*Cfl. aurantiacus* has chimeric photosynthetic components, which contain characteristics of green sulfur bacteria (e.g., the chlorosomes) and purple photosynthetic bacteria (e.g., the integral-membrane antenna core complex surrounding a type II (quinone-type) reaction center). As the first FAP bacterium to be discovered, the excitation energy transfer and electron transfer processes in *Cfl. aurantiacus* have been investigated extensively [3]. During phototrophic growth of *Cfl. aurantiacus*, the light energy is first absorbed by its peripheral light-harvesting antenna, the chlorosome, which is a self-assembled bacteriochlorophyll complex (the major bacteriochlorophyll in chlorosomes is bacteriochlorophyll *c* (BChl *c*)) and encapsulated by a lipid monolayer. Energy is then transferred to the B808-866 light-harvesting core antenna complex, which is a protein-pigment complex associated with two spectral types of bacteriochlorophyll *a* (BChl *a*) (B808 and B866),



**Figure 1 Phylogenetic tree of photosynthetic bacteria.** The tree was constructed using the phylogenetic software MEGA4.1 with un-rooted neighbor joining 16S rRNA dendrogram from five phyla of photosynthetic microbes, including cyanobacteria, heliobacteria, purple bacteria, green sulfur bacteria and filamentous anoxygenic phototrophs (FAPs) (each phylum of bacteria highlighted in different color). Bacterial names and accession numbers of 16S rRNA genes are listed as follows: (1) purple bacteria: *Roseobacter denitrificans* OCh114 (CP000362), *Roseobacter litoralis* (X78312), *Rhodobacter capsulatus* (D16428), *Rhodobacter sphaeroides* 2.4.1 (X53853), *Rhodopseudomonas faecalis* strain gc (AF123085), *Rhodopseudomonas palustris* (D25312), *Rhodopseudomonas acidophila* (FR733696), *Rhodopseudomonas viridis* DSM 133 (AF084495), *Rubrivivax gelatinosus* (D16213); (2) heliobacteria: *Heliobacterium gestii* (AB100837), *Heliobacterium modesticaldum* (CP000930); (3) cyanobacteria: *Oscillatoria amphigranulata* strain 19-2 (AF317504), *Oscillatoria amphigranulata* strain 11-3 (AF317503), *Oscillatoria amphigranulata* strain 23-3 (AF317505), *Microcystis aeruginosa* NIES-843 (AP009552), *Nostoc azollae* 0708 (NC\_014248); (4) green sulfur bacteria: *Chlorobaculum thiosulfatophilum* DSM 249 (Y08102), *Pelodictyon luteolum* DSM 273 (CP000096), *Chlorobium limicola* DSM 245 (CP001097), *Chlorobaculum tepidum* TLS (M58468), *Chlorobium vibrioforme* DSM 260 (M62791); and (5) FAPs: *Chloroflexus aurantiacus* J-10-fl (M34116), *Chloroflexus aggregans* (D32255), *Oscillochloris trichoides* (AF093427), *Roseiflexus castenholzii* DSM 13941 (AB041226). Archaea (*Archaeoglobus profundus* DSM 5631 (NC\_013741) and *Methanocaldococcus jannaschii* DSM 2661 (NC\_000909)) were used as an out-group.

**Table 1 Organism information and genome statistics of *Chloroflexus aurantiacus* J-10-fl**

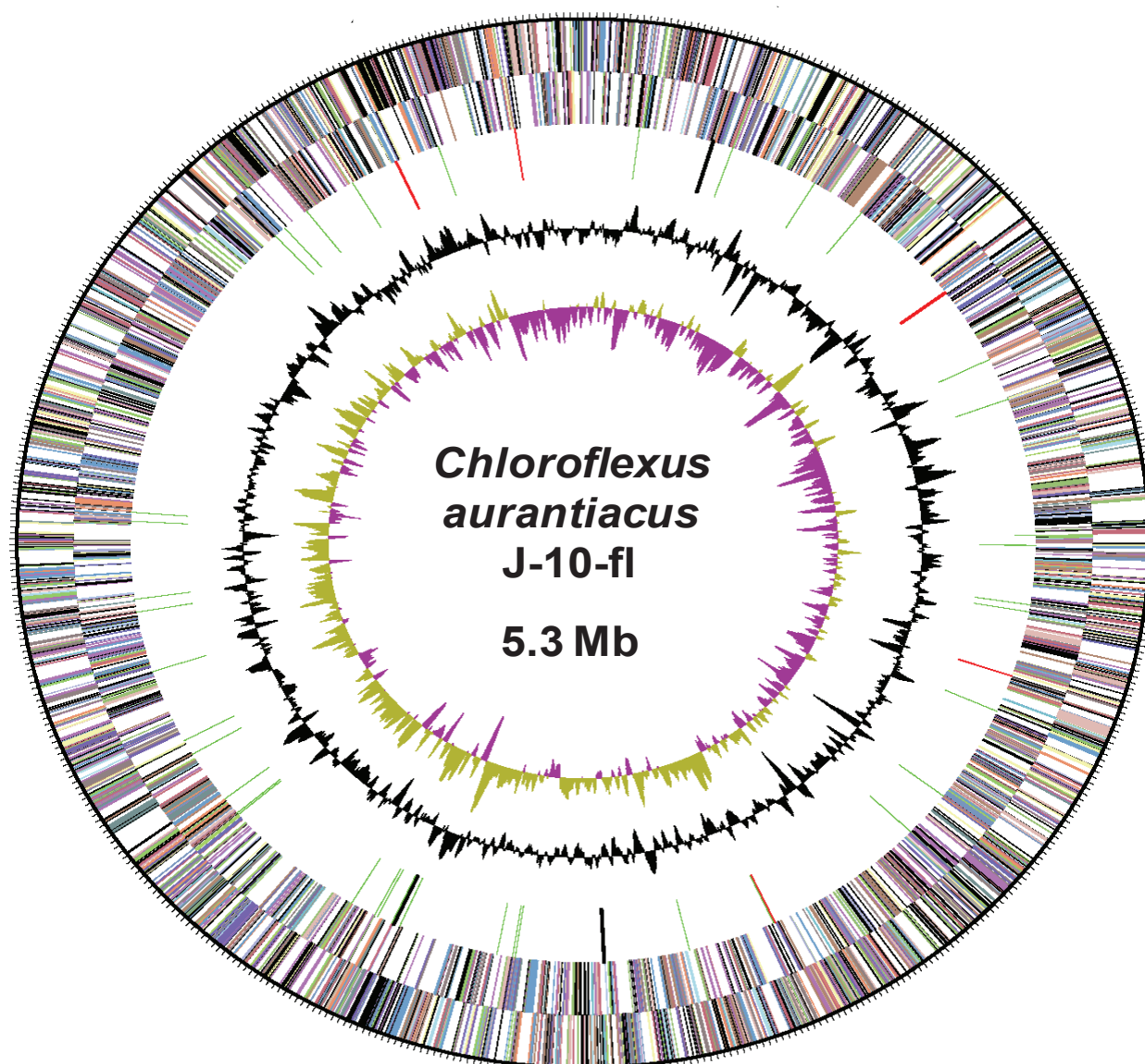
Organism Information		
Organism name	Chloroflexus aurantiacus J-10-fl	
External links	NCBI/RefSeq:NC_010175	
Lineage	Bacteria; Chloroflexi; Chloroflexi; Chloroflexales; Chloroflexaceae; Chloroflexus; aurantiacus	
Sequencing status	Finished	
Sequencing center	DOE Joint Genome Institute	
Oxygen requirement	Anaerobes	
Isolation	Hakone hot spring area in Japan	
Habitat	Fresh water, hot spring	
Motility	Mobile	
Temperature range (temperature optimum)	Thermophile (52-60°C)	
Cell shape	Filament-shaped	
Cell arrangement	filaments	
Gram staining	Gram-negative	
Phenotype	Green non-sulfur	
Energy source	Light (phototrophic growth) and organic carbon sources (chemotrophic growth)	
carbon assimilation	Photoautotrophy, photoheterotrophy, and chemoheterotrophy	
Genome Statistics		
	Number	Percentage of total genes or base pairs
DNA, total number of bases	5258541	100.00%
DNA G + C number of bases	2981443	56.70%
Total number of genes	3914	100.00%
Protein coding genes	3853	98.44%
- with function prediction	2845	72.69%
- without function prediction with similarity	1004	25.65%
- without function prediction without similarity	4	0.10%
Genes coding enzymes	934	23.86%
Genes coding fusion proteins	331	8.46%
Genes coding signal peptides	653	16.68%
Genes coding transmembrane proteins	789	20.16%
Pseudo genes	0	0.00%
RNA genes	61	1.56%
rRNA genes	9	0.23%
tRNA genes	49	1.25%

through the baseplate of chlorosomes. The baseplate is a CsmA chlorosome protein-bacteriochlorophyll *a* (BChl *a*)-carotenoid complex (i.e. a protein-pigment complex) [23]. Finally, the excitons are transferred to the reaction center (RC), in which photochemical events occur. While both purple photosynthetic proteobacteria and *Cfl. aurantiacus* have a type II RC [24], the *Cfl. aurantiacus* RC is simpler than the purple bacterial RC [2] and contains only the L- and M-subunits (PufL and PufM), and not the H-subunit [25,26].

All of the genes encoding the B808-866 core complex ( $\alpha$ -subunit (Caur\_2090) and  $\beta$ -subunit (Caur\_2091)) and RC (*pufM* (Caur\_1051) and *pufL* (Caur\_1052)) are present in the *Cfl. aurantiacus* genome. The *pufL* and

*pufM* genes are fused in the *Roseiflexus castenholzii* genome. The arrangement of genes for the core structural proteins of the photosynthetic complexes is significantly different from that found in purple bacteria, where the *puf* (photosynthetic unit fixed) operon invariably contains the LH complex genes, the RC genes encoding for the L and M subunits and the tetraheme cytochrome associated with the reaction center (if present) [27].

Although proteins are not required for BChl *c* self-assemblies in chlorosomes, various proteins have been identified to be associated with the lipid monolayer of the *Cfl. aurantiacus* chlorosomes [3]. In addition to the baseplate protein CsmA, the chlorosome proteins CsmM and CsmN have been characterized [3], and



**Figure 2 Circular genome map of the 5.2-Mb *Cfl. aurantiacus* chromosome.** From outside to the center: Genes on forward strand (color by COG categories); Genes on reverse strand (color by COG categories); RNA genes (tRNAs, green; rRNAs, red; other RNAs, black); GC content; GC skew.

used to be considered as the only two proteins associated with the chlorosome mono-lipid layer. Other chlorosome proteins have also been reported, either through biochemical characterization (CsmP (unpublished results in Blankenship lab from in-solution trypsin digestion of the *Cfl. aurantiacus* chlorosomes) and AcsF [28]) or genomic analysis by analogy to green sulfur bacteria (CsmO, CsmP, CsmY) [29]. Among these proteins, AcsF, a protein responsible for chlorophyll biosynthesis under aerobic and semi-aerobic growth conditions, was unexpectedly identified from chlorosome fractions during anaerobic and

photoheterotrophic growth of *Cfl. aurantiacus* [28]. There has been some discussions as to whether AcsF is obligated to be associated with the chlorosomes [21,30], and the role of AcsF under anaerobic growth condition remains to be addressed, because it is an oxygen-dependent enzyme in other systems [31,32]. Although more chlorosome proteins have been identified recently, it is clear that CsmA, CsmM and CsmN are the most abundant proteins of the *Cfl. aurantiacus* chlorosomes. Genes encoding the experimentally identified and proposed chlorosome proteins are *csmA* (Caur\_0126), *csmM* (Caur\_0139), *csmN* (Caur\_0140),



*csnP* (Caur\_0142), *csnO* (Caur\_1311), and *csnY* (Caur\_0356).

## B. Electron transport complex genes

Figure 3A shows the proposed pathway of photosynthetic electron transport in *Cfl. aurantiacus* and purple photosynthetic proteobacteria. Similar to the purple photosynthetic proteobacteria, a cyclic electron transport pathway in *Cfl. aurantiacus* is also proposed. Nevertheless, some protein complexes in the electron transport chain of *Cfl. aurantiacus* are recognized to be substantially different from those of purple bacteria. *Cfl. aurantiacus* uses menaquinone as liposoluble electron and proton carrier [33-36], and purple proteobacteria use either ubiquinone [37,38] or menaquinone [39] as the mobile carrier in light-induced cyclic electron transport chain. The genetic information, analyses, and possible roles in photosynthesis and respiration for the complexes are described below.

### (I) Alternative complex III (ACIII)

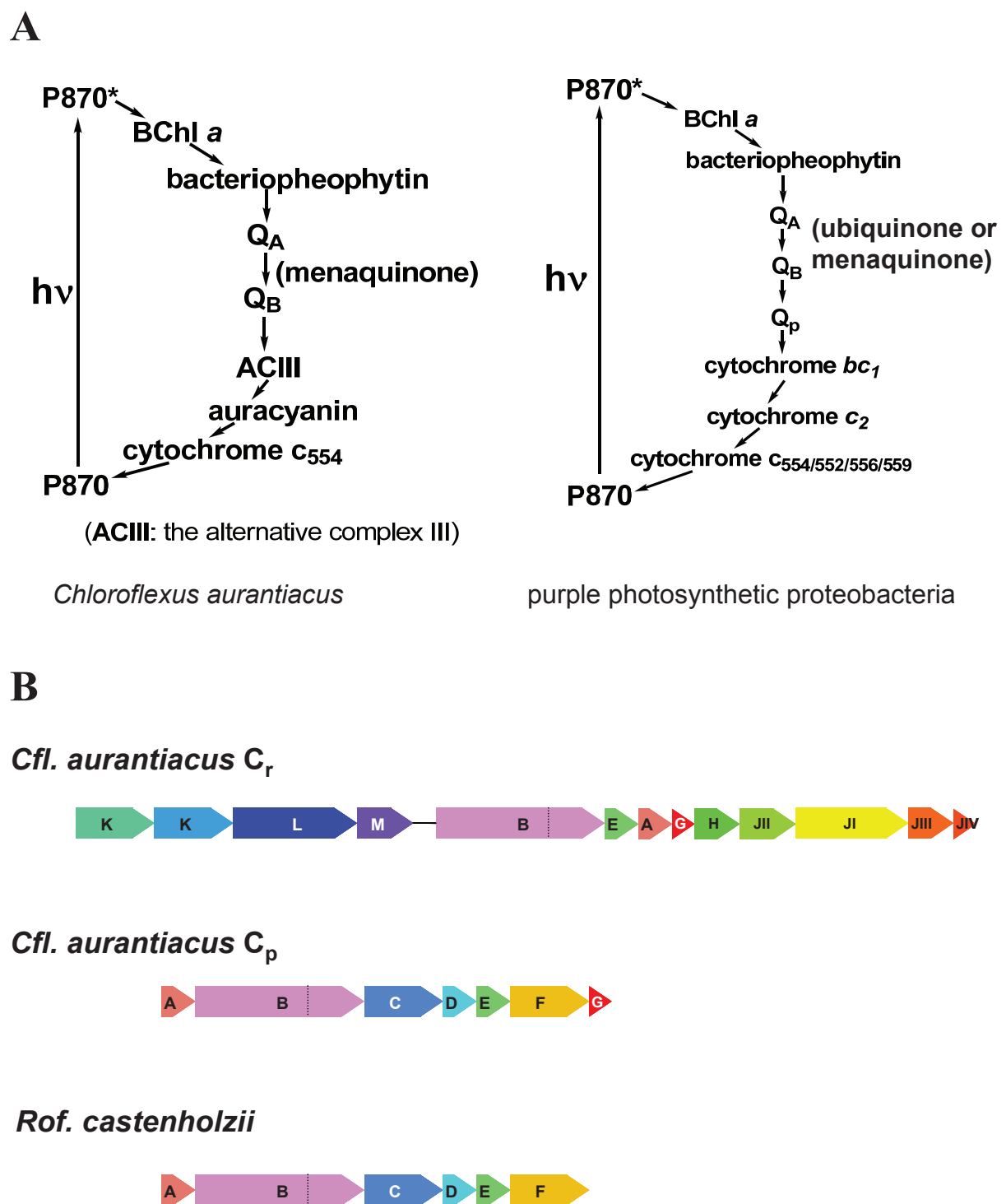
Integral membrane oxidoreductase complexes are essential for energy metabolism in all bacteria. In phototrophic bacteria, these almost invariably include the photoreaction center and a variant of respiratory Complex III, either the cytochrome *bc<sub>1</sub>* complex (anoxygenic) or cytochrome *b<sub>6</sub>f* complex (oxygenic). No homolog of the Complex III has been identified biochemically in *Cfl. aurantiacus*, and no genes with significant homology to Complex III are found in the *Cfl. aurantiacus* genome. Previous experimental evidence indicated that alternative complex III (ACIII) complexes, identified in *Cfl. aurantiacus* and some non-phototrophic bacteria, function in electron transport [34,40-44]. Genes encoding an ACIII have also been identified in the genome of *Candidatus Chloracidobacterium thermophilum* [21], an aerobic phototrophic *Acidobacterium* [45]. In the *Cfl. aurantiacus* genome, two ACIII operons have been identified: one encodes the C<sub>p</sub> (subscript p stands for photosynthesis) ACIII complex for anaerobic photosynthesis, and the other encodes the C<sub>r</sub> (subscript r stands for respiration) ACIII complex for aerobic respiration (Table 2). The C<sub>p</sub> operon is similar to a seven-gene *nrf* operon in *E. coli* strain K-12. Hussain et al. suggested that the *nrf* operon in *E. coli* is essential for reducing nitrate to ammonia [46]. The *Cfl. aurantiacus* C<sub>p</sub> operon (Caur\_0621 to Caur\_0627) contains genes encoding two types of cytochrome *c*; a multi-heme cytochrome *c* (component A, *actA*, Caur\_0621), which has recently been identified experimentally to be a penta-heme component [44], and a mono-heme cytochrome *c* (component E, *actE*, Caur\_0625), which forms a homodimer in the ACIII complex [44], a putative FeS-cluster-hydrogenase component-like protein (component B, *actB*, Caur\_0622), a polysulfide reductase (component

C, *actC*, Caur\_0623), similar to NrfD and likely involved in the transfer of electrons from the quinone pool to cytochrome *c*, an integral membrane protein (component F, *actF*, Caur\_0626) and two uncharacterized proteins (component D (*actD*, Caur\_0624) and component G (*actG*, Caur\_0627)) (Figure 3B).

The proposed C<sub>r</sub> ACIII operon contains 12 genes (Caur\_2133 to 2144) encoding a putative FAD-dependent oxidase (component K, *actK*, Caur\_2133), D-lactate dehydrogenase (component L, *actL*, Caur\_2134), a Cys-rich protein with Fe-S binding motifs (component M, *actM*, Caur\_2135), components B (*actB*, Caur\_2136), E (*actE*, Caur\_2137), A (*actA*, Caur\_2138), and G (*actG*, Caur\_2139) in the C<sub>p</sub> ACIII operon, an electron transport protein SCO1/SenC (Caur\_2140), and four subunits of cytochrome *c* oxidase (component J, Caur\_2141 - 2144). The cytochrome *c* oxidase (COX, or complex IV, EC 1.9.3.1) genes in the C<sub>r</sub> operon are part of complex IV, so the C<sub>r</sub> ACIII operon clustered with complex IV genes could create a respiratory superoperon (Figure 3B). Additionally, a gene cluster encoding a putative SCO1/SenC electron transport protein (Caur\_2423) and two COX subunits (Caur\_2425 (subunit II) and Caur\_2426 (subunit I)) is 300 genes away from the putative C<sub>r</sub> ACIII operon. Note that genes encoding components C (*actC*), D (*actD*) and F (*actF*) in the C<sub>p</sub> ACIII operon are absent in the C<sub>r</sub> ACIII operon. Whether these three components are required for the formation of the ACIII complex under aerobic respiratory growth will be addressed with biochemical studies.

### (II) Auracyanin

Two type I blue copper proteins have been isolated and proposed to function as the mobile electron carriers in photosynthetic electron transport of photosynthetic organisms: one is plastocyanin in cyanobacteria, photosynthetic algae and higher plants and the other is auracyanin in *Chloroflexus* and *Roseiflexus*. The type I blue copper protein auracyanin, which has a single copper atom coordinated by two histidine, one cysteine and one methionine residues at the active site, is proposed to participate in the electron transfer from ACIII to the reaction center in *Cfl. aurantiacus* [35,47-49], and it has also been recently characterized in *Roseiflexus castenholzii* [50]. Additionally, an auracyanin gene (*trd\_0373*) has been identified in the genome of the non-photosynthetic bacterium *Thermomicrobium roseum* DSM 5159, which is evolutionally related to *Cfl. aurantiacus* [51]. Two ACIII operons are proposed in *Cfl. aurantiacus*, and the two auracyanin proteins of *Cfl. aurantiacus*, AuraA and AuraB, which share 38% sequence identity, have been suggested to function with the two variant ACIII complexes [35]. AuraA, a water-soluble protein, can only be detected during phototrophic growth, whereas AuraB, a membrane-tethered protein, is synthesized during both



**Figure 3 Schematic representation of the proposed photosynthetic electron transport and the proposed ACIII operons.** The proposed photosynthetic electron transport in *Cfl. aurantiacus* (left) and in purple photosynthetic proteobacteria (right) (A), and the proposed ACIII operons in anaerobic photosynthesis ( $C_p$ ) and aerobic respiration ( $C_r$ ) in *Cfl. aurantiacus*, as well as the ACIII operon in *Roseiflexus* (*Rof. castenholzii*) (B). The characterized and putative proteins in the ACIII operon are listed as follows: **A**, multi-heme cytochrome *c*; **B**, MoCo Subunit (left) and FeS subunit (right); **C**, Integral membrane protein (polysulfide reductase, NrfD); **D**, uncharacterized protein; **E**, mono-heme cytochrome *c*; **F**, integral membrane protein; **G**, uncharacterized protein; **H**, electron transport SC01/SenC; **J**, cytochrome *c* oxidase subunits I-IV; **K**, FAD-linked oxidase; **L**, D-lactate dehydrogenase; and **M**, Cys and FeS rich domains. Abbreviation: BChl *a*, bacteriochlorophyll *a*;  $Q_A$ ,  $Q_B$ ,  $Q_p$ , quinone-type molecules.

**Table 2 Duplicate gene clusters of alternative complex III (ACIII) and NADH:quinone oxidoreductase (complex I) identified in the *Cfl. aurantiacus* genome**

Gene products	Gene cluster 1	Gene cluster 2
Alternative complex III (ACIII)	<i>actABCDEFG</i> (Caur_0621 to Caur_0627) encoding C <sub>p</sub> (subscript p stands for photosynthesis) ACIII	Caur_2133 to Caur_2144 (12 genes) encoding C <sub>r</sub> (subscript r stands for respiration) ACIII
NADH:quinone oxidoreductase (complex I, EC 1.6.5.3)	<i>nuoA</i> to <i>nuoN</i> (Caur_2896 to Caur_2909)	<i>nuoA</i> (Caur_1987), <i>nuoB</i> (Caur_1986), <i>nuoC</i> (Caur_1985), <i>nuoDE</i> (Caur_1984), <i>nuoF</i> (Caur_1185), <i>nuoH</i> (Caur_1982), <i>nuoI</i> (Caur_1983), <i>nuoJ</i> (Caur_1981), <i>nuoK</i> (Caur_1980), <i>nuoL</i> (Caur_1979), <i>nuoM</i> (Caur_1977 and Caur_1978) and <i>nuoN</i> (Caur_1976)

phototrophic and dark growth [35]. It has been hypothesized that AuraA transports electron from the C<sub>p</sub> ACIII during photosynthesis and AuraB from the C<sub>r</sub> ACIII during respiration. The *auraA* (Caur\_3248) and *auraB* (Caur\_1950) genes are distant from the C<sub>p</sub> operon (Caur\_0621 to Caur\_0627) and C<sub>r</sub> operon (Caur\_2132 to Caur\_2144). In addition to *auraA* and *auraB*, two more genes encoding auracyanin-like proteins (or type I blue-copper proteins) (Caur\_2212 and Caur\_2581) have also been found in the *Cfl. aurantiacus* genome. In contrast, *Roseiflexus castenholzii* has only one copy of the ACIII operon (a six-gene cluster, Rcas\_1462 to Rcas\_1467), in which the gene encoding the component G of the *Cfl. aurantiacus* C<sub>p</sub> ACIII complex is missing (Figure 3B), and one *auraA*-like gene (Rcas\_3112).

#### (III) NADH:quinone oxidoreductase

Two operons encoding the enzymes for NADH:quinone oxidoreductase (Complex I, EC 1.6.5.3) are present in the genome. Complex I catalyzes electron transport in the oxidative phosphorylation pathway. Many bacteria have 14 genes (*nuoA* to *nuoN*) encoding Complex I, and some photosynthetic bacteria, such as the purple photosynthetic proteobacteria *Rhodobacter sphaeroides* and *Rhodospseudomonas palustris*, contain two Complex I gene clusters. In *Cfl. aurantiacus*, two putative Complex I gene clusters have been identified, one with all of the 14 gene subunits arranging in order (*nuoA* to *nuoN*, Caur\_2896 - 2909), and the other has genes loosely arranged (with *nuoE* and *nuoF* 800 genes apart), duplicated *nuoM* genes, and the lack of *nuoG* (Table 2). It is possible that either *nuoG* is shared with the two putative Complex I gene clusters or an alternative gene with less sequence similarity functions as *nuoG*. For example, two gene loci (Caur\_0184 and Caur\_2214) encoding gene products that have ~24% sequence identity with NuoG, which is a molybdopterin oxidoreductase. To date, there have been no biochemical studies on the Complex I from *Cfl. aurantiacus* or any FAP bacteria.

#### (IV) Other electron transport proteins

In addition to the electron transport proteins described above, the sequence has been determined of cytochrome *c*<sub>554</sub>, which is also a subunit of the reaction center of *Cfl. aurantiacus* [52-54]. The sequence of the cytochrome *c* subunit in the *Roseiflexus castenholzii* RC has

also been reported [55]. The gene encoding cytochrome *c*<sub>554</sub> (*pufC*, Caur\_2089) is in an operon flanked with two genes encoding the bacteriochlorophyll biosynthesis enzymes, *bchP* (Caur\_2087) and *bchG* (Caur\_2088) at the 5'-end, and two genes encoding the B808-866 complex (Caur\_2090 (α-subunit) and Caur\_2091 (β-subunit)) at the 3'-end (Figure 4C).

### C. Aerobic/anaerobic enzyme pairs

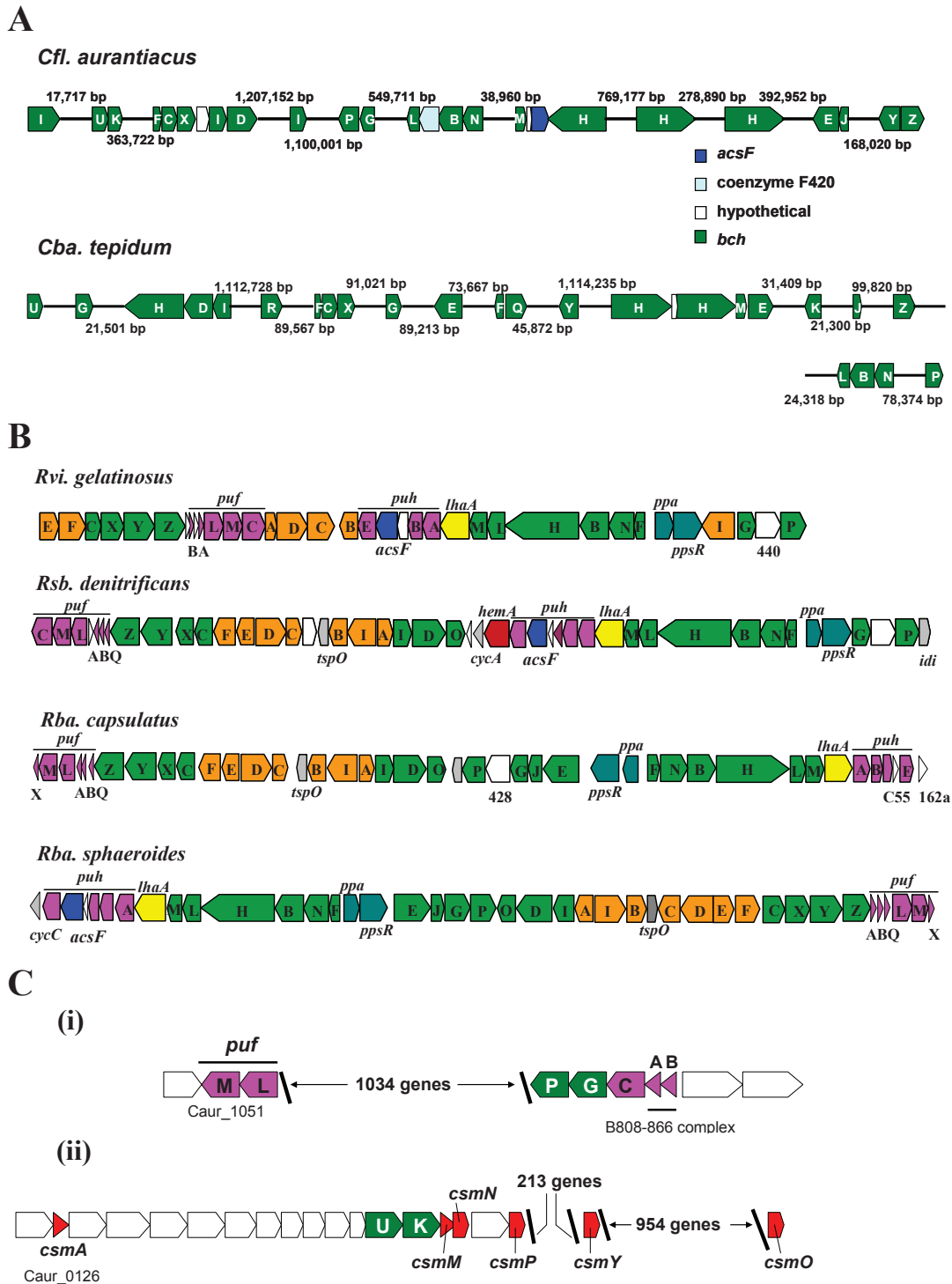
#### (I) Tetrapyrroles

In the biosynthesis of heme and chlorophyll (Chl), three aerobic/anaerobic enzyme pairs participate: coproporphyrinogen III decarboxylase (aerobic, HemF (EC 1.3.3.3); anaerobic, HemN (EC 1.3.99.22)) and protoporphyrinogen IX oxidase (anaerobic and aerobic) in heme biosynthesis, and Mg-protoporphyrin IX monomethyl ester cyclase (aerobic, AcsF; anaerobic, BchE) in chlorophyll (Chl) biosynthesis. Both aerobic and anaerobic gene pairs can be found in *Cfl. aurantiacus*, for example, *acsF* (Caur\_2590) and *bchE* (Caur\_3676) [28] as well as *hemF* (Caur\_2599) and *hemN* (Caur\_0209 and/or Caur\_0644) gene pairs. The *acsF* and *hemF* genes cannot be found in the green sulfur bacterium *Chlorobaculum tepidum* [56] and other strictly anaerobic bacteria. The genes involved in the biosynthesis of tetrapyrroles, as well as proposed gene replacements, are further elaborated below.

#### (a) Cobalamin

The gene replacements during the anaerobic to aerobic transitions are best known in the biosynthesis of cobalamin, in which the genes in the anaerobic pathway up to cobalt insertion into the corrin ring are completely replaced in the aerobic pathway [57]. Different strategies are used to generate cobyrinate diamide, the end product of both anaerobic and aerobic pathways, in which cobalt is introduced into the corrin ring at the dihydrosobacteriochlorin stage (early stage) of the anaerobic pathway and at the late stage of the aerobic pathway. The genomic information of *Cfl. aurantiacus* reveals a large cobalamin biosynthesis and cobalt transporter operon (Caur\_2560 - 2580), containing genes in both aerobic and anaerobic biosynthesis pathways, suggesting that *Cfl. aurantiacus* can synthesize cobalamin under various growth conditions. Genes encoding anaerobic





**Figure 4 Schematic representation of photosynthetic genes in photosynthetic bacteria.** Photosynthetic genes in *Cfl. aurantiacus*, *Cba. tepidum* (A), *Rba. sphaeroides*, *Rba. capsulatus*, *Rvi. gelatinosus*, and *Rsb. denitrificans* (B), an operon consist of the *puf* genes (the L- and M-subunits of the reaction center and cytochrome *c<sub>S54</sub>*), genes encoding two subunits of the B808-866 complex and two bacteriochlorophyll *a* biosynthesis enzymes (*bchP* and *bchG*) (i) and genes encoding bacteriochlorophyll *c* biosynthesis enzymes (*bchU* and *bchK*) and characterized and putative chlorosome proteins (*csmA*, *csmM*, *csmN*, *csmP*, *csmY* and *csmO*) (ii) (C). Genes are colored as listed: chlorosome proteins (*csm*) in red, chlorophyll biosynthesis (*bch*) in green, carotenoid biosynthesis (*crt*) in orange, reaction centers and light-harvesting complexes (*puf* and *puh*) in purple, regulatory proteins in blue (only in panel B), and uncharacterized and non-photosynthetic genes in white. All other gene colors are unique for clarity.

cobalt chelatase (EC 4.99.1.3) (*cbiK*, Caur\_2572) and aerobic cobalt chelatase (EC 6.6.1.2) (*cobNST*, *cobN* (Caur\_2579), *cobS* (Caur\_1198) and *cobT* (Caur\_2578)) have been identified (Tables 3 and 4). The aerobic cobalt chelatase (EC 6.6.1.2), containing three subunits (CobN, CobS and CobT), is a close analog to Mg-chelatase (also containing three subunits, BchH, BchI and BchD) that catalyzes the Mg-insertion in the chlorin ring in chlorophyll biosynthesis. It is known that aerobic cobalt chelatase subunits CobN and CobS are homologous to Mg-chelatase subunits BchH and BchI, respectively, and that CobT has also been found to be remotely related to the third subunit of Mg-chelatase, BchD. Compared to other strictly aerobic and anaerobic photosynthetic bacteria, the aerobic anoxygenic phototrophic proteobacterium *Roseobacter denitrificans* only carries the *cobNST* genes [27], and the strictly anaerobic bacterium *Heliobacterium modesticaldum* has only the *cbiK* gene [58]. The presence of *cobNST* and *cbiK* gene pairs in the *Cfl. aurantiacus* genome suggests a gene replacement in cobalamin biosynthesis by *Cfl. aurantiacus* under different growth conditions.

## (b) Heme

The heme operon (Caur\_2593 to Caur\_2599: *hemA*, *hemC*, *hemD*, *hemB*, *hemE*, *hemL* and *hemF*) is downstream of the cobalamin operon (Caur\_2560 to Caur\_2580). Except for *hemF*, other genes in the heme operon are utilized for synthesizing heme under both aerobic and anaerobic conditions. Furthermore, Caur\_0209 and Caur\_0644, encoding the putative O<sub>2</sub>-independent coproporphyrinogen III decarboxylase/oxidase (HemN, EC 1.3.99.22), and Caur\_0645, encoding O<sub>2</sub>-dependent protoporphyrinogen oxidase (HemG/HemY, EC 1.3.3.4), are ~ 2000 genes away from the heme gene cluster. It is interesting to note that two *hemN* genes have been identified in the *Cfl. aurantiacus* genome, while no gene encoding the O<sub>2</sub>-independent protoporphyrinogen oxidase in *Cfl. aurantiacus* has been characterized. Because heme can be synthesized by *Cfl. aurantiacus* in both aerobic and anaerobic environments, it is possible that one of the *hemN* genes in *Cfl. aurantiacus* may encode an O<sub>2</sub>-independent protoporphyrinogen oxidase. Finally, in addition to protoheme (heme *b*), the genome predicts that heme *o* and heme *a*

**Table 3 Aerobic and anaerobic gene pairs identified in the *Cfl. aurantiacus* genome**

(Putative) gene product and the EC number for aerobic and anaerobic gene product	Reaction catalyzed	Gene symbol and locus for aerobic genes	Gene symbol and locus for the anaerobic genes
<b>Tetrapyrrole (heme, cobalamin and chlorophyll) biosynthesis</b>			
coproporphyrinogen III decarboxylase/oxidase - aerobic (EC 1.3.3.3) - anaerobic (EC 1.3.99.22)	convert coproporphyrinogen III to protoporphyrinogen IX in heme biosynthesis	<i>hemF</i> (Caur_2599)	<i>hemN</i> (Caur_0209, Caur_0644)
Mg-protoporphyrin IX monomethyl ester oxidative cyclase (EC 1.14.13.81)	the isocyclic ring formation in chlorophyll biosynthesis	<i>acsF</i> (Caur_2590)	<i>bchE</i> (Caur_3676)
cobalt chelatase - aerobic (EC 6.6.1.2) - anaerobic (EC 4.99.1.3)	cobalt insertion on the corrin ring in cobalamin biosynthesis	<i>cobT</i> (Caur_2578), <i>cobN</i> (Caur_2579) <i>cobS</i> (Caur_1198)	<i>cbiK</i> (Caur_2572)
<b>Central carbon metabolism</b>			
$\alpha$ -ketoglutarate dehydrogenase E1 (EC 1.2.4.2), E2 (EC 2.3.1.61), and E3 (EC 1.8.1.4)	convert $\alpha$ -ketoglutarate to succinyl-CoA in the TCA cycle	E1 ( <i>sucA</i> , Caur_3727) E2 ( <i>olst</i> , Caur_1691, Caur_3726) E3 ( <i>dld</i> , Caur_0170, Caur_2840)	
$\alpha$ -ketoglutarate:ferredoxin oxidoreductase (or $\alpha$ -ketoglutarate synthase) (EC 1.2.7.3)	convert succinyl-CoA to $\alpha$ -keto-glutarate in the TCA cycle		<i>korA</i> (Caur_0249, Caur_0952) <i>korB</i> (Caur_0250, Caur_0953)
pyruvate dehydrogenase E1 (EC 1.2.4.1), E2 (EC 2.3.1.12), and E3 (EC 1.8.1.4)	convert pyruvate to acetyl-CoA in pyruvate metabolism	E1 (Caur_1334, 1335, 1972, 1973, 2805, 3121, 3671, and 3672), E2 (Caur_1333, Caur_1974), E3 ( <i>dld</i> , Caur_0170, Caur_2840)	
<b>Central carbon metabolism</b>			
pyruvate:ferredoxin oxidoreductase (or pyruvate synthase) (EC 1.2.7.1)	convert acetyl-CoA to pyruvate in pyruvate metabolism		<i>porA</i> (Caur_0249, Caur_0952, Caur_2080)
<b>Nucleic acid biosynthesis</b>			
ribonucleoside-diphosphate reductase aerobic (EC 1.17.4.1) anaerobic (EC 1.17.4.2)	convert ribonucleotide into deoxyribonucleotide in nucleic acids biosynthesis	<i>nrdB</i> (Caur_3331, $\beta$ -subunit)	<i>nrdJ</i> (Caur_1750)
dihydroorotate oxidase (EC 1.3.3.1, aerobic), and dihydroorotate dehydrogenase (EC 1.3.99.11, anaerobic),	conversion of dihydroorotate to orotate in pyrimidine biosynthesis	Caur_2081 and Caur_3923	Caur_2338

**Table 4 Selected genes and gene clusters in metabolic pathways of *Cfl. aurantiacus***

Selected metabolic pathways and/or gene products	Gene symbol, gene loci, and/or gene products
CO oxidation	<i>coxSML</i> and <i>coxG</i> (Caur_3467 - 3470)
CO <sub>2</sub> -anaplerotic pathways	<i>tme</i> (Caur_1614), <i>pckA</i> (Caur_2331), <i>ppc</i> (Caur_3161, Caur_3888)
3-hydroxypropionate cycle	<i>accC</i> (Caur_1378, Caur_3421), <i>accA</i> (Caur_1647), <i>accD</i> (Caur_1648), <i>accB</i> (Caur_3739), <i>mcr</i> (Caur_2614), <i>pcs</i> (Caur_0613), <i>pccB</i> (Caur_2034, Caur_3435), <i>mcee</i> (Caur_3037), <i>mut</i> (Caur_1844, Caur_2508, Caur_2509), <i>smtA</i> (Caur_0179), <i>smtB</i> (Caur_0178), <i>sdhBAC</i> (Caur_1880 to Caur_1882), <i>fh</i> (Caur_1443), <i>mcl</i> (Caur_0174), <i>mch</i> (Caur_0173), <i>mct</i> (Caur_0174), <i>meh</i> (Caur_0180)
Glycolate assimilation and glyoxylate cycle	<i>glcDEF</i> (Caur_1144 - 1145, Caur_2133, Caur_2135), <i>glyr</i> (Caur_0825), <i>icl</i> (Caur_3889) and <i>mas</i> (Caur_2969)
Cobalamin biosynthesis	<i>cobS</i> (Caur_1198), <i>cobQ/cbiP</i> (Caur_2560), <i>N</i> -transferase (Caur_2561), <i>cobP/cobU</i> (Caur_2562), <i>cobB/cbiA</i> (Caur_2563), <i>cobA</i> (Caur_2564, Caur_0687), <i>cobJ</i> (Caur_2565), <i>cobM/cbiF</i> (Caur_2566), <i>cobl/cbiL</i> (Caur_2567), <i>cobL/cbiET</i> (Caur_2568), <i>cbiD</i> (Caur_2569), <i>cobH/cbiCH</i> (Caur_2570), <i>cbiK</i> (Caur_2572), <i>cobD/cbiB</i> (Caur_2573), <i>cbiMNQO</i> (Caur_2574 to Caur_2577), <i>cobU/cobT</i> (Caur_2578), <i>cobN</i> (Caur_2579), <i>cobO</i> (Caur_2580), <i>cysG</i> (Caur_0688) and <i>cbiMNQ</i> (Caur_3680 to Caur_3682) (Genes shared with aerobic and anaerobic pathways shown as aerobic/anaerobic)
Heme biosynthesis	<i>hemA</i> , <i>hemC</i> , <i>hemD</i> , <i>hemB</i> , <i>hemE</i> , <i>hemL</i> , <i>hemF</i> (Caur_2593 - 2599), <i>hemN</i> (Caur_0209 and Caur_0644), <i>hemG/Y</i> (Caur_0645), protoheme IX farnesyltransferase (Caur_0029)
Chlorophyll biosynthesis	<i>bchl</i> (Caur_0117, Caur_0419, Caur_1255), <i>bchU</i> (Caur_0137), <i>bchK</i> (Caur_0138), <i>bchF</i> (Caur_0415), <i>bchC</i> (Caur_0416), <i>bchX</i> (Caur_0417), <i>bchD</i> (Caur_0420), <i>bchP</i> (Caur_2087), <i>bchG</i> (Caur_2088), <i>bchBNL</i> (Caur_2554 to Caur_2557), <i>bchM</i> (Caur_2588), <i>acsF</i> (Caur_2590), <i>bchH</i> (Caur_2591, Caur_3151, Caur_3371), <i>bchE</i> (Caur_3676), <i>bchJ</i> (Caur_3677), and <i>bchYZ</i> (Caur_3805, Caur_3806)
Nitrogen metabolism and amino acid biosynthesis	<i>hal</i> (Caur_0974), <i>tpl</i> (Caur_0573), <i>aspg</i> (Caur_3511), <i>glud1</i> (Caur_1698, Caur_2070), <i>glul</i> (Caur_0844, Caur_1448, Caur_3395), <i>ilvA</i> (Caur_2585, Caur_3892), isoleucine/leucine/valine biosynthesis (Caur_0041, Caur_0163 - 0169, Caur_0329 - 0331, Caur_0488, Caur_1435, and Caur_2851)
Sulfur metabolism	Sulfur reduction operon (Caur_0686 - 0692), a sulfate adenylyl-transferase/adenylylsulfate kinase (Caur_2113), <i>cysK</i> (Caur_1341), <i>cysM</i> (Caur_3489), <i>sqr</i> (Caur_3894, type II SQR), sulfotransferase (Caur_2114).
B808-866 light-harvesting complex	Caur_2090 ( $\alpha$ -subunit) and Caur_2091 ( $\beta$ -subunit)
Reaction center	<i>pufM</i> (Caur_1051) and <i>pufL</i> (Caur_1052)
Chlorosome proteins	<i>csmA</i> (Caur_0126), <i>csmM</i> (Caur_0139), <i>csmN</i> (Caur_0140), <i>csmP</i> (Caur_0142), <i>csmO</i> (Caur_1311) and <i>csmY</i> (Caur_0356).
Auracyanins (type I blue-copper proteins)	<i>auraA</i> (Caur_3248), <i>auraB</i> (Caur_1950), Caur_2212 and Caur_2581
Cytochrome <i>c</i> <sub>554</sub> complex	<i>pufC</i> (Caur_2089)
Succinate dehydrogenase/fumarate reductase (complex II, EC 1.3.99.1)	<i>sdhBAC</i> (Caur_1880 - 1882)
Cytochrome <i>b</i> <sub>c</sub> , or <i>b</i> <sub>6</sub> / <i>f</i> complex	Not annotated
Cytochrome <i>c</i> oxidase (complex IV, EC 1.9.3.1)	COX I - IV (Caur_2141 - 2144), COX I (Caur_2426) and COX II (Caur_2425 and Caur_2582)

can be synthesized respectively by the gene products of Caur\_0029 (encoding protoheme IX farnesyltransferase) and Caur\_1010 (encoding a cytochrome *aa3* biosynthesis protein), consistent with the spectral evidence provided by Pierson that protoheme and heme derivatives can be identified [59].

### (c) (Bacterio)chlorophylls

The anaerobic to aerobic transitions are particularly intriguing in chlorophyll (Chl) biosynthesis and photosynthesis, in which molecular oxygen is lethal for photosynthetic anaerobes but is required for the life of aerobic phototrophs. Contrary to the cobalamin and heme biosynthesis, no gene cluster for (B)Chl biosynthesis is recognized in the *Cfl. aurantiacus* genome, whereas photosynthesis gene clusters are present in purple photosynthetic proteobacteria [27,60,61] (Figure 4B)

and heliobacteria [58]. The photosynthetic genes of *Cfl. aurantiacus* are rather spread out in the chromosome, similar to the distribution of photosynthetic genes in *Cba. tepidum* (Figure 4A and Additional file 1: **Table S1**). Both aerobic and anaerobic genes, *acsF* (Caur\_2590) and *bchE* (Caur\_3676), have been identified in the *Cfl. aurantiacus* genome (Table 3), as well as in the genome of other phototrophic *Chloroflexi* species, including *Roseiflexus castenholzii* [28]. *AcsF* and *BchE* catalyze the isocyclic ring (or the E-ring) formation of Chl under aerobic and anaerobic growth conditions, respectively. *AcsF* (aerobic cyclase) catalyzes the formation of the isocyclic ring and reduces O<sub>2</sub> into H<sub>2</sub>O. *BchE* requires cobalamin for catalytic activity, and has putative cobalamin and [4Fe-4S] cluster/*S*-adenosyl-*L*-methionine binding motifs [28,31,32,62-64].

Other anaerobic to aerobic transitions may also be found in the biosynthesis of BChls in *Cfl. aurantiacus*. For example, Mg-insertion to the porphyrin ring is the first committed step of BChl biosynthesis, and three *bchH* (Caur\_2591, Caur\_3151, and Caur\_3371), three *bchI* (Caur\_0117, Caur\_0419 and Caur\_1255) and one *bchD* (Caur\_0420) have been annotated for the Mg-chelatase (BchHID) of *Cfl. aurantiacus*, whereas one each for *Hbt. modesticaldum* [58], *Rsb. denitrificans* [27] and several strictly anaerobic and aerobic bacteria. On the other hand, three *bchH*, and one copy of *bchD* and *bchI*, have been identified in the green sulfur bacterium *Cba. tepidum* (Additional file 1: **Table S1**), and Eisen et al. proposed [56] that different BchH gene products may contribute to synthesize different isoforms of BChl (BChl *a*, BChl *c*, and Chl *a* can be synthesized in *Cba. tepidum*). In comparison, two types of BChls, BChl *a* and BChl *c* can be synthesized by *Cfl. aurantiacus*. It is also possible that different *bchH* and *bchI* genes catalyze Mg-chelation to the BChl in various growth conditions of *Cfl. aurantiacus*.

Two *bchG*-like genes (*bchG* and *bchK*) encoding chlorophyll synthases that attach the tail into (bacterio) chlorophylls are present in the genome, as shown in Additional file 1: **Table S1**. Because the tails of BChl *a* (mostly phytol- or geranylgeranyl-substituted) and BChl *c* (mainly stearyl-substituted) are rather distinct, it was suggested that one *bchG*-like gene encodes the enzyme synthesizing BChl *a* and the other homolog synthesizes BChl *c* [65]. The *bchG* gene sequence reported by Lopez et al. [65] was proposed to be BChl *a* synthase, since the encoding protein sequence is analogous to the sequence of chlorophyll synthase from *Rhodobacter capsulatus*. The proposed gene function was later verified [66]. The *bchK* gene encoding BChl *c* synthase was later confirmed with the *bchK*-knockout *Cba. tepidum* mutant [67]. Thus, *bchG* (Caur\_2088) and *bchK* (Caur\_0138) encode enzymes synthesizing BChl *a* and BChl *c*, respectively, in *Cfl. aurantiacus*. Although genes responsible for chlorophyll biosynthesis are rather spread out in the *Cfl. aurantiacus* genome, two genes responsible for BChl *c* biosynthesis, *bchU*, encoding C-20 methyltransferase [68], and *bchK*, are clustered with the genes encoding chlorosome proteins, and two BChl *a* biosynthesis genes, *bchP*, encoding geranylgeranyl hydrogenase [69], and *bchG*, are in the operon containing genes encoding cytochrome *c*<sub>554</sub> (*pufC*) and the B808-866 light-harvesting complex (Figure 4C).

### (III) Nucleic acids

The level of oxygen tolerance in *Cfl. aurantiacus* may be suggested from the presence of genes encoding ribonucleotide reductase (RNR), which is essential for DNA synthesis. Three classes of RNR have been reported, in which the class I is a diiron oxygen-dependent (NrdB,

EC 1.17.4.1), class II is coenzyme B<sub>12</sub>-dependent (NrdJ, EC 1.17.4.2), and class III is *S*-adenosyl-*L*-methionine/[4Fe-4S] cluster-dependent (NrdG, EC 1.17.7.1). It has been suggested that biosynthesis of dNTP is catalyzed by NrdB, NrdJ, and NrdG in aerobic, aerobic and anaerobic, and strictly anaerobic environments, respectively. The activity of NrdJ in *Cfl. aurantiacus* has been reported [70]. Genes encoding NrdB and NrdJ, but not NrdG, have been found in the *Cfl. aurantiacus* genome, suggesting that NrdG and NrdJ produce dNTP for *Cfl. aurantiacus* in response to the oxygen level (Table 3). Moreover, in the fourth step of the biosynthesis of pyrimidine, the conversion of dihydroorotate to orotate, dihydroorotate oxidase (EC 1.3.3.1, aerobic) versus dihydroorotate dehydrogenase (EC 1.3.99.11, anaerobic) are expressed in aerobic versus anaerobic conditions and genes encoding these enzymes have been identified (Table 3). Together, different classes of RNR and dihydroorotate oxidoreductase in the nucleic acid biosynthesis also suggest adaptation or evolution from anaerobic to aerobic conditions.

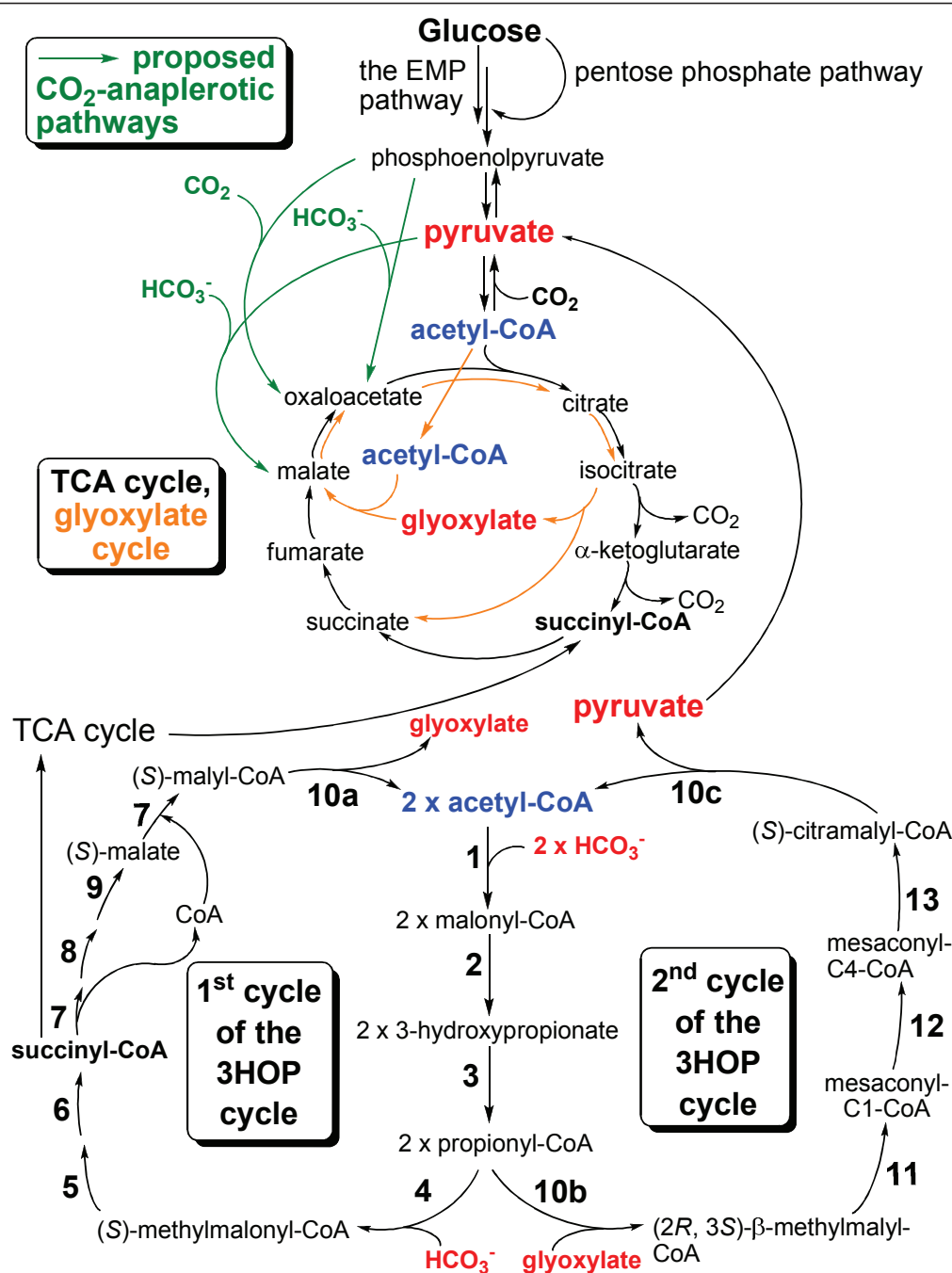
### (III) Central carbon metabolism

Genes encoding several aerobic/anaerobic enzyme pairs in the central carbon metabolism, such as genes encoding pyruvate dehydrogenase (PDH, EC 1.2.4.1) and  $\alpha$ -ketoglutarate dehydrogenase (KDH, EC 1.2.4.2), as well as pyruvate:ferredoxin oxidoreductase (or pyruvate synthase) (PFOR, EC 1.2.7.1) and  $\alpha$ -ketoglutarate:ferredoxin oxidoreductase (or  $\alpha$ -ketoglutarate synthase) (KFOR, EC 1.2.7.3) are present in the genome (Table 3). PFOR and KFOR, which are essential for pyruvate metabolism and energy metabolism through the TCA cycle, are commonly found in anaerobic organisms, whereas PDH and KDH are more widely spread and have been found in all aerobic organisms.

## D. CO<sub>2</sub> assimilation and carbohydrate, nitrogen and sulfur metabolisms

### (I) Carbon fixation and metabolism

Genes encoding carbon monoxide dehydrogenase (*coxG* and *coxSML*) have been found in the genome, suggesting that *Cfl. aurantiacus* can use CO as an electron source during aerobic or semi-aerobic growth. A similar mechanism has been suggested in the aerobic anoxygenic phototrophic proteobacterium *Rsb. denitrificans* [27]. CO<sub>2</sub> generated from CO oxidation can be assimilated by *Cfl. aurantiacus* via the autotrophic carbon fixation cycle and/or the CO<sub>2</sub>-anaplerotic pathways. Under autotrophic growth conditions *Cfl. aurantiacus* is known to use a unique carbon fixation pathway: the 3-hydroxypropionate (3HOP) autotrophic cycle [4,5,71-74]. Three inorganic carbon molecules are assimilated into the 3HOP cycle to produce one molecule of pyruvate (Figure 5). A similar carbon fixation



**Figure 5** Proposed autotrophic and anaplerotic CO<sub>2</sub> assimilation and central carbon metabolic pathways of *Cfl. aurantiacus*. The enzymes required for each reaction step of the 3-hydroxypropionate (3HOP) autotrophic CO<sub>2</sub>-fixation cycle were described in the Results and Discussions.

pathway called 3-hydroxypropionate/4-hydroxybutyrate (3HOP/4HOB) cycle was reported recently in archaea (Crenarchaeota) [75-78]. Several enzymes responsible for the 3HOP and 3HOP/4HOB cycles, including CO<sub>2</sub>-fixing enzymes (e.g., acetyl-CoA carboxylase and propionyl-CoA carboxylase), are common to the two pathways.

When the genome was first available, some genes required for the 3HOP cycle could not be found [5], whereas some of the missing genes/enzymes for the 3HOP cycle were later characterized experimentally [4]: (1) acetyl-CoA carboxylase (*accC* (Caur\_1378 and Caur\_3421, biotin carboxylase), *accA* (Caur\_1647,



$\alpha$ -subunit), *accD* (Caur\_1648,  $\beta$ -subunit) and *accB* (Caur\_3739) [79], (2) malonyl-CoA reductase (*mcr*, Caur\_2614) [80], (3) propionyl-CoA synthase (*pcs*, Caur\_0613) [81], (4) propionyl-CoA carboxylase (*pccB*, Caur\_2034, Caur\_3435), (5) methylmalonyl-CoA epimerase (*mcee*, Caur\_3037), (6) *L*-methylmalonyl-CoA mutase (MCM) (*mut*, Caur\_1844, Caur\_2508, Caur\_2509), (7) succinyl-CoA:(*S*)-methyl-CoA transferase (*smtA* (Caur\_0179), *smtB* (Caur\_0178)), (8) succinate dehydrogenase (*sdhBAC*, Caur\_1880 to Caur\_1882), (9) fumarate hydratase (*fhl*, Caur\_1443), (10) (*S*)-methyl-CoA/ $\beta$ -methylmalonyl-CoA/(*S*)-citramalonyl-CoA (MMC) lyase (*mcl*, Caur\_0174), (11)  $\beta$ -methylmalonyl-CoA dehydratase (*mch*, Caur\_0173), (12) mesaconyl-C1-CoA-C4-CoA transferase (*mct*, Caur\_0174) and (13) mesaconyl-C4-CoA hydratase (*meh*, Caur\_0180) (Figure 5). The enzymes responsible for the 4<sup>th</sup> step to the 6<sup>th</sup> step of the 3HOP cycle are involved in fatty acid oxidation and amino acid metabolism. Finally, no genes encoding ribulose 1,5-bisphosphate carboxylase (RuBisCO) (the Calvin-Benson cycle), ATP citrate lyase (the reductive TCA cycle) and acetyl-CoA synthase (the Wood-Ljungdahl pathway) are present, strongly suggesting that these autotrophic carbon fixation pathways are not present in *Cfl. aurantiacus*.

Additionally, genes encoding malic enzyme (*tme*), phosphoenolpyruvate (PEP) carboxykinase (*pckA*) and PEP carboxylase (*ppc*) have been identified, suggesting that *Cfl. aurantiacus* can assimilate some CO<sub>2</sub> and replenish the metabolites in the TCA cycle through the CO<sub>2</sub>-anaplerotic pathways. The active CO<sub>2</sub>-anaplerotic pathways have been identified experimentally in other anoxygenic phototrophs during autotrophic, mixotrophic and heterotrophic growth [82–87], and the activities of PEP carboxylase and malic enzyme have also been detected in cell extracts during photoheterotrophic growth of *Cfl. aurantiacus* (Tang and Blankenship, unpublished results). Moreover, all of the genes in the TCA cycle are present in the *Cfl. aurantiacus* genome.

In central carbon metabolism, all of the genes in the TCA cycle as well as the glyoxylate cycle have been identified. The glyoxylate cycle is one of the anaplerotic pathways for assimilating acetyl-CoA, thus lipids can be converted to carbohydrates. Glyoxylate is synthesized and also assimilated in the 3HOP cycle (Figure 5), and is also produced by isocitrate lyase (EC 4.1.3.1) (*icl*, Caur\_3889) and consumed by malate synthase (EC 2.3.3.9) (*mas*, Caur\_2969) in the glyoxylate cycle. With acetate as the sole organic carbon source to support the photoheterotrophic growth of *Cfl. aurantiacus*, higher activities of isocitrate lyase and malate synthase have been reported [88]. Further, some FAP bacteria have been shown to assimilate glycolate from their habitat [89]. As glycolate can be converted to glyoxylate by

glycolate oxidase (*glcDEF*, EC 1.1.3.15) and glyoxylate reductase (*glyr*, EC 1.1.1.26), the glyoxylate shunt and the 3HOP cycle can be employed by *Cfl. aurantiacus* for assimilating glycolate. Together, genes encoding central carbon metabolism, 3HOP cycle, glycolate assimilation, the glyoxylate shunt and CO oxidation are listed in Table 4.

### (II) Carbohydrate metabolism

Three carbohydrate metabolism pathways are utilized by various bacteria: the Embden-Meyerhof-Parnas (EMP) pathway (glycolysis), the Entner-Doudoroff (ED) pathway, and the pentose phosphate (PP) pathway. *Cfl. aurantiacus* does not have genes in the ED pathway, but has genes in the oxidative PP pathway, in agreement with the activities reported for the essential enzymes (glucose-6-phosphate dehydrogenase and 6-phosphogluconate dehydrogenase) in the oxidative PP pathway [90]. Genes in the non-oxidative pathway have also been found. The gene encoding fructose-1,6-bisphosphate (FBP) aldolase (EC 4.1.2.13), catalyzing the reaction of D-fructose-1,6-bisphosphate (FBP)  $\leftrightarrow$  glyceraldehyde-3-phosphate (GAP) + dihydroxyacetone phosphate (DHAP) in the EMP/gluconeogenic pathway, is missing in the genomes of *Chloroflexi* species (e.g., *Cfl. aurantiacus*, *Chloroflexus* sp. Y-400-fl and *Chloroflexus aggregans*). If *Cfl. aurantiacus* were unable to synthesize FBA, an active pentose phosphate pathway would be required for the interconversion of D-glucose-6-phosphate and GAP, so glucose and other sugars can be converted to pyruvate and other energy-rich species, and vice versa. However, *Cfl. aurantiacus* has been reported to grow well in glucose and a number of sugars during aerobic respiration [91], and uses the EMP pathway for carbohydrate catabolism [90]. Also, higher activities of phosphofructokinase and FBP aldolase have been found in the cells grown with glucose than with acetate [1,92,93]. Note that *Roseiflexi* species (e.g., *Roseiflexus* sp. RS-1 and *Roseiflexus castenholzii* DSM 13941), which are closely related to *Chloroflexi* species, have a putative bifunctional FBP aldolase/phosphatase gene identified [94], and genes encoding various types of aldolase have been found in the *Cfl. aurantiacus* genome. Taken together, *Cfl. aurantiacus* and other *Chloroflexi* species may employ either a novel FBP aldolase or more than one carbohydrate metabolic pathway. Further efforts will be needed to clarify this picture.

### (III) Nitrogen metabolism and amino acid biosynthesis

*Cfl. aurantiacus* is known to use ammonia as the sole nitrogen source, and several amino acids (nitrogenous compounds), but not nitrate, can enhance the growth. Neither nitrogenase nor nitrogen fixation has been reported in *Cfl. aurantiacus* [95]. Consistent with the physiological studies, genes encoding enzymes in ammonia production, such as histidine ammonia lyase (*hal*),

tyrosine phenol-lyase (*tpl*), asparaginase (*aspg*), glutamate dehydrogenase (*glud1*), and glutamate ammonia-lyase (*glul*), but not nitrogen metabolism (*nifDHK*) and nitrate reduction, are present in the genome. Note that *Cfl. aurantiacus* has genes encoding a copper-containing nitrite reductase (EC 1.7.2.1) (Caur\_1570) and the  $\alpha$ -subunit (*narG*, Caur\_3201), but not other subunits (*narHII*) and the catalytic subunit (*nasA*), of nitrate reductase. Two threonine/serine dehydratases (EC 4.3.1.19), one of which is inhibited by isoleucine may be related to the isoleucine biosynthesis, and other key enzymes in isoleucine biosynthesis have been reported [96]. Consistent with the biochemical studies, two *ilvA* genes (Caur\_2585 and Caur\_3892) encoding threonine dehydratases, and genes in the isoleucine/leucine/valine biosynthesis pathway have been identified (Table 4). The biosynthesis of isoleucine has recently been discovered through the citramalate pathway in several microbes [97,98], while no gene encoded citramalate synthase (CimA, EC 2.3.1.182), required for the citramalate pathway, has been found in the *Cfl. aurantiacus* genome.

#### (IV) Sulfur assimilation and sulfate reduction

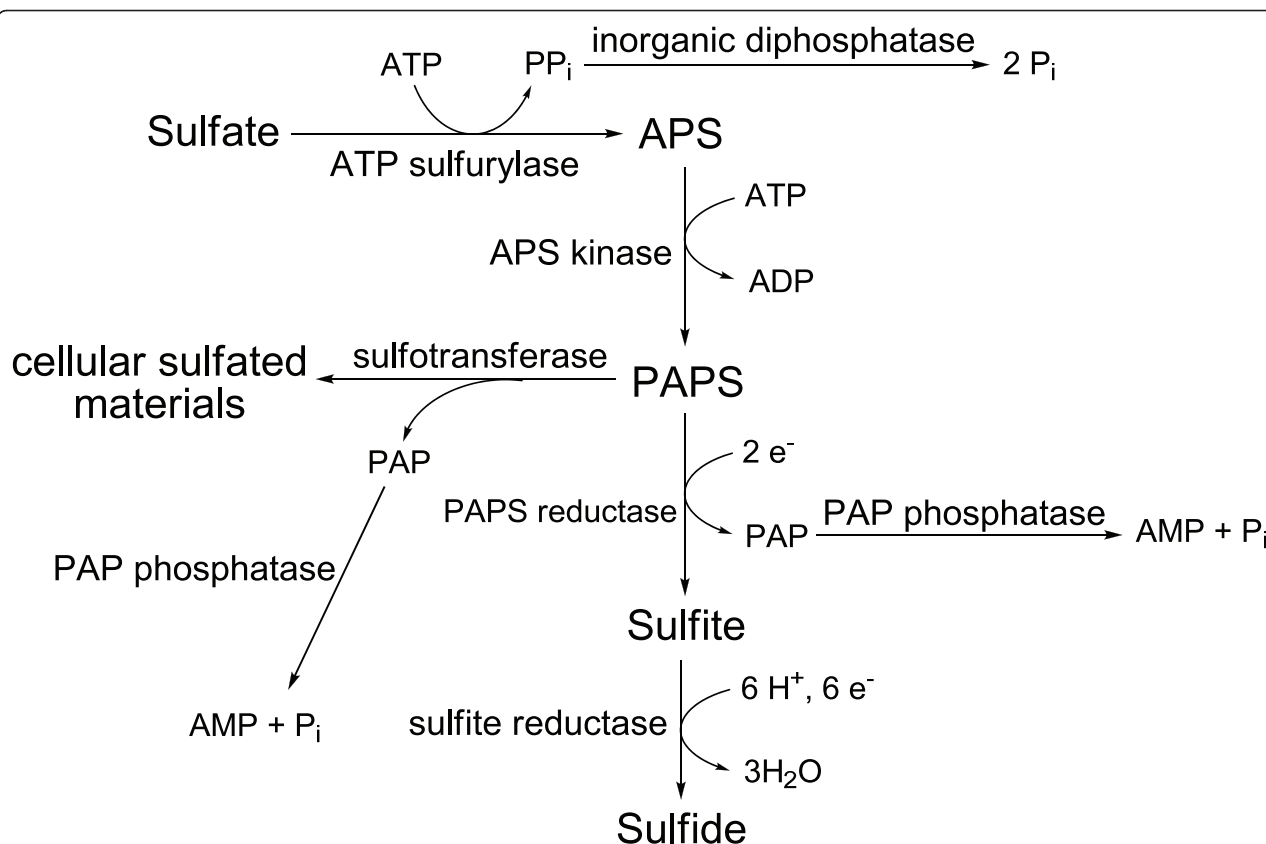
*Cfl. aurantiacus* can use a variety of compounds as sulfur sources, including cysteine, glutathione, methionine, sulfide and sulfate, during photoheterotrophic or photoautotrophic growth [99,100]. When *Cfl. aurantiacus* uses sulfate as the sulfur source, high activity of ATP sulfurylase has been reported [99]. Sulfate is reduced to sulfide during photoautotrophic and photoheterotrophic growth for synthesizing cysteine and cofactors. Consistent with the experimental data, a complete sulfur reduction pathway with a sulfur reduction operon (Caur\_0686 - 0692) has been identified (Table 4). Genes encoding two ATP sulfurylases ( $\text{ATP} + \text{sulfate} \rightarrow \text{adenosine } 5'\text{-phosphosulfate (APS)} + \text{PPi}$ ) can be identified: sulfate adenylyltransferase (EC 2.7.7.4, Caur\_0690) and a bifunctional sulfate adenylyltransferase/adenylylsulfate kinase (Caur\_2113). Pyrophosphate (PPi) produced in the reaction of ATP sulfurylase is hydrolyzed to inorganic phosphate (Pi) via inorganic diphosphatase (EC 3.6.1.1) (Caur\_3321). The bifunctional enzyme or/and adenylylsulfate kinase (EC 2.7.1.25, Caur\_0692) converts APS to 3'-phosphoadenosine 5'-phosphosulfate (PAPS), which is reduced to sulfite and PAP (adenosine 3',5'-diphosphate) by PAPS reductase (EC 1.8.4.8, *cysH*, Caur\_0691). While many organisms reduce APS instead of PAPS to sulfite, it is unknown if *Cfl. aurantiacus* carries out this reaction as genes encoding APS reductase (EC 1.8.99.2) have not been identified in the genome. In addition to the proposed pathway, sulfotransferase (Caur\_2114) can also transfer the sulfate group from PAPS, which serves as a sulfur donor, to an alcohol or amine acceptor for generating various cellular sulfate compounds. PAP is generated as a by-product in the

reactions catalyzed by PAPS reductase and sulfotransferase, and has no known functions in metabolism and is likely hydrolyzed to AMP and Pi via PAP phosphatase (unidentified yet). Sulfite reductase (EC 1.8.1.2, Caur\_0686) reduces sulfite to sulfide, which is incorporated into cysteine by cysteine synthase A (*cysK*, Caur\_1341) or cysteine synthase B (*cysM*, Caur\_3489). The overall proposed sulfur reduction and assimilation pathways are illustrated in Figure 6.

Other than using sulfide as the sulfur source during photoheterotrophic growth, *Cfl. aurantiacus* grows photoautotrophically in the presence of sufficient sulfide [100-102]. Under these circumstances, sulfide likely functions as electron donor by replacing organic carbon sources contributed from cyanobacteria. In agreement with these physiological and ecological studies, the gene encoding a type II sulfide:quinone oxidoreductase (SQR) (*sqr*, Caur\_3894), has been found in the genome. SQRs belong to the members of the disulfide oxidoreductase flavoprotein superfamily. Other than type II SQRs, type I and type III SQRs with distinct sequences and structures and cofactor requirements have also been reported [103]. Although all of the characterized SQRs catalyze oxidization of sulfide to elemental sulfur ( $\text{E(ox)} + \text{H}_2\text{S} \rightarrow \text{E}_{\text{H}_2}(\text{red}) + \text{S}^0$ ), different types of SQR have been identified in various classes of photosynthetic bacteria [103]: type I, purple non-sulfur anoxygenic photosynthetic proteobacteria (such as *Rhodobacter capsulatus*) [104]; type II, *Cfl. aurantiacus* and cyanobacteria (such as *Synechocystis* PCC 6803) [103]; and type III, green sulfur bacteria [105]. In addition to being characterized in phototrophic microbes, type II SQRs have also been identified in various non-photosynthetic bacteria as well as in the mitochondria of animals, and are suggested to be involved in sulfide detoxification, heavy metal tolerance, sulfide signaling, and other essential cellular processes [103].

#### E. Evolution perspectives

Our paper reports numerous aerobic/anaerobic gene pairs and oxygenic/anoxygenic metabolic pathways in the *Cfl. aurantiacus* genome. As suggested by phylogenetic analyses [6-8] and comparisons to the genome and reports of other photosynthetic bacteria, one can propose lateral or horizontal gene transfers between *Cfl. aurantiacus* and other photosynthetic bacteria. Some proposed lateral gene transfers are listed below and also illustrated in Figure 7. Note that horizontal/lateral gene transfers suggested below are important in the evolution of photosynthesis. It is important to remember that it has not yet been generally accepted which organisms were donors and which were acceptors during gene transfers. The proposed gene transfers below remain to be verified with more sequenced genomes and biochemical studies in photosynthetic organisms.



**Figure 6** Proposed sulfur reduction and assimilation pathways in *Cfl. aurantiacus*. All of the genes, except PAP phosphatase, have been identified in the genome.

#### (I) Photosynthetic components

Chlorosomes were transferred between *Cfl. aurantiacus* and the green sulfur bacteria (GSBs). The GSBs have larger chlorosomes and more genes encoding chlorosome proteins [21,29]. The integral membrane core antenna complex and a type II (quinone-type) reaction center were transferred either to or from the purple photosynthetic bacteria.

#### (II) (Bacterio)chlorophyll biosynthesis

AcsF and BchE are proposed to be responsible for biosynthesis of the isocyclic ring of (bacterio)chlorophylls under aerobic and anaerobic growth conditions, respectively [28,31,32,62-64]. The *acsF* gene was transferred either to or from purple bacteria and cyanobacteria, and the *bchE* gene either to or from heliobacteria, purple bacteria and GSBs.

#### (III) Electron transfer complexes

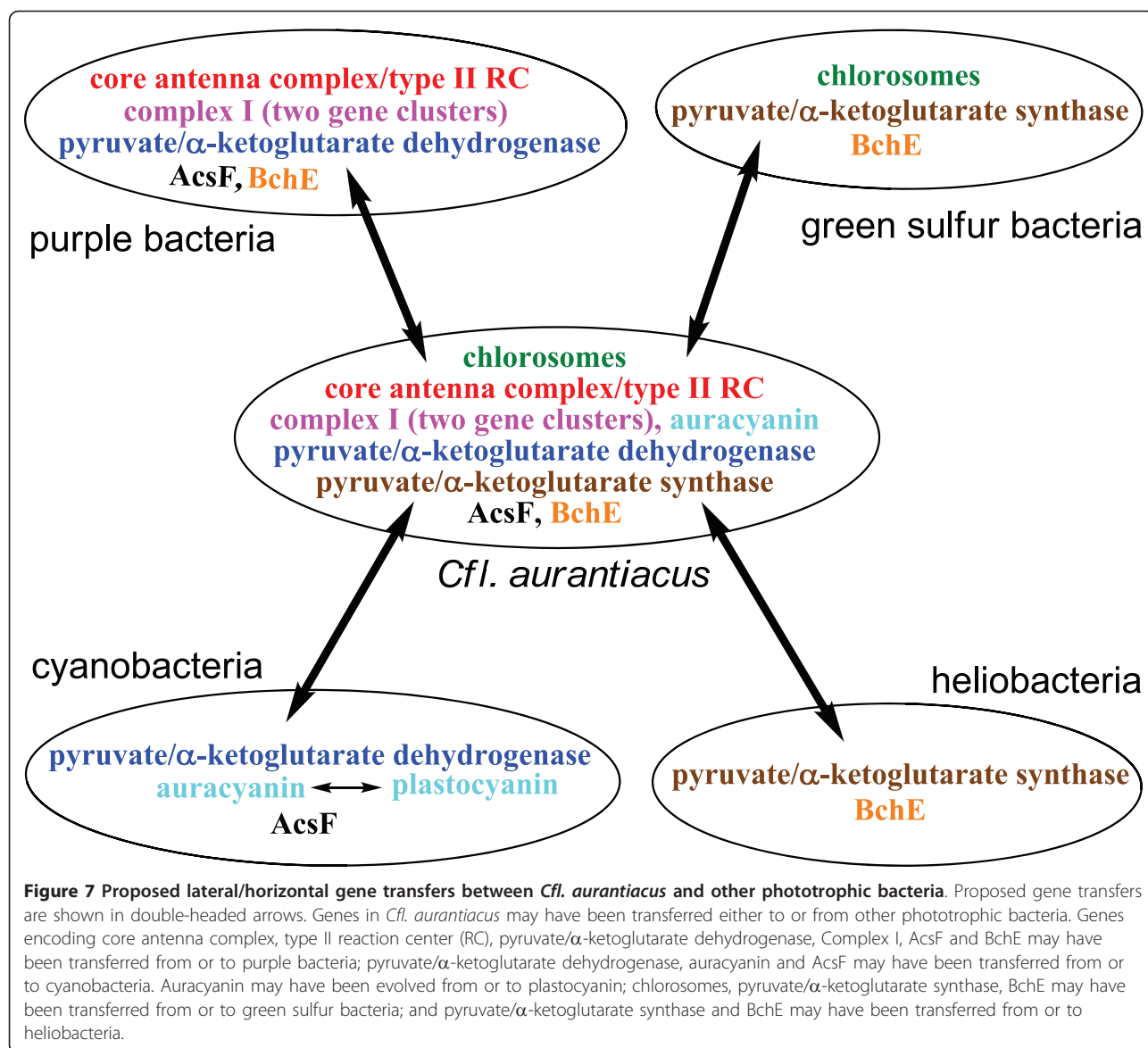
Two gene clusters of the complex I genes were transferred to (some) purple bacteria. Genes encoding auracyanin may have been transferred either to or from cyanobacteria where the type I blue copper protein plastocyanin is found. Alternative complex III (ACIII) may have evolved from or to the cytochrome *bc<sub>1</sub>* or *b<sub>6</sub>/f* complex.

#### (IV) Central carbon metabolism

Genes encoding pyruvate dehydrogenase and  $\alpha$ -ketoglutarate dehydrogenase were transferred either to or from purple bacteria and cyanobacteria, and genes encoding PFOR (or pyruvate synthase) and KFOR (or  $\alpha$ -ketoglutarate synthase) to or from heliobacteria and GSBs. GSBs may have acquired the ATP citrate lyase gene to complete the reductive TCA (RTCA) cycle for CO<sub>2</sub> fixation, and heliobacteria obtained the gene encoding (*Re*)-citrate synthase for synthesizing citrate and operating the partial oxidative TCA (OTCA) cycle [84]. Since *Cfl. aurantiacus* operates the OTCA cycle, the RTCA cycle in GSBs may have evolved from the OTCA cycle [106].

#### Conclusions

The filamentous anoxygenic phototrophic (FAP) bacteria (or the green non-sulfur bacteria) have been suggested to be a critical group in the evolution of photosynthesis. As the first characterized FAP bacterium, the thermophilic bacterium *Chloroflexus aurantiacus* is an amazing organism. It has a chimerical photosystem that comprises characteristic types of green sulfur bacteria and



purple photosynthetic bacteria. It is metabolically versatile, and can grow photoautotrophically and photoheterotrophically under anaerobic growth conditions, and chemotrophically under aerobic growth conditions. Consistent with these physiological and ecological studies, the *Cfl. aurantiacus* genome has duplicated genes and aerobic/anaerobic enzyme pairs in (photosynthetic) electron transport chain, central carbon metabolism, and biosynthesis of tetrapyrroles and nucleic acids. In particular, duplicate genes and gene clusters for two unique proteins and protein complexes in *Cfl. aurantiacus* and several FAP bacteria, the alternative complex III (ACIII) and type I blue copper protein auracyanin, have been identified in *Cfl. aurantiacus* genome. The genomic information and previous biochemical studies also

suggest that *Cfl. aurantiacus* operates diverse carbon assimilation pathways. In contrast to the purple photosynthetic bacteria, the photosynthetic genes are rather spread out in the *Cfl. aurantiacus* chromosome. Overall, the genomic analyses presented in this report, along with previous physiological, ecological and biochemical studies, indicate that *Cfl. aurantiacus* has many interesting and certain unique features in its metabolic pathways.

## Methods

### Gene sequencing

The genome of *Chloroflexus aurantiacus* J-10-fl was sequenced at the Joint Genome Institute (JGI) using a combination of 8-kb and 14-kb DNA libraries. All general



aspects of library construction and sequencing performed at the JGI can be found at <http://www.jgi.doe.gov/>. Draft assemblies were based on 58246 total reads. Both libraries provided 10x coverage of the genome. The Phred/Phrap/Consed software package (<http://www.phrap.com>) was used for sequence assembly and quality assessment [107-109]. After the shotgun stage, reads were assembled with parallel phrap (High Performance Software, LLC). Possible mis-assemblies were corrected with Dupfinisher [110] or transposon bombing of bridging clones (Epicentre Biotechnologies, Madison, WI). Gaps between contigs were closed by editing in Consed, custom primer walk or PCR amplification (Roche Applied Science, Indianapolis, IN). A total of 1893 additional reactions were necessary to close gaps and to raise the quality of the finished sequence. The completed *Cfl. aurantiacus* genome sequence contains 61248 reads, achieving an average of 11-fold sequence coverage per base with an error rate less than 1 in 100,000.

#### Annotation

The genome of *Chloroflexus aurantiacus* J-10-fl has been annotated by the default JGI annotation pipeline. Genes were identified using two gene modeling programs, Glimmer [111] and Critica [112] as part of the Oak Ridge National Laboratory genome annotation pipeline. The two sets of gene calls were combined using Critica as the preferred start call for genes with the same stop codon. Briefly, structural RNAs were predicted using BLASTn and tRNAscan-SE [113] with default prokaryotic settings. Protein-coding genes were identified using gene modeling program Prodigal [114]. Genes with less than 80 amino acids which were predicted by only one of the gene callers and had no Blast hit in the KEGG database at 1e-05 were deleted. Predicted gene models were analyzed using GenePRIMP pipeline [115], and erroneous gene models were manually curated. The revised gene-protein set was searched by BLASTp against the KEGG GENES database [116] and GenBank NR using e-value of 1.0e-05, a minimum of 50% identity and alignment length of at least 80% of both the query and subject protein. These BLASTp hits were used to perform the initial automated functional assignments. In addition, protein sequences were searched against Pfam [117] and TIGRFAM [118] databases using HMMER2 package and trusted cutoffs for each model. Protein sequences were also searched against COG database [119] using RPS-BLAST search with e-value of 1.0e-05 and retaining the best hit. These data sources were combined to assert a product description for each predicted protein. Non-coding genes and miscellaneous features were predicted using tRNAscan-SE [113], TMHMM [120], and signalP [121]. The annotated genome sequence was submitted to GenBank and

loaded into the Integrated Microbial Genomes (IMG) database [122].

#### Phylogenetic analyses

The 16S rRNA gene sequences of various photosynthetic bacteria were obtained from NCBI. The sequences of 16S rRNA genes were aligned using the program Bioedit [123], and the phylogenetic tree was constructed using the program MEGA 4.1 [124]. The tree is an unrooted neighbor joining tree.

#### Abbreviations of phototrophic bacteria

Three-letter abbreviation for the generic name of phototrophic bacteria follows the information listed on LPSN (List of Prokaryotic names with Standing in Nomenclature), an on-line database curated by professor Jean P. Euzéby (<http://www.bacterio.cict.fr/index.html>).

#### Additional material

**Additional file 1: Table S1.** Annotation of photosynthetic genes in *Cfl. aurantiacus* and *Cba. tepidum*.

#### Acknowledgements and Funding

The work conducted by the U.S. Department of Energy Joint Genome Institute is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. The authors thank the contribution of Alex Copeland, Susan Lucas, Tijana Glavina del Rio, Nancy Hammon, Hope N. Tice, Jeremy Schmutz, Thomas S. Brettin, David Bruce, Chris Detter, Nikos C. Kyrpides, and Paul Richardson for gene sequencing, assembling and annotation.

#### Author details

<sup>1</sup>Department of Biology and Department of Chemistry, Campus Box 1137, Washington University in St. Louis, St. Louis, MO 63130, USA. <sup>2</sup>Lawrence Berkeley National Laboratory & Production Genomics Facility, The DOE Joint Genome Institute, Walnut Creek, CA 94598, USA. <sup>3</sup>The DOE Joint Genome Institute and Bioscience Division, M888, Los Alamos National Laboratory, Los Alamos, NM 87544, USA. <sup>4</sup>Computational Biology and Bioinformatics Group, Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA. <sup>5</sup>The DOE Joint Genome Institute, Walnut Creek, CA 94598, USA. <sup>6</sup>Department of Biology, CMB 1088, University of Puget Sound, Tacoma, WA 98416, USA. <sup>7</sup>Current: Baylor College of Medicine, Houston, TX 77030.

#### Authors' contributions

BKP and REB designed research and provided DNA; KB, OC, ED, CSH, LJH, MLL, AL, FWL, NM and SP conducted sequencing, assembly and automated annotation; KHT, BMH, LEK and REB analyzed data; and KHT and REB wrote the paper. All authors read and approved the final manuscript.

#### Competing interests

The authors declare that they have no competing interests.

Received: 1 March 2011 Accepted: 29 June 2011

Published: 29 June 2011

#### References

1. Hanada S, Pierson BK: **The family chloroflexaceae.** In *The Prokaryotes. Volume 7.* 3 edition. Springer, New York; 2006:815-842.
2. Tang KH, Urban VS, Wen J, Xin Y, Blankenship RE: **SANS investigation of the photosynthetic machinery of *Chloroflexus aurantiacus*.** *Biophys J* 2010, **99**:2398-2407.



3. Blankenship RE, Matsuura K: **Antenna complexes from green photosynthetic bacteria.** In *Anoxygenic Photosynthetic Bacteria*. Edited by: Green BR, Parson WW. Kluwer Academic Publishers, Dordrecht; 2003:195-217.
4. Zarzycki J, Brecht V, Muller M, Fuchs G: **Identifying the missing steps of the autotrophic 3-hydroxypropionate CO<sub>2</sub> fixation cycle in *Chloroflexus aurantiacus*.** *Proc Natl Acad Sci USA* 2009, **106**:21317-21322.
5. Klatt CG, Bryant DA, Ward DM: **Comparative genomics provides evidence for the 3-hydroxypropionate autotrophic pathway in filamentous anoxygenic phototrophic bacteria and in hot spring microbial mats.** *Environ Microbiol* 2007, **9**:2067-2078.
6. Pace NR: **A molecular view of microbial diversity and the biosphere.** *Science* 1997, **276**:734-740.
7. Oyaizu H, Debrunner-Vossbrinck B, Mandelco L, Studier JA, Woese CR: **The green non-sulfur bacteria: a deep branching in the eubacterial line of descent.** *Syst Appl Microbiol* 1987, **9**:47-53.
8. Woese CR: **Bacterial evolution.** *Microbiol Rev* 1987, **51**:221-271.
9. Blankenship RE: **Early evolution of photosynthesis.** *Plant Physiol* 2010, **154**:434-438.
10. Blankenship RE, Hartman H: **The origin and evolution of oxygenic photosynthesis.** *Trends Biochem Sci* 1998, **23**:94-97.
11. Ragan MA, Beiko RG: **Lateral genetic transfer: open issues.** *Philos Trans R Soc Lond B Biol Sci* 2009, **364**:2241-2251.
12. Hohmann-Marriott MF, Blankenship RE: **Evolution of photosynthesis.** *Annu Rev Plant Biol* 2011, **62**:515-548.
13. Gupta RS: **Molecular signatures for the main phyla of photosynthetic bacteria and their subgroups.** *Photosynth Res* 2010, **104**:357-372.
14. Olson JM, Blankenship RE: **Thinking about the evolution of photosynthesis.** *Photosynth Res* 2004, **80**:373-386.
15. Olson JM, Pierson BK: **Evolution of reaction centers in photosynthetic prokaryotes.** *Int Rev Cytol* 1987, **108**:209-248.
16. Blankenship RE: **Origin and early evolution of photosynthesis.** *Photosynth Res* 1992, **33**:91-111.
17. Raymond J, Zhaxybayeva O, Gogarten JP, Gerdes SY, Blankenship RE: **Whole-genome analysis of photosynthetic prokaryotes.** *Science* 2002, **298**:1616-1620.
18. Raymond J, Segre D: **The effect of oxygen on biochemical networks and the evolution of complex life.** *Science* 2006, **311**:1764-1767.
19. Falkowski PG: **Evolution. Tracing oxygen's imprint on earth's metabolic evolution.** *Science* 2006, **311**:1724-1725.
20. Raymond J, Blankenship RE: **Biosynthetic pathways, gene replacement and the antiquity of life.** *Geobiology* 2004, **2**:199-203.
21. Bryant DA, Klatt CG, Frigaard NU, Liu Z, Li T, Zhao F, Garcia Costas AM, Overmann J, Ward DM: **Comparative and functional genomics of anoxygenic green bacteria from the taxa *Chlorobi*, *Chloroflexi*, and *Acidobacteria*.** In *Advances in Photosynthesis and Respiration, Functional Genomics and Evolution of Photosynthetic Systems*. Edited by: Burnap RL, Vermaas W. Springer, Dordrecht; 2011; Vol. 33.
22. Hanada S, Takaichi S, Matsuura K, Nakamura K: ***Roseiflexus castenholzii* gen. nov., sp. nov., a thermophilic, filamentous, photosynthetic bacterium that lacks chlorosomes.** *Int J Syst Evol Microbiol* 2002, **52**:187-193.
23. Montaño GA, Xin YY, Lin S, Blankenship RE: **Carotenoid and bacteriochlorophyll energy transfer in the B808-866 complex from *Chloroflexus aurantiacus*.** *J Phys Chem B* 2004, **108**:10607-10611.
24. Blankenship RE: **Molecular mechanisms of photosynthesis.** Blackwell Science Ltd, Oxford; 2002.
25. Pierson BK, Thornber JP: **Isolation and spectral characterization of photochemical reaction centers from the thermophilic green bacterium *Chloroflexus aurantiacus* strain J-10-f1.** *Proc Natl Acad Sci USA* 1983, **80**:80-84.
26. Shiozawa JA, Lottspeich F, Feick R: **The photochemical reaction center of *Chloroflexus aurantiacus* is composed of two structurally similar polypeptides.** *Eur J Biochem* 1987, **167**:595-600.
27. Swingle WD, Sadekar S, Mastrian SD, Matthies HJ, Hao J, Ramos H, Acharya CR, Conrad AL, Taylor HL, Dejesa LC, et al: **The complete genome sequence of *Roseobacter denitrificans* reveals a mixotrophic rather than photosynthetic metabolism.** *J Bacteriol* 2007, **189**:683-690.
28. Tang KH, Wen J, Li X, Blankenship RE: **Role of the AcsF protein in *Chloroflexus aurantiacus*.** *J Bacteriol* 2009, **191**:3580-3587.
29. Frigaard NU, Chew AGM, Maresca JA, Bryant DA: **Bacteriochlorophyll Biosynthesis in Green Bacteria.** In *Chlorophylls and Bacteriochlorophylls*. Edited by: Grimm B, Porra RJ, Rudiger W, Scheer H. Springer Academic Publishers, Dordrecht; 2006:201-221.
30. Li H, Bryant DA: **Envelope proteins of the CsmB/CsmF and CsmC/CsmD motif families influence the size, shape, and composition of chlorosomes in *Chlorobaculum tepidum*.** *J Bacteriol* 2009, **191**:7109-7120.
31. Ouchane S, Steunou AS, Picaud M, Astier C: **Aerobic and anaerobic Mg-protoporphyrin monomethyl ester cyclases in purple bacteria: a strategy adopted to bypass the repressive oxygen control system.** *J Biol Chem* 2004, **279**:6385-6394.
32. Pinta V, Picaud M, Reiss-Husson F, Astier C: **Rubrivivax gelatinosus acsF (previously orf358) codes for a conserved, putative binuclear-iron-cluster-containing protein involved in aerobic oxidative cyclization of Mg-protoporphyrin IX monomethylester.** *J Bacteriol* 2002, **184**:746-753.
33. Hale MB, Blankenship RE, Fuller RC: **Menaquinone Is the sole quinone in the facultatively aerobic green photosynthetic bacterium *Chloroflexus aurantiacus*.** *Biochim Biophys Acta* 1983, **723**:376-382.
34. Gao X, Xin Y, Blankenship RE: **Enzymatic activity of the alternative complex III as a menaquinol:auracyanin oxidoreductase in the electron transfer chain of *Chloroflexus aurantiacus*.** *FEBS Lett* 2009, **583**:3275-3279.
35. Lee M, del Rosario MC, Harris HH, Blankenship RE, Guss JM, Freeman HC: **The crystal structure of auracyanin A at 1.85 Å resolution: the structures and functions of auracyanins A and B, two almost identical "blue" copper proteins, in the photosynthetic bacterium *Chloroflexus aurantiacus*.** *J Biol Inorg Chem* 2009, **14**:329-345.
36. Shuvalov VA, Vasmel H, Ames J, Duysens LNM: **Picosecond spectroscopy of the charge separation in reaction centers of *Chloroflexus aurantiacus* with selective excitation of the primary electron donor.** *Biochim Biophys Acta* 1986, **851**:361-368.
37. Okamura MY, Isaacson RA, Feher G: **Primary acceptor in bacterial photosynthesis: obligatory role of ubiquinone in photoactive reaction centers of *Rhodospseudomonas spheroides*.** *Proc Natl Acad Sci USA* 1975, **72**:3491-3495.
38. McAuley KE, Fyfe PK, Ridge JP, Cogdell RJ, Isaacs NW, Jones MR: **Ubiquinone binding, ubiquinone exclusion, and detailed cofactor conformation in a mutant bacterial reaction center.** *Biochemistry* 2000, **39**:15032-15043.
39. Schoepp-Cothenet B, Lieutaud C, Baymann F, Vermeglio A, Friedrich T, Kramer DM, Nitschke W: **Menaquinone as pool quinone in a purple bacterium.** *Proc Natl Acad Sci USA* 2009, **106**:8549-8554.
40. Yanyushin MF: **Fractionation of cytochromes of phototrophically grown *Chloroflexus aurantiacus*. Is there a cytochrome bc complex among them?** *FEBS Lett* 2002, **512**:125-128.
41. Yanyushin MF, del Rosario MC, Brune DC, Blankenship RE: **New class of bacterial membrane oxidoreductases.** *Biochemistry* 2005, **44**:10037-10045.
42. Pereira MM, Refojo PN, Hreggvidsson GO, Hjelreidsdottir S, Teixeira M: **The alternative complex III from *Rhodothermus marinus* - a prototype of a new family of quinol:electron acceptor oxidoreductases.** *FEBS Lett* 2007, **581**:4831-4835.
43. Refojo PN, Sousa FL, Teixeira M, Pereira MM: **The alternative complex III: a different architecture using known building modules.** *Biochim Biophys Acta* 2010, **1797**:1869-1876.
44. Gao X, Xin Y, Bell PD, Wen J, Blankenship RE: **Structural analysis of alternative complex III in the photosynthetic electron transfer chain of *Chloroflexus aurantiacus*.** *Biochemistry* 2010, **49**:6670-6679.
45. Bryant DA, Costas AM, Maresca JA, Chew AG, Klatt CG, Bateson MM, Tallon LJ, Hostetler J, Nelson WC, Heidelberg JF, Ward DM: ***Candidatus Chloracidobacterium thermophilum*: an aerobic phototrophic *Acidobacterium*.** *Science* 2007, **317**:523-526.
46. Hussain H, Grove J, Griffiths L, Busby S, Cole J: **A seven-gene operon essential for formate-dependent nitrite reduction to ammonia by enteric bacteria.** *Mol Microbiol* 1994, **12**:153-163.
47. McManus JD, Brune DC, Han J, Sanders-Loehr J, Meyer TE, Cusanovich MA, Tollin G, Blankenship RE: **Isolation, characterization, and amino acid sequences of auracyanins, blue copper proteins from the green photosynthetic bacterium *Chloroflexus aurantiacus*.** *J Biol Chem* 1992, **267**:6531-6540.
48. Trost JT, Mcmanus JD, Freeman JC, Ramakrishna BL, Blankenship RE: **Auracyanin, a blue copper protein from the green photosynthetic bacterium *Chloroflexus aurantiacus*.** *Biochemistry* 1988, **27**:7858-7863.
49. Bond CS, Blankenship RE, Freeman HC, Guss JM, Maher MJ, Selvaraj FM, Wilce MC, Willingham KM: **Crystal structure of auracyanin, a "blue" copper**

- protein from the green thermophilic photosynthetic bacterium *Chloroflexus aurantiacus*. *J Mol Biol* 2001, **306**:47-67.
50. Tsukatani Y, Nakayama N, Shimada K, Mino H, Itoh S, Matsuura K, Hanada S, Nagashima KV: Characterization of a blue-copper protein, auracyanin, of the filamentous anoxygenic phototrophic bacterium *Roseiflexus castenholzii*. *Arch Biochem Biophys* 2009, **490**:57-62.
  51. Wu D, Raymond J, Wu M, Chatterji S, Ren Q, Graham JE, Bryant DA, Robb F, Colman A, Tallon LJ, et al: Complete genome sequence of the aerobic CO-oxidizing thermophile *Thermomicrobium roseum*. *PLoS One* 2009, **4**: e4207.
  52. Dracheva S, Williams JC, Van Driessche G, Van Beeumen JJ, Blankenship RE: The primary structure of cytochrome c554 from the green photosynthetic bacterium *Chloroflexus aurantiacus*. *Biochemistry* 1991, **30**:11451-11458.
  53. Meyer TE, Tollin G, Cusanovich MA, Freeman JC, Blankenship RE: In vitro kinetics of reduction of cytochrome c554 isolated from the reaction center of the green phototrophic bacterium, *Chloroflexus aurantiacus*. *Arch Biochem Biophys* 1989, **272**:254-261.
  54. Freeman JC, Blankenship RE: Isolation and characterization of the membrane-bound cytochrome c554 from the thermophilic green photosynthetic bacterium *Chloroflexus aurantiacus*. *Photosynth Res* 1990, **23**:29-38.
  55. Yamada M, Zhang H, Hanada S, Nagashima KV, Shimada K, Matsuura K: Structural and spectroscopic properties of a reaction center complex from the chlorosome-lacking filamentous anoxygenic phototrophic bacterium *Roseiflexus castenholzii*. *J Bacteriol* 2005, **187**:1702-1709.
  56. Eisen JA, Nelson KE, Paulsen IT, Heidelberg JF, Wu M, Dodson RJ, Deboy R, Gwinn ML, Nelson WC, Haft DH, et al: The complete genome sequence of *Chlorobium tepidum* TLS, a photosynthetic, anaerobic, green-sulfur bacterium. *Proc Natl Acad Sci USA* 2002, **99**:9509-9514.
  57. Roth JR, Lawrence JG, Bobik TA: Cobalamin (coenzyme B12): synthesis and biological significance. *Annu Rev Microbiol* 1996, **50**:137-181.
  58. Sattley WM, Madigan MT, Swingle WD, Cheung PC, Clocksin KM, Conrad AL, Dejesa LC, Honchak BM, Jung DO, Karbach LE, et al: The genome of *Heliobacterium modesticaldum*, a phototrophic representative of the *Firmicutes* containing the simplest photosynthetic apparatus. *J Bacteriol* 2008, **190**:4687-4696.
  59. Pierson BK: Cytochromes in *Chloroflexus aurantiacus* grown with and without oxygen. *Arch Microbiol* 1985, **143**:260-265.
  60. Igarashi N, Harada J, Nagashima S, Matsuura K, Shimada K, Nagashima KV: Horizontal transfer of the photosynthesis gene cluster and operon rearrangement in purple bacteria. *J Mol Evol* 2001, **52**:333-341.
  61. Beatty JT: Organization of photosynthesis gene transcripts. In *Anoxygenic Photosynthetic Bacteria*. Volume 2. Edited by: Blankenship R. E., Madigan M. T., Bauer C. E. Kluwer Academic Publishers, Dordrecht, The Netherlands; 1995:1209-1219.
  62. Minamizaki K, Mizoguchi T, Goto T, Tamiaki H, Fujita Y: Identification of two homologous genes, *chlAI* and *chlAII*, that are differentially involved in isocyclic ring formation of chlorophyll *a* in the cyanobacterium *Synechocystis* sp. PCC 6803. *J Biol Chem* 2008, **283**:2684-2692.
  63. Tottey S, Block MA, Allen M, Westergren T, Albrieux C, Scheller HV, Merchant S, Jensen PE: Arabidopsis CHL27, located in both envelope and thylakoid membranes, is required for the synthesis of protochlorophyllide. *Proc Natl Acad Sci USA* 2003, **100**:16119-16124.
  64. Gough SP, Petersen BO, Duus JO: Anaerobic chlorophyll isocyclic ring formation in *Rhodobacter capsulatus* requires a cobalamin cofactor. *Proc Natl Acad Sci USA* 2000, **97**:6908-6913.
  65. Lopez JC, Ryan S, Blankenship RE: Sequence of the *bchG* gene from *Chloroflexus aurantiacus*: relationship between chlorophyll synthase and other polyprenyltransferases. *J Bacteriol* 1996, **178**:3369-3373.
  66. Adlessee HA, Fiedor L, Hunter CN: Physical mapping of *bchG*, *orf427*, and *orf177* in the photosynthesis gene cluster of *Rhodobacter sphaeroides*: functional assignment of the bacteriochlorophyll synthetase gene. *J Bacteriol* 2000, **182**:3175-3182.
  67. Frigaard NU, Voigt GD, Bryant DA: *Chlorobium tepidum* mutant lacking bacteriochlorophyll *c* made by inactivation of the *bchK* gene, encoding bacteriochlorophyll *c* synthase. *J Bacteriol* 2002, **184**:3368-3376.
  68. Maresca JA, Gomez Maqueo Chew A, Ponsati MR, Frigaard NU, Ormerod JG, Bryant DA: The *bchU* gene of *Chlorobium tepidum* encodes the c-20 methyltransferase in bacteriochlorophyll *c* biosynthesis. *J Bacteriol* 2004, **186**:2558-2566.
  69. Gomez Maqueo Chew A, Frigaard NU, Bryant DA: Identification of the *bchP* gene, encoding geranylgeranyl reductase in *Chlorobaculum tepidum*. *J Bacteriol* 2008, **190**:747-749.
  70. Jordan A, Torrents E, Jeanthon C, Eliasson R, Hellman U, Wernstedt C, Barbe J, Gilbert I, Reichard P: B12-dependent ribonucleotide reductases from deeply rooted eubacteria are structurally related to the aerobic enzyme from *Escherichia coli*. *Proc Natl Acad Sci USA* 1997, **94**:13487-13492.
  71. Tabita FR: The hydroxypropionate pathway of CO<sub>2</sub> fixation: *Fait accompli*. *Proc Natl Acad Sci USA* 2009, **106**:21015-21016.
  72. Strauss G, Fuchs G: Enzymes of a novel autotrophic CO<sub>2</sub> fixation pathway in the phototrophic bacterium *Chloroflexus aurantiacus*, the 3-hydroxypropionate cycle. *Eur J Biochem* 1993, **215**:633-643.
  73. Eisenreich W, Strauss G, Werz U, Fuchs G, Bacher A: Retrobiosynthetic analysis of carbon fixation in the phototrophic eubacterium *Chloroflexus aurantiacus*. *Eur J Biochem* 1993, **215**:619-632.
  74. Holo H: *Chloroflexus aurantiacus* secretes 3-hydroxypropionate, a possible intermediate in the assimilation of CO<sub>2</sub> and acetate. *Arch Microbiol* 1989, **151**:252-256.
  75. Huber H, Gallenberger M, Jahn U, Eylert E, Berg IA, Kockelkorn D, Eisenreich W, Fuchs G: A dicarboxylate/4-hydroxybutyrate autotrophic carbon assimilation cycle in the hyperthermophilic Archaeum *Ignicoccus hospitalis*. *Proc Natl Acad Sci USA* 2008, **105**:7851-7856.
  76. Berg IA, Kockelkorn D, Buckel W, Fuchs G: A 3-hydroxypropionate/4-hydroxybutyrate autotrophic carbon dioxide assimilation pathway in Archaea. *Science* 2007, **318**:1782-1786.
  77. Alber BE, Kung JW, Fuchs G: 3-Hydroxypropionyl-coenzyme A synthetase from *Metallosphaera sedula*, an enzyme involved in autotrophic CO<sub>2</sub> fixation. *J Bacteriol* 2008, **190**:1383-1389.
  78. Hugler M, Huber H, Stetter KO, Fuchs G: Autotrophic CO<sub>2</sub> fixation pathways in archaea (Crenarchaeota). *Arch Microbiol* 2003, **179**:160-173.
  79. Menendez C, Bauer Z, Huber H, Gad'on N, Stetter KO, Fuchs G: Presence of acetyl coenzyme A (CoA) carboxylase and propionyl-CoA carboxylase in autotrophic Crenarchaeota and indication for operation of a 3-hydroxypropionate cycle in autotrophic carbon fixation. *J Bacteriol* 1999, **181**:1088-1098.
  80. Hugler M, Menendez C, Schagger H, Fuchs G: Malonyl-coenzyme A reductase from *Chloroflexus aurantiacus*, a key enzyme of the 3-hydroxypropionate cycle for autotrophic CO(2) fixation. *J Bacteriol* 2002, **184**:2404-2410.
  81. Alber BE, Fuchs G: Propionyl-coenzyme A synthase from *Chloroflexus aurantiacus*, a key enzyme of the 3-hydroxypropionate cycle for autotrophic CO<sub>2</sub> fixation. *J Biol Chem* 2002, **277**:12137-12143.
  82. Feng X, Tang KH, Blankenship RE, Tang YJ: Metabolic flux analysis of the mixotrophic metabolisms in the green sulfur bacterium *Chlorobaculum tepidum*. *J Biol Chem* 2010, **285**:39544-39550.
  83. Tang KH, Blankenship RE: Both forward and reverse TCA cycles operate in green sulfur bacteria. *J Biol Chem* 2010, **285**:35848-35854.
  84. Tang KH, Feng X, Zhuang WQ, Alvarez-Cohen L, Blankenship RE, Tang YJ: Carbon flow of heliobacteria is related more to *clostridia* than to the green sulfur bacteria. *J Biol Chem* 2010, **285**:35104-35112.
  85. Tang KH, Yue H, Blankenship RE: Energy metabolism of *Heliobacterium modesticaldum* during phototrophic and chemotrophic growth. *BMC Microbiol* 2010, **10**:150.
  86. Tang KH, Feng X, Tang YJ, Blankenship RE: Carbohydrate metabolism and carbon fixation in *Roseobacter denitrificans* OCH114. *PLoS One* 2009, **4**: e7233.
  87. Evans MC, Buchanan BB, Arnon DI: New cyclic process for carbon assimilation by a photosynthetic bacterium. *Science* 1966, **152**:673.
  88. Loken O, Sirevag R: Evidence for the presence of the glyoxylate cycle in *Chloroflexus aurantiacus*. *Arch Microbiol* 1982, **132**:276-279.
  89. Bateson MM, Ward DM: Photoexcretion and fate of glycolate in a hot spring cyanobacterial mat. *Appl Environ Microbiol* 1988, **54**:1738-1743.
  90. Krasilnikova EN, Keppen OI, Gorlenko VM, Kondratieva EN: Growth of *Chloroflexus aurantiacus* on media with different organic compounds and pathways of their metabolism. *Microbiology* 1986, **55**:325-329.
  91. Madigan MT, Petersen SR, Brock TD: Nutritional studies on *Chloroflexus*, a filamentous photosynthetic, gliding bacterium. *Arch Microbiol* 1974, **100**:97-103.
  92. Kondratieva EN, Ivanovsky RN, Krasilnikova EN: Carbon metabolism in *Chloroflexus aurantiacus*. *FEMS Microbiol Lett* 1992, **100**:269-271.

93. Krasilnikova EN, Kondratieva EN: **Growth of *Chloroflexus aurantiacus* under anaerobic conditions in the dark and the metabolism of organic substrates.** *Microbiology* 1987, **56**:281-285.
94. Say RF, Fuchs G: **Fructose 1,6-bisphosphate aldolase/phosphatase may be an ancestral gluconeogenic enzyme.** *Nature* 2010, **464**:1077-1081.
95. Heda GD, Madigan MT: **Utilization of amino acids and lack of diazotrophy in the thermophilic anoxygenic phototroph *Chloroflexus aurantiacus*.** *J Gen Microbiol* 1986, **132**:2469-2473.
96. Laakmannidiges G, Klemme JH: **Amino acid metabolism in the thermophilic phototroph, *Chloroflexus aurantiacus* - properties and metabolic role of 2 L-threonine (L-serine) dehydratases.** *Arch Microbiol* 1988, **149**:249-254.
97. Howell DM, Xu H, White RH: **(R)-citramalate synthase in methanogenic archaea.** *J Bacteriol* 1999, **181**:331-333.
98. Risso C, Van Dien SJ, Orloff A, Lovley DR, Coppi MV: **Elucidation of an alternate isoleucine biosynthesis pathway in *Geobacter sulfurreducens*.** *J Bacteriol* 2008, **190**:2266-2274.
99. Krasil'nikova EN: **ATP sulfurylase activity in *Chloroflexus aurantiacus* and other photosynthesizing bacteria as a function of temperature.** *Microbiology* 1987, **55**:418-421.
100. Madigan MT, Brock TD: **Photosynthetic sulfide oxidation by *Chloroflexus aurantiacus*, a filamentous, photosynthetic, gliding bacterium.** *J Bacteriol* 1975, **122**:782-784.
101. Castenholz RW: **The possible photosynthetic use of sulfide by the filamentous phototrophic bacteria of hot springs.** *Limnol Oceanogr* 1973, **18**:863-876.
102. Castenholz RW, Pierson BK: **Ecology of thermophilic anoxygenic phototrophs.** In *Anoxygenic Photosynthetic Bacteria. Volume 2*. Edited by: Blankenship RE, Madigan MT, Bauer CE. Kluwer Academic Publishers, Dordrecht, The Netherlands; 1995:87-103.
103. Shahak Y, Hauska G: **Sulfide oxidation from cyanobacteria to humans: Sulfide-quinone oxidoreductase (SQR).** In *Advances in Photosynthesis and Respiration. Volume 27*. Edited by: Hell R, Dahl C, Knaff DB, Leustek T. Springer, Heidelberg; 2008:319-335.
104. Griesbeck C, Schutz M, Schodl T, Bathe S, Nausch L, Mederer N, Vielreicher M, Hauska G: **Mechanism of sulfide-quinone reductase investigated using site-directed mutagenesis and sulfur analysis.** *Biochemistry* 2002, **41**:11552-11565.
105. Sakurai H, Ogawa T, Shiga M, Inoue K: **Inorganic sulfur oxidizing system in green sulfur bacteria.** *Photosynth Res* 2010, **104**:163-176.
106. Aoshima M: **Novel enzyme reactions related to the tricarboxylic acid cycle: phylogenetic/functional implications and biotechnological applications.** *Appl Microbiol Biotechnol* 2007, **75**:249-255.
107. Ewing B, Green P: **Base-calling of automated sequencer traces using phred. II. Error probabilities.** *Genome Res* 1998, **8**:186-194.
108. Ewing B, Hillier L, Wendl MC, Green P: **Base-calling of automated sequencer traces using phred. I. Accuracy assessment.** *Genome Res* 1998, **8**:175-185.
109. Gordon D, Abajian C, Green P: **Consed: a graphical tool for sequence finishing.** *Genome Res* 1998, **8**:195-202.
110. Han CS, Chain P: **Finishing repeat regions automatically with Dupfinisher.** In *Proceeding of the 2006 international conference on bioinformatics & computational biology*. Edited by: Arabnia HR, Valafar H. CSREA Press; 2006:141-146.
111. Delcher AL, Bratke KA, Powers EC, Salzberg SL: **Identifying bacterial genes and endosymbiont DNA with Glimmer.** *Bioinformatics* 2007, **23**:673-679.
112. Badger JH, Olsen GJ: **CRITICA: coding region identification tool invoking comparative analysis.** *Mol Biol Evol* 1999, **16**:512-524.
113. Lowe TM, Eddy SR: **tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence.** *Nucleic Acids Res* 1997, **25**:955-964.
114. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ: **Prodigal: prokaryotic gene recognition and translation initiation site identification.** *BMC Bioinformatics* 2010, **11**:119.
115. Pati A, Ivanova NN, Mikhailova N, Ovchinnikova G, Hooper SD, Lykidis A, Kyrpides NC: **GenePRIMP: a gene prediction improvement pipeline for prokaryotic genomes.** *Nat Methods* 2010, **7**:455-457.
116. Kanehisa M, Goto S, Kawashima S, Okuno Y, Hattori M: **The KEGG resource for deciphering the genome.** *Nucleic Acids Res* 2004, **32**:D277-280.
117. Bateman A, Coin L, Durbin R, Finn RD, Hollich V, Griffiths-Jones S, Khanna A, Marshall M, Moxon S, Sonnhammer EL, et al: **The Pfam protein families database.** *Nucleic Acids Res* 2004, **32**:D138-141.
118. Haft DH, Loftus BJ, Richardson DL, Yang F, Eisen JA, Paulsen IT, White O: **TIGRFAMs: a protein family resource for the functional identification of proteins.** *Nucleic Acids Res* 2001, **29**:41-43.
119. Tatusov RL, Fedorova ND, Jackson JD, Jacobs AR, Kiryutin B, Koonin EV, Krylov DM, Mazumder R, Mekhedov SL, Nikolskaya AN, et al: **The COG database: an updated version includes eukaryotes.** *BMC Bioinformatics* 2003, **4**:41.
120. Krogh A, Larsson B, von Heijne G, Sonnhammer EL: **Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes.** *J Mol Biol* 2001, **305**:567-580.
121. Bendtsen JD, Nielsen H, von Heijne G, Brunak S: **Improved prediction of signal peptides: SignalP 3.0.** *J Mol Biol* 2004, **340**:783-795.
122. Markowitz VM, Korzeniewski F, Palaniappan K, Szeto E, Werner G, Padki A, Zhao X, Dubchak I, Hugenholtz P, Anderson I, et al: **The integrated microbial genomes (IMG) system.** *Nucleic Acids Res* 2006, **34**:D344-348.
123. Hall TA: **BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT.** *Nucl Acids Symp Ser* 1999, **41**:95-98.
124. Tamura K, Dudley J, Nei M, Kumar S: **MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0.** *Mol Biol Evol* 2007, **24**:1596-1599.

doi:10.1186/1471-2164-12-334

**Cite this article as:** Tang et al.: Complete genome sequence of the filamentous anoxygenic phototrophic bacterium *Chloroflexus aurantiacus*. *BMC Genomics* 2011 **12**:334.

**Submit your next manuscript to BioMed Central and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
www.biomedcentral.com/submit

