

SCIENTIFIC REPORTS



OPEN

Comparative genome analysis of marine purple sulfur bacterium *Marichromatium gracile* YL28 reveals the diverse nitrogen cycle mechanisms and habitat-specific traits

Bitong Zhu¹, Xiaobo Zhang¹, Chungui Zhao¹, Shicheng Chen² & Suping Yang¹

Mangrove ecosystems are characteristic of the high salinity, limited nutrients and S-richness. *Marichromatium gracile* YL28 (YL28) isolated from mangrove tolerates the high concentrations of nitrite and sulfur compounds and efficiently eliminates them. However, the molecular mechanisms of nitrite and sulfur compounds utilization and the habitat adaptation remain unclear in YL28. We sequenced YL28 genome and further performed the comparative genome analysis in 36 purple bacteria including purple sulfur bacteria (PSB) and purple non-sulfur bacteria (PNSB). YL28 has 6 nitrogen cycle pathways (up to 40 genes), and possibly removes nitrite by denitrification, complete assimilation nitrate reduction and fermentative nitrate reduction (DNRA). Comparative genome analysis showed that more nitrogen utilization genes were detected in PNSB than those in PSB. The partial denitrification pathway and complete assimilation nitrate reduction were reported in PSB and DNRA was reported in purple bacteria for the first time. The three sulfur metabolism genes such as oxidation of sulfide, reversed dissimilatory sulfite reduction and *sox* system allowed to eliminate toxic sulfur compounds in the mangrove ecosystem. Several unique stress response genes facilitate to the tolerance of the high salinity environment. The CRISPR systems and the transposon components in genomic islands (GIs) likely contribute to the genome plasticity in purple bacteria.

Purple bacteria are anoxygenic and phototrophic bacteria, including purple sulfur bacteria (PSB) and purple non-sulfur bacteria (PNSB). They are ubiquitously distributed in different natural environments with versatile metabolism potentials. Purple bacteria also serve as models for clarifying the biogeochemistry of C, N, S and Fe in the earth evolution¹. YL28, a member of PSB, was isolated from mangrove special ecosystem². It synthesized abundant rhodopin (not spirilloxanthin) carotenoid component under anaerobic in the light condition. It was also capable of using reduced sulfur compounds, nitrogen compounds or molecular hydrogen as electron donors². Moreover, YL28 utilized ammonium, nitrite or nitrate as the sole nitrogen source for phototrophic growth. It is interesting that it was found to completely remove nitrite (up to 200 mg/L)^{3,4} and perform simultaneous heterotrophic nitrification and denitrification under anaerobic conditions⁵.

Nitrogen is one of the most abundant elements on earth. It comprises the majority of earth's atmosphere and functions as one of the primary nutrients⁶. Nitrogen cycle is the most complex biogeochemical one in the biosphere^{7,8}. However, the nitrogen cycle has been drastically disrupted because of overexploitation and modern agricultural activities⁹. Nearly half of the nitrogen reaches the coastal ocean via river input and/or atmospheric deposition¹⁰. This leads to extensive eutrophication of waters and coastal zones and increases inventories of the

¹Department of Bioengineering and Biotechnology, Huaqiao University, Xiamen, 361021, China. ²Department of Microbiology and Molecular Genetics, Michigan State University, East Lansing, Michigan, 48863, USA. Correspondence and requests for materials should be addressed to S.C. (email: shicheng@msu.edu) or S.Y. (email: yangsuping@hqu.edu.cn)

potent greenhouse gas (such as nitrous oxide). The elevated concentrations of $\text{NH}_3\text{-N}$ and $\text{NO}_2\text{-N}$ (the major pollutants) are problematic in aquatic ecosystems¹¹.

Up to 10 nitrogen cycle pathways were reported⁹. For examples, purple bacteria have three nitrogen reduction cycles (reduced nitrogen compounds as electron acceptor, such as nitrogen fixation, nitrate assimilation, denitrification) and one nitrogen oxidation cycle (nitrite oxidation to nitrate as electron donor for photosynthesis)^{12–15}. However, nitrification or dissimilatory nitrate reduction to ammonium (DNRA) pathway has not been reported in purple bacteria yet. Denitrification and assimilation nitrate reduction cycles were only reported in PNSB, but unknown in PSB¹⁶. Currently, the molecular mechanism of nitrogen cycles in purple bacteria remain unclear^{17–19}. To elucidate nitrogen cycles in purple bacteria, we sequenced the genome of YL28 as a representative strain, compared the nitrogen cycle genes to other 35 sequenced purple bacteria. The genes involved in the sulfur metabolism, salt tolerance and the stress response were also investigated in details to gain the insights of the surviving mechanisms of YL28 in the highly selective environments.

Results

Genome features and phylogenetic inference. The general features of genome of YL28 and the other 35 sequenced purple bacteria were presented in Table 1. Removal of short contigs and sequences with potential contamination resulted in 120 contigs with a N_{50} of 93,809 bp. The largest contig size was 176,489 bp. The assembled genome comprised 3.8 million nucleotides with a GC content of 68.84%. Gene annotation was carried out by the NCBI Prokaryotic Genome Annotation Pipeline (PGAP). A total of 3,361 genes yielded a coding capacity of 3.3 million nucleotides (genes/genome, 86.4%). There were at least 69 RNA sequences, including 56 tRNAs, 9 rRNAs, and 4 function unknown RNAs in the genome. Among the protein-encoding genes, 1,752 of them could be assigned putative functions while 1,612 were predicted to encode hypothetical proteins. At least 1,671 proteins were assigned to 25 different functional categories with 368 subsystems using SEED subsystems by RAST analysis (Supplementary Fig. SF1). 40 genes involved in nitrogen metabolism were found (Supplementary Fig. SF1). Compared to its relative *M. purpuratum* 984, YL28 has more genes for nitrogen metabolism, photosynthesis and sulfur metabolism (Supplementary Fig. SF2, Tables 2 and 3).

No plasmid sequences were discovered in YL28. However, at least one was found in *Rhodovulum* spp, *Rhodobacter* spp, *Allochrochromatium* spp, *R. rubrum* ATCC 11170, *T. mobilis* 8321 and *R. palustris* CGA009 (Table 1). Moreover, most of the *Rhodobacter* strains possesses up to two chromosomes. The genomic size of the purple bacteria ranges from 2.6 to 5.4 Mb. Except for *T. mobilis* 8321 and *T. violascens* DSM 198, the genome size in PSB is smaller than that in PNSB (Table 1). Overall, genome size in most of PSB is smaller than 4.0 MB, with the smallest one (2.68 Mb) in *Halorhodospira halophila* SL1. Except for *Blastochloris*, genome size of PNSB is generally larger than 4.0 MB; among them, *Rhodoplanes* sp. Z2-YC6860 has the largest genome size (8.19 Mb). Collectively, a high GC content (>55%) is often observed in purple bacteria.

Phylogenetic inference. The phylogenetic analysis based on 16S rDNA sequence (Fig. 1A) showed that PSB and PNSB derived from a common ancestor and formed two clades, which was consistent with the previous studies²⁰. 28 PNSB strains were clustered into three subdivisions: the first one consisted of *Rhodopseudomonas*, *Blastochloris*, *Rhodoplanes* and *Rhodomicrobium vannielii* ATCC 17100; the second one included *Rhodobacter* (freshwater species) and *Rhodovulum* (marine species); the third one had *Rhodospirillum* and *Rhodocista*. 7 PSB species were clustered together on the same clade. Both the phylogenetic trees based on whole genome or core genome (Fig. 1B,C) showed that some PSB species were included in PNSB. Interestingly, *Rubrivivax gelatinosus* IL144 (a member of PNSB) was the closest phylogenetic neighbors of PSB based on 16S rDNA sequence and core genome.

Gene repertoire of PSB. The power trend line has not reached the plateau (Fig. 2A), demonstrating that PSB displays an open accessory-genome. The core genome analysis of PSB shows that the numbers of shared genes decreases with the addition of the input genomes and is predicted to converge against 550 (Fig. 2A). Singleton development plot data indicates that up to 1594 new genes are expected with each newly added genome (Fig. 2B). The core genomes for the seven PSB species contain at least 539 CDSs per genome. The current core genome represents the PSB quite well. YL28 and *M. purpuratum* 984 share at least 2869 genes (data not shown); only 241 and 210 genes uniquely occur in YL28 and 984, respectively. The shared gene numbers in genome between YL28 and other six PSB species range from 732 to 2869 genes (data not shown).

Nitrogen cycle genes in YL28 and purple bacteria. YL28 possesses six nitrogen cycle pathways (Fig. 3 and Table 2). Two genes encoding NirS and NorBC (the key enzymes for nitrite reduction) exists in YL28. These two enzymes in YL28 possibly contribute to transform toxic nitrite into lower cytotoxic N_2O by denitification⁵. The genes encoding NasAB and NirA (the key enzymes for assimilation nitrite reduction) were also detected in YL28, suggesting that it had the complete nitrate reduction ability. YL28 has the genes encoding Nrf which is critical in fermentative nitrate reduction (DNRA). However, the DNRA pathway in YL28 may be incomplete due to the lack of entire complement enzymes. Ammonium assimilation and ammonification related genes were detected in all purple bacteria. Additional four nitrogen cycle pathways display diverse metabolism traits in examined strains (Table 2). The presence of genes encoding NifDKH key enzymes in all purple bacteria shows this group of bacteria has nitrogen fixation potentials. Interestingly, the alternative nitrogenase (AnfG) from *Methanosarcina* was only found in PNSB but not in PSB²¹.

Sulfur metabolism genes in YL28. YL28 has genes involved in at least three sulfur metabolism pathways such as oxidation of sulfide, reversed dissimilatory sulfite reduction and sox system (Table 3). The *sqr* and the *dsr* gene clusters (*dsrABEFHCKMLJOPNR*) are the key enzyme genes for the oxidation of sulfide and the reversed dissimilatory sulfite reduction. These process oxidize toxic sulfide into S^0 or sulfate²⁰. In addition, *sox* gene clusters

Organisms	Chromosome	Plasmid	Size (Mb)	Chr 1 (Mb)	Chr 2 (Mb)	G + C %	Protein	gene	Accession
YL28	1	—	3.84	3.84	—	68.9	3,188	3,360	LSYU00000000.1
<i>Thiocystis violascens</i> DSM 198	1	—	5.02	5.02	—	62.6	4,261	4,465	CP003154.1
<i>Allochrochromatium vinosum</i> DSM 180	1	2	3.67	3.53	—	64.4	3,148	3,262	CP001896.1
<i>Marichromatium purpuratum</i> 984	1	—	3.78	3.78	—	67.9	3,159	3,272	EU850807.1
<i>Ectothiorhodospira</i> sp. BSL-9	1	—	3.55	3.55	—	63	3,118	3,249	CP011994.1
<i>Thioflavicoccus mobilis</i> 8321	1	1	4.13	4.05	—	65.5	3,462	3,602	NR102479.1
<i>Halorhodospira halophila</i> SL1 extremely halophilic	1	—	2.68	2.68	—	68	2,375	2,458	CP000544.1
<i>Rhodopseudomonas palustris</i> CGA009	1	1	5.47	5.46	—	65	4,820	4,898	BX572608.1
<i>Rhodopseudomonas palustris</i> HaA2	1	—	5.33	5.33	—	66	4,683	4,772	CP000250.1
<i>Rhodopseudomonas palustris</i> BisB18	1	—	5.51	5.51	—	65	4,886	5,016	CP000301.1
<i>Rhodopseudomonas palustris</i> BisB5	1	—	4.89	4.89	—	64.8	4,397	4,492	CP000283.1
<i>Rhodopseudomonas palustris</i> BisA53	1	—	5.51	5.51	—	64.4	4,878	4,972	CP000463.1
<i>Rhodopseudomonas palustris</i> TIE-1	1	—	5.74	5.74	—	64.9	5,246	5,382	CP001096.1
<i>Rhodopseudomonas palustris</i> DX-1	1	—	5.4	5.4	—	65.4	4,917	5,082	EU221583.1
<i>Rhodoplanes</i> sp. Z2-YC6860	1	—	8.19	8.19	—	63.5	7,443	7,606	GQ369128.3
<i>Blastochloris viridis</i>	1	—	3.73	3.72	—	67.9	3,107	3,230	D14430.2
<i>Blastochloris viridis</i> ATCC19567	1	—	3.73	3.72	—	67.9	3,104	3,238	D2 5314.1
<i>Blastochloris viridis</i> DSM_133	1	—	3.72	3.72	—	67.9	3,114	3,236	HQ009851.1
<i>Rhodospirillum rubrum</i> ATCC 11170	1	1	4.41	4.35	—	65.3	3,838	3,917	CP000230.1
<i>Rhodospirillum rubrum</i> F11	1	—	4.35	4.35	—	65.4	3,878	3,946	CP003046.1
<i>Rhodospirillum centenum</i> SW (Rhodocista centenaria)	1	—	4.36	4.36	—	70.5	4,003	4,102	CP000613.2
<i>Pararhodospirillum photometricum</i> DSM 122	1	—	3.88	3.88	—	64.7	3,117	3,434	HE663493.1
<i>Rhodomicrobium vannielii</i> ATCC 17100	1	—	4.01	4.01	—	62.2	3,565	3,793	CP002292.1
<i>Rhodobacter sphaeroides</i> 2.4.1	2	5	4.6	3.19	0.94	68.8	4,242	4,370	CP000144.2
<i>Rhodobacter sphaeroides</i> MBTLJ-8	2	4	4.66	3.19	0.94	68.7	4,204	4,362	CP012960.1
<i>Rhodobacter sphaeroides</i> MBTLJ-13	2	5	4.65	3.19	0.94	68.8	4,200	4,358	CP015210.1
<i>Rhodobactersphaeroides</i> ATCC 17029	2	1	4.49	3.15	1.22	69	4,131	4,254	CP000577.1
<i>Rhodobacter sphaeroides</i> ATCC 17025	1	5	4.56	3.22	—	68.2	4,333	4,466	CP000661.1
<i>Rhodobacter sphaeroides</i> KD131	2	2	4.71	3.15	1.3	69.1	4,569	4,638	CP001151.1
<i>Rhodoferax ferrireducens</i> T118	1	1	4.97	4.71	—	59.6	4,436	4,554	CP000267.1
<i>Rhodobacter capsulatus</i> SB 1003	1	1	3.87	3.74	—	66.6	3,642	3,740	CP001312.1
<i>Rhodobacter</i> sp. LPB0142	1	3	3.78	3.46	—	66	3,530	3,660	KX066852.2
<i>Rhodovulum sulfidophilum</i> DSM1374	1	2	4.34	4.13	—	66.9	3,876	4,002	CP015418.1
<i>Rhodovulum sulfidophilum</i> DSM2351	1	3	4.73	4.45	—	66.9	4,155	4,340	D16430.1
<i>Rhodovulum sulfidophilum</i> SNK001	1	1	4.19	4.08	—	66.8	3,746	3,873	CP015421.1
<i>Rubrivivax gelatinosus</i> IL144	1	—	5.04	5.02	—	71.2	4,505	4,591	AB016167.1

Table 1. The genome characteristics of purple bacteria used in this study.

in YL28 shows that thiosulfate is possibly converted into S^0 or sulfate by the truncated Sox system. Moreover, the presence of *aprAB* and *sat* genes (encoding adenylylsulfate reductase, sulfate adenylyltransferase, respectively) suggests that YL28 possibly possesses an alternative sulfite oxiditon pathway (converting toxic sulfite into sulfate).

Halo-tolerance. YL28 possessed a gene cluster involved in the salt-alkali tolerance (*nhaABCDEFGF*, encoding Na^+/H^+ antiporter). Moreover, a novel putative Na^+/H^+ antiporter gene (*duf2062*) exist in YL28. The unique stress response subsystem in YL28 was listed in Table 4. There are six unique proteins participating in choline and betaine uptake and betaine biosynthesis (betaine aldehyde dehydrogenase, choline dehydrogenase), synthesis of osmoregulated periplasmic glucans (phosphoglycerol transferase I, cation channel protein), heavy metal resistance (DNA-binding heavy metal response regulator, heavy metal sensor histidine kinase).

Genome islands (GIs) of YL28. At least 9 GIs were identified with YL28 (Fig. 4A,B and Supplemental Table ST1) by both IslandPick and IslandPath-DIMOB methods. The GI size ranges from 8 to 31.3 Kb (Supplemental Table ST1). These GIs were noted as follows: the mobile element proteins, transposase, phage structure proteins, integration host factor, proteins involved in CRISPR system, 5-methylcytosine-specific restriction related enzyme, cephalosporin hydroxylase, flagellum synthesis component (CheABR), carbohydrate/nitrogen metabolism proteins (phenylalanyl-tRNA metabolism, Threonyl-tRNA biosynthesis, amyloamylase, asparagine synthetase, glycosyl transferase, aspartate aminotransferase), DNA replication and proofreading

Species	Fermentative nitrate reduction		Assimilatory nitrate reduction		Denitrification				Nitrogen fixation
	<i>nar</i> GHIJ/ <i>nap</i> AB	<i>nir</i> BD/ <i>nrf</i> AH	<i>nar</i> B/NR/ <i>nas</i> AB	NIT-6/ <i>nir</i> A	<i>nar</i> GHIJ/ <i>nap</i> AB	<i>nir</i> K/ <i>nir</i> S	<i>nor</i> BC	<i>nos</i> Z	<i>nif</i> DKH/ <i>anf</i> G
<i>R. palustris</i> CGA009				<i>nir</i> A		<i>nir</i> K	<i>nor</i> BC	<i>nos</i> Z	<i>nif</i> DKH/ <i>anf</i> G
<i>R. palustris</i> HaA2		<i>nir</i> BD	<i>nas</i> AB	<i>nir</i> A					<i>nif</i> DKH
<i>R. palustris</i> BisB18	<i>nap</i> AB				<i>nap</i> AB			<i>nos</i> Z	<i>nif</i> DKH/ <i>anf</i> G
<i>R. palustris</i> BisB5									<i>nif</i> DKH/ <i>anf</i> G
<i>R. palustris</i> BisA53		<i>nir</i> BD	<i>nas</i> AB	<i>nir</i> A		<i>nir</i> K	<i>nor</i> BC	<i>nos</i> Z	
<i>R. palustris</i> TIE-1				<i>nir</i> A		<i>nir</i> K	<i>nor</i> BC	<i>nos</i> Z	
<i>R. palustris</i> DX-1				<i>nir</i> A		<i>nir</i> K	<i>nor</i> BC	<i>nos</i> Z	<i>nif</i> DKH/ <i>anf</i> G
<i>Rhodoplanes</i> sp. Z2-YC6860	<i>nar</i> GHIJ	<i>nir</i> BD	<i>nas</i> AB						
<i>B. viridis</i>									<i>nif</i> DKH/ <i>anf</i> G
<i>B. viridis</i> ATCC19567									<i>nif</i> DKH/ <i>anf</i> G
<i>B. viridis</i> DSM_133									<i>nif</i> DKH/ <i>anf</i> G
<i>R. vannielii</i> ATCC 17100	<i>nap</i> AB				<i>nap</i> AB		<i>nor</i> BC		<i>nif</i> DKH
<i>R. sphaeroides</i> 2.4.1	<i>nap</i> AB				<i>nap</i> AB	<i>nir</i> K	<i>nor</i> BC		<i>nif</i> DKH
<i>R. sphaeroides</i> ATCC17029							<i>nor</i> BC		<i>nif</i> DKH
<i>R. sphaeroides</i> ATCC17025					<i>nap</i> AB	<i>nir</i> K	<i>nor</i> BC	<i>nos</i> Z	<i>nif</i> DKH
<i>R. sphaeroides</i> KD131						<i>nir</i> K	<i>nor</i> BC	<i>nos</i> Z	
<i>R. capsulatus</i> SB1003								<i>nos</i> Z	<i>nif</i> DKH/ <i>anf</i> G
<i>R. sphaeroides</i> MBTLJ-13	<i>nap</i> AB				<i>nap</i> AB		<i>nor</i> BC		<i>nif</i> DKH
<i>R. sphaeroides</i> MBTLJ-8	<i>nap</i> AB				<i>nap</i> AB		<i>nor</i> BC		<i>nif</i> DKH
<i>R. rubrum</i> ATCC11170									<i>nif</i> DKH/ <i>anf</i> G
<i>R. rubrum</i> F11									<i>nif</i> DKH/ <i>anf</i> G
<i>R. centenum</i> SW	<i>nap</i> AB				<i>nap</i> AB		<i>nor</i> BC	<i>nos</i> Z	<i>nif</i> DKH
<i>P. photometricum</i> DSM 122									<i>nif</i> DKH/ <i>anf</i> G
<i>R. vannielii</i> ATCC17100							<i>nor</i> BC		<i>nif</i> DKH/ <i>anf</i> G
<i>R. ferrireducens</i> T118	<i>nar</i> GHIJ	<i>nir</i> BD	<i>nas</i> AB		<i>nar</i> GHIJ		<i>nor</i> BC	<i>nos</i> Z	
<i>R. sulfidophilum</i> DSM1374									<i>nif</i> DKH
<i>R. sulfidophilum</i> DSM2351									<i>nif</i> DKH
<i>R. sulfidophilum</i> SNK001									<i>nif</i> DKH
<i>R. gelatinosus</i> IL144						<i>nir</i> K	<i>nor</i> BC	<i>nos</i> Z	<i>nif</i> DKH
<i>T. violascens</i> DSM198	<i>nap</i> AB				<i>nap</i> AB				<i>nif</i> DKH
<i>A. vinosum</i> DSM180				<i>nir</i> A			<i>nor</i> BC		<i>nif</i> DKH
<i>M. purpuratum</i> 984		<i>otr</i>	<i>nas</i> AB			<i>nir</i> K	<i>nor</i> BC		<i>nif</i> DKH
YL28		<i>nrf</i> AH	<i>nas</i> AB	<i>nir</i> A		<i>nir</i> S	<i>nor</i> BC		<i>nif</i> DKH
<i>H. halophila</i> SL1	<i>nar</i> GHIJ	<i>nrf</i> AH			<i>nar</i> GHIJ		<i>nor</i> BC		<i>nif</i> DKH
<i>Ectothiorhodospira</i> sp. BSL-9		<i>nrf</i> AH							<i>nif</i> DKH
<i>T. mobilis</i> 8321									<i>nif</i> DKH

Table 2. The key enzymes responsible for different nitrogen metabolism in purple bacteria.

systems (DNA recombination, repair DNA recombination and repair protein RecF Type II restriction enzyme, ATP-dependent endonuclease of the OLD family, ATP-dependent DNA helicase *pcrA*, DNA processing chain A) (Fig. 4A,B and Supplemental Table ST1).

Among these predicted GIs, GI-IX has the largest size (31.3 kb) in YL28. Notably, GI-IX has many DNA metabolic related genes including DNA recombination and repair protein RecF, ATP-dependent endonuclease of the OLD family, ATP-dependent DNA helicase *pcrA*, ATP-dependent DNA helicase RecQ, DNA processing chain A and type II restriction enzyme. It also contains some genes encoding diverse enzymes such as aspartate aminotransferase, methylase and ATP-dependent protease. The second largest GI (GI-II) carries a chemoreceptor gene cluster including the *cheA*, *cheB*, *cheR* and a gene encoding the methyl-accepting chemotaxis protein. Moreover, GI-II has the predicted two-component hybrid sensor and regulator, 4- α -glucanotransferase (amylomaltase), and mobile element protein (Supplemental Table ST1). Immediately downstream of GI-II, there are genes involved in the DNA metabolism such as 5-methylcytosine-specific restriction related enzyme, DNA helicase, phosphatase 2C homolog and 5-methylcytosine-specific restriction related enzyme. The size of other GIs is smaller than GI-IX.

There are five GIs in *Rhodopseudomonas*, *Blastochloris* and *Rhodoplanes* (Supplemental Table ST2). These GIs carry nitrogen-fixing genes (nitrogenase gene and alternative nitrogenase gene), LysR family of proteins transcriptional regulation, glycosyltransferase protein family and arsenic resistance (*ArsH*). There is an alternative nitrogenase (*AnfG*) in a GI region of *R. palustris*, which replaces nitrogenase (*NifDKH*) for nitrogen fixation²². At least 10 GIs are predicted in *Rhodobacter* (Supplemental Table ST3). Among these GIs, there are many genes involved in nitrogen metabolism including a nitrogen-fixing island (nitrogenase genes, oxidoreductase/

Species	Oxidation of sulfide	Reversed dissimilatory sulfite reductase			Sox system			
	FCC/SQR	DsrAB	AprAB	Sat	SoxB	SoxY/SoxZ	SoxA/SoxX	SoxC/SoxD
<i>A. vinosum</i> DSM180	SQR	DsrAB	AprAB	Sat	SoxB	SoxY/SoxZ	SoxA/SoxX	—
<i>Ectothiorhodospira</i> sp. BSL-9	FCC	—	—	Sat	—	—	—	—
<i>H. halophila</i> SL1	SQR	DsrAB	—	—	SoxB	SoxY/SoxZ	SoxA	—
YL28	SQR	DsrAB	AprAB	Sat	SoxB	SoxY/SoxZ	SoxA/SoxX	SoxD
<i>M. purpuratum</i> 984	SQR	DsrAB	—	—	SoxB	SoxY/SoxZ	SoxA/SoxX	SoxD
<i>T. violascens</i> DSM198	FCC	DsrAB	AprAB	Sat	SoxB	SoxY/SoxZ	SoxA/SoxX	—
<i>T. mobilis</i> 8321	FCC	DsrAB	AprAB	Sat	—	—	—	—

Table 3. The key enzymes responsible for different sulfur metabolism in purple sulfur bacteria (PSB).

nitrogenase component), a nitrogen-rich²³ island (nitrate/sulfonate/bicarbonate ABC transporter, nitrogen regulatory protein P-II), and a solid-island²⁴ (metallophosphoesterase). Moreover, some GIs were found to have several genes that involved in sulfur metabolism, biosynthesis of flagella, phage infection area and BadM/Rrf2 family of transcriptional regulatory protein. *Rhodobacter* possibly acquires nitrogen-fixation and the sulfur metabolism by the horizontal gene transfer (HGT). At least 11 GIs were predicted in genus *Rhodospirillum* (Supplemental Table ST4). Among those GIs, *Rhodospirillum* has gene elements including CRISPR, sulfide metabolism, arsenic resistance and the DNA and ribosomal protein synthesis, allowing this bacterium to survive in toxic environments under certain concentration of arsenic and sulfide.

Synten analysis. A good collinearity relationship was shown in PSB using the synten plots analysis (Supplementary Fig. SF3, YL28 as the reference genome). YL28, *T. violascens* DSM 198, *A. vinosum* DSM 180, *T. mobilis* 8312 and *M. purpuratum* 984 have the closer relationship than do *Ectothiorhodospira* sp. BSL-9 and *H. halophila* SL1. When *R. palustris* CGA009, *R. sphaeroides* 2.4.1 or *R. rubrum* ATCC 11170 was selected as a reference sequence, PNSB members showed a diverse collinearity relationship (Supplementary Figs SF4–SF6). Three groups are divided: the first one consists of the 4 genera including *Rhodopseudomonas*, *Blastochloris*, *Rhodoplanes*, and *Rhodomicrobium* (Fig. SF4); the second one consists of the 2 genera *Rhodobacter* and *Rhodovillum* (Fig. SF5); the third one only contains *Rhodospirillum* (except *R. centenum* SW) (Fig. SF6). Interestingly, a good collinearity was found between *R. ferrireducens* T118 (a non-phototrophic member) and PSB. Instead, a relatively poor collinearity was presented among PSB, *R. gelatinosus* IL144 and *R. centenum* SW. Collectively, the synten pattern is consistent with the whole genome phylogenetic tree (Fig. 1B). The poorest collinearity relationship was shown between YL28 and *R. gelatinosus* IL144.

Discussion

The interfaces of mangrove locate between land and sea in the tropical and the sub-tropical latitudes²⁵. PSB and PNSB are predominantly detected in the mangrove ecosystems. The two bacterial groups significantly contribute to the primary productivity of coastal seaboards and the food-web dynamics of various tropical coastline ecosystems^{26,27}. The high salinity, limited nutrients and S-richness sulfate concentrations often occur in the niche²⁵. However, it remains unclear about how the mangrove-associated microorganisms survive in the mangrove ecosystems and contribute to the host physiology. Particularly, the nitrogen utilization mechanisms in PSB and PNSB have not well studied. In this study, we sequenced the genome of YL28, investigated the nitrogen cycle pathways and compared them to other purple bacteria. Our study will contribute to elucidate bacterial surviving mechanisms in the special mangrove ecosystem.

YL28 utilized ammonium, nitrite or nitrate as the sole nitrogen source for phototrophic growth^{3–5}. In the present study, our results reveal that the strong ability for nitrate and nitrite utilization may be due to the six nitrogen metabolic pathways. Ammonium is preferred nitrogen sources by all species of purple bacteria²⁸. YL28 may grow on ammonium by the ammonium assimilation pathway, and converted organic nitrogen into ammonium by ammonification. Moreover, YL28 utilized N₂ as nitrogen source to support cell growth by nitrogen fixation when the absence of ammonium happened. It is interesting that YL28 grew chemoheterotrophically using the nitrate and/or nitrite as electron acceptor. The denitrification and/or fermentative nitrate reduction and/or assimilation nitrate reduction pathways may contribute to this process under the anoxic condition. Diverse nitrogen utilization pathways allow YL28 to grow on the different nitrogen compounds. YL28 can convert nitrite (toxic compounds in the mangrove ecosystems) into non-toxicity or low-toxicity products by fermentative nitrate reduction or assimilation nitrate reduction pathways. Denitrification may also contribute to this process⁵.

In this study, our bioinformatics studies further show that purple bacteria have 6 nitrogen cycle pathways (Table 3). The alternative nitrogenase genes are frequently observed in purple bacteria. The genes (encoding two nitrogenase subunits) are highly conserved in many phyla of bacteria and archaea, suggesting that nitrogen fixation genes evolve once and subsequently spread by vertical inheritance or by HGT⁹. Nitrate/nitrite reduction-related genes were found only in the accessory genomes. The observation agrees with the previous studies that show that nitrate/nitrite is not preferred nitrogen sources²⁹. The partial denitrification pathway was widely found in PNSB^{30,31}. To best of our knowledge, our results first demonstrated that the partial denitrification pathway existed in PSB, suggesting that some PSB possibly also played important roles in the nitrogen cycle (e.g., *Rhodopseudomonas* and *Rhodobacter*). The partial denitrification pathway contributes to remove excess redox and mitigates the toxicity from the certain nitrogen oxide intermediates²⁰. The fermentative nitrate reduction is less known in purple bacteria and assimilation nitrate reduction is only limited to a member of PNSB (*Rhodobacter*

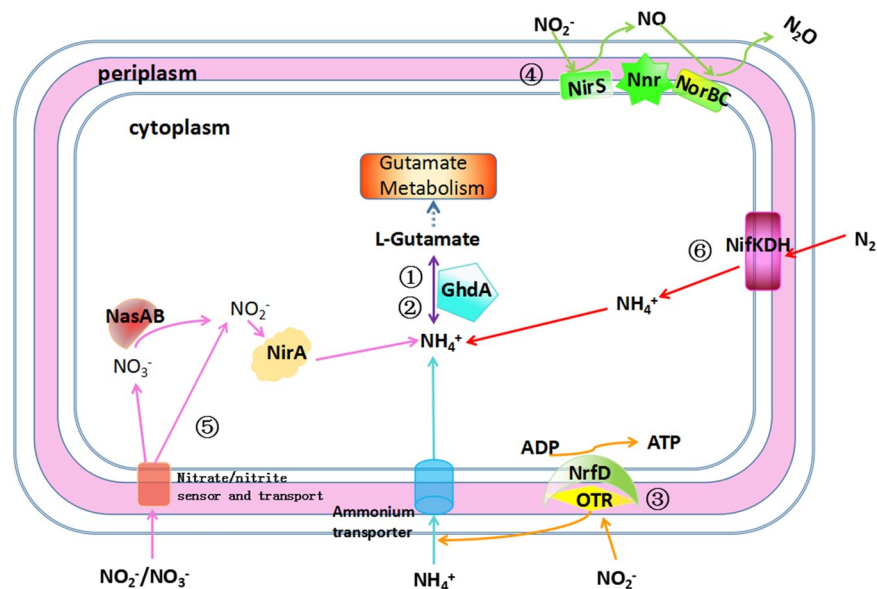


Figure 3. The model of nitrogen metabolism in *M. gracile* YL28. There are six nitrogen metabolic pathways in YL28 whose key enzyme confirmed by CDART and CDD. (A) Ammonification pathway; (B) ammonium assimilation pathway; (C) fermentative nitrate reduction; (D) denitrification; (E) assimilation nitrate reduction; (F) nitrogen fixation.

Category	Subcategory	Subsystem	Role
Virulence, Disease and Defense	Resistance to antibiotics and toxic compounds	Cobalt-zinc-cadmium resistance	DNA-binding heavy metal response regulator
	Resistance to antibiotics and toxic compounds	Cobalt-zinc-cadmium resistance	Heavy metal sensor histidine kinase
	Resistance to antibiotics and toxic compounds	Cobalt-zinc-cadmium resistance	cation channel protein
Stress Response	Osmotic stress	Choline and Betaine Uptake and Betaine Biosynthesis	Betaine aldehyde dehydrogenase (EC 1.2.1.8)
	Osmotic stress	Choline and Betaine Uptake and Betaine Biosynthesis	Choline dehydrogenase (EC 1.1.99.1)
	Osmotic stress	Synthesis of osmoregulated periplasmic glucans	Phosphoglycerol transferase I (EC 2.7.8.20)

Table 4. The unique stress response protein in YL28.

sulfite reduction (key enzyme gene *sqr* and *dsr*) possibly contribute to detoxification of toxic sulfide. Thiosulfate may be converted into S^0 or sulfate by truncated Sox system. Moreover, an alternative pathway (key genes *apr* and *sat*) also allows to detoxifying the toxic sulfite. However, sulfite may not be directly converted into sulfate via a two-electron transfer because purple bacteria lack of the sulfite:cytochrome c oxidoreductase³³. Without *soxC/soxD*, sulfane sulfur atom cannot be subsequently oxidized³⁴. However, YL28 has *soxD*. The study on the complete sulfur oxidation in YL28 by *soxD* needs to be investigated. Some genes involved in classical sulfite metabolism were not detected in most purple bacteria³⁴ (except for some freshwater PSB species such as *A. vinosum* DSM180)³⁵. However, YL28 has both sulfide: quinone reductase gene (*sqr*, YL28 could directly oxidize S^{2-} to S^0) and APS reductase gene (*apr*). This suggests that YL28 may oxidize SO_3^{2-} to SO_4^{2-} and reduce the sulfite toxicity to cells. Sulfite reductase gene (*dsr*) was observed in all of the selected PSB (not PNSB). Gene *dsr* may be a useful taxonomic or systematic marker for PSB phylogeny.

Our studies revealed that there was a distinctive difference in halo-tolerance traits between freshwater and salt-dependent species by the 16S rDNA, core or whole genome sequence analysis. The salt-dependent or salt-requiring species showed different halo-tolerance. For example, YL28 possesses unique stress response genes (*betA*, *duf* 2062, *mrp*, glucan and heavy metal response genes etc.). These genes possibly contribute to the tolerance to in high salinity environment^{36,37}.

The size of genome, the numbers and size of GIs in PSB were smaller than those in PNSB, implying that PSB had more flexible response to environment change. A gene transfer agent (GTA) is an unusual bacteriophage-like element which transfers a random host genomic DNA fragments (4–14 kb in size) between closely related bacteria^{38–40}. Genes involved in the photosynthesis may be horizontally transferred between the same phyla by GTA⁴¹. Our results revealed that GTA gene clusters including ICEs (the self-transmissible mobile genetic elements, integrative and conjugative elements) were presented in all examined genomes. PNSB may acquire more genes by HGT to survive in the niches. The flagellum biosynthesis and methyl accepting chemotaxis genes (such as *cheA*,

system with a 500 bp pair-end library. The reads were assembled using SOAPdenovo v2.04. Putative protein-encoding genes were identified using Glimmer 3.02⁴³. Annotation was performed by BLAST +2.2.24 searching against databases, including the National Center for Biotechnology Information (NCBI), Clusters of Orthologous Groups of Proteins (COG), the Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene ontology (GO). The genome sequence of *M. gracile* YL-28 was deposited in the GenBank database under the accession number LSU000000000⁴⁴.

Acquisition and re-annotation of the selected genome sequences. The genome sequences of other 35 sequenced purple bacteria were obtained from NCBI database which assessed level was complete. To avoid the possible deviations due to different annotation methods, we used Rapid Annotation using Subsystem Technology (RAST) server for reannotation⁴⁵ of the selected genomes. Glimmer algorithm was used for gene calling.

Phylogeny. Three different methods were used for constructing the phylogeny trees. The 16S rDNA sequences of the selected microorganisms were first used to infer phylogenies using the Neighbor-Joining (NJ) method of MEGA 6.0.6⁴⁶. The core genomes for these selected genomes were next clustered by USEARCH and the phylogeny trees of core and accessory genomes were constructed by Neighbor-Joining algorithm using the BPGA-1.3⁴⁷. The 36 genome sequences were aligned by the progressive MAUVE to generate a phylogenetic guide tree⁴⁸.

Comparisons of conserved and variable regions. The genome of YL28 was used as a reference sequence to align all other genomes of 35 purple bacteria. Similarly, the genome of *R. palustris* CGA009 was selected as a reference to align those in *Rhodospseudomonas*, *Blastochloris* and *Rhodoplanes*; the genome of *R. sphaeroides* 2.4.1 was used as a reference to align the genomes of two genera *Rhodobacter* and *Rhodovulum*; *R. rubrum* ATCC11170 was selected as a reference to align the genomes of *Rhodospirillum* genus. The multiple whole-genome alignment was conducted using the progressive alignment algorithm implemented in MAUVE. Syntheny plots were generated by aligning regions of the predicted and reference genomes that differed by default parameter. All regions were aligned and displayed in MAUVE version 2.3.1⁴⁸. The circular genomic map was generated by the BLAST Ring Image Generator (BRIG, version 0.95)⁴⁹ using alignment reference genome on a local BLAST+ basis, with standard parameters (50% lower – 70% upper cut-off for identity and E-value of 10). The ring color gradients correspond to varying degrees of identity of BLAST matches. Circular genomic maps also include the information on GC skew and GC content.

Comparative analyses of core genome, accessory genome and unique genome. In order to depict the core and accessory genome in each genus, a reciprocal best hit search using the BPGA software was performed⁴⁷. Orthologous clusters (OCs) were assigned by grouping all protein sequences in the 36 genomes using USEARCH based on their sequence similarity (E-value < 10⁻⁵, >50% coverage). A series of built-in scripts were used to (i) parse, (ii) upload to the MySQL relational database, (iii) perform a reciprocal best hit analysis to form pairs of sequences, and (iv) normalize the E-values for all the pairs formed. Normalization of E-values was done by averaging all recent ortholog (in paralogs) and dividing each pair of ortholog by the average. Pan-core plot against combinations will give core and pan genome boxplot and dot plot generated using desired number of unique combinations of genomes. Atypical GC analysis will give sequences of core, accessory and unique genes with atypical (extreme) GC content. COG and KEGG distributions of the core, accessory and unique gene families were calculated based on representative sequences. Keywords were used to query for nitrogen metabolism functional in the orthologous families and to calculate the number of matches of those functions, using custom bash commands.

Genomic island prediction. Genomic islands (GIs) were predicted by using IslandViewer3⁵⁰ including IslandPick, IslandPath-DIMOB and SIGI-HMM.

Nitrogen metabolism, sulfur metabolism and stress response analysis. We chose three RAST subsystems for further analyze the nitrogen metabolism, sulfur metabolism and stress response pathways. The key protein sequences^{51–53} responsible for nitrogen metabolism, sulfur metabolism was used as reference sequences to search against the 36 genomes of purple bacteria by BlastP. The obtained sequence by BLAST was further evaluated by the combination of KEGG, RAST, CDART (conserved domain architecture retrieval tool⁵⁴ and CDD (the conserved domain database)⁵⁵.

References

- Trüper, H. G. & Pfennig, N. Characterization and identification of the anoxygenic phototrophic bacteria, 1. [Starr, M. P., Stolp, H., Trüper, H. G., Balows, A. & Schlegel, H. G. (eds)] *The prokaryotes* **18**, 299–312. (Springer, Berlin, Heidelberg, 1981).
- Zhao, J. Y. *et al.* Identification and characterization of a purple sulfur bacterium from mangrove with rhodopin as predominant carotenoid. *Acta Microbiologica Sinica*. **51**(10), 1318–1325 (2011).
- Jiang, P., Hong, X., Zhao, C. G. & Yang, S. P. Reciprocal transformation of Inorganic nitrogen by Resting Cells of *Marichromatium gracile* YL28. *Journal of Huaqiao University*. **36**(2), 45–52 (2015).
- Jiang, P., Zhao, C. G., Jia, Y. Q. & Yang, S. P. Inorganic nitrogen removal by a marine purple sulfur bacterium capable of growth on nitrite as sole nitrogen source. *Microbiology China*. **41**(5), 824–831 (2014).
- Xuan, H., Chen, Z. W., Zhao, C. G. & Yang, S. P. Nitrogen transformation under different dissolved oxygen levels by the anoxygenic phototrophic bacterium *Marichromatium gracile*. *World Journal of Microbiology and Biotechnology*. **33**, 113 (2017).
- Hungate, B. A., Duker, J. S., Shaw, M. R., Luo, Y. & Field, C. B. Nitrogen and climate change. *Science*. **302**(5650), 1512–1513 (2003).
- Stein, L. Y. & Klotz, M. G. The nitrogen cycle. *Current Biology*. **26**(3), 94–98 (2016).
- Gruber, N. & Galloway, J. N. An earth-system perspective of the global nitrogen cycle. *Nature*. **451**(7176), 293 (2008).

9. Canfield, D. E., Glazer, A. N. & Falkowski, P. G. The evolution and future of Earth's nitrogen cycle. *Science*. **330**(6001), 192–196 (2010).
10. Duce, R. A. *et al.* Impacts of atmospheric anthropogenic nitrogen on the open ocean. *Science*. **320**(5878), 893–897 (2008).
11. Luo, S., Wu, B. L., Xiong, X. Q. & Wang, J. W. Short-term toxicity of ammonia, nitrite, and nitrate to early life stages of the rare minnow (*Gobiocypris rarus*). *Environmental Toxicology and Chemistry*. **35**(6), 1422–1427 (2015).
12. Dixon, R. & Kahn, D. Genetic regulation of biological nitrogen fixation. *Nature Reviews Microbiology*. **2**(8), 220–229 (1998).
13. Dubbs, J. M. & Tabita, F. R. Regulators of nonsulfur purple phototrophic bacteria and the interactive control of CO₂ assimilation, nitrogen fixation, hydrogen metabolism and energy generation. *FEMS Microbiology Reviews*. **28**(3), 353–376 (2004).
14. Griffin, B. M., Schott, J. & Schink, B. Nitrite, an electron donor for anoxygenic photosynthesis. *Science*. **316**(5833), 1870–1870 (2007).
15. Hülsen, T. *et al.* Domestic wastewater treatment with purple phototrophic bacteria using a novel continuous photo anaerobic membrane bioreactor. *Water Research*. **100**, 486–495 (2016).
16. Imhoff, J. F. Taxonomy, phylogeny, and general ecology of anoxygenic phototrophic bacteria, 1. [Mann, Nicholas H. & Noel G. (eds)] *Photosynthetic Prokaryotes 2*, 53–92. (Springer Science and Business Media, 2012).
17. Lee, T. I. *et al.* Transcriptional regulatory networks in *Saccharomyces cerevisiae*. *Science*. **298**, 799–804 (2002).
18. Pino, C. *et al.* The assimilatory nitrate reduction system of the phototrophic bacterium *Rhodobacter capsulatus* E1F1. *Biochemical Society Transactions*. **34**(5594), 127–129 (2006).
19. Olmo-Mira, M. F. *et al.* Expression and characterization of the assimilatory NADH-nitrite reductase from the phototrophic bacterium *Rhodobacter capsulatus* E1F1. *Archives of Microbiology*. **186**(4), 339–344 (2006).
20. Swingle, W. D., Blankenship R. E. & Raymond, J. Evolutionary relationships among purple photosynthetic bacteria and the origin of proteobacterial photosynthetic systems, 1. [Hunter, C. N., Daldal, F., Thurnauer, M. C. & Beatty, J. T. (eds)] *The purple phototrophic bacteria 2*, 17–29. (Springer Netherlands, 2008).
21. Young, J. P. W. The phylogeny and evolution of nitrogenases, 1. [Palacios R. & Newton W. E. (eds)] *Genomes and genomics of nitrogen-fixing organisms 14*, 221–241. (Springer Netherlands, 2005).
22. Bishop, P. E. & Joergers, R. D. Genetics and molecular biology of alternative nitrogen fixation systems. *Annual review of plant biology*. **41**(1), 109–125 (1990).
23. Yan, Y. *et al.* Nitrogen fixation island and rhizosphere competence traits in the fenome of root-associated *Pseudomonas stutzeri* A1501. *Proceedings of the National Academy of Sciences of the United States of America*. **105**(21), 7564 (2008).
24. Ramachandran, A., Yun, H., Wen-Mei, W. H., Ma, J. & Chen, D. FPGA accelerated DNA error correction. *Design, Automation and Test in Europe Conference and Exhibition*. **1**, 1371–1376 (2015).
25. Lin, X. L. *et al.* Mangrove sediment microbiome: adaptive microbial assemblages and their routed biogeochemical processes in Yunxiao mangrove national nature reserve, China. *Microbial Ecology*, <https://doi.org/10.1007/s00248-018-1261-6> (2018).
26. Kobayashi, M., Fujii, K., Shimamoto, I. & Maki, T. Treatment and reuse of industrial waste water by phototrophic bacteria. *Progress in Water Technology*. **11**(1), 279–284 (1978).
27. Vethanayagam, R. R. Purple photosynthetic bacteria from a tropical mangrove environment. *Marine Biology*. **110**(1), 161–163 (1991).
28. Masepohl, B. & Hallenbeck, P. C. Nitrogen and molybdenum control of nitrogen fixation in the phototrophic bacterium *Rhodobacter capsulatus*, 1. [Hallenbeck, P. C. (ed.)] *Recent advances in phototrophic prokaryotes 4*, 49–70. (Springer New York, 2010).
29. Sun, Y. *et al.* Ammonium assimilation: An important accessory during aerobic denitrification of *Pseudomonas stutzeri* T13. *Bioresource Technology*. **234**, 264–272 (2017).
30. Shen, J. Y. *et al.* Improvement of fisheries water quality and fish growth of photosynthetic bacteria HZPSB. *Bulletin of Science and Technology*. **15**, 16–31 (2004).
31. Qiu, H. D., Shi, X. A., Guo, Y. H., Fen-Wu, W. U. & Zhao, J. P. The breeding and application of salt-resistant and high NH₄⁺ or NO₂⁻ reducible bacteria of Rhodospirillaceae. *Journal of Fisheries of China* **23**, 92–96 (1999).
32. Dahl, C. Sulfur metabolism in phototrophic bacteria, 1. [Hallenbeck, P. C. (ed.)] *Modern Topics in the Phototrophic Prokaryotes 2*, 27–66. (Springer International, 2017).
33. Kappler, U. & Dahl, C. Enzymology and molecular biology of prokaryotic sulfite oxidation. *FEMS Microbiology Letters*. **203**(1), 1–9 (2001).
34. Frigaard, N. U. & Dahl, C. Sulfur metabolism in phototrophic sulfur bacteria. *Advances in Microbial Physiology*. **54**, 103–200 (2008).
35. Dahl, C. Insertional gene inactivation in a phototrophic sulphur bacterium: APS-reductase-deficient mutants of *Chromatium vinosum*. *Microbiology*. **142**(12), 3363–3372 (1996).
36. Jiang, J. *et al.* *Halomonas zhaodongensis* sp. nov., a slightly halophilic bacterium isolated from saline-alkaline soils in Zhaodong, China. *Antonie Van Leeuwenhoek*. **104**(5), 685–694 (2013).
37. Meng, L. *et al.* Cloning and identification of group 1 mrp operon encoding a novel monovalent cation/proton antiporter system from the moderate halophile *Halomonas zhaodongensis*. *Extremophiles*. **18**(6), 963–972 (2014).
38. Marrs, B. Genetic recombination in *Rhodospseudomonas capsulata*. *Proceedings of the National Academy of Sciences of the United States of America*. **71**(3), 971–973 (1974).
39. Lang, A. S. & Beatty, J. T. Importance of widespread gene transfer agent genes in alpha-proteobacteria. *Trends in Microbiology*. **15**(2), 54–62 (2007).
40. Zheng, Q., Jiao, N. Z., Zhang, R., Wei, J. J. & Zhang, F. The evolutionary divergence of psbA gene in *synechococcus* and their myoviruses in the east China Sea. *Plos One*. **9**(1), e86644 (2014).
41. Raymond, J. The role of horizontal gene transfer in photosynthesis, oxygen production, and oxygen tolerance, 1. [Gogarten, M. B., Gogarten, J. P. & Olendzenski, L. C. (eds)] *Horizontal Gene Transfer 19*, 323–338. (Humana Press, 2009).
42. Armitage, J. P. Swimming and behavior in purple non-sulfur bacteria, 1. [Hunter, C. N., Daldal, F., Thurnauer, M. C. & Beatty, J. T. (eds)] *The purple phototrophic bacteria 2*, 643–654. (Springer Netherlands, 2009).
43. Silva, A. C. R. D. *et al.* Comparison of the genomes of two *Xanthomonas* pathogens with differing host specificities. *Nature*. **417**(6887), 459–463 (2002).
44. Zhang, X. B., Zhao, C. G., Hong, X., Chen, S. C. & Yang, S. P. Genome sequence of *Marichromatium gracile* YL-28, a purple sulfur bacterium with bioremediation potential. *Genome Announcements*. **4**(3), e00288–00216 (2016).
45. Aziz, R. K. *et al.* The RAST Server: rapid annotations using subsystems technology. *BMC Genomics*. **9**(1), 75–75 (2008).
46. Tamura, K., Stecher, G., Peterson, D., Filipowski, A. & Kumar, S. MEGA6: molecular evolutionary genetics analysis Version 6.0. *Molecular Biology and Evolution*. **30**(12), 2725–2729 (2013).
47. Chaudhari, N. M., Kumar, G. V. & Chitra, D. BPGA an ultra-fast pan-genome analysis pipeline. *Scientific reports*. **6**, 24373 (2016).
48. Darling, A. E., Mau, B. & Perna, N. T. Progressivemaue: multiple genome alignment with gene gain, loss and rearrangement. *Plos One*. **5**(6), e11147 (2010).
49. Beatson, S. A., Ben, Z. N. L., Petty, N. K. & Nabil-Fareed, A. BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics*. **12**(1), 1 (2011).
50. Dhillon, B. K. *et al.* IslandViewer 3: more flexible, interactive genomic island discovery, visualization and analysis. *Nucleic Acids Research*. **43**(1), 104–108 (2015).

51. Kranz, R. G. & Haselkorn, R. Anaerobic regulation of nitrogen-fixation genes in *Rhodospseudomonas capsulata*. *Proceedings of the National Academy of Sciences of the United States of America*. **83**(18), 6805–6809 (1986).
52. Cabello, P. *et al.* Hydroxylamine assimilation by *Rhodobacter capsulatus* E1F1 requirement of the hcp gene (hybrid cluster protein) located in the nitrate assimilation nas gene region for hydroxylamine reduction. *Journal of Biological Chemistry*. **279**(44), 45485–45494 (2004).
53. Bartnikas, T. B. *et al.* Characterization of a member of the NnrR regulon in *Rhodobacter sphaeroides* 2.4.3 encoding a haem-copper protein. *Microbiology*. **148**(3), 825 (2002).
54. Geer, L. Y., Domrachev, M., Lipman, D. J. & Bryant, S. H. CDART: Protein homology by domain architecture. *Genome Research*. **12**(10), 1619–1623 (2002).
55. Marchlerbauer, A. *et al.* CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Research*. **45**(1), 200–203 (2017).

Acknowledgements

This work was supported by National Marine Public Industry Research (201505026); by Natural Science Foundation of Fujian Province (Fujian Provincial Natural Science Foundation) (2018J01049); by Subsidized Project for Cultivating Postgraduates Innovative Ability in Scientific Research of Huaqiao University.

Author Contributions

B.Z., X.Z., C.Z., S.Y. and S.C. conceived the study, analyzed the data and wrote the paper.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-018-36160-2>.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018