Article

Nitrogen Fixation Activity and Genome Analysis of a Moderately Haloalkaliphilic Anoxygenic Phototrophic Bacterium *Rhodovulum tesquicola*

Anastasia V. Komova [1,*], Elizaveta D. Bakhmutova [1], Anna O. Izotova [1], Evelina S. Kochetova [1], Stepan V. Toshchakov [1], Zorigto B. Namsaraev [1], Maxim V. Golichenkov [2,3] and Aleksei A. Korzhenkov [1]

[1] Kurchatov Centre for Genome Research, NRC Kurchatov Institute, 123098 Moscow, Russia
[2] Soil Biology Department, Soil Science Faculty, Moscow State University, 119991 Moscow, Russia
[3] Department of Biology, Faculty of Pediatrics, Pirogov Russian National Research Medical University, 117997 Moscow, Russia
[*] Correspondence: komovaav@gmail.com

Abstract: The genome of the moderately haloalkaliphilic diazotrophic anoxygenic phototrophic bacterium *Rhodovulum tesquicola* A-36s[T] isolated from an alkaline lake was analyzed and compared to the genomes of the closest species *Rhodovulum steppense* A-20s[T] and *Rhodovulum strictum* DSM 11289[T]. The genomic features of three organisms are quite similar, reflecting their ecological and physiological role of facultative photoheterotrophs. Nevertheless, the nitrogenase activity of the pure cultures of the studied bacteria differed significantly: the highest rate (4066 nmoles C_2H_2/mg of dry weight per hour) was demonstrated by *Rhodovulum strictum* while the rates in *Rhodovulum tesquicola* and *Rhodovulum steppense* were an order of magnitude lower (278 and 523 nmoles C_2H_2/mg of dry weight per hour, respectively). This difference can be attributed to the presence of an additional nitrogenase operon found exclusively in *R. strictum* and to the structural variation in nitrogenase operon in *R. tesquicola*.

Keywords: anoxygenic phototrophs; halophiles; alkaliphiles; nitrogen fixation; diazotroph; *Rhodovulum tesquicola*; genome sequence

1. Introduction

The genus *Rhodovulum* includes anoxygenic phototrophic bacteria that require sodium chloride for growth [1,2]. The members of the genus are widely distributed in marine and saline environments [3–5], and due to their metabolic capacities may be used in a variety of biotechnological applications [6–8]. Additionally, recent studies show a wider spread of photoferrotrophy and phototrophic extracellular electron uptake in *Rhodovulum* that previously thought [9]. Thus, the genome analysis of the representatives of the genus appears to provide new insights into the genetic and evolutionary basis of metabolic processes in halotolerant anoxygenic phototrophic *Proteobacteria*. To date, *Rhodovulum* comprises 20 validly described species, for eight of which no genome sequence has been published so far.

Bacteria of the genus *Rhodovulum* preferably grow photoheterotrophically and are able to fix nitrogen. Thus, they serve as an important mediator of the carbon and nitrogen cycles as it links the organic matter decomposition and nitrogen fixation. Saline and soda lakes are among the most productive ecosystems in the world [10,11]. The alkaline conditions favour the volatilization of ammonium from the lakes; therefore the regular supply of nitrogen is needed to provide the functioning of ecosystem [12].

The process of nitrogen fixation in saline and soda lakes was studied mostly for the natural ecosystems and mesocosms [13–17]. The nitrogen fixation in pure cultures of phototrophic *Alphaproteobacteria* was studied in details for the neutrophilic and some

of the marine and true halophilic representatives [18–20]. Nevertheless, little is known about this process in pure cultures of halotolerant and alkaliphilic anoxygenic phototrophic *Alphaproteobacteria* that are widespread in the lakes with relatively low salinity and increased alkalinity [5,21].

Rhodovulum tesquicola A-36s[T] was isolated from an alkaline lake Sul'fatnoe (Buryat Republic, Siberia, Russia; pH 9.2, mineralization 7.7 g/L) [5] and later described as a type strain of a new species within *Rhodovulum* genus [22]. It is a typical anoxygenic phototrophic bacterium able to grow photoheterotrophically and chemoheterotrophically with a range of organic substrates, and to grow photolithoautotrophically with sulfur compounds as electron donors. However, its ability to fix nitrogen has not been studied before. Here we present the analysis of genome of *R. tesquicola* in comparison with the genomes of the closest species, *R. steppense* A-20s[T] and *R. strictum* DSM 11289[T], and the data on the nitrogen fixation activity by the pure cultures of these organisms.

2. Materials and Methods

2.1. Microbial Cultivation

R. tesquicola A-36s[T] (=ATCC BAA-1573[T]), *R. steppense* A-20s[T] (=DSM 21153[T]) (both isolated previously by A.V.K. and kept at the laboratory collection) and *R. strictum* DSM 11289[T] obtained from the Leibniz-Institut DSMZ, Braunschweig, Germany were grown photoheterotrophically at 30 °C on the previously described medium [22] (pH 8.5–9).

2.2. Nitrogenase Activity Measurement

Nitrogenase activity measurement was performed by the acetylene reduction assay [23]. 1 mL of acetylene (about 10% of gaseous headspace in the vial) was injected into the hermetically sealed 23 mL vials containing liquid photoheterotrophically grown cultures of the investigated bacteria with argon filled headspace. The vials were incubated at 25 °C for 20 h. After that, 1 mL of gaseous phase was sampled and the detection of ethylene was performed on Crystal-2000 chromatograph (LLC NPF Meta-Chrom, Yoshkar-Ola, Russia) with a heated flame ionization detector (HFID). Instrument characteristics: column length—1 m, diameter—3 mm, filler—Porapak N 80/100, column temperature—60 °C, detector temperature—160 °C, evaporator temperature—100 °C, carrier gas flow (N_2)—50 mL/min, air—280 mL/min, hydrogen—28 mL/min.

2.3. Genome Sequencing and Assembly

Genomic DNA was isolated from cell culture of *R. tesquicola* A-36s[T] using the QIAamp® DNA Mini kit following the manufacturer's recommendations (Qiagen, Dusseldorf, Germany).

DNA was fragmented by ultrasound on a Covaris M220 (Covaris, Woburn, MA, USA). A paired-end genomic library (average insertion size 300 bp) was made using the NEBNext® Ultra[TM] II DNA Library Prep Kit (New England Biolabs, Ipswich, MA, USA). The DNA library was sequenced using an Illumina NovaSeq 6000 System (Illumina, San Diego, CA, USA) with a 2 × 250 bp paired-end read.

Genome assembly was performed using ZGA software pipeline v0.1 [24]. Sequencing reads were processed using BBTools software toolkit v38.96 [25]. Reads with low quality bases and adapter sequences were trimmed from sequencing reads with BBduk v38.96 (ktrim = r qtrim = rl trimq = 21 k = 20 mL = 33 trimpolyg = 4 trimpolya = 4). Trimmed reads were corrected with Tadpole (mode = correct, cecc = t). Corrected reads were normalized with BBnorm v38.96 to 250× depth coverage. Normalized reads were assembled to scaffolds with SPAdes v3.14 [26] using default parameters. The resulting assembly was polished with Pilon v1.24 [27] using trimmed reads.

Primary annotation was performed using DFAST v1.2.15 [28], CDS were predicted with Prodigal. Additional annotation was performed with NCBI Prokaryotic Genome Annotation Pipeline (PGAP) v6.1 [29], proteins were predicted using GeneMarkS-2+ [30] and best-placed reference protein set. KEGG orthologs (KO) annotation was performed

using KofamKOALA [31]. Completeness for a given KEGG module [32] was calculated using the program anvi-estimate-metabolism from anvi'o toolkit v7.1 [33].

All available on the 1 July 2022 *Rhodovulum* genomes including the genomes of *R. steppense* A-20s[T] and *R. strictum* DSM 11289[T] were downloaded from NCBI Assembly database. Genome completeness was assessed with CheckM v1.2.0 [34] taxonomy workflow using *Rhodobacteraceae* marker set. Pairwise average nucleotide identity (ANI) was estimated using FastANI [35]. Pangenomic analysis was conducted using Proteinortho v6 [36] with BLASTP+ [37] algorithm. Nitrogen fixation genes in *Rhodovulum* genomes were searched using gene annotations in NCBI protein and Uniprot and BLASTP+ [37] search in NCBI nr and Swissprot databases. Phi29-like polymerase and viral single-stranded DNA-binding protein were searched using BLASTP+ [37] (e-value 0.001) with P03680 and Q38504 reference sequences, respectively.

Amino acid sequences of nifH gene homologs were found via BLASTP+ [37] search in NBCI nr database. The first 250 hits were downloaded and clustered using CD-HIT [38] with 95% identity threshold. Representative sequences of 16S rRNA and *nifH* genes were aligned using MAFFT v7.475 with the FFT-NS-i algorithm [39]. The resulting alignments were manually curated and phylogenetic trees were inferred using FastTree v2.1.10 [40] in auto mode. The trees were visualized using ITOL web-server [41].

3. Results and Discussion

3.1. General Characteristics of Genome

The sequencing yielded in 6,272,843 pair-end reads with a length of 251 bp, totaling to 41.39 Gbp. The draft genome of *R. tesquicola* A-36s[T] consists of 80 sequences resulting in 3,521,822 bp with GC-content of 67.68% and N50 of 156,513 bp (Table 1). The most similar to *R. tesquicola* A-36s[T] genomes of cultivated *Rhodovulum* species were *R. steppense*—ANI value 94.86% and *R. strictum*—88.91% (Figures 1 and 2). These values support the novelty of *R. tesquicola*.

Table 1. Genome features of *R. tesquicola*, *R. steppense* and *R. strictum*, according to NCBI database.

	Rhodovulum tesquicola A-36s[T]	*Rhodovulum steppense* A-20s[T]	*Rhodovulum strictum* DSM 11289[T]
Genome size	3,521,822	3,665,949	3,825,097
Total number of genes	3490	3634	3756
Proteins	3375	3582	3636
rRNA genes	1, 1, 1 (5S, 16S, 23S)	1, 1, 2 (5S, 16S, 23S partial)	1, 1, 1 (5S, 16S, 23S)
tRNA genes	45	51	42
Genome completeness, %	99.41	99.12	98.83
Genome contamination, %	0	0.15	0.59

The completeness of the KEGG modules was assessed for *R. tesquicola*, *R. steppense* and *R. strictum* genomes. The completeness differs only in 28 out of 204 modules. Only 18 differing modules are complete for at least 50% for at least one of the genomes and differ for at least 10% (Figure 3, Supplementary Table S1).

Pangenome analaysis of three *Rhodovulum* genomes (Figure 4) showed that *R. tesquicola* has the smallest genome and more common protein clusters comparing to *R. steppense* and *R. strictum*. *R. tesquicola* and *R. steppense* have more in common than any other pair of genomes that emphasize their proximity. The core genome of the three *Rhodovulum* genomes consists of 2646 protein clusters corresponding to 79% of *R. tesquicola* genome. Among genes lacking in *R. tesquicola* genome there are mobile elements, genes of prophage loci and hypothetical proteins without any functional annotation. Phi29-like DNA polymerase and phage single-stranded DNA-binding protein genes were not found in any of the three genomes.
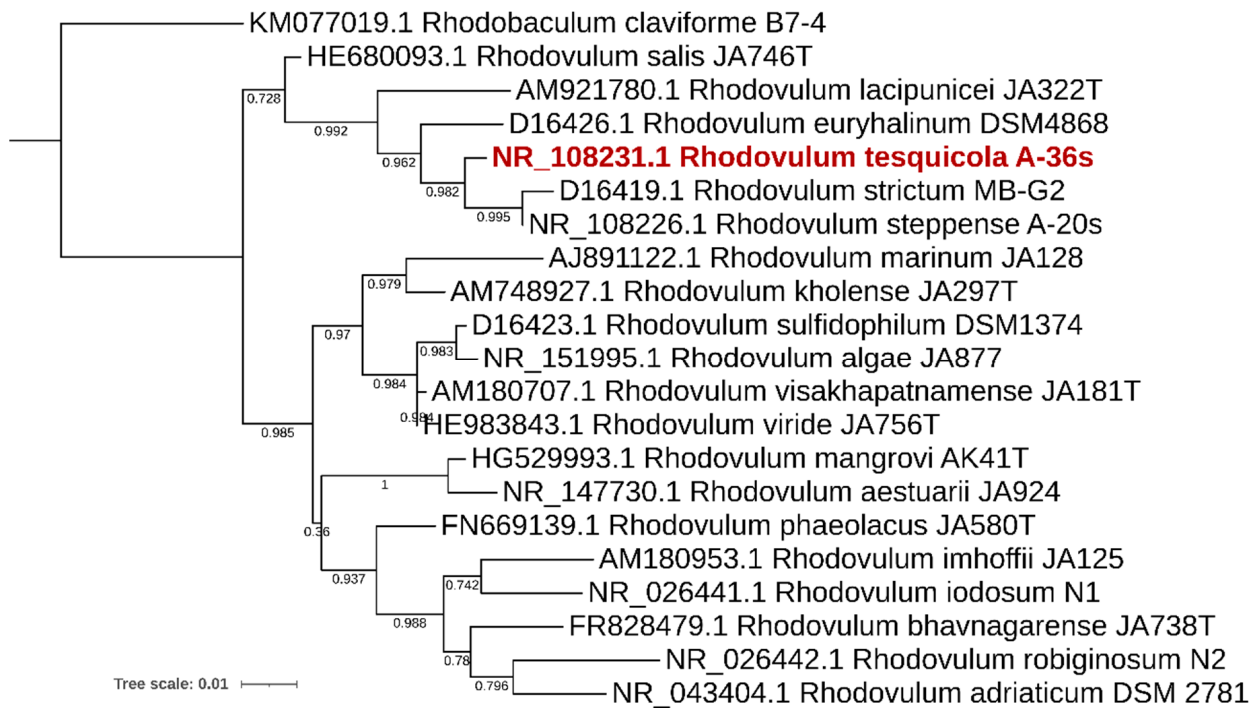
Figure 1. 16S rRNA gene phylogeny of the genus *Rhodovulum*. The tree is scaled by evolutionary distance (count of substitutions per site).
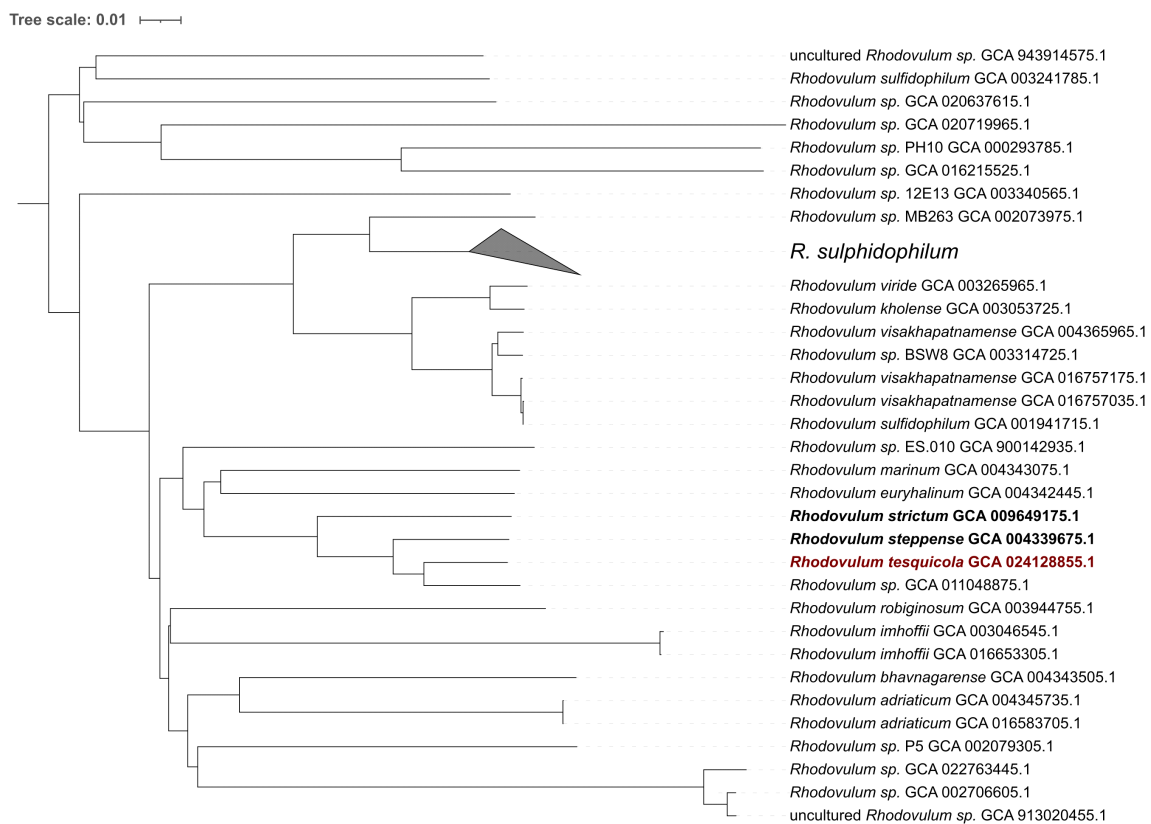

Figure 2. ANI dissimilarity dendrogram of high quality public *Rhodovulum* genomes. The dendrogram is rooted at the midpoint and scaled by genomic distance between strains.

KEGG module	Module name	Subcategory	Module completeness		
			R. tesquicola	*R. steppense*	*R. strictum*
M00135	GABA biosynthesis, eukaryotes, putrescine => GABA	Polyamine biosynthesis	0.25	0.5	0.5
M00033	Ectoine biosynthesis, aspartate => ectoine	Serine and threonine metabolism	1	1	0.4
M00555	Betaine biosynthesis, choline => betaine	Serine and threonine metabolism	0	0.5	0.4
M00793	dTDP-L-rhamnose biosynthesis	Polyketide sugar unit biosynthesis	0.06	1	0.06
M00552	D-galactonate degradation, De Ley-Doudoroff pathway, D-galactonate => glycerate-3P	Other carbohydrate metabolism	0.4	0.4	0.67
M00565	Trehalose biosynthesis, D-glucose 1P => trehalose	Other carbohydrate metabolism	0.67	0.67	0.04
M00173	Reductive citrate cycle (Arnon-Buchanan cycle)	Carbon fixation	0.8	0.7	0.8
M00579	Phosphate acetyltransferase-acetate kinase pathway, acetyl-CoA => acetate	Carbon fixation	1	0.5	1
M00115	NAD biosynthesis, aspartate => quinolinate => NAD	Cofactor and vitamin metabolism	0.8	0.4	1
M00896	Thiamine biosynthesis, archaea, AIR (+ NAD+) => TMP/TPP	Cofactor and vitamin metabolism	0.6	0.4	0.4
M00912	NAD biosynthesis, tryptophan => quinolinate => NAD	Cofactor and vitamin metabolism	0.63	0.5	0.63
M00546	Purine degradation, xanthine => urea	Purine metabolism	0.67	0.67	0.17
M00545	Trans-cinnamate degradation, trans-cinnamate => acetyl-CoA	Aromatics degradation	0.17	0.17	1
M00568	Catechol ortho-cleavage, catechol => 3-oxoadipate	Aromatics degradation	0.25	0.25	0.67
M00649	Multidrug resistance, efflux pump AdeABC	Drug resistance	0.83	0.83	0
M00714	Multidrug resistance, efflux pump QacA	Drug resistance	0.5	0.5	0
M00564	Helicobacter pylori pathogenicity signature, cagA pathogenicity island	Pathogenicity	0.04	0.04	0.5
M00616	Sulfate-sulfur assimilation	Metabolic capacity	0.33	0.33	0.71
M00618	Acetogen	Metabolic capacity	0.71	0.46	0.71

Figure 3. Completeness of the most different KEGG modules for three *Rhodovulum* genomes. Colors indicate the completeness of the module: red for the least complete, green for the most complete.
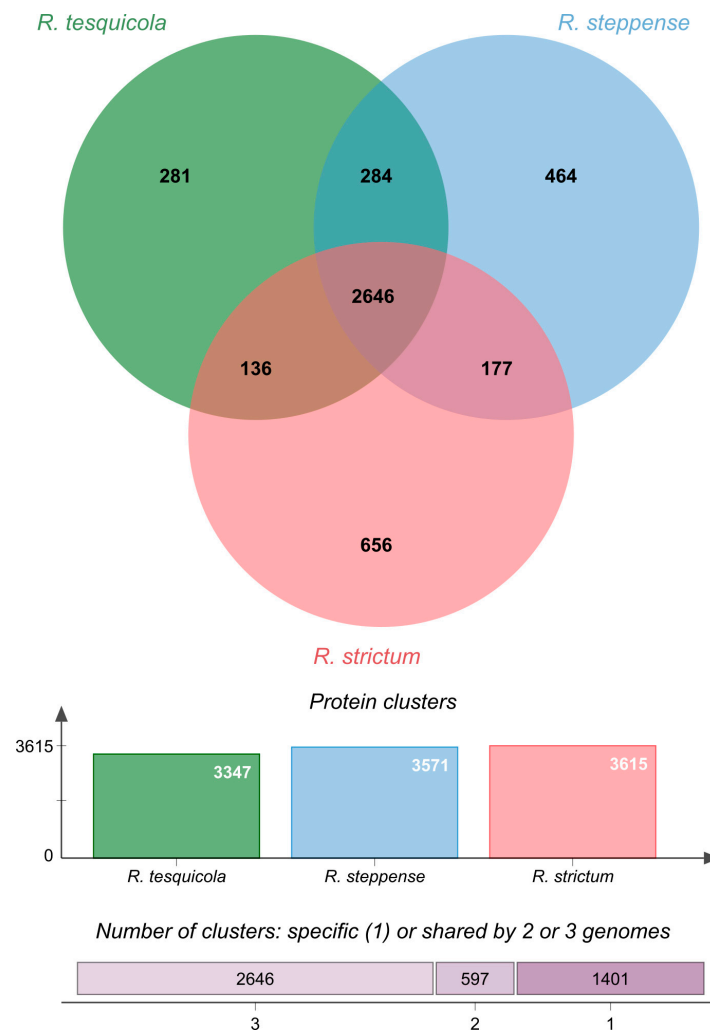

Figure 4. Venn diagram summarizing pangenome analysis of *R. tesquicola*, *R. steppense* and *R. strictum*.

3.2. Carbon Metabolism

Genes responsible for a complete TCA cycle, non-oxidative phase of pentose phosphate pathway, Entner-Doudoroff pathway, glycolysis and gluconeogenesis are present in all the genomes studied (Supplementary Table S1). The anoxygenic phototrophic bacteria can assimilate acetate as a sole organic substrate in a number of anaplerotic pathways to replenish the oxaloacetate pool [42,43]. Among them, the genes for the full ethylmalonyl pathway were found in the genome of *R. tesquicola*. The methylaspartate and glyoxylate pathways are incomplete. Although there is evidence for the isocitrate lyase presence in *R. steppense* [44], we have found no corresponding or homologous genes in any of the three genomes studied. Thus, the detection of the potential presence and functioning of glyoxylate cycle is a subject of further research. The absence of gene for acetyl-CoA synthetase in *R. tesquicola* shows that only the two-step mechanism of acetate assimilation via acetyl phosphate may function in *R. tesquicola*, while *R. steppense* and *R. strictum* have genes for both ways [45]. On the opposite, the gene for malonyl-CoA decarboxylase potentially involved in malonyl/methylmalonyl utilization [46,47] was detected only in *R. tesquicola*, thus expanding the variety of organic substrates for utilization.

3.3. Nitrogen Metabolism

During photoheterotrophic growth, the nitrogenase activity of *R. strictum* was significantly higher than that of *R. steppense* and *R. tesquicola* (Tables 2 and 3).

Table 2. Physiological features of *R. tesquicola*, *R. steppense* and *R. strictum* ([22], this study). b, Biotin; paba, p-aminobenzoate; t, thiamine; n, niacin; +, high nitrogenase activity rate; +−, moderate nitrogenase activity rate.

	R. tesquicola **A-36s**[T]	*R. steppense* **A-20s**[T]	*R. strictum* **DSM 11289**[T]
pH range (optimum)	7.5–10.0 (8.5–9.0)	7.5–10.0 (8.5)	7.5–9.0 (8.0)
NaCl range (optimum) (%)	0.3–10.0 (1–3)	0.5–10.0 (1–5)	0.25–3.00 (0.8)
Vitamin requirement	b, paba, t, n	b, n, t	b, paba, t
Nitrogenase activity	+−	+−	+

Table 3. Nitrogenase activity in the pure cultures of anoxygenic phototrophic *Alphaproteobacteria* (strain designation is indicated in case if only one strain was tested).

Species	Nitrogenase Activity (nmoles C_2H_2/mg Dry Weight per Hour)	Reference
Rhodobacter capsulatus	1430–4900	[19]
Rhodobacter capsulatus B10	3650	[19]
Cereibacter sphaeroides	1110–2755	[19]
Rhodovulum sulfidophilum W4	640	[19]
Rhodothalassium salexigens DSM 2172	630	[19]
Rhodomicrobium vannielii	660–1570	[19]
Rhodopseudomonas palustris	480–700	[19]
Rhodoblastus acidophilus	500–900	[19]
Rhodovulum tesquicola A-36s[T]	278	This work
Rhodovulum steppense A-20s[T]	523	This work
Rhodovulum strictum DSM 11289[T]	4066	This work

The genomes of all three investigated bacteria contain the genes necessary for the biosynthesis of Mo-nitrogenase (*nifHDK*) and iron–molybdenum cofactor FeMoco (*nifB*, *nifEN*, *nifV*); no alternative nitrogenases were found.

All three genomes share a common nitrogenase operon ("slow" operon) (Figure 5). These genes (GH815_10410-GH815_10390 in case of *R. strictum*) have close homologues in several representatives of the genera *Rhodobacter* and *Cereibacter*, as well as in a small number of *Alphaprotobacteria* mainly from *Hyphomicrobiales*. Relatively close homologues

are also found in the representatives of *Firmicutes* and *Chloroflexi* (Figure S1), suggesting a lateral gene transfer event in case of the "slow" operon. In addition, we have found the second nitrogenase operon in *R. strictum* genome (Figure 5). Only a limited number of *Alphaproteobacteria* including *R. strictum* and *Rhodobacter capsulatus* possesses second operon with putative higher nitrogen fixation activity ("fast" operon) (Figure 6).
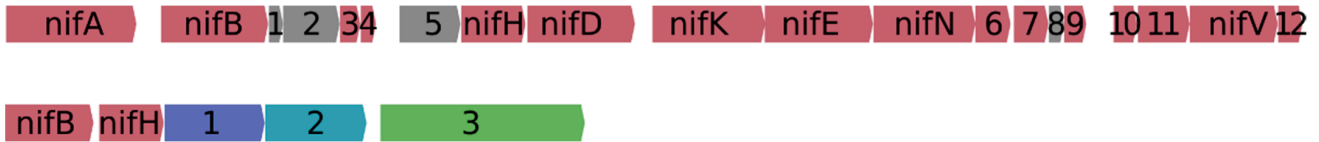

Figure 5. Nitrogen fixation operons. Top: common nitrogen fixation operon of three studied *Rhodovulum* species ("slow" operon): 1—4Fe-4S binding protein, 2—leucine rich repeat (LRR) protein, 3—*nifZ*, 4—*nifT*, 5—SIR2-like protein, 6—*nifX*, 7—*nifX*-associated nitrogen fixation protein, 8—hypothetical protein, 9—*nif*-specific ferredoxin III, 10—iron-sulfur cluster assembly accessory protein, 11—*nifU*, 12—*nifW*. Bottom: additional operon for nitrogen fixation present in *R. strictum* ("fast" operon): 1—oxidoreductase, 2—nitrogenase, 3—molybdopterin-dependent oxidoreductase (internal stop codon).
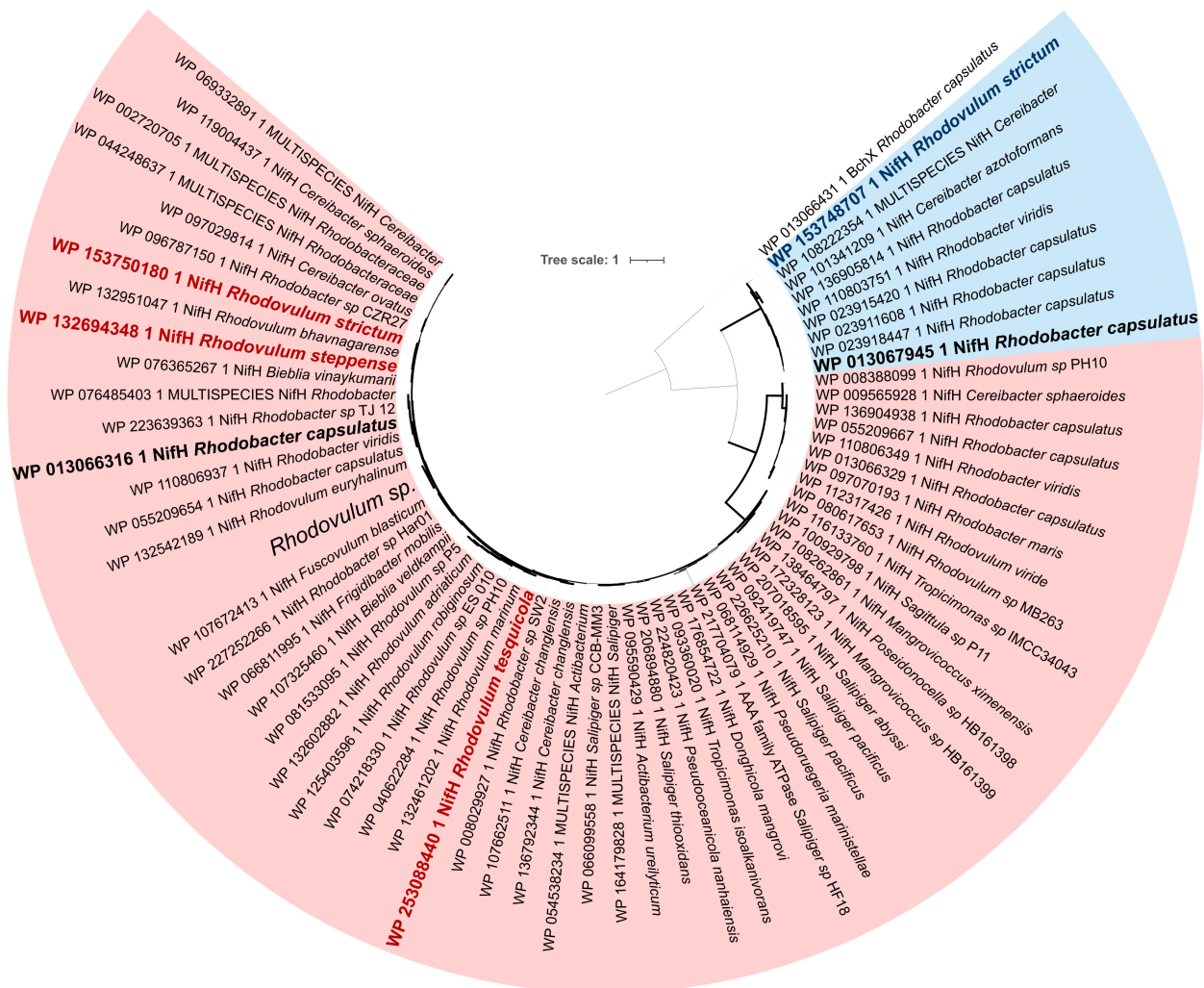

Figure 6. Phylogeny of *nifH* gene across *Rhodobacterales*. The tree includes *nifH* homologues missing in *R. steppense* and *R. tesquicola* but found in *R. strictum* and *Rhodobacter capsulatus* strain SB1003 (blue background) and *nifH* homologues appearing in all three *Rhodovulum* strains (red background). Ultrafast bootstrap values are displayed as branch width. The tree is scaled by evolutionary distance (count of substitutions per site).

The *nif* operon in *R. tesquicola* is broken into two parts. 16 genes from *nifA* (NHN26_05215) to *fdxB* (NHN26_05140) are separated from *nifU* (NHN26_05060), *nifV* (NHN26_05055) and *nifW* (NHN26_05050) by a cluster of 15 genes coding mainly for *RnfABCDGE* type electron transport complex. This break may lead to a decreased nitrogenase activity in comparison with *R. steppense*.

Nitrogen fixing phototrophic microorganisms play an important role in the nitrogen uptake by the phototrophic communities of the saline and soda lakes. It has been previously demonstrated that in soda lakes, nitrogen is fixed mainly at low salinities with the nitrogen fixation rates in the light much higher than those in the dark at the salinities of less than 100 g/L [17,48]. The investigated communities consisted mostly of the photoautotrophic heterocystous cyanobacteria with *nifH*-containing *Ectothiorhodospira* sp. present to a lesser extent. In the saline and soda lakes, *Ectothiorhodospiraceae* usually occur at the salinities up to the saturation, while *Rhodobacteraceae* reach high numbers at the mineralization below 100 g/L [21], and thus may be responsible for the nitrogen fixation activity at these conditions. However, in spite of wide distribution of *nif*-related genes among the phototrophic *Alphaproteobacteria* inhabiting saline and soda environments [49], the actual rates of nitrogen fixation in pure cultures as well as *in situ* may vary strongly (Table 3).

The product of the *nifH* gene is dinitrogenase reductase, also known as the Fe protein or component II. The Fe protein has an obligate redox role and is involved in the process of maturation of the P-clusters of the apo-MoFe protein to their catalytically active forms [50–52]. Probably the "fast" and "slow" operon function differs in the expression rate of *nifH* genes, leading to the different rates of nitrogen fixation. As *R. strictum* and *Rba. capsulatus* contain 2 operons with different *nifH* variants (Figure 5) while *R. steppense* and *R. tesquicola* have a single operon, we hypothesize that either the "fast" operon or both of them are responsible for high nitrogenase activity. Phylogenetically, the "fast" operon cluster (Figure S1) includes not only *Rba. capsulatus* and *R. strictum* performing high nitrogenase activity but also *Rhodoblastus sphagnicola*, *Rhodomicrobium vannielii* and *Rhodopseudomonas palustris* with moderate nitrogenase activity, so it is likely that both "fast" and "slow" operons may be required to achieve high activities.

The structure and organization of genes for molybdenum-only and alternative nitrogenases in *Rba. capsulatus* and other organisms are well studied [50,52–55]. The presence of multiple copies of *nif* genes in *Rba. capsulatus* was demonstrated [56], but little is known about the nitrogenase activity depending on the number and structure of nitrogenase operon. The phylogenetic studies of nitrogenase on the phyla level show that most probably nitrogenase first appeared in the methanogenic archaea and later *nif* genes entered bacteria via horizontal gene transfer [57,58]. Nevertheless, the taxa within phylum can have different nitrogenases that can function in a different way depending on the environmental conditions and physiological features [58]. For example, it was demonstrated that *Anabaena variabilis* possesses two types of nitrogenase that function under different environmental conditions and express in different types of cells [59]. In our study, we show that two types of nitrogenase operons are present in the *Rhodovulum* species that probably results in different nitrogen fixation rates, so it is necessary to further investigate this phenomenon.

Unlike *R. tesquicola* and *R. steppense*, both isolated from saline soda lakes, *R. strictum* was isolated from a coastal colored bloom [60]. In spite of the fact that purple nonsulfur bacteria do not predominate in the coastal colored blooms, the members of *Rhodovulum* may represent a significant proportion of the phototrophic bacterial population in such environments [3]. It is probably due to the high nitrogenase activity that gives them the competitive advantage in case of nitrogen removal during the bloom.

3.4. Vitamin Requirements

The vitamin requirements of the investigated bacteria according to the physiological tests are given in Table 2. Biotin requirement corresponds with the fact that biotin biosynthesis pathways are incomplete in the genome of all three bacteria. In contrast to *R. steppense* and *R. strictum*, *R. tesquicola* contains a gene for phosphomethylpyrimidine

synthase which may allow it to start the biosynthesis of thiamine monophosphate. The physiological tests, however, demonstrate that *R. tesquicola* requires additional thiamine. Some of the genes involved into thyamine synthesis are probably not expressed for reasons yet to be discovered. The niacin requirement in *R. steppense* and *R. tesquicola* may be due to the absence of *nadAB* genes which in turn are present in *R. strictum*.

3.5. Osmotic Adaptation

R. tesquicola and *R. steppense* are more halotolerant than *R. strictum* (Table 2). Apart from the complete pathway of ectoin biosynthesis and a partial trehalose synthesis pathway in genomes of the former two bacteria (Figure 3), the ectoine/hydroxyectoine TRAP transporter genes *TeaABC* [61] and glycine betaine uptake system *proVWX* [62] were found. The distribution of the compatible solutes biosynthesis and uptake genes in the investigated moderately halophilic bacteria correlates with the previously described pattern [63].

4. Conclusions

The genomes of *R. tesquicola*, *R. steppense* and *R. strictum* demonstrate significant similarity, except for some features including the vitamin synthesis pathways and the repertoire of the nitrogen fixation genes. The analysis of genome and nitrogen fixation rates shows that the studied *Rhodovulum* members are capable of fixing nitrogen with different rates, and thus may play an important role in nitrogen fixation in high-salt and alkaline conditions. Nevertheless, it becomes evident that the presence of representatives of this group does not necessarily indicate the effective nitrogen fixation in these ecosystems. Moreover, the presence of the "fast" operon does not predict high rates of nitrogen fixation in an organism. Literature data and our results show that the relationship between nitrogen fixation rate and the nitrogenase operon structure is still poorly studied, and thus further studies in this direction appear promising.

References

1. Imhoff, J.F. Genus XIV Rhodovulum. In *Bergey's Manual of Systematic Bacteriology, Volume 2: The Proteobacteria, Part C: The Alpha-, Beta-, Delta-, and Epsilonproteobacteria*; Garrity, G., Brenner, D.J., Krieg, N.R., Staley, J.T., Eds.; Springer: New York, NY, USA, 2005.
2. Hiraishi, A.; Ueda, Y. Intrageneric structure of the genus Rhodobacter: Transfer of Rhodobacter sulfidophilus and related marine species to the genus *Rhodovulum* gen. nov. *Int. J. Syst. Evol. Microbiol.* **1994**, *44*, 15–23. [CrossRef]

3. Hiraishi, A.; Nagao, N.; Yonekawa, C.; Umekage, S.; Kikuchi, Y.; Eki, T.; Hirose, Y. Distribution of phototrophic purple nonsulfur bacteria in massive blooms in coastal and wastewater ditch environments. *Microorganisms* **2020**, *8*, 150. [CrossRef] [PubMed]

4. Imhoff, J.F. True marine and halophilic anoxygenic phototrophic bacteria. *Arch. Microbiol.* **2001**, *176*, 243–254. [CrossRef]

5. Kompantseva, E.I.; Komova, A.V.; Krauzova, V.I.; Kolganova, T.V.; Panteleeva, A.N. Purple nonsulfur bacteria in weakly and moderately mineralized soda lakes of the southern Transbaikal Region and northeastern Mongolia. *Mikrobiologia* **2009**, *78*, 246–253. [CrossRef]

6. Fujimoto, H.; Wakabayashi, M.; Yamashiro, H.; Maeda, I.; Isoda, K.; Kondoh, M.; Kawase, M.; Miyasaka, H.; Yagi, K. Whole-cell arsenite biosensor using photosynthetic bacterium *Rhodovulum sulfidophilum*. *Appl. Microbiol. Biot.* **2006**, *73*, 332–338. [CrossRef]

7. Cai, J.; Wang, G. Hydrogen production by a marine photosynthetic bacterium, Rhodovulum sulfidophilum P5, isolated from a shrimp pond. *Int. J. Hydrogen. Energy* **2012**, *37*, 15070–15080. [CrossRef]

8. Higuchi-Takeuchi, M.; Morisaki, K.; Toyooka, K.; Numata, K. Synthesis of high-molecular-weight polyhydroxyalkanoates by marine photosynthetic purple bacteria. *PLoS ONE* **2016**, *11*, e0160981. [CrossRef] [PubMed]

9. Gupta, D.; Guzman, M.S.; Rengasamy, K.; Stoica, A.; Singh, R.; Ranaivoarisoa, T.O.; Davenport, E.J.; Bai, W.; McGinley, B.; Meacham, J.M.; et al. Photoferrotrophy and phototrophic extracellular electron uptake is common in the marine anoxygenic phototroph *Rhodovulum sulfidophilum*. *ISME* **2021**, *15*, 3384–3398. [CrossRef] [PubMed]

10. Melack, J.M.; Kilham, P. Photosynthetic rates of phytoplankton in East African alkaline, saline lakes 1. *Limnol. Oceanogr.* **1974**, *19*, 743–755. [CrossRef]

11. Cloern, J.E.; Cole, B.E.; Oremland, R.S. Autotrophic processes in meromictic Big Soda Lake, Nevada. *Limnol. Oceanogr.* **1983**, *28*, 1049–1061. [CrossRef]

12. Jellison, R.; Miller, L.G.; Melack, J.M.; Dana, G.L. Meromixis in hypersaline Mono Lake, California. 2. Nitrogen fluxes. *Limnol. Oceanogr.* **1993**, *38*, 1020–1039. [CrossRef]

13. Oremland, R.S. Nitrogen fixation dynamics of two diazotrophic communities in Mono Lake, California. *Appl. Environ. Microbiol.* **1990**, *56*, 614–622. [CrossRef]

14. Dubinin, A.V.; Gerasimenko, L.M.; Zavarzin, G.A. Nitrogen-fixation by cyanobacterium Microcoleus chthonoplastes from hypersaline lagoons od Lake Sivash. *Mikrobiologia* **1992**, *61*, 593–597.

15. Pinckney, J.L.; Paerl, H.W. Anoxygenic photosynthesis and nitrogen fixation by a microbial mat community in a Bahamian hypersaline lagoon. *Appl. Environ. Microbiol.* **1997**, *63*, 420–426. [CrossRef]

16. Herbst, D.B. Potential salinity limitations on nitrogen fixation in sediments from Mono Lake, California. *Int. J. Salt Lake Res.* **1998**, *7*, 261–274. [CrossRef]

17. Namsaraev, Z.; Samylina, O.; Sukhacheva, M.; Borisenko, G.; Sorokin, D.Y.; Tourova, T. Effect of salinity on diazotrophic activity and microbial composition of phototrophic communities from Bitter-1 soda lake (Kulunda Steppe, Russia). *Extremophiles* **2018**, *22*, 651–663. [CrossRef]

18. Munson, T.O.; Burris, R.H. Nitrogen fixation by Rhodospirillum rubrum grown in nitrogen-limited continuous culture. *J. Bacteriol.* **1969**, *97*, 1093–1098. [CrossRef]

19. Madigan, M.; Cox, S.S.; Stegeman, R.A. Nitrogen fixation and nitrogenase activities in members of the family Rhodospirillaceae. *J. Bacteriol.* **1984**, *157*, 73–78. [CrossRef]

20. Oda, Y.; Samanta, S.K.; Rey, F.E.; Wu, L.; Liu, X.; Yan, T.; Zhou, J.; Harwood, C.S. Functional genomic analysis of three nitrogenase isozymes in the photosynthetic bacterium *Rhodopseudomonas palustris*. *J. Bacteriol.* **2005**, *187*, 7784–7794. [CrossRef]

21. Kompantseva, E.I.; Komova, A.V.; Sorokin, D.Y. Communities of anoxygenic phototrophic bacteria in the saline soda lakes of the Kulunda Steppe (Altai Krai). *Mikrobiologia* **2010**, *79*, 89–95. [CrossRef]

22. Kompantseva, E.I.; Komova, A.V.; Novikov, A.A.; Kostrikina, N.A. Rhodovulum tesquicola sp. nov., a haloalkaliphilic purple non-sulfur bacterium from brackish steppe soda lakes. *Int. J. Syst. Evol. Microbiol.* **2012**, *62 Pt 12*, 2962–2966. [CrossRef]

23. Hardy, R.W.; Holsten, R.D.; Jackson, E.K.; Burns, R. The acetylene-ethylene assay for N2 fixation: Laboratory and field evaluation. *Plant Physiol.* **1968**, *43*, 1185–1207. [CrossRef]

24. Korzhenkov, A. ZGA: A flexible pipeline for read processing, de novo assembly and annotation of prokaryotic genomes. *bioRxiv* **2021**. [CrossRef]

25. BBMap—Bushnell B. Available online: Sourceforge.net/projects/bbmap/ (accessed on 12 April 2022).

26. Bankevich, A.; Nurk, S.; Antipov, D.; Gurevich, A.A.; Dvorkin, M.; Kulikov, A.S.; Lesin, V.M.; Nikolenko, S.I.; Pham, S.; Prjibelski, A.D.; et al. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **2012**, *19*, 455–477. [CrossRef] [PubMed]

27. Walker, B.J.; Abeel, T.; Shea, T.; Priest, M.; Abouelliel, A.; Sakthikumar, S.; Cuomo, C.A.; Zeng, Q.; Wortman, J.; Young, A.K.; et al. Pilon: An integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS ONE* **2014**, *9*, e112963. [CrossRef]

28. Tanizawa, Y.; Fujisawa, T.; Nakamura, Y. DFAST: A flexible prokaryotic genome annotation pipeline for faster genome publication. *Bioinformatics* **2018**, *34*, 1037–1039. [CrossRef] [PubMed]

29. Tatusova, T.; DiCuccio, M.; Badretdin, A.; Chetvernin, V.; Nawrocki, E.P.; Zaslavsky, L.; Lomsadze, A.; Pruitt, K.D.; Borodovsky, M.; Ostell, J. NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res.* **2016**, *44*, 6614–6624. [CrossRef] [PubMed]

30. Lomsadze, A.; Gemayel, K.; Tang, S.; Borodovsky, M. Modeling leaderless transcription and atypical genes results in more accurate gene prediction in prokaryotes. *Genome Res.* **2018**, *29*, 1079–1089. [CrossRef]

31. Aramaki, T.; Blanc-Mathieu, R.; Endo, H.; Ohkubo, K.; Kanehisa, M.; Goto, S.; Ogata, H. KofamKOALA: KEGG ortholog assignment based on profile HMM and adaptive score threshold. *Bioinformatics* **2019**, *36*, 2251–2252. [CrossRef]

32. Kanehisa, M.; Furumichi, M.; Tanabe, M.; Sato, Y.; Morishima, K. KEGG: New perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* **2017**, *45*, D353–D361. [CrossRef]

33. Eren, A.M.; Kiefl, E.; Shaiber, A.; Veseli, I.; Miller, S.E.; Schechter, M.S.; Fink, I.; Pan, J.N.; Yousef, M.; Fogarty, E.C.; et al. Community-led, integrated, reproducible multi-omics with anvi'o. *Nat. Microbiol.* **2021**, *6*, 3–6. [CrossRef] [PubMed]

34. Parks, D.H.; Imelfort, M.; Skennerton, C.T.; Hugenholtz, P.; Tyson, G.W. CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* **2015**, *25*, 1043–1055. [CrossRef]

35. Jain, C.; Rodriguez-R, L.M.; Phillippy, A.M.; Konstantinidis, K.T.; Aluru, S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat. Commun.* **2018**, *9*, 5114. [CrossRef]

36. Lechner, M.; Findeiß, S.; Steiner, L.; Marz, M.; Stadler, P.F.; Prohaska, S.J. Proteinortho: Detection of (Co-)orthologs in large-scale analysis. *BMC Bioinform.* **2011**, *12*, 124. [CrossRef] [PubMed]

37. Camacho, C.; Coulouris, G.; Avagyan, V.; Ma, N.; Papadopoulos, J.; Bealer, K.; Madden, T.L. BLAST+: Architecture and applications. *BMC Bioinform.* **2009**, *10*, 1–9. [CrossRef]

38. Fu, L.; Niu, B.; Zhu, Z.; Wu, S.; Li, W. CD-HIT: Accelerated for clustering the next-generation sequencing data. *Bioinformatics* **2012**, *28*, 3150–3152. [CrossRef] [PubMed]

39. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. [CrossRef]

40. Price, M.N.; Dehal, P.S.; Arkin, A.P. FastTree 2–approximately maximum-likelihood trees for large alignments. *PLoS ONE* **2010**, *5*, e9490. [CrossRef]

41. Letunic, I.; Bork, P. Interactive Tree of Life (iTOL) v4: Recent updates and new developments. *Nucleic Acids Res.* **2019**, *47*, W256–W259. [CrossRef]

42. Tang, K.H.; Tang, Y.J.; Blankenship, R.E. Carbon metabolic pathways in phototrophic bacteria and their broader evolutionary implications. *Front. Microbiol.* **2011**, *2*, 165. [CrossRef] [PubMed]

43. Petushkova, E.; Mayorova, E.; Tsygankov, A. TCA Cycle Replenishing Pathways in Photosynthetic Purple Non-Sulfur Bacteria Growing with Acetate. *Life* **2021**, *11*, 711. [CrossRef]

44. Eprintsev, A.T.; Falaleeva, M.I.; Lyashchenko, M.S.; Toropygin, I.Y.; Igamberdiev, A.U. Oligomeric forms of bacterial malate dehydrogenase: A study of the enzyme from the phototrophic non-sulfur bacterium *Rhodovulum steppense* A-20s. *Biosci. Biotechnol. Biochem.* **2018**, *82*, 81–89. [CrossRef]

45. Priefert, H.O.R.S.T.; Steinbüchel, A. Identification and molecular characterization of the acetyl coenzyme a synthetase gene (acoE) of Alcaligenes eutrophus. *J. Bacteriol.* **1992**, *174*, 6590–6599. [CrossRef]

46. Crosby, H.A.; Rank, K.C.; Rayment, I.; Escalante-Semerena, J.C. Structure-guided expansion of the substrate range of methyl-malonyl coenzyme a synthetase (MatB) of *Rhodopseudomonas palustris*. *Appl. Environ. Microbiol.* **2012**, *78*, 6619–6629. [CrossRef]

47. Wang, Z.; Wen, Q.; Harwood, C.S.; Liang, B.; Yang, J. A disjointed pathway for malonate degradation by *Rhodopseudomonas palustris*. *Appl. Environ. Microbiol.* **2020**, *86*, e00631-20. [CrossRef]

48. Samylina, O.S.; Namsaraev, Z.B.; Grouzdev, D.S.; Slobodova, N.V.; Zelenev, V.V.; Borisenko, G.V.; Sorokin, D.Y. The patterns of nitrogen fixation in haloalkaliphilic phototrophic communities of Kulunda Steppe soda lakes (Altai, Russia). *FEMS Microbiol. Ecol.* **2019**, *95*, fiz174. [CrossRef]

49. Tourova, T.P.; Slobodova, N.V.; Bumazhkin, B.K.; Sukhacheva, M.V.; Sorokin, D.Y. Diversity of diazotrophs in the sediments of saline and soda lakes analyzed with the use of the nifH gene as a molecular marker. *Mikrobiologia* **2014**, *83*, 634–647. [CrossRef]

50. Roberts, G.P.; MacNeil, T.; MacNeil, D.; Brill, W.J. Regulation and characterization of protein products coded by the nif (nitrogen fixation) genes of Klebsiella pneumoniae. *J. Bacteriol.* **1978**, *136*, 267–279. [CrossRef]

51. Rangaraj, P.; Ryle, M.J.; Lanzilotta, W.N.; Ludden, P.W.; Shah, V.K. In vitro biosynthesis of iron-molybdenum cofactor and maturation of the nif-encoded apodinitrogenase: Effect of substitution for NifH with site-specifically altered forms of NifH. *J. Biol. Chem.* **1999**, *274*, 19778–19784. [CrossRef]

52. Ribbe, M.W.; Bursey, E.H.; Burgess, B.K. Identification of an Fe protein residue (Glu146) of Azotobacter vinelandii nitrogenase that is specifically involved in FeMo cofactor insertion. *J. Biol. Chem.* **2000**, *275*, 17631–17638. [CrossRef]

53. Avtges, P.; Kranz, R.G.; Haselkorn, R. Isolation and organization of genes for nitrogen fixation in *Rhodopseudomonas capsulata*. *Mol. Gen. Genet.* **1985**, *201*, 363–369. [CrossRef] [PubMed]

54. Schüddekopf, K.; Hennecke, S.; Liese, U.; Kutsche, M.; Klipp, W. Characterization of anf genes specific for the alternative nitrogenase and identification of nif genes required for both nitrogenases in *Rhodobacter capsulatus*. *Mol. Microbiol.* **1993**, *8*, 673–684. [CrossRef]

55. Mus, F.; Alleman, A.B.; Pence, N.; Seefeldt, L.C.; Peters, J.W. Exploring the alternatives of biological nitrogen fixation. *Metallomics* **2018**, *10*, 523–538. [CrossRef]

56. Scolnik, P.A.; Haselkorn, R. Activation of extra copies of genes coding for nitrogenase in *Rhodopseudomonas capsulata*. *Nature* **1984**, *307*, 289–292. [CrossRef] [PubMed]

57. Raymond, J.; Siefert, J.L.; Staples, C.R.; Blankenship, R.E. The natural history of nitrogen fixation. *Mol. Biol. Evol.* **2004**, *21*, 541–554. [CrossRef] [PubMed]

58. Koirala, A.; Brözel, V.S. Phylogeny of Nitrogenase Structural and Assembly Components Reveals New Insights into the Origin and Distribution of Nitrogen Fixation across Bacteria and Archaea. *Microorganisms* **2021**, *9*, 1662. [CrossRef] [PubMed]

59. Thiel, T.; Pratte, B.S. Regulation of three nitrogenase gene clusters in the cyanobacterium *Anabaena variabilis* ATCC 29413. *Life* **2014**, *4*, 944–967. [CrossRef] [PubMed]

60. Hiraishi, A.; Ueda, Y. Isolation and characterization of Rhodovulum strictum sp. nov. and some other purple nonsulfur bacteria from colored blooms in tidal and seawater pools. *Int. J. Syst. Evol. Microbiol.* **1995**, *45*, 319–326.

61. Grammann, K.; Volke, A.; Kunte, H.J. New type of osmoregulated solute transporter identified in halophilic members of the bacteria domain: TRAP transporter TeaABC mediates uptake of ectoine and hydroxyectoine in *Halomonas elongata* DSM 2581T. *J. Bacteriol.* **2002**, *184*, 3078–3085. [CrossRef] [PubMed]

62. Tsuzuki, M.; Moskvin, O.V.; Kuribayashi, M.; Sato, K.; Retamal, S.; Abo, M.; Zeilstra-Ryalls, J.; Gomelsky, M. Salt stress-induced changes in the transcriptome, compatible solutes, and membrane lipids in the facultatively phototrophic bacterium *Rhodobacter sphaeroides*. *Appl. Environ. Microbiol.* **2011**, *77*, 7551–7559. [CrossRef] [PubMed]

63. Imhoff, J.F.; Rahn, T.; Künzel, S.; Keller, A.; Neulinger, S.C. Osmotic adaptation and compatible solute biosynthesis of phototrophic bacteria as revealed from genome analyses. *Microorganisms* **2020**, *9*, 46. [CrossRef] [PubMed]