SCIEPublish

Article

Analysis of a σ^{54} Transcription Factor L420P Mutation in Context of Increased Organic Nitrogen Tolerance of Photofermentative Hydrogen Production in *Cereibacter sphaeroides* Strain 2.4.1 Substrain H2

Nadine Wappler and Röbbe Wünschiers *

Faculty of Computer and Life Sciences, University of Applied Sciences Mittweida, Technikumplatz 17, 09648 Mittweida, Germany; wappler.nadine@gmail.com (N.W.)

* Corresponding author. E-mail: wuenschi@hs-mittweida.de (R.W.)

Received: 14 December 2023; Accepted: 19 January 2024; Available online: 23 January 2024

ABSTRACT: Photofermentative hydrogen production with non-sulfur purple bacteria like *Cereibacter sphaeroides* (formerly *Rhodobacter sphaeroides*) is a promising and sustainable process to convert organic waste into the energy carrier hydrogen gas. However, this conversion is inhibited by elevated organic nitrogen concentrations in the substrate, which limits its applicability to nitrogen-poor organic waste. We present genomic and transcriptomic insights into a substrain of *Cereibacter sphaeroides* strain 2.4.1 that shows unexpected high levels of photofermentative hydrogen evolution when fed with glutamate. Genome sequencing revealed 222 single nucleotide variances (SNVs) between the reference genome of *C. sphaeroides* strain 2.4.1 and the analyzed substrain H2. These affect 61 protein coding genes. A leucine-proline exchange is present in the σ^{54} factor (*rpoN2* gene), a global hydrogen and nitrogen metabolism regulator. We propose a model how this mutation alters DNA-binding properties that explain the unexpected organic nitrogen tolerance of hydrogen production. Transcriptomic analyses under varying glutamate concentrations support this finding. Thus, we present the first thorough genomic and transcriptomic analysis of a *Cereibacter* strain that shows promising metabolic characteristics for biotechnological hydrogen gas production from organic waste. These results suggest a potential target for strain optimization. Possibly, our key finding can be transferred to other hydrogen producing microorganisms.

Keywords: Bioenergy; Photobiological hydrogen production; Re-sequencing; Transcriptomic analysis; RNASeq; σ^{54}; RpoN; Sigma factor

© 2024 by the authors; licensee SCIEPublish, SCISCAN co. Ltd. This article is an open access article distributed under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

Currently, about 80% of the worldwide energy demand is covered by non-renewable fossil fuels. This results in the emission of pollutants and causes global warming. Hydrogen gas (H_2) is an ideal future energy carrier with a has a high energy density per unit mass [1]. In the presence of oxygen, it can be converted to either electricity by fuel cells [2] or mechanical work by combustion engines, water being the sole by-product. This makes hydrogen gas a clean energy carrier. Currently, dominant technologies for hydrogen production use fossil fuels. These include natural gas steam reforming (50% world's production), oil reforming (30%), and coal gasification (18%) [1,3]. Besides electrolytic hydrogen generation from renewable energy sources, biologically produced hydrogen gas may contribute to a green economy. A particular field of interest is wastewater treatment, since it is worldwide established with a very long technological history [4]. Over 50% of lost waste resources are contained in wastewater. Therefore, major drivers are pushing to recover and regain these substances. The European Union is currently investing substantial resources into bioeconomy funded by the European Commission under the Horizon 2020 framework.

Anaerobic phototrophs, such as the purple non-sulfur bacterium *Cereibacter sphaeroides* (formerly *Rhodobacter sphaeroides* [5,6]), are promising candidates for the photobiotechnological production of hydrogen gas from organic waste and waste water [7–9]. *Rhodobacteraceae* belong to the metabolically versatile group of alpha- or beta-proteobacteria and primarily grow photoheterotrophically under anaerobic conditions [10]. They are able to fix nitrogen

https://doi.org/10.35534/sbe.2024.10001

and grow (photo-)heterotrophic via fermentation or via aerobic or anaerobic respiration on a number of organic substrates. Hydrogen production is catalyzed by nitrogenase activity during nitrogen-limited cultivation conditions. The specificity of purple bacteria is their ability to form the metabolic energy carrier ATP by anoxygenic photosynthesis using cyclic photosynthesis. The electrons can be derived from organic acids, e.g., from biowaste [11], tofu [12], whey [13], congee [14], apple pomace [15], and wheat starch [16]. The amount of H_2 produced is depended on the substrate and mediated by various metabolic pathways [17].

The metabolism in general [18] and photofermentative hydrogen production capabilities of *Cereibacter sphaeroides* in particular [19–21] have been analyzed in great detail. Usually, purple non-sulfur bacteria evolve molecular hydrogen as a by-product from ammonia synthesis catalyzed by the nitrogenase. However, ammonia synthesis takes only place under nitrogen-deficient conditions [21]. Hence, hydrogen evolution is inhibited by high concentrations of nitrogenous substrates. This repression of the hydrogen evolving nitrogenase by fixed nitrogen is a major issue when using waste water as a substrate [22–24].

Here we characterize *Cereibacter sphaeroides* strain 2.4.1 substrain H2 that produces hydrogen gas in the presence of relatively high substrate organic nitrogen concentrations. Recently, we presented a first draft of its genome structure [25]. Here, we present our results on genomic mutations, probably acquired during lab cultivation, that cause transcriptomic changes which help to explain decoupling of the cellular organic nitrogen status and hydrogen evolution. Thus, we present a deeper genomic and the first transcriptomic characterization of this substrain.

2. Materials and Methods

2.1. Bacterial Growth and Harvest

Cereibacter sphaeroides strain 2.4.1 substrain H2 was kindly provided by the research group of Jost Weber from the Technical University Dresden, Germany. It was originally bought from the Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures (Braunschweig, Germany) as *Rhodobacter sphaeroides* DSM-No. 158. However, after years of cultivation in Weber's group it accumulated the SNVs detailed in this work. The substrain was cultivated in a stirred tank reactor with a working volume of one liter in continuous mode with a dilution rate D = 0.1 h^{-1} under conditions outlined in [20]. Samples for both, DNA and RNA analysis, were harvested after about 20 h of steady state conditions, as determined by hydrogen production rate and biomass concentration.

2.2. Nucleic Acid Preparation

Cells were separated from the culture medium by centrifugation (13,000× *g*) and frozen at −80 °C. Cells for transcriptome analyses were resuspended in 1:5 RNAlater (life technologies, Darmstadt, Germany). Nucleic acids were extracted and isolated using the MasterPure™ Complete DNA and RNA Purification Kit (epicentre, Hessisch Oldendorf, Germany) as described by the manufacturer. The nucleic acid pellet was resuspended in 50 µL TE-buffer, incubated for 15 min at 65 °C and subsequently prepared for sequencing.

2.3. DNA Sequencing

The isolated DNA was exempted from RNA contamination with RNase Cocktail™ Enzyme Mix (Life Technologies). 2 µL were added to the tube followed by an incubation at 37 °C for 30 min. Sequencing was performed with Illumina MiSeq (1 × 150 nt) by Eurofins Genomics GmbH (Ebersberg, Germany). All sequence data are available at the NCBI sequence read archive (SRA) under BioProject ID PRJNA392388.

2.4. RNA Sequencing

DNA was degraded by TURBO DNA-free™ Kit (life technologies) following the standard protocol. Ribosomal RNA was depleted with mRNA-ONLY™ Prokaryotic mRNA Isolation Kit (epicentre, Hessisch Oldendorf, Germany), using the standard protocol. Sequencing was performed with Illumina MiSeq (1 × 150 nt) by Eurofins Genomics GmbH. All sequence data are available at the NCBI sequence read archive (SRA) under BioProject ID PRJNA392388.

2.5. Data Processing and Analysis

After a first assessment of sequence quality by FastQC (Version 0.11.2) we trimmed and clipped the sequence files using the FastX-Toolkit (Version 0.0.13). Indexing and mapping to the reference genome from NCBI

(NZ_AKVW01000001.1) was performed with Bowtie2 (Version 2.1.0; [26]). Samtools (Version 0.1.19; [27]) converted the incurred SAM- to a BAM-file, which was sorted and indexed afterwards.

2.6. Genomic Characterization

From the resulting FastA file of the re-sequencing data, a consensus sequence was created with BLAST (Version 2.2.29), which delivers a text file with SNVs. The resulting sequences were compared with the reference genome (http://www.ncbi.nlm.nih.gov/nuccore/484336764). A multiple genome sequence alignment of all available *C. sphaeroides* strains from the NCBI nucleotide database (ATCC 17025: CP000661.1; ATCC 17029: NC_009049.1, CP000579.1, CP000578.1; KD131: CP001152.1, NC_011963.1, CP001153.1, CP001151.1; WS8N: NZ_CM001161.1, NZ_CM001164.1, NZ_CM001163.1, NZ_CM001162.1; and 2.4.1: NZ_AKVW01000001.1, NZ_AKVW01000002.1, NZ_AKVW01000003.1, NZ_AKVW01000004.1, NZ_AKVW01000005.1, NZ_AKVW01000006.1, NC_009008.1) was performed with Mauve (Version 2.3.1; [28]). For a better overview of evolutionary relationships between all six strains a phylogenetic tree from the Mauve output was generated using MegAlign (Version 11). The SIFT BLink tool (Version 1.03; [29]) was used to predict potential effects of SNVs on enzyme functionality.

2.7. Transcriptomic Characterization

A read count analysis from the transcriptome data was carried out with Artemis [30]. After normalization of the data by the R-package DESeq [31], we analyzed the expression values in terms of different expression and genes of interest.

2.8. Homology Modelling

As a template to map the SNV affecting the σ^{54} protein as part of the RNA polymerase holoenzyme, the X-ray structure from *PDB* was used *5NSR* [32]. Tertiary protein structures were rendered using JMOL (Version 1.3; [33]).

3. Results and Discussion

3.1. Genomic Characterization

The reference strain *Cereibacter sphaeroides* strain 2.4.1 has two circular chromosomes (3,188,524 and 943,018 bp) and five plasmids (A–E with 114,179, 105,284, 100,827, 37,100 bp, respectively) and a GC-content of 66% [10]. The complete genome encodes for 4,280 proteins and 67 RNAs. It was sequenced for the first time by the *University of Texas* in 2001 [34]. To our knowledge, eight commonly used strains, namely 2.4.1 (two genome sequences: NCBI BioProjects PRJNA57653, PRJNA40077; [10,35]), MBTLJ-13 (PRJNA316780, unpublished), MBTLJ-8 (PRJNA293258, unpublished), WS8N (PRJNA63267; [36]), ATCC 17025 (PRJNA15755; [35]), ATCC 17029 (PRJNA15754; [35]) and KD131 (PRJNA31111; [37]), have been sequenced. They are characterized by a wide genomic divergence. While the size of chromosome 1 is relatively constant among all strains, the size of chromosome 2 varies [35].

We sequenced the genome of a *Cereibacter sphaeroides* strain 2.4.1 substrain H2 (named *C.s.* substrain H2 throughout this publication; BioProject ID PRJNA392388) that is a promising candidate for photofermentative hydrogen production from organic waste [25].

Figure 1 shows the comparative genome structure and evolutionary diversity for all eight genome sequences. No genetic rearrangements between strain 2.4.1 and substrain H2 can be observed, which proofs the close genetic relatedness albeit phenotypic differences.

Choudhary et al. [35] postulated that the divergence between *Cereibacter* strains, in particular of chromosome II sequences, is based on inclusion of genetic material, horizontal DNA transfer and rearrangements of genetic loci, respectively. They focused their analysis on strains 2.4.1, ATCC 17029 and ATCC 17025 and postulated a closer relationship of the former two. The variations can be explained by adaptions to different ecological niches, which results in formation of variable metabolic pathways [35]. Our results support this finding. Most importantly, the genome sequence comparison shows no obvious structural difference between the reference strain *C.s.* strain 2.4.1 and the substrain H2 (Figure 1).

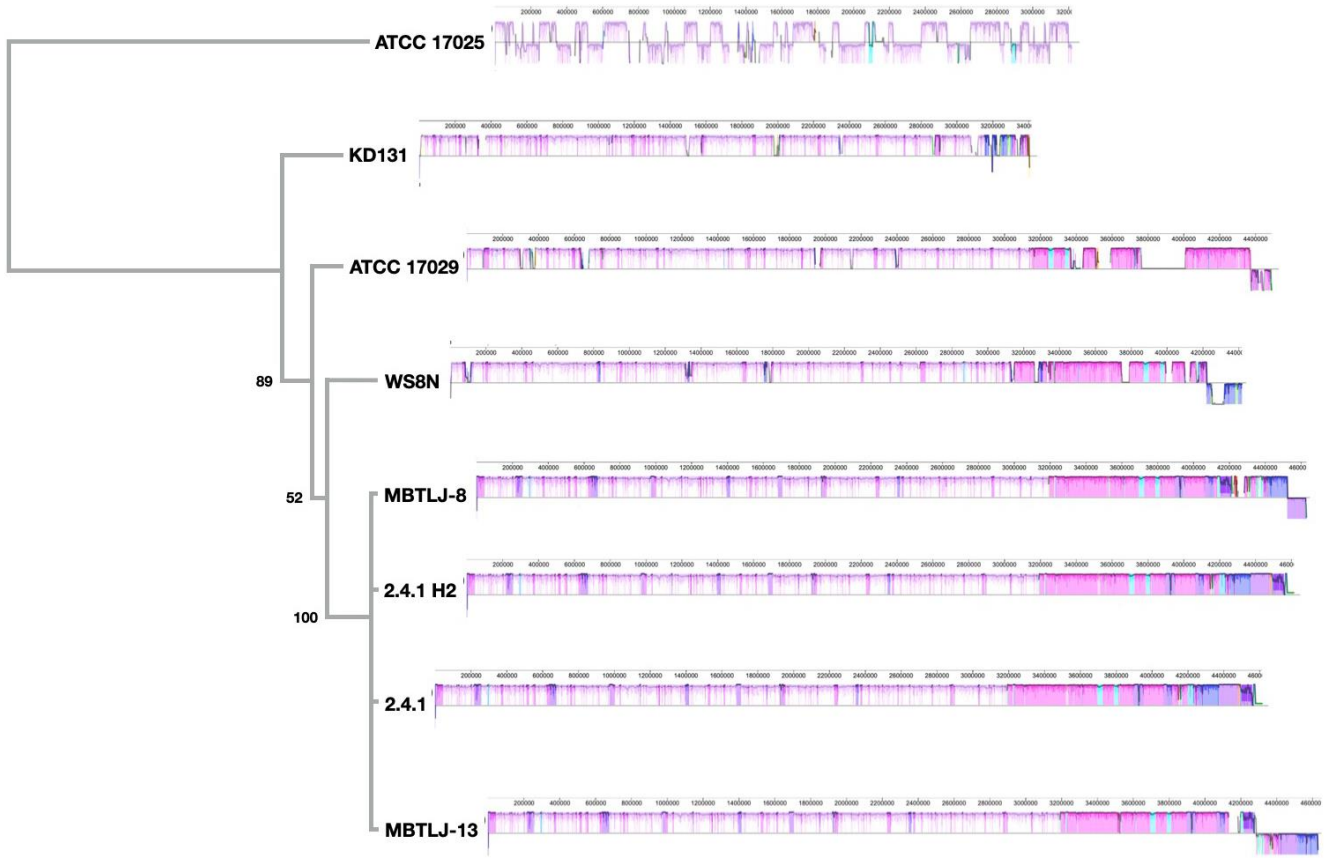

Figure 1. Multiple genome alignment and bootstrap consensus genomic distance tree of available Cereibacter sphaeroides genome assemblies. Each color block identifies a region that aligns to part of another genome. Parts of the alignment which are colored mauve are conserved among all genomes, while portions colored differently are segments conserved only among individual strains. If drawn below the horizontal line, it aligns to the reverse complement. The consensus distance tree of this genomic alignment tree is inferred by neighbor-joining from 100 replicates based on the genomic alignment from. The bootstrap values are indicated at the branch points.

3.2. Single Nucleotide Variance Analysis

Since *C.s.* substrain H2 shows no obvious genome-structure deviation from the reference strain, we analyzed the occurrences and distribution of genomic single nucleotide variances (SNVs). Therefore, we established a next generation sequencing data processing pipeline and identified 222 SNVs. The distribution of the SNVs is summarized in Figure 2. Intergenic SNVs can either be neutral or have an influence on gene-expression if they are located, e.g., in a promotor region. Intragenic SNVs can be either neutral too, or lead to the substitution of amino acids that may affect enzyme functionality. However, since intragenic SNVs can lead to *trans*-acting mutations, which can then potentially be transferred to other microorganisms, this study focuses on them. Of course, it cannot be ruled out that a intragenic SNV acts in *cis*. Since we are focusing on intragenic SNVs that lead to amino acid exchanges, this would mean that the causative SNV has two functions. We consider this to be very unlikely.

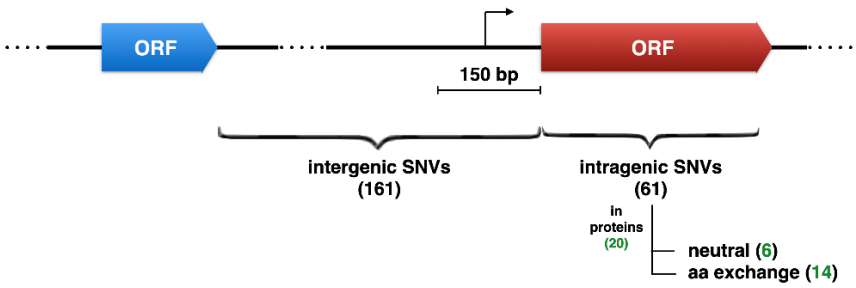

Figure 2. Allocation and classification of SNVs detected in the C.s. substrain H2 described in this work. Green numbers indicate SNVs that affect protein coding nucleotides. Of the 20 intragenic SNVs, 14 cause an amino acid exchange. See Tables 1 and 2 for details.

Of all 222 SNVs 61 are within genomic open reading frames and protein coding sequences (Table 1).

Table 1. Overview about identified genomic SNVs in open reading frames (ORFs) and protein coding sequences in C.s. strain 2.4.1 substrain H2. The latter is further specified in Table 2.

Locus	Genes	SNVs	SNVs in ORFs	SNV in Enzymes
Chr. 1	6022	128	48	18
Chr. 2	1708	0	0	0
pA	217	5	0	0
pB	209	0	0	0
pC	177	2	0	0
pD	200	83	12	2
pE	68	4	1	0
total	8601	222	61	20

Thereof, 41 SNVs are located within non-enzymatic proteins that are not obviously connected to metabolism. Of the remaining 20 enzyme coding sequence affecting SNVs, 15 cause amino acid exchanges within 14 enzymes (see Table 2 for details). A bioinformatic prediction of the effect on enzyme functionality or protein function is positive for 5 enzymes.

Table 2. SNV-based amino acid substitutions within enzymes in C.s. strain 2.4.1 substrain H2. SNVs that are predicted to affect enzyme functionality are printed in bold. Genes expressed under the conditions examined are underlined. The effect on enzyme functionality was predicted using SIFT BLink Tool as described in material and methods.

Locus	Gene-ID COC_...	Enzyme	EC-Number	AA-Exchange	Effect on Functionality
chr. 1	RS0100455	uroporphyrinogen-III synthase	4.2.1.75	P28L	no
chr. 1	RS0101150	peptidase S49	3.4.21.-	T266I	no
chr. 1	RS0102200	pyruvate kinase	2.7.1.40	P100R	no
chr. 1	RS0102425	inositol monophosphatase	3.1.3.25	–	–
chr. 1	RS0103215	DNA methyltransferase	2.1.1.72	S225X	not defined
chr. 1	**RS0105170**	**ATP-dependent DNA ligase**	**6.5.1.1**	**R46C**	**yes**
chr. 1	**RS0105350**	**malto-oligosyltrehalose trehalohydrolase**	**3.2.1.141**	**T326A**	**yes**
chr. 1	RS0105690	NADH dehydrogenase subunit D	1.6.5.3	–	–
chr. 1	RS0106065	ATPase	3.6.1.3	S317T	no
chr. 1	RS0106630	tRNA Δ2-isopentenyl-PP-transferase	2.5.1.75	–	–
chr. 1	RS0106805	sulfurtransferase	2.8.1.1.	–	–
chr. 1	RS0106985	exopolyphosphatase	3.6.1.11	H309Y	no
chr. 1	RS0107295	cobalt-precorrin-6x reductase	1.3.1.54	–	–
chr. 1	**RS0108540**	**RNA polymerase subunit sigma-54**	**2.7.7.6**	**L420P**	**yes**
chr. 1	**RS0109915**	**cellulose synthase**	**2.4.1.12**	**H249R**	**yes**
chr. 1	**RS0110470**	**hydrolase**	**3.-.-.-**	**T187C**	**yes**
chr. 1	RS0110710	carbamoyltransferase	2.1.3.3.	M1L, G300R	no
chr. 1	RS0114895	ATPase	3.6.1.3	T599P	no
pD	RS0121750	diguanylate cyclase	2.7.7.65	V249I	no
pD	RS0121870	transposase	2.7.7.-	–	–

3.3. Transcriptome Analysis

With the genomic SNV analysis, we detected no obvious evidence, e.g., mutations in enzymes that are known to be involved in either hydrogen or nitrogen metabolism, that could explain elevated hydrogen evolutions rates in the presence of high nitrogen concentration. However, the genome-wide differential gene expression at three different substrate glutamate concentrations revealed global effects (Figure 3).

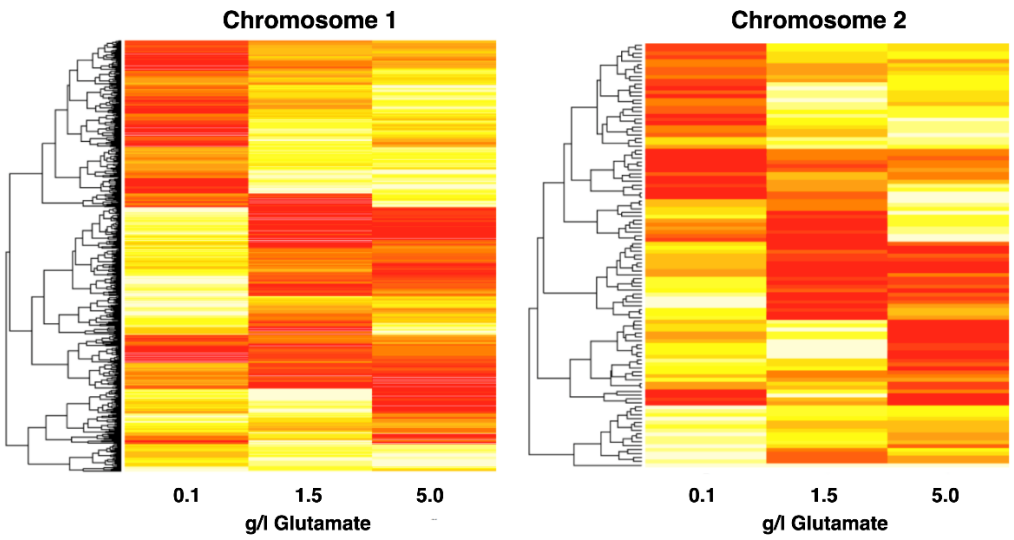

Figure 3. Heatmap of clustered gene expression data from chromosomes I and II under three different glutamate concentrations and anaerobic phototrophic growth. Red corresponds to high, white to low expression values. Expression data were elucidated from RNA-Seq read counts.

In order to concentrate on the effect of the genotype on the phenotypic observations, we checked which of the proteins from Table 2 are expressed at all at high organic nitrogen concentrations (Figure 4). Only the sigma-54 (σ54) subunit of RNA-polymerase is both expressed under the conditions examined and the SNV exhibits a potential functional effect (Table 2).

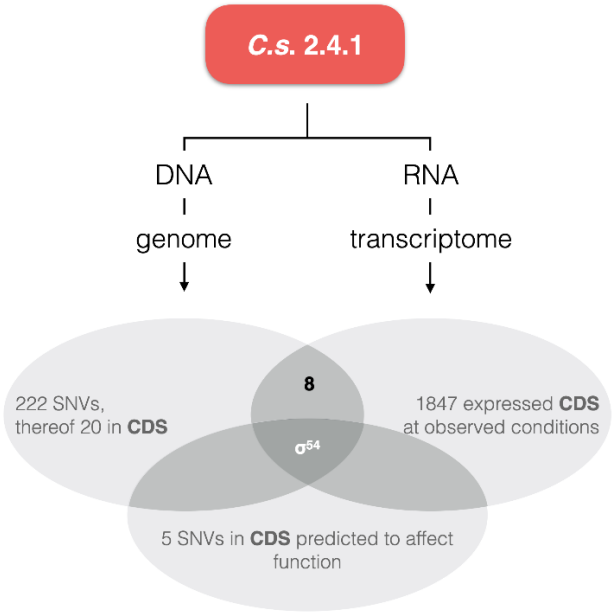

Figure 4. Result of combined genomic SNV, transcriptomic and predicted SNV effect analyses. Only the σ54 subunit of the RNA-polymerase holoenzymes matches all criteria.

3.4. Sequence Analyses of σ^{54}

In bacteria, the RNA polymerase core enzyme requires sigma factors for promoter recognition and to initiate transcription [38,39]. Sigma factors can be classified into two major families, the σ^{70} and σ^{54} family. Most sigma factors belong to the extensive σ^{70} family, including housekeeping sigma factor σ^{70} (RpoD) and alternative sigma factors (σ^{38} (RpoS), σ^{32} (RpoH) and σ^{24} (RpoE)). They direct the binding of the RNA polymerase to the promoter consensus -10 (TATAAT) and -35 (TTGACA) sequences for transcription initiation [39,40]. The σ^{54} family contains only one single member, RpoN (σ^{54}), which is structurally and functionally distinct from all other sigma factors. It directs binding of the RNA polymerase to the conserved -12 (TGC) and -24 (GG) promoter elements [40].

The σ^{54} factor (RpoN) is highly conserved in a large number of bacterial species. While it is not essential for growth, it is required for the expression of a wide variety of genes involved in many diverse functions including dicarboxylic acid transport, hydrogenase biosynthesis, and nitrogen fixation [41]. In *C. sphaeroides* at least the nitrogen fixation genes and the flagellar genes are transcribed from σ^{54}-dependent promoters. Most bacteria have a single σ^{54} factor (encoded by *rpoN*) that interacts with multiple bacterial enhancer-binding proteins which recognize specific binding sites to achieve the transcription of a specific set of genes [42]. In contrast, it was previously shown that *C. sphaeroides* encodes four σ^{54} copies, RpoN1 to RpoN4, respectively [43]. The similarity among these four copies is around 50% at the protein level. The σ^{54} identified in our analyses (Figure 4) shows 100 % sequence identity (except the L420P amino acid exchange) to the RpoN2 protein (σ^{54}-2) of *C.s.* strain WS8N (ID EGJ21523.1; [36]) and high similarity with RpoN proteins of other strains (Figure 5). Strikingly, the L420P (leucine to proline) amino acid exchange is specific to *C.s.* 2.4.1 substrain H2 analyzed in this work.

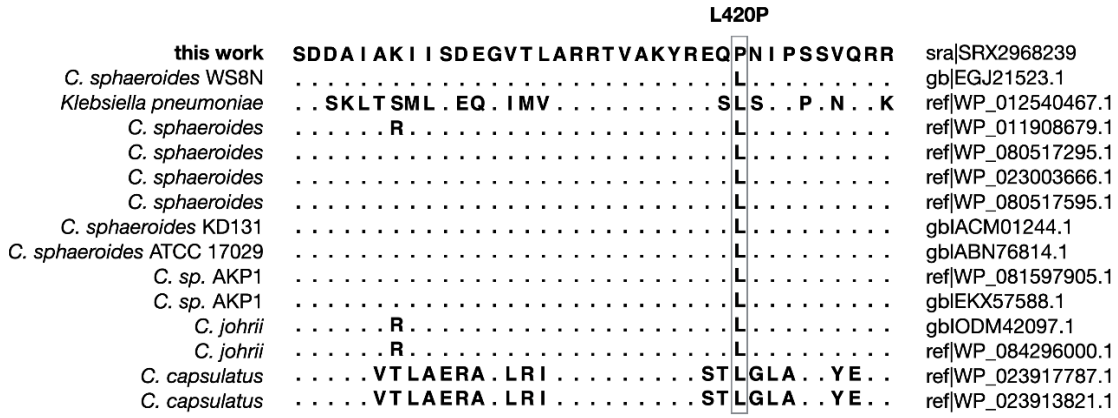

Figure 5. Excerpt of aligned Cereibacter RpoN protein sequences. Amino acids identical to the top sequence are represented by a dot. The L420P amino acid exchange is specific to C.s. 2.4.1 substrain H2 analyzed in this work. To the right, database IDs are shown.

3.5. Structural Analysis of σ54

The before mentioned unique L420P mutation in the RpoN2 protein prompted us to analyze how this amino acid might affect σ^{54} function. Glyde et al. were able to analyze the structure of the *Escherichia coli* K12 RNA polymerase associated with the of *Klebsiella pneumoniae* σ^{54} factor bound to the DNA promoter sequence [32]. Part of the protein sequence of this sigma factor is shown in the alignment in Figure 5. The identity and similarity of the *K. pneumoniae* and *C.s.* 2.4.1 H2 sequences are 32 % and 50 %, respectively. The 3D-molecule structure, shown in Figure 6, clearly shows the binding of the helix-turn-helix motif of σ^{54} to the DNA. Strikingly, the leucine at position 465, that is homologous to the leucine at position 420 in *C.s.* 2.4.1 H2, is in close proximity to the DNA-binding helix-turn-helix motif (Figure 6, red arrow). Under the assumption that the *C.s.* 2.4.1 H2 RpoN protein resembles the structure of the *K. pneumoniae* σ^{54} protein and binds in a similar way to DNA, the exchange of leucine with proline will most likely affect the binding properties.

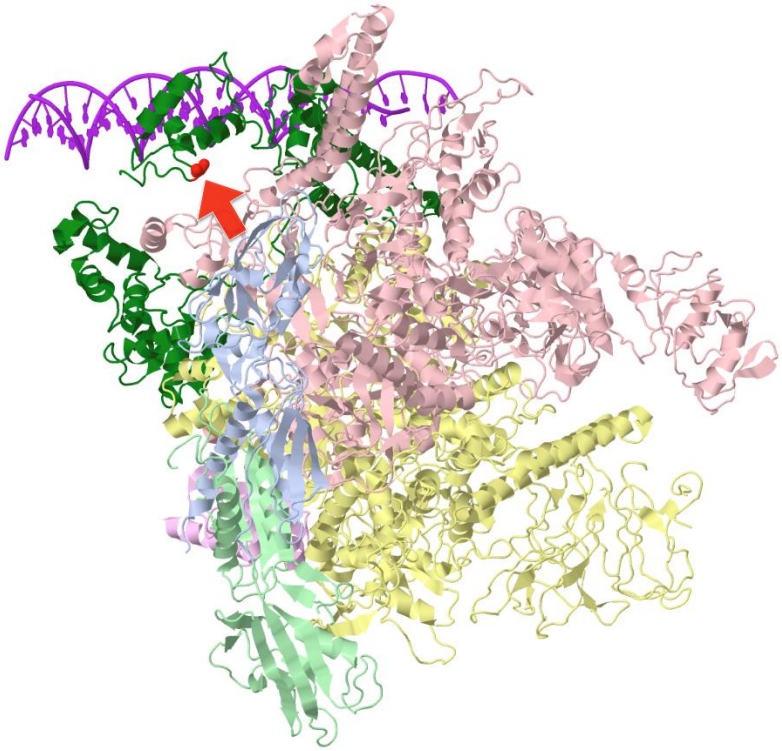

Figure 6. Structural model of the RNA polymerase holoenzyme bound to DNA based on PDB 5NSR [32]. Only the protein backbone with colored subunits is shown. The σ^{54} protein is colored entirely in dark green. The leucine at position 465 in close proximity to the DNA-binding helix-turn-helix motif is shown as red wireframe (marked by the red arrow).

3.6. A Regulatory Model

The σ^{54}-RNA polymerase holoenzyme complex forms a closed, transcriptionally silent loop that requires bacterial enhancer-binding proteins to start transcription [44,45]. Although initially identified for its role in nitrogen assimilation, σ^{54} has been found to control many other physiological processes through different enhancer-binding proteins [46]. Among those are the nitrogen regulatory protein C (NtrC) and the nitrogen fixation regulatory protein (NifA).
As already mentioned, *C. sphaeroides* encodes four σ^{54} copies. The L420P SNV is located in the *rpo*N2 gene encoding the σ^{54}-2 protein. The factor σ^{54}-1 was found to be activated by the specific enhancer-binding protein NifA [43,47,48]. NifA is an important nitrogenase activator under nitrogen limiting conditions and a lesion in σ^{54}-1 was shown to impair growth on nitrogen free medium [43]. The same study revealed that intact σ^{54}-2 is needed for the expression of flagellar genes. In contrary, Meijer and Tabita found no deviation from the wildtype with respect to nitrogen assimilation in a σ^{54}-1 disruption mutant [49] under photofermentative growth in the presence of glutamate (the condition used in our experiments). This shows the strong dependence of sigma-factor and enhancer-binding protein effects on the metabolic state. Our findings point to an involvement of σ^{54}-2 in the regulation of the nitrogenase, as summarized in Figure 7.

At high glutamate concentrations, this nitrogenous substrate can be converted by the glutamate synthetase to ammonium (NH_4^+) and α-ketoglutarate. Compared to the glutamate synthetase, we measured a higher expression level of the glutamine synthetase that forms glutamine from glutamate and ammonium. In accordance with the elevated hydrogen evolution rate at high glutamate concentration we propose that the ammonia is derived from nitrogenase activity. Usually, the nitrogenase and uptake hydrogenase are co-regulated. This way, the cell is able to recycle the energy from hydrogen that was otherwise lost. Nitrogenases deviate at least one third of all electrons to hydrogen formation. The observed hydrogen evolution rates can only be explained by an impaired activity of the uptake hydrogenase in addition to sustained nitrogenase activity. Indeed, for the close relative *R. capsulatus* it was shown that the uptake hydrogenase is independently regulated via a σ^{70}-RNA polymerase and a hydrogenase specific two-component signal transduction system (HupT/HupR) (Figure 7) [50]. HupR is an important member of the NtrC family of σ^{54} enhancer-binding proteins. Both, HupR and HupT are encoded and expressed in the substrain H2.

As shown above, the L420P mutation most likely affects σ^{54}-2–DNA binding properties. From our observations, we conclude that this in turn uncouples nitrogenase expression from the nitrogen state of the cell. In contrast, regulation of the hydrogenase remains unchanged, which ultimately leads to the observed nitrogen resistance of hydrogen production.

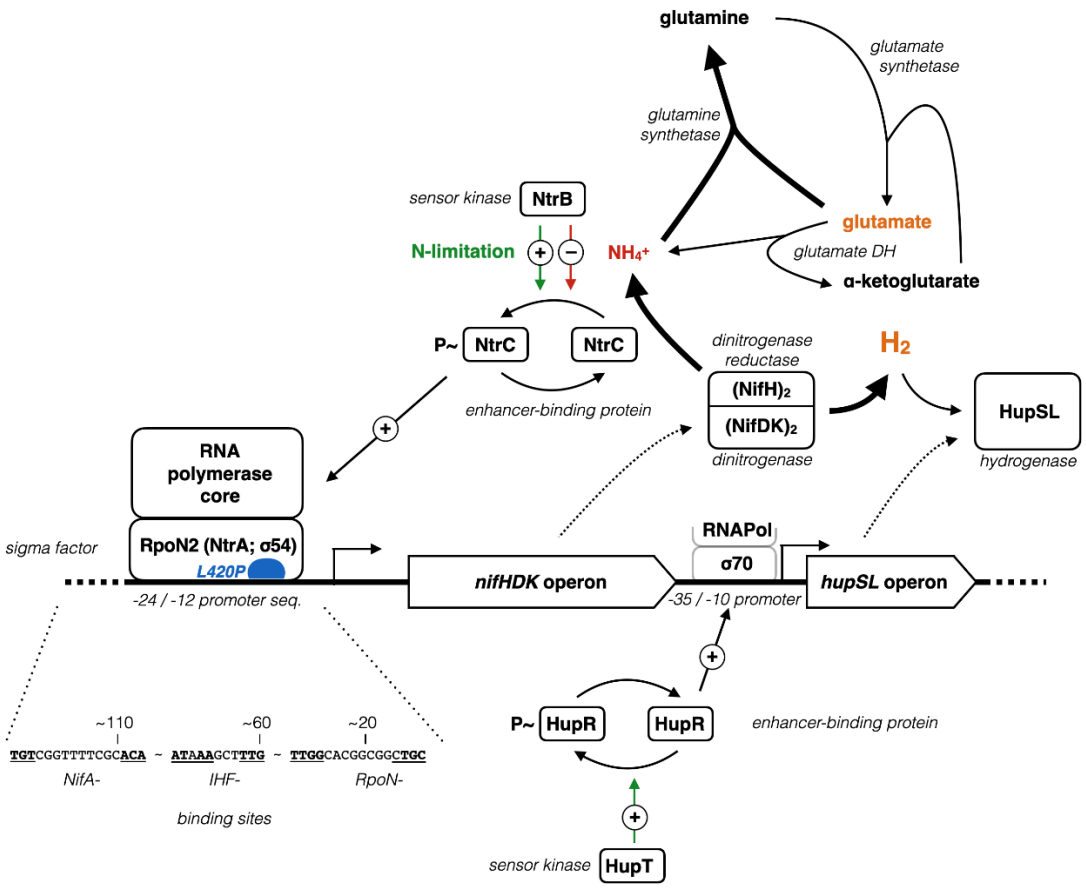

Figure 7. Effect of the L420P SNV in σ^{54} on hydrogen evolution. We propose that the down regulation of nitrogenase expression at elevated glutamate concentrations is impaired by the SNV. In contrast, expression of the hydrogen consuming uptake hydrogenase, usually regulated in concert with the nitrogenase holoenzyme, is down regulated. In total, this leads to a net yield of hydrogen gas that is released into the environment. Our data show a high expression level of glutamine synthase. The conserved (bold) binding site motifs (underlined) for the nitrogen fixation regulatory protein NifA, the integration host factor (IHF) and σ^{54} (RpoN) are shown.

4. Conclusions

We present the first genomic and transcriptomic characterization of the non-sulfur purple bacterium *Cereibacter sphaeroides* strain 2.4.1 substrain H2. It is closely related to *C. sphaeroides* strains MBTLJ-8, MBTLJ-13, and the type strain 2.4.1 DSM 158. From the latter it only distinguishable by 222 genomic SNVs at chromosome 1 and plasmids A, C, D, and E as well as by hydrogen gas evolution at high organic bound nitrogen in the substrate. This makes this substrain a promising candidate for wastewater treatment and prompted us to scout for affecting SNVs. A leucine-proline exchange at position 420 in the σ^{54}-2 factor encoding *rpoN2* gene is most likely responsible for the phenotypic behavior. Our proposed regulatory model based on genome sequence and expression data reveal an uncoupling of the otherwise co-expressed hydrogen producing nitrogenase and hydrogen consuming uptake hydrogenase that leads to a net release of hydrogen gas.

Our data (BioProject ID PRJNA392388) and findings may help to engineer improved *Cereibacter* derivatives that digest nitrogen loaded organic wastewater and simultaneously produce hydrogen gas by anaerobic photofermentation [9,51,52]. Unfortunately, CRISPR/Cas-based gene editing is not as straight-forward with *Cereibacter* as expected [53,54]. Thus, we are currently evaluating different molecular genetic techniques to edit σ^{54}.

Data Availability

The datasets analyzed during the current study are available in the NCBI sequence read archive (SRA) under BioProject ID PRJNA392388 repository, http://www.ncbi.nlm.nih.gov/bioproject/PRJNA392388/

Acknowledgments

We thank Karsten Helbig and Jost Weber from the Technical University Dresden, Germany, for providing the strain, sharing hydrogen yield data, and discussions.

Author Contributions

Conceptualization, R.W.; Methodology, N.W. and R.W.; Software, N.W.; Validation, N.W. and R.W.; Formal Analysis, R.W.; Investigation, N.W.; Resources, R.W.; Data Curation, N.W.; Writing – Original Draft Preparation, R.W.; Writing – Review & Editing, N.W. and R.W.; Visualization, N.W. and R.W.; Supervision, R.W.; Project Administration, R.W.; Funding Acquisition, R.W.

Ethics Statement

Not applicable.

Informed Consent Statement

Not applicable.

Funding

This work was kindly supported by an SAB-funded (Central Development Agency of the Free State of Saxony, Germany) PhD fellowship to NW. For financial support, we thank the SMWK (State Ministry of Higher Education, Research and the Arts, Free State of Saxony, Germany) and by European Social Fund and the Free State of Saxony in the context of the "Europa fördert Sachsen" initiative (Project number: 100098189).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Roy S, Das D. Biohythane production from organic wastes: Present state of art. *Environ. Sci. Pollut. Res. Int*. **2016**, *23*, 9391–9410. https://doi.org/10.1007/s11356-015-5469-4.
2. Levin D. Biohydrogen production: Prospects and limitations to practical application. *Int. J. Hydrogen Energy* **2004**, *29*, 173–185. https://doi.org/10.1016/s0360-3199(03)00094-6.
3. Dincer I, Acar C. Review and evaluation of hydrogen production methods for better sustainability. *Int. J. Hydrogen Energy* **2015**, *40*, 11094–11111. https://doi.org/10.1016/j.ijhydene.2014.12.035.
4. van Loosdrecht MC, Brdjanovic D. Water treatment. Anticipating the next century of wastewater treatment. *Science* **2014**, *344*, 1452–1453. https://doi.org/10.1126/science.1255183.
5. Suresh G, Sasikala C, Ramana CV. Reclassification of *Gemmobacter changlensis* to a new genus as *Cereibacter changlensis* gen. nov., comb. nov. *Int. J. Syst. Evol. Microbiol*. **2015**, *65*, 794–798. https://doi.org/10.1099/ijs.0.000016.
6. Hördt A, López MG, Meier-Kolthoff JP, Schleuning M, Weinhold LM, Tindall BJ, et al. Analysis of 1,000+ Type-Strain Genomes Substantially Improves Taxonomic Classification of. *Front. Microbiol*. **2020**, *11*, 468. https://doi.org/10.3389/fmicb.2020.00468.
7. Weber J, Krujatz F, Hilpmann G, Grützner S, Herrmann J, Thierfelder S, et al. Biotechnological hydrogen production by photosynthesis. *Eng. Life Sci.* **2014**, *14*, 592–606. https://doi.org/10.1002/elsc.201400056.
8. Adessi A, De Philippis R. Photobioreactor design and illumination systems for H_2 production with anoxygenic photosynthetic bacteria: A review. *Int. J. Hydrogen Energy* **2014**, *39*, 3127–3141. https://doi.org/10.1016/j.ijhydene.2013.12.084.
9. Orsi E, Beekwilder J, Eggink G, Kengen S. Weusthuis RA. The transition of *Rhodobacter sphaeroides* into a microbial cell factory. *Biotechnol. Bioeng*. **2021**, *118*, 531–541. https://doi.org/10.1002/bit.27593.
10. Kontur WS, Schackwitz WS, Ivanova N, Martin J, Labutti K, Deshpande S, et al. Revised sequence and annotation of the *Rhodobacter sphaeroides* 2.4.1 genome. *J. Bacteriol*. **2012**, *194*, 7016–7017. https://doi.org/10.1128/JB.01214-12.
11. Yasin NH, Mumtaz T, Hassan MA, Rahman N. Food waste and food processing waste for biohydrogen production: A review. *J. Environ. Manag*. **2013**, *130*, 375–385. https://doi.org/10.1016/j.jenvman.2013.09.009.

12. Kim M-S, Lee D-Y, Kim D-H. Continuous hydrogen production from tofu processing waste using anaerobic mixed microflora under thermophilic conditions. *Int. J. Hydrogen Energy* **2011**, *36*, 8712–8718. https://doi.org/10.1016/j.ijhydene.2010.06.040.
13. Davila-Vazquez G, Cota-Navarro CB, Rosales-Colunga LM, de León-Rodríguez A, Razo-Flores E. Continuous biohydrogen production using cheese whey: Improving the hydrogen production rate. *Int. J. Hydrogen Energy* **2009**, *34*, 4296–4304. https://doi.org/10.1016/j.ijhydene.2009.02.063.
14. Fang H, Zhu H, Zhang T. Phototrophic hydrogen production from glucose by pure and co-cultures of *Clostridium butyricum* and *Rhodobacter sphaeroides*. *Int. J. Hydrogen Energy* **2006**, *31*, 2223–2230. https://doi.org/10.1016/j.ijhydene.2006.03.005.
15. Doi T, Matsumoto H, Abe J, Morita S. Application of rice rhizosphere microflora for hydrogen production from apple pomace. *Int. J. Hydrogen Energy* **2010**, *35*, 7369–7376. https://doi.org/10.1016/j.ijhydene.2010.04.173.
16. Argun H, Kargi F. Effects of sludge pre-treatment method on bio-hydrogen production by dark fermentation of waste ground wheat. *Int. J. Hydrogen Energy* **2009**, *34*, 8543–8548. https://doi.org/10.1016/j.ijhydene.2009.08.049.
17. Kontur WS, Ziegelhoffer EC, Spero MA, Imam S, Noguera DR, Donohue TJ. Pathways involved in reductant distribution during photobiological H_2 production by *Rhodobacter sphaeroides*. *Appl. Environ. Microbiol*. **2011**, *77*, 7425–7429. https://doi.org/10.1128/AEM.05273-11.
18. Hunter CN, Daldal F, Thurnauer MC, Beatty JT, Eds. *The Purple Phototrophic Bacteria*; Springer: Dordrecht, The Netherlands, 2009. https://doi.org/10.1007/978-1-4020-8815-5.
19. Obeid J, Magnin J, Flaus J, Adrot O, Willison J, Zlatev R. Modelling of hydrogen production in batch cultures of the photosynthetic bacterium *Rhodobacter capsulatus*. *Int. J. Hydrogen Energy* **2009**, *34*, 180–185. https://doi.org/10.1016/j.ijhydene.2008.09.081.
20. Krujatz F, Härtel P, Helbig K, Haufe N, Thierfelder S, Bley T, et al. Hydrogen production by *Rhodobacter sphaeroides* DSM 158 under intense irradiation. *Bioresour. Technol.* **2015**, *175*, 82–90. https://doi.org/10.1016/j.biortech.2014.10.061.
21. Gabrielyan L, Sargsyan H, Trchounian A. Novel properties of photofermentative biohydrogen production by purple bacteria *Rhodobacter sphaeroides*: Effects of protonophores and inhibitors of responsible enzymes. *Microb. Cell Factor.* **2015**, *14*, 131. https://doi.org/10.1186/s12934-015-0324-3.
22. McKinlay JB, Harwood CS. Photobiological production of hydrogen gas as a biofuel. *Curr. Opin. Biotechnol*. **2010**, *21*, 244–251. https://doi.org/10.1016/j.copbio.2010.02.012.
23. Androga DD, Özgür E, Eroglu I, Gündüz U, Yücel M. Significance of carbon to nitrogen ratio on the long-term stability of continuous photofermentative hydrogen production. *Int. J. Hydrogen Energy* **2011**, *36*, 15583–15594. https://doi.org/10.1016/j.ijhydene.2011.09.043.
24. Keskin T, Abo-Hashesh M, Hallenbeck PC. Photofermentative hydrogen production from wastes. *Bioresour. Technol.* **2011**, *102*, 8557–8568. https://doi.org/10.1016/j.biortech.2011.04.004.
25. Leidenfrost RM, Wappler N, Wünschiers R. Draft Genome Assembly of *Rhodobacter sphaeroides* 2.4.1 Substrain H2 from Nanopore Data. *Microbiol. Resour. Announc*. **2020**, *9*, e00414-20. https://doi.org/10.1128/MRA.00414-20.
26. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **2012**, *9*, 357–359. https://doi.org/10.1038/nmeth.1923.
27. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **2009**, *25*, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352.
28. Darling AC, Mau B, Blattner FR, Perna NT. Mauve: Multiple alignment of conserved genomic sequence with rearrangements. *Genome Res*. **2004**, *14*, 1394–1403. https://doi.org/10.1101/gr.2289704.
29. Ng PC, Henikoff S. Predicting Deleterious Amino Acid Substitutions. *Genome Res.* **2001**, *11*, 863–874. https://doi.org/10.1101/gr.176601.
30. Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, et al. Artemis: Sequence visualization and annotation. *Bioinformatics* **2000**, *16*, 944–945. https://doi.org/10.1093/bioinformatics/16.10.944.
31. Anders S, Huber W. Differential expression analysis for sequence count data. *Genome Biol*. **2010**, *11*, R106. https://doi.org/10.1186/gb-2010-11-10-r106.
32. Glyde R, Ye F, Darbari VC, Zhang N, Buck M, Zhang X. Structures of RNA Polymerase Closed and Intermediate Complexes Reveal Mechanisms of DNA Opening and Transcription Initiation. *Mol. Cell* **2017**, *67*, 106–116.e4. https://doi.org/10.1016/j.molcel.2017.05.010.
33. Hanson RM. Jmol—A paradigm shift in crystallographic visualization. *J. Appl. Crystallogr.* **2010**, *43*, 1250–1260. https://doi.org/10.1107/s0021889810030256.

34. Mackenzie C, Choudhary M, Larimer FW, Predki PF, Stilwagen S, Armitage JP, et al. The home stretch, a first analysis of the nearly completed genome of *Rhodobacter sphaeroides* 2.4.1. *Photosynth. Res*. **2001**, *70*, 19–41. https://doi.org/10.1023/A:1013831823701.
35. Choudhary M, Zanhua X, Fu YX, & Kaplan S. Genome Analyses of Three Strains of *Rhodobacter sphaeroides*: Evidence of Rapid Evolution of Chromosome II. *J. Bacteriol.* **2007**, *189*, 1914–1921. https://doi.org/10.1128/jb.01498-06.
36. Porter SL, Wilkinson DA, Byles ED, Wadhams GH, Taylor S, et al. Genome sequence of *Rhodobacter sphaeroides* Strain WS8N. *J. Bacteriol*. **2011**, *193*, 4027–4028. https://doi.org/10.1128/JB.05257-11.
37. Lim SK, Kim SJ, Cha SH, Oh YK, Rhee HJ, Kim MS, et al. Complete genome sequence of *Rhodobacter sphaeroides* KD131. *J. Bacteriol.* **2009**, *191*, 1118–1119. https://doi.org/10.1128/JB.01565-08.
38. Ghosh T, Bose D, Zhang X. Mechanisms for activating bacterial RNA polymerase. *FEMS Microbiol. Rev*. **2010**, *34*, 611–627.
39. Österberg S, del Peso-Santos T, Shingler V. Regulation of alternative sigma factor use. *Annu. Rev. Microbiol*. **2011**, *65*, 37–55. https://doi.org/10.1146/annurev.micro.112408.134219.
40. Gruber TM, Gross CA. Multiple sigma subunits and the partitioning of bacterial transcription space. *Annu. Rev. Microbiol*. **2003**, *57*, 441–466. https://doi.org/10.1146/annurev.micro.57.030502.090913.
41. Merrick MJ. In a class of its own—The RNA polymerase sigma factor sigma 54 (sigma N). *Mol. Microbiol*. **1993**, *10*, 903–909. https://doi.org/10.1111/j.1365-2958.1993.tb00961.x.
42. Francke C, Groot Kormelink T, Hagemeijer Y, Overmars L, Sluijter V, Moezelaar R, et al. Comparative analyses imply that the enigmatic Sigma factor 54 is a central controller of the bacterial exterior. *BMC Genomics* **2011**, *12*, 385. https://doi.org/10.1186/1471-2164-12-385.
43. Poggio S, Osorio A, Dreyfus G, Camarena L. The four different σ54 factors of *Rhodobacter sphaeroides* are not functionally interchangeable. *Mol. Microbiol.* **2002**, *46*, 75–85. https://doi.org/10.1046/j.1365-2958.2002.03158.x.
44. Bush M, Dixon R. The role of bacterial enhancer binding proteins as specialized activators of σ54-dependent transcription. *Microbiol. Mol. Biol. Rev*. **2012**, *76*, 497–529. https://doi.org/10.1128/MMBR.00006-12.
45. Schumacher J, Joly N, Rappas M, Zhang X, Buck M. Structures and organisation of AAA+ enhancer binding proteins in transcriptional activation. *J. Struct. Biol.* **2006**, *156*, 190–199. https://doi.org/10.1016/j.jsb.2006.01.006.
46. Studholme DJ, Dixon R. Domain Architectures of σ54-Dependent Transcriptional Activators. *J. Bacteriol.* **2003**, *185*, 1757–1767. https://doi.org/10.1128/jb.185.6.1757-1767.2003.
47. Poggio S, Osorio A, Dreyfus G, Camarena L. The flagellar hierarchy of *Rhodobacter sphaeroides* is controlled by the concerted action of two enhancer-binding proteins. *Mol. Microbiol*. **2005**, *58*, 969–983. https://doi.org/10.1111/j.1365-2958.2005.04900.x.
48. Poggio S, Osorio A, Dreyfus G, Camarena L. Transcriptional specificity of RpoN1 and RpoN2 involves differential recognition of the promoter sequences and specific interaction with the cognate activator proteins. *J. Biol. Chem*. **2006**, *281*, 27205–27215. https://doi.org/10.1074/jbc.M601735200.
49. Meijer WG, Tabita FR. Isolation and characterization of the *nifUSVW-rpoN* gene cluster from *Rhodobacter sphaeroides*. *J. Bacteriol.* **1992**, *174*, 3855–3866. https://doi.org/10.1128/jb.174.12.3855-3866.1992.
50. Dischert W, Vignais PM, Colbeau A. The synthesis of *Rhodobacter capsulatus* HupSL hydrogenase is regulated by the two-component HupT/HupR system. *Mol. Microbiol.* **1999**, *34*, 995–1006. https://doi.org/10.1046/j.1365-2958.1999.01660.x.
51. Franchi E, Tosi C, Scolla G, Penna GD, Rodriguez F, Pedroni PM. Metabolically engineered *Rhodobacter sphaeroides* RV strains for improved biohydrogen photoproduction combined with disposal of food wastes. *Mar. Biotechnol.* **2004**, *6*, 552–565. https://doi.org/10.1007/s10126-004-1007-y.
52. Ryu M-H, Hull NC, Gomelsky M. Metabolic engineering of *Rhodobacter sphaeroides* for improved hydrogen production. *Int. J. Hydrogen Energy* **2014**, *39*, 6384–6390. https://doi.org/10.1016/j.ijhydene.2014.02.021.
53. Mougiakos I, Orsi E, Ghiffary MR, Post W, de Maria A, Adiego-Perez B, et al. Efficient Cas9-based genome editing of *Rhodobacter sphaeroides* for metabolic engineering. *Microb. Cell Fact*. **2019**, *18*, 204. https://doi.org/10.1186/s12934-019-1255-1.
54. Zhang Y, Yuan J. CRISPR/Cas12a-mediated genome engineering in the photosynthetic bacterium *Rhodobacter capsulatus*. *Microb. Biotechnol*. **2021**, *14*, 2700–2710. https://doi.org/10.1111/1751-7915.13805.