



Cultivation driven transcriptomic changes in the wild-type and mutant strains of *Rhodospirillum rubrum*

Katerina Jureckova^{a,1}, Marketa Nykrynova^{a,1}, Eva Slaninova^b, Hugo Fleuriot-Blitman^c, Véronique Amstutz^c, Kristyna Hermankova^a, Matej Bezdicsek^{d,e}, Katerina Mrazova^{b,f}, Kamila Hrubanova^f, Manfred Zinn^c, Stanislav Obruca^b, Karel Sedlar^{a,*}

^a Department of Biomedical Engineering, Faculty of Electrical Engineering and Communication, Brno University of Technology, Brno, Czech Republic

^b Department of Food Chemistry and Biotechnology, Faculty of Chemistry, Brno University of Technology, Brno, Czech Republic

^c Institute of Life Technologies, University of Applied Sciences and Arts Western Switzerland Valais-Wallis (HES-SO Valais-Wallis), Sion, Switzerland

^d Department of Internal Medicine – Haematology and Oncology, University Hospital Brno, Brno, Czech Republic

^e Department of Internal Medicine – Haematology and Oncology, Faculty of Medicine, Masaryk University, Brno, Czech Republic

^f Institute of Scientific Instruments of the Czech Academy of Sciences, v.v.i., Brno, Czech Republic

ARTICLE INFO

Keywords:

Rhodospirillum rubrum
RNA-Seq
Transcriptome
Genome
Depolymerase knock-out
Gene ontology
Metabolism
Fructose
Acetate
Polyhydroxyalkanoates

ABSTRACT

Purple photosynthetic bacteria (PPB) are versatile microorganisms capable of producing various value-added chemicals, e.g., biopolymers and biofuels. They employ diverse metabolic pathways, allowing them to adapt to various growth conditions and even extreme environments. Thus, they are ideal organisms for the Next Generation Industrial Biotechnology concept of reducing the risk of contamination by using naturally robust extremophiles. Unfortunately, the potential of PPB for use in biotechnology is hampered by missing knowledge on regulations of their metabolism. Although *Rhodospirillum rubrum* represents a model purple bacterium studied for polyhydroxyalkanoate and hydrogen production, light/chemical energy conversion, and nitrogen fixation, little is known regarding the regulation of its metabolism at the transcriptomic level. Using RNA sequencing, we compared gene expression during the cultivation utilizing fructose and acetate as substrates in case of the wild-type strain *R. rubrum* DSM 467^T and its knock-out mutant strain that is missing two polyhydroxyalkanoate synthases PhaC1 and PhaC2. During this first genome-wide expression study of *R. rubrum*, we were able to characterize cultivation-driven transcriptomic changes and to annotate non-coding elements as small RNAs.

1. Introduction

Bacteria represent a remarkably diverse group of organisms, which is not surprising, as, according to estimations, there are 10⁶ to 10⁸ separate prokaryotic genospecies on Earth [1]. While the majority of these microorganisms living in a wide range of natural environments seem to be uncultivable with current techniques, some organisms are very versatile and can prosper under various cultivation conditions and grown on various substrates. The ideal example is *Rhodospirillum rubrum*, a purple, non-sulfur, Gram-negative facultative anaerobe from the class of *Alphaproteobacteria*, which was observed for the first time by Esmarch in 1887 [2]. It was shown to grow both aerobically and anaerobically. The absence of oxygen triggering the photosynthesis apparatus for synthesis of membrane proteins, bacteriochlorophylls, and the carotenoids

turning the culture purple. Its metabolic versatility is further supported by the fact that it prospers under both heterotrophic and autotrophic conditions. Besides utilizing various organic substrates as saccharides or organic ions, *R. rubrum* can fix and metabolize inorganic compounds such as CO and CO₂ and is therefore considered a prospective strain for valorization of waste gases such as combustion products or syngas [3,4]. On the other hand, it is capable of producing other gaseous products utilizable as fuels, particularly hydrogen [5]. Last but not least, *R. rubrum* is a potent producer of biopolymers of the class of polyhydroxyalkanoates (PHAs) in the form of intracellular granules from many carbon sources. When supplemented by butyrate, for instance, it can accumulate up to 50 w/w % of dry weight as PHA [6]. Hence, it hosts a large variety of metabolic pathways that can be leveraged to produce sustainable carbon substrates.

* Corresponding author.

E-mail address: sedlar@vut.cz (K. Sedlar).

¹ Contributed equally

<https://doi.org/10.1016/j.csbj.2024.06.023>

Received 15 March 2024; Received in revised form 11 June 2024; Accepted 18 June 2024

Available online 20 June 2024

2001-0370/© 2024 The Authors. Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

R. rubrum is considered to be a model strain for studying the conversion of light energy to chemical energy [7], hydrogen biosynthesis [8], formation of photosynthetic membranes (PM) [9], and regulatory pathways of the nitrogen fixation system [10,11]. Therefore, it is not surprising that there are currently 10 genome assemblies of several *R. rubrum* in the GenBank database (accessed March 8th, 2024), including the genome presented in this study. The genes involved in various metabolic pathways are therefore known. Apart from synthetic pathways mentioned above, PHA synthesis occurs, and the three genes coding PHA synthases were identified in *R. rubrum* genome [12]. While one of these genes is as part of biosynthetic *phaCAB* operon, the other two are located separately in different locations in the genome. The main challenges related to *R. rubrum* and its capacity to produce value-added chemicals are a low specific growth rate, a low biomass and PHA volumetric productivity, and the need for an organic co-substrate to increase productivity of the autotrophic pathways in the case of PHAs synthesis. These hurdles could be overcome by adopting genome engineering strategies, such as the one already applied to *R. rubrum*, for example, consisting of overexpressing genes coding PHA synthases [13].

Despite relatively good genome characterization and even availability of various mutant strains, only a little is known on gene regulations in *R. rubrum* as studies exploring gene expression on a genome-wide scale using RNA sequencing (RNA-Seq) are missing and only two studies based on microarrays are available [14,15]. Thus, in our study, we compared *R. rubrum* transcriptomes from cultivation on fructose and on acetate to describe basic changes in various pathways observed along the cultivation or between substrates. Moreover, to explore the stability of expression in engineered strains, we also analyzed transcriptomes of a mutant strain with two PHA synthases deleted and compared it to the wild-type strain for both same substrates. Furthermore, as the genome assembly of the type strain was created relatively long time ago, we re-sequenced the genome of *R. rubrum* DSM 467^T, which was used for experiments to exert influence of possible mutations that might be accumulated over time. We also sequenced a $\Delta phaC1\Delta phaC2$ mutant strain to confirm deletions of PHA synthases and to capture other genomic changes. Additionally, we improved genome annotation by small RNA (sRNA) gene inference using RNA-Seq data.

2. Material and methods

2.1. Growth conditions and experiments

A freeze-dried bacterial culture of the type strain *Rhodospirillum rubrum* DSM 467^T (WT strain) was purchased from the Leibniz Institute DSMZ-German Collection of Microorganism and Cell Cultures, Braunschweig, Germany. The double mutant *R. rubrum* $\Delta phaC1\Delta phaC2$ knock-out (KO) strain was obtained from the team of Professor Kevin E. O'Connor and Professor Tanja Narancic (University College Dublin, Ireland). This PHA-negative mutant was designed and constructed as reported previously in [16].

The first cultivation step was incubation of both wild-type and mutant strains on LB broth (tryptone 10.0 g/L, yeast extract 5.0 g/L, NaCl 5.0 g/L) in Petri dishes at 30 °C in the dark for five days. The second part of cultivation was inoculated by loop of the bacterial culture from Petri dishes and the cultivation was performed in 500 mL Erlenmeyer flasks containing 100 mL of LB broth at 30 °C in the dark under shaking at 160 rpm for approximately 72 h till $OD_{660} = 1.5$. During the main part of the cultivation, the cultures were inoculated to $OD_{660} = 0.1$ by culture grown in liquid LB broth and cultivated in SYN medium. Its composition for 1 L of medium was: 250 mg of $MgSO_4 \cdot 7 H_2O$, 132 mg of $CaCl_2 \cdot 2 H_2O$, 10 g of NH_4Cl , 21 g of MOPS buffer, 10 mL $NiCl_2$ (20 μM), 100 mL of a chelated iron-molybdenum solution (0.28 g H_3BO_3 , 2.1 g of Na_2EDTA , 0.7 g of $FeSO_4 \cdot 7 H_2O$ and 0.1 g of Na_2MoO_4 per Liter of distilled water). In this medium, two carbon sources were used: 2.4 mL of 1.5 M fructose (18 mM in total) with 4 mL of 50 g/L yeast extract (1 g/L in total) and 11 mL of 1 M acetate (55 mM in total) with 4 mL of 50 g/L

yeast extract (1 g/L in total). They will be referred to as FY and AY, respectively. Cultivation was performed in triplicates in 200 mL of medium in 1 L Erlenmeyer flask at 30 °C in the dark under 160 rpm until the last sample of culture in stationary phase was taken. All the cultivations were performed under aerobic conditions.

During the cultivation of wild-type *R. rubrum* (WT) and *R. rubrum* knock-out (KO), samples were taken at three time-points, at different growth phases. In each sampling time-point, the biological triplicates were spectrophotometrically screened at $\lambda = 660$ nm and used for further DNA and RNA analysis. The culture-specific growth rates and doubling time were calculated using optical density data. The OD_{660} measurements for calculating μ_{max} were taken every hour. The μ_{max} values were calculated using a standard method: OD values were transformed into $\ln(OD_{660})$, and the linear part of the $\ln(OD_{660})$ vs. time curve was used for regression analysis to determine μ_{max} .

2.2. Transmission electron microscopy

Cultures of wild-type *R. rubrum* cultivated on both fructose and acetate as carbon sources were fixed using the high-pressure freezing method. Samples were pipetted on the 200 μm side of 3 mm copper-gold carrier type A, which was closed using the flat side of the type B carrier. Both carriers were pretreated with 1 % solution of lecithin in chloroform. Vitrification of the samples was performed using a high-pressure freezer EM ICE (Leica Microsystems, Austria). Frozen samples were transferred under liquid nitrogen into a freeze substitution unit (EM AFS2, Leica Microsystems, Austria). The substitution solution contained 1.5 % OsO_4 in acetone and the protocol was set as previously described in [17]. Freeze substitution was followed by resin embedding (Epoxy Embedding Medium kit, Sigma Aldrich, Germany) and curing at 62 °C for 48 h. Cured samples were cut to ultrathin sections (~75 nm) using a diamond knife (ultra 45°, DiATOME, Switzerland) and ultramicrotome (EM UC7, Leica Microsystems, Vienna, Austria). Since ultrathin sections were imaged using a low-voltage transmission electron microscope (LVEM 25 Delong Instruments, Czech Republic) at 25 kV voltage of electron beam, no post-staining procedure was necessary to achieve sufficient contrast [18].

2.3. DNA and RNA extraction and sequencing

High molecular weight genomic DNA of the WT strain for long-read sequencing was extracted using a MagAttract HMW DNA kit (Qiagen, NL) in accordance with the manufacturer's instructions. The DNA purity was checked using NanoDrop (Thermo Fisher Scientific, USA), while the concentration was measured using Qubit 4.0 Fluorometer (Thermo Fisher Scientific, USA), and the length was measured using Agilent 4200 TapeStation (Agilent Technologies, USA). The Ligation Sequencing 1D Kit (Oxford Nanopore Technologies, UK) was used for library preparation, and the sequencing was performed using R9.4.1 Flow Cell on the Oxford Nanopore Technologies (ONT) MinION platform.

Genomic DNA of WT and KO strains for the high-throughput short-read sequencing was extracted using GenElute Bacterial Genomic DNA Kit (Sigma-Aldrich, USA) in accordance with the manufacturer's instructions. Sequencing libraries were prepared using the KAPA HyperPlus kit, and sequencing was carried out using MiSeq Reagent kit v2 (500 cycles) on the MiSeq platform (Illumina, USA).

RNA isolation followed an optimized extraction protocol consisting of a combination of procedures focused on RNA isolation, where the crucial point was addition of 1 mL of TRIzol per 40 mg of wet biomass followed by incubation for 5 min to permit complete dissociation of the nucleoproteins complex. Then, 0.1 mL of 1-bromo-3-chloropropane per 1 mL of TRIzol™ Reagent used for cell lysis was added, and the tube was securely capped and incubated for 2–3 min at room temperature. Afterwards, the samples were centrifuged at $11,000 \times g$ at 4 °C for 15 min. Subsequently, the supernatant containing the RNA was transferred to a new tube where 70 v/v % EtOH was added at a ratio 1:1. Then, the

samples were transferred to spin columns and the procedure continued according to the manual of the NucleoSpin RNA Plus isolation kit (Macherey-Nagel) with washing and drying steps of silica membrane and elution of RNA that were stored at -80°C till the sequencing. Ribodepletion was performed with QIAseq FastSelect $-5\text{S}/16\text{S}/23\text{S}$ Kit (Qiagen, NL) for WT samples cultivated on fructose (no. 1 – 9, see [Supplementary Table S1](#)) or with RiboCop rRNA Depletion Kit for Bacteria Mixed bacterial samples (Lexogen, AT) for remaining samples (no. 10 – 33). Strand-specific sequencing libraries were prepared with NEBNext Ultra II Directional RNA Library Prep Kit (New England Biolabs, USA) and sequenced with Illumina NextSeq550 to produce reversely stranded reads. For samples no. 10 – 33, Unique Molecular Identifiers (UMIs) were added using xGen Duplex Seq Adapters (IDT, USA).

2.4. Genome assembly

Nanopore sequencing data of wild-type strain were basecalled using Guppy (v3.4.4), and the data quality was checked using PycoQC (v2.2.3, [19]). *De novo* assembly by Flye (v2.8.1, [20]) was performed, and the obtained contigs were polished using Minimap2 (v2.17, [21]) combined with Racon (v1.4.13, [22]) and then final polishing was performed by Medaka (v1.1.2).

Both WT and KO strains were sequenced using Illumina MiSeq. Obtained reads were firstly checked for quality using FastQC (v0.11.5) combined with MultiQC (v1.7, [23]) and secondly, the low-quality ends of reads together with sequencing adapters were removed by Trimmomatic (v0.36, [24]) with following settings: ILLUMINACLIP:TruSeq3-PE.fa:2:30:10 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36. Next, the set of reads was checked for contamination by human DNA, and detected reads were removed using BBMap (v39.01) and human genome GCF_000001405.40 available in the NCBI RefSeq database. In the case of the WT strain, the reads were mapped to the nanopore contigs using BWA (v0.7.17, [25]), and for managing the obtained assembly Samtools (v1.10, [26]) was employed. Final assembly polishing was made by Pilon (v1.24, [27]). In the last step, the assembled genome and plasmid were rearranged so that the *DnaA* gene was the first gene in the genome and *repB* was the first gene in the plasmid. For the KO strain assembly, BWA and Samtools were used again, but in this case, the pre-processed reads were mapped to the previously assembled WT genome.

The variant calling was conducted to find differences between both (WT and KO) analyzed strains. For this purpose, GATK4 (v4.3) was used. Underrepresented variants, variants with low coverage and false positive calls were filtered out, and the remaining ones underwent additional analysis to ascertain their presence within coding regions and, if applicable, their impact on the phenotype.

2.5. Genome annotation

NCBI Prokaryotic Genome Annotation Pipeline (PGAP) [28] was used for the wild-type strain chromosome and plasmid annotation. Protein coding genes functional annotation was performed by classifying them into clusters of orthologous groups (COG) categories from the eggNOG database via the eggNOGmapper (v2.1.6, [29]). The DNAPlotter [30], which is integrated into Artemis (v18.2.0, [31]), was used to create the chromosome and plasmid circular maps including GC content and CG skew plots, which were calculated in a window of size 10,000 bp with 200 bp step. The interspaced short palindromic repeats (CRISPR) arrays were searched by the CRISPRDetect tool (v2.4, [32]), and Cas genes were manually searched in the annotated genome. Phypsy (v4.2.6, [33]) and online tools Prophage Hunter and Phaster were used for prophage DNA identification. The restriction-modification systems were located using REBASE database [34].

The GenBank database was searched for genomes of *R. rubrum* using Entrez [35]. Then Roary (v3.13, [36]) was used to identify the core

genome, which consists of genes located in all analyzed strains, the accessory genome formed by genes presented in at least two strains but not in all of them, and the number of unique genes. The minimum percentage identity to assess whether two genes are similar was 95 %, and all other parameters for Roary were left at their default settings.

2.6. Transcriptomic analysis

The raw RNA-Seq reads for both strains were checked for their quality using FastQC and MultiQC. Next, the reads trimming was performed to discard low-quality bases and adapters using Trimmomatic with following parameters for samples no. 10 – 33: ILLUMINACLIP:TruSeq3-SE.fa:2:30:10 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36 and for samples no. 1 – 9 HEADCROP parameter was added with value 5, because first five bases of the reads contained randomized 5 bp long adapters. Remaining contamination by rRNA was removed using SortMeRNA (v4.3.4, [37]) together with the default database (smr_v4.3_default_db.fasta) and the SILVA database [38] with known 16S and 23S rRNA sequences. Processed reads were mapped to the genome of the wild-type strain using STAR (v2.7.10a, [39]), and mapped reads were deduplicated using UMI-tools (v1.1.4, [40]). Finally, the reads were counted using featureCounts function [41] from Rsubread package (R/Bioconductor). Reads counting considered two options: Uniquely mapped reads and multimapping reads. For the multimapping reads the contribution of those reads to the final count was always divided by the number of genomic loci to which the read was mapped, therefore the number of reads remained unchanged.

Created count tables for all samples were further normalized by calculating RPKM (Reads Per Kilobase per Million mapped reads) and by using built-in function in R/Bioconductor package DESeq2 [42], which was also used for differential expression analysis. Normalized count tables were used for dimension reduction using Barnes-Hut t-SNE implemented in R package Rtsne [43]. The results were visualized using ggplot2 [44] R package, which was also employed for creation of Volcano plots. Gene ontology (GO) enrichment analysis was performed by topGO [45] R/Bioconductor package together with GO map, that was created with custom Python script based on available GO annotation of wild-type strain in NCBI RefSeq database (NZ_CP077803.1 for chromosome and NZ_CP077804.1 for plasmid). Finally, the expression profiles of selected genes were visualized using heatmaps created using R packages gplots, RColorBrewer, magick and openxlsx.

To gain insight into the regulatory processes of *R. rubrum*, mapped reads were further used for non-coding RNAs (ncRNA) prediction with baerhunter (v0.9, [46]). The sample depths were normalized using sizeFactors function from DESeq2 package and visualized using boxplots. Subsequently, to obtain a count table of a collapsed annotation file with newly predicted features, baerhunter's count_features script using featureCounts function was applied. Furthermore, attention was paid to putative sRNAs, and their length distribution was visualized by a histogram. Also, these predictions were categorized as trans, respectively cis-acting elements using IRanges R package [47]. Finally, differential expression analysis was performed on the count table using DESeq2. Counts of the differentially expressed (up or down) sRNAs between specific conditions were visualized by a bar chart using R package ggplot2 as well as for all mentioned graphs.

3. Results

3.1. Genome assembly

The Oxford Nanopore Technologies MinION produced 135,804 reads; from them, 104,955 had $Q > 7$ and were used for further WT strain assembly. The mean read's length was approximately 3.5 kbp. The Illumina MiSeq provided 2.5 million 250 bp-long paired reads with an average Phred score of 35. From them, 296 reads were mapped to the human genome; thus, they were discarded from further analysis. The

assembly process of wild-type strain resulted in the final assembly of one circular chromosome and one circular plasmid with a coverage of 370×. The sequences were deposited under accession numbers CP077803.1 for chromosome and CP077804.1 for plasmid at DDBJ/EMBL/GenBank.

The obtained chromosome sequence length was 4,352,570 bp with a GC content of about 65.4 %, and the plasmid was 53,835 bp long with a GC content of around 59.8 %. In total, 3,968 open reading frames (ORFs) divided into 2,146 operons were identified for both sequences. Most of the genes were protein coding; however, 49 pseudogenes were also found. Furthermore, 521 genes overlapped with another neighboring gene, and 3 overlaps were found between a gene and a pseudogene. The overlap size was in 390 cases equal to three nucleotides and the longest overlap was 87 bp between genes KUL73_19715 and KUL73_19720 on plasmid. See [Table 1](#) for complete statistics for chromosome and plasmid.

Illumina MiSeq sequencing of KO strain provided about 2.2 million 250-bp long reads with an average Phred score of 33. Of these, 260 reads that mapped to the human genome were removed. Variant calling of the KO strain confirmed deletion of PHA polymerases. Moreover, it revealed three changes in its genome: two single nucleotide mutations and one insertion. The first mutation was found in position 823,861 (A → G) in the promoter of *cimA* gene encoding citramalate synthase, the second mutation was identified in position 1,305,962 (C → T) in the *rpoH* gene encoding RHA polymerase sigma factor RpoH and the last third change was localized in position 3,782,354 (C → CGCTTCAGGGGAAA-CACGTTATGAAG) in the promoter of gene encoding hypothetical protein.

Ten genomes obtained from GenBank were used for *R. rubrum* core genome determination. In total, 1,732 genes were found in all analyzed strains and thus formed the core genome. Another 2,859 genes were identified in at least two genomes, and thus, they comprised accessory genome. Together these 4,591 genes form the *R. rubrum* pangenome. In addition, 920 unique genes only identified in one genome were found. The number of unique genes ranged from 4 to 467, with a median of 44 (see [Supplementary Table S2](#)).

3.2. Wild-type strain functional annotation

Protein coding genes and pseudogenes were classified according to COG into 20 categories. Only 2 CDSs were not assigned to any COG; however, 1,093 CDSs were classified to class S with unknown function. The remaining 2,804 CDSs (out of all 3,850 protein coding genes and 49 pseudogenes) were classified to COG class. Individual chromosomal and plasmidic features are shown in [Fig. 1](#) together with GC content and GC skew plots for the whole genome. Substantial drop in the average CG content was observed around position 3,810,000 bp on chromosome, where are located 5S rRNA (KUL73_17050), 23S rRNA (KUL73_17055) and 16S rRNA (KUL73_17070) genes. More detailed results can be seen in [Supplementary Table S3](#).

The genome of *R. rubrum* DSM 467 was annotated in terms of gene ontology. Together, 3,047 GO terms were assigned to 1,312 genomic loci on the chromosome, and 30 GO terms were connected to 14 genomic elements on the plasmid. The most common terms for the three

Table 1
Genomic features of *Rhodospirillum rubrum* DSM 467^T.

	Chromosome	Plasmid
Length [bp]	4,352,570	53,835
GC content [%]	65.4	59.8
Number of ORFs	3,919	49
Number of operons	2,116	30
Protein coding genes	3,807	43
Pseudogenes	43	6
rRNA genes (5S, 16S, 23S)	4, 4, 4	-
tRNA	54	-
ncRNA	3	-

GO categories were: GO:0006412 “translation” for biological process assigned 58 times, for molecular function GO:0005524 “ATP binding” with 86 genomic loci and for cellular component GO:0016020 “membrane” was assigned 107 times.

CRISPR analysis showed 11 arrays with lengths from 337 to 2,467 bp and a number of spacers from 4 to 40 (see [Supplementary Table S4](#)). CRISPRDetect tool did not identify any *Cas* genes, but a manual search of the annotated genes found 34 *Cas*-like genes, but no *Cas9* protein. In the case of prophage identification, Phypsy did not identify any viral DNA, Prophage Hunter found one ambiguous prophage candidate, and Phaster localized four incomplete prophages, all of them presenting a low score (see [Supplementary Table S5](#)). The restriction-modification systems analysis revealed two systems containing restriction endonuclease and methylase where one belonged to type II and one to type III. In addition, another gene coding for methylase type II was also found (see [Supplementary Table S6](#)).

3.3. Cultivation & growth kinetics

To achieve cultivation-driven transcriptome changes in the wild-type strain of *R. rubrum* and its knock-out strain, these microorganisms were cultivated on two substrates, namely acetate and fructose. Growth curves were determined throughout cultivation, with samples taken over time and their optical density measured at 660 nm. Except for the cultivation of the KO strain on acetate, samples were collected in each cultivation in the mid-exponential phase, at the end of the exponential phase, and in the stationary phase (for precise sampling times see [Supplementary Table S1](#)). For the KO strain on acetate, only two time points were selected for further characterization (see [Fig. 2](#)).

From the obtained growth curves, the growth rate and doubling time for each cultivation were also calculated (see [Table 2](#) and [Supplementary Fig. S1](#)).

Additionally, pH was of course also analyzed during the cultivations. The pH of both cultures cultivated on fructose quickly reached values of 7.5 - 7.7 and remained almost constant until the end of cultivation. In the case of acetate, the pH reached more alkaline values of about 8.6 - 8.8 during the initial stage of cultivation and remained constant until the end of the experiment.

3.4. Ultrastructural analysis

The presence of cytoplasmic PHA granules in the cultures cultivated on two different carbon sources as well as the overall ultrastructure of the cells of *R. rubrum* wild-type was determined using low voltage transmission electron microscopy. As seen in [Fig. 3](#), both cultures contain in their spiral-shaped cells a substantial amount of small electron-lucent granules – chromatophores. However, only cultures grown on acetate as a carbon source were able to also produce PHA in their cells, which formed bigger electron-lucent granules.

3.5. RNA-Seq transcriptome

RNA sequencing ([Supplementary Table S1](#)) produced 713 millions of single-end reads averaging on 21.6 million reads per sample. After the first part of pre-processing (adapter and quality trimming), the reads' lengths were either 70 bp (samples no. 1 – 9) or 66 bp (samples no. 10 – 33) based on the used library construction method, and the average Phred score was 35 for all samples. The subsequent pre-processing step removed the remaining contamination of rRNA from the data. The proportion of reads corresponding to 16S and 23S differed among samples, e.g., all three samples for WT strain grown on fructose in the third time-point contained around 46 % of rRNA reads, and on the other hand 23 samples had less than 5 % of rRNA contamination (see [Supplementary Fig. S2](#)).

Finally, reads were mapped to the reference genome of the wild-type strain. The majority of reads (81 - 99 %) were mapped uniquely, yet

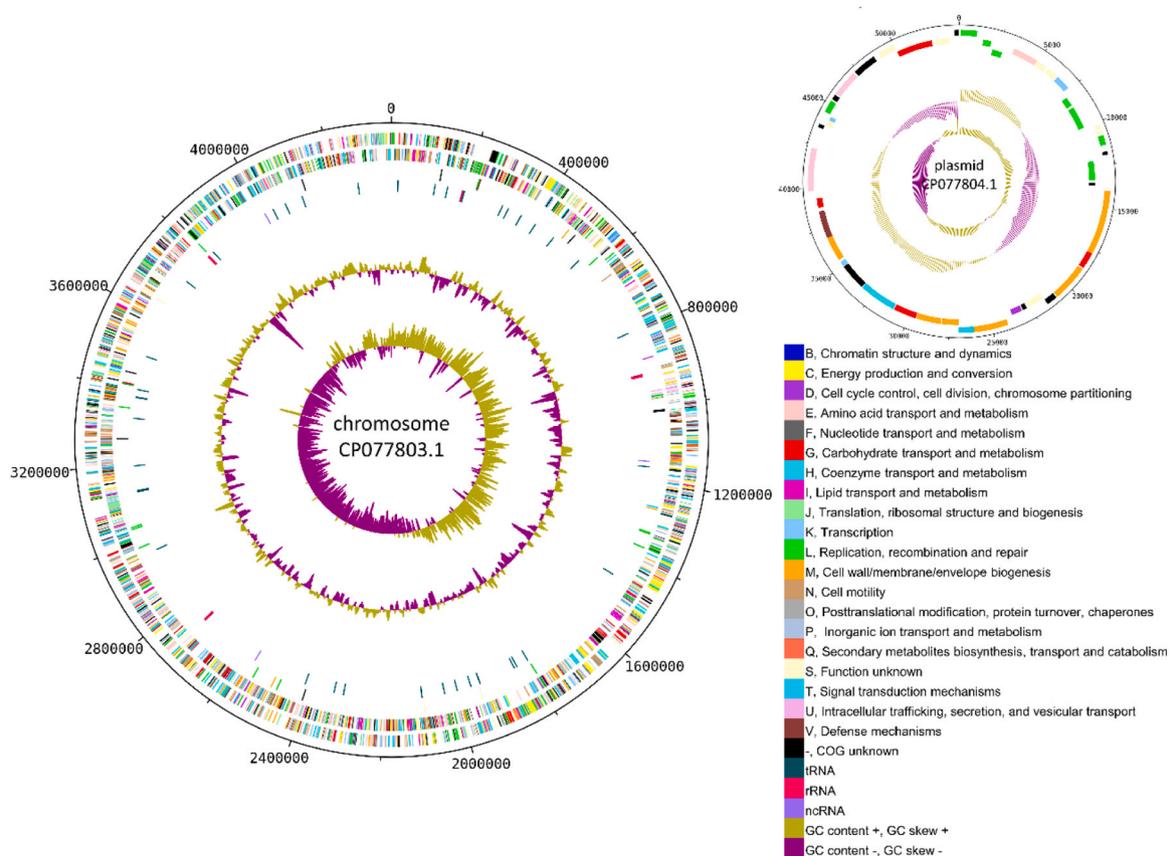


Fig. 1. Chromosomal maps of *Rhodospirillum rubrum* DSM 467^T chromosome and plasmid. The first, second, and third outermost circles represent CDSs on the forward and backward strands, and pseudogenes, respectively. Classification of COGs is represented by colors. Next, RNA genes, distinguishing among tRNA, rRNA, and ncRNA, are represented in the fourth outermost circle. The inner area represents the GC content and GC skew (window size 10,000 bp, step size 200 bp).

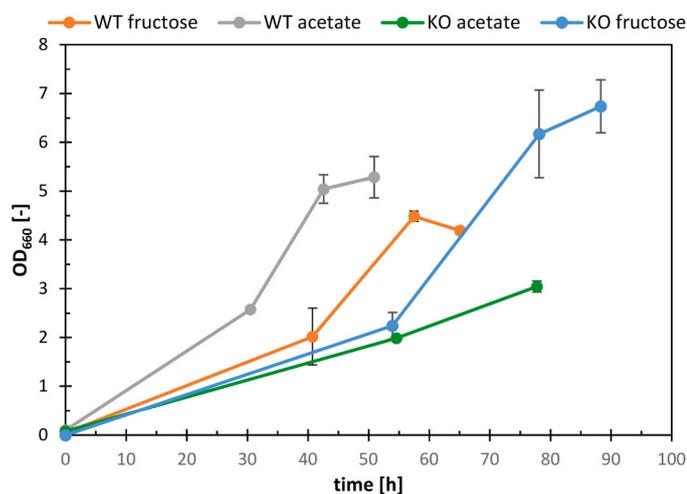


Fig. 2. Growth dynamics of *R. rubrum* WT and KO on acetate and fructose, obtained in SYN medium at 30 °C, 160 rpm in dark.

Table 2

Growth rate and doubling time for *R. rubrum* WT and KO strains grown on acetate (AY) and fructose (FY).

Sample	μ max [h ⁻¹]	Td [h]
WT FY	0.057 ± 0.001	12.082 ± 0.280
WT AY	0.094 ± 0.000	7.356 ± 0.039
KO FY	0.037 ± 0.002	18.992 ± 1.299
KO AY	0.017 ± 0.001	40.312 ± 1.946

multimapping reads, persisting deduplication, were also considered due to the existence of overlapping genes in the genome (see [Supplementary Fig. S3](#)). The multimapping reads were in further analysis down heightened by the number of associated genomic features; therefore, the original number of reads in each sample remained the same. Overall, only two pseudogenes located on the plasmid (KUL73_19640 and KUL73_19670) remained completely silent with RPKM < 1 in all tested conditions.

Reproducibility of our experiments was verified by three biological replicates for each sampling time-point and all tested condition, and visualized through dimension reduction of normalized read counts by t-Distributed Stochastic Neighbor Embedding (t-SNE) method. In most cases, the data points formed individual clusters representing specific cultivation condition and sampling time-point (see [Fig. 4](#)). Exceptions can be found in WT AY and KO AY samples, where two bigger clusters were formed based on the type of cultivated strain and used substrate, yet within them smaller clusters can be located and they are representing the time-points.

The main differences between cultivation conditions were identified through differential expression analysis, which was performed between cultivations on fructose and on acetate separately for wild-type (WT AY vs WT FY) and for knock-out strain (KO AY vs KO FY), and between wild-type and knock-out strain, separately for cultivation on fructose (KO FY vs WT FY) and on acetate (KO AY vs WT AY). Therefore, four different combinations of cultivation conditions were tested and for each of them, comparisons between available time-points were conducted, e.g., for WT AY vs WT FY differences were found between WT AY T1 vs WT FY T1, WT AY T2 vs WT FY T2, and WT AY T3 vs WT FY T3, and similarly for the rest of comparisons (see [Supplementary Figs. S4-S13](#) for corresponding Volcano plots). Results for each combination of cultivation

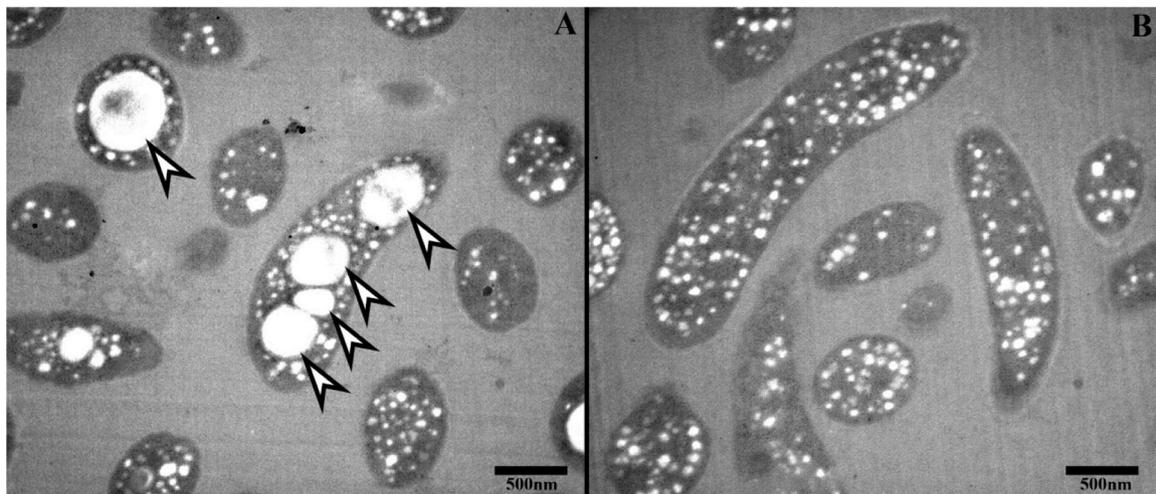


Fig. 3. Morphology of *R. rubrum* wild-type grown on A) acetate or B) fructose as carbon source, imaged using low voltage transmission electron microscopy. PHA granules are marked with arrows, smaller electron-lucent granules represent chromophores.

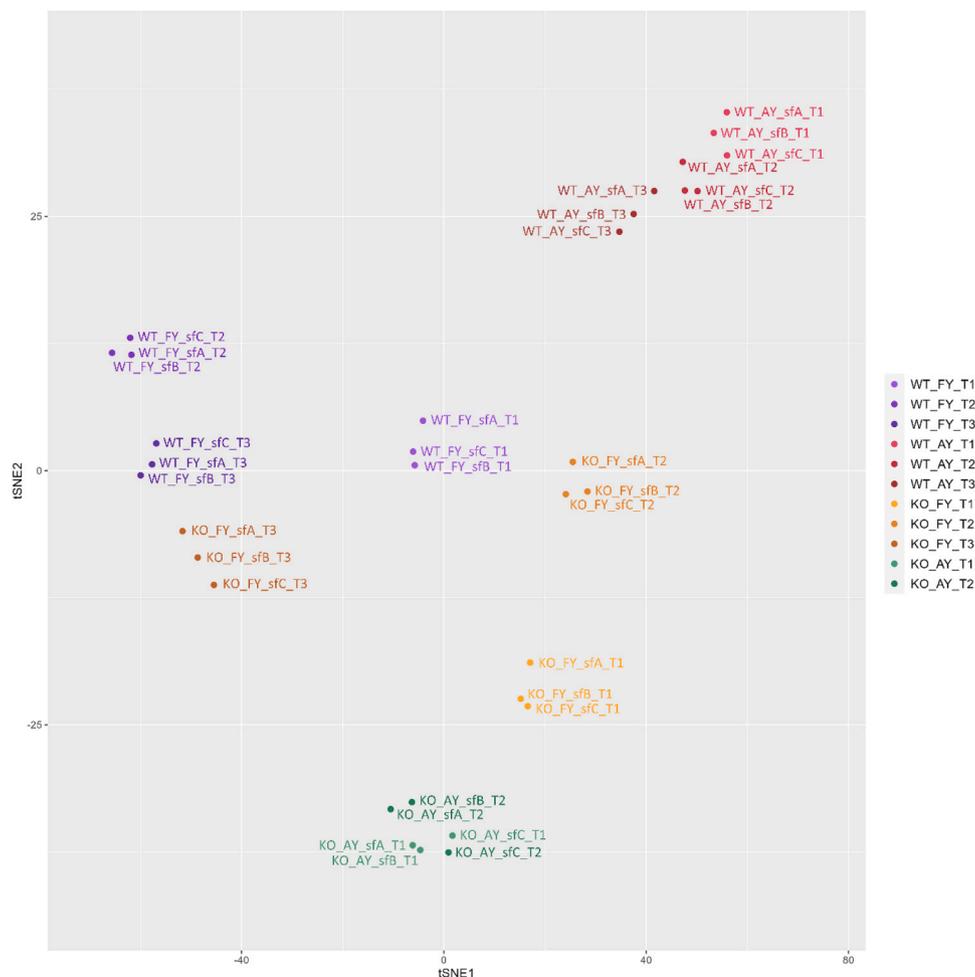


Fig. 4. Comparison of RNA-Seq samples through dimensionality reduction by t-SNE. All samples are represented as points color-coded according to strain (WT or KO), substrate (FY or AY), time-point (T1, T2 or T3) and text label indicates biological replicate (sfA, sfB or sfC).

conditions were then separately searched for genes with any statistically significant changes in the gene expression levels (either up- or down-regulated) (adjusted p value < 0.05 , Benjamini-Hochberg correction) and they were sorted into two groups based on the number of significant regulations. The first group contained genes with at least one regulation,

i.e., statistically significant differential expression, among tested pairs of time-points. The second group consisted of genes that had statistically significant change in every performed comparison of available time-points. The first group was used as gene universum in GO enrichment analysis and the second group was used as a set of interesting genes, that

were used to identify GO terms that were significantly enriched (p value < 0.05, Fisher's exact test) between tested conditions.

GO enrichment analysis revealed that for the comparison WT AY vs WT FY 18 GO terms were enriched in biological process (BP) category and 11 terms in molecular function (MF) category. BP terms were related to “disaccharide metabolic process”, “cell cycle”, “DNA recombination” and more, and MF terms corresponded to “acetyltransferase activity”, “transposase activity” or “isomerase activity” (see [Supplementary File 1](#) and [Table 3](#) for results of differential expression analysis of selected genes).

Comparison between KO AY and KO FY showed 10 enriched BP GO terms connected to “regulation of biological process”, “maintenance of DNA repeat elements” and “organic acid catabolic process” terms and the only MF term “intramolecular transferase activity” (see [Supplementary File 2](#)).

The difference between WT and KO strains grown on fructose were characterized with 12 BP terms and 5 MF terms, that were significantly enriched. The main BP terms described “archaeal or bacterial-type flagellum-dependent cell motility” or “vitamin metabolic process” processes and MF terms were related to “NAD binding”, “methyltransferase activity” or “oxidoreductase activity, acting on CH-OH group of donors” (see [Supplementary File 3](#) and [Table 4](#) for results of differential expression analysis of selected genes).

Finally, GO enrichment analysis between KO AY and WT AY identified 21 BP and 6 MF enriched terms. The main BP terms belonged to “aspartate family amino acid metabolic process”, “regulation of

biological process”, “regulation of cellular metabolic process” terms. MF terms were “pyrophosphatase activity” or “hydrolase activity, acting on acid anhydride” (see [Supplementary File 4](#) and [Table 5](#) for results of differential expression analysis of selected genes).

3.6. Small RNA prediction

For the prediction of small RNAs from the strand-specific RNA-Seq data, a coverage-based detection with baerhunter tool, which requires setting three input values, was performed. While the *low_cut_off* threshold was left on default value of 5, the *high_cut_off* threshold was set to 25 as inferred from our dataset (see [Supplementary Fig. S14](#)). The last parameter, *min_sRNA_length* of predicted elements was set to 40 bp. The length distribution of putative sRNAs (see [Supplementary Fig. S15](#)) contains a wide variety of lengths ranging from ten to thousands bp. [Table 6](#) then contains counts of predicted elements in the chromosome and plasmid sequences of *R. rubrum*. Small RNAs were further classified into two groups based on their overlap with other CDSs. Those predictions that did not overlap with the CDS on either strand, thus were in intergenic regions, were labeled as *trans*-encoded sRNAs. On the contrary, if there was even a slight overlap with any annotated element on the opposite strand as an inferred sRNA, the prediction was considered *cis*-encoded.

Additionally, differential expression analysis revealed that majority (>99%) of predicted sRNA were at least once differentially expressed among various combinations of available conditions, see [Table 6](#). The

Table 3

Differential expression analysis results of selected genes related to significantly enriched GO terms for comparison between cultivation on fructose and on acetate for wild-type strain (WT AY vs WT FY).

Gene abbr.	Putative physiological function	Locus tag	WT AY T1 vs WT FY T1		WT AY T2 vs WT FY T2		WT AY T3 vs WT FY T3	
			log2 fold change	p-adj	log2 fold change	p-adj	log2 fold change	p-adj
MF GO term: GO:0009055 electron transfer activity								
<i>nuoB</i>	NADH-quinone oxidoreductase subunit NuoB	KUL73_07380	-1.65	4.95E-02	-1.52	5.11E-02	-1.04	2.43E-01
	NADH-quinone oxidoreductase subunit C	KUL73_08070	-0.33	3.31E-02	0.54	2.38E-04	-1.08	2.95E-13
<i>nuoF</i>	NADH-quinone oxidoreductase subunit NuoF	KUL73_08085	0.66	6.42E-04	2.15	5.34E-31	1.51	3.40E-15
	NADH-quinone oxidoreductase subunit J	KUL73_08105	-0.84	7.79E-03	1.00	1.18E-03	3.49	4.78E-25
<i>nuoL</i>	NADH-quinone oxidoreductase subunit L	KUL73_08115	-1.36	2.39E-09	0.61	8.42E-03	1.98	2.17E-17
	NADH-quinone oxidoreductase subunit M	KUL73_08120	-2.27	7.70E-22	-0.35	1.58E-01	1.23	6.61E-07
<i>nuoN</i>	NADH-quinone oxidoreductase subunit NuoN	KUL73_08125	-1.12	1.87E-06	0.87	1.75E-04	2.02	6.22E-17
<i>grxD</i>	Grx4 family monothiol glutaredoxin	KUL73_03670	-1.67	1.44E-06	-3.93	1.46E-31	-2.82	7.84E-17
	c-type cytochrome	KUL73_05315	-1.51	1.16E-12	-1.72	1.44E-16	-2.66	1.14E-37
<i>petA</i>	ubiquinol-cytochrome c reductase iron-sulfur subunit	KUL73_06225	0.14	7.21E-01	-1.77	2.11E-08	-2.18	9.80E-12
	cytochrome b/b6 domain-containing protein	KUL73_06485	-0.55	8.36E-02	-2.70	5.84E-20	-2.79	4.51E-21
<i>ccoP</i>	cytochrome-c oxidase, cbb3-type subunit III	KUL73_17205	-0.74	7.63E-03	1.55	6.66E-09	1.93	3.56E-12
<i>ccoO</i>	cytochrome-c oxidase, cbb3-type subunit II	KUL73_17215	-1.91	1.81E-09	1.25	8.00E-05	1.03	1.97E-03
<i>ccoN</i>	cytochrome-c oxidase, cbb3-type subunit I	KUL73_17220	-0.79	9.68E-03	1.73	3.21E-09	1.90	3.12E-10
	pseudoazurin	KUL73_05930	-0.51	1.35E-01	3.25	1.64E-22	1.42	1.02E-03
BP GO term: GO:0006950 response to stress								
<i>ahpC</i>	peroxiredoxin	KUL73_07360	-1.56	2.05E-06	-1.80	2.39E-08	-1.56	1.57E-06
<i>msrA</i>	peptide-methionine (S)-S-oxide reductase MsrA	KUL73_12215	-1.64	1.64E-04	-5.21	1.09E-35	-4.50	5.38E-27
<i>nth</i>	endonuclease III	KUL73_00795	-1.41	8.15E-10	-3.08	5.56E-43	-2.03	2.29E-16

Table 4

Differential expression analysis results of selected genes related to significantly enriched GO terms for comparison between wild-type vs knock-out strain for cultivation on fructose (KO FY vs WT FY).

Gene abbr.	Putative physiological function	Locus tag	KO FY T1 vs WT FY T1		KO FY T2 vs WT FY T2		KO FY T3 vs WT FY T3	
			log2 fold change	p-adj	log2 fold change	p-adj	log2 fold change	p-adj
BP GO term: GO:0009110 vitamin biosynthetic process								
<i>thiE</i>	thiamine phosphate synthase	KUL73_05625	-1.19	1.32E-06	-0.44	8.13E-02	0.88	4.37E-03
<i>thiD</i>	bifunctional hydroxymethylpyrimidine kinase/ phosphomethylpyrimidine kinase	KUL73_05735	0.09	4.91E-01	0.75	5.81E-10	-0.74	2.96E-04
<i>thiL</i>	thiamine-phosphate kinase	KUL73_09425	1.17	1.44E-07	3.08	5.67E-43	1.23	1.96E-04
<i>thiC</i>	phosphomethylpyrimidine synthase ThiC	KUL73_10385	1.06	1.49E-04	0.34	2.45E-01	-0.65	3.02E-02
<i>pdxY</i>	pyridoxal kinase	KUL73_06260	1.05	4.41E-04	1.91	5.21E-11	1.74	3.95E-07
	pyridoxine 5'-phosphate synthase	KUL73_09595	1.05	4.41E-04	1.91	5.21E-11	1.74	3.95E-07
<i>pdxH</i>	pyridoxamine 5'-phosphate oxidase	KUL73_13555	0.05	8.79E-01	0.89	5.05E-04	-0.71	1.15E-02
<i>cobS</i>	cobaltochelatae subunit CobS	KUL73_01095	-0.40	3.57E-03	-0.79	3.57E-09	-0.97	2.89E-11
<i>cobN</i>	cobaltochelatae subunit CobN	KUL73_17390	0.46	8.99E-02	2.40	5.11E-21	1.07	1.19E-04
<i>cobS</i>	adenosylcobinamide-GDP ribazoletransferase	KUL73_03465	1.58	1.63E-08	2.38	5.65E-18	1.42	1.66E-04
<i>cobU</i>	bifunctional adenosylcobinamide kinase/adenosylcobinamide-phosphate guanylyltransferase	KUL73_03475	0.42	5.80E-02	2.44	7.94E-29	0.52	1.77E-01
<i>cobW</i>	cobalamin biosynthesis protein CobW	KUL73_03480	0.33	3.57E-02	1.68	1.83E-29	-0.93	1.43E-08
<i>cobF</i>	precorrin-6A synthase (deacetylating)	KUL73_15320	-0.13	6.83E-01	0.73	1.12E-02	0.22	6.50E-01
	cobalt-precorrin-6A reductase	KUL73_15410	0.36	1.54E-01	0.90	2.06E-04	-0.79	1.60E-02
	precorrin-2 C(20)-methyltransferase	KUL73_15420	1.33	2.04E-06	2.30	1.57E-16	2.27	4.18E-08
<i>cobJ</i>	precorrin-3B C(17)-methyltransferase	KUL73_15415	1.49	1.35E-17	2.75	2.30E-53	2.19	2.69E-13
BP GO term: GO:0071973 bacterial-type flagellum-dependent cell motility								
BP GO term: GO:0097588 archaeal or bacterial-type flagellum-dependent cell motility								
<i>fliJ</i>	flagellar export protein FliJ	KUL73_02755	0.84	5.36E-03	1.88	1.39E-10	1.28	2.55E-04
	flagellin	KUL73_13080	-1.35	1.00E-05	-2.56	5.22E-18	-2.46	2.30E-16
	flagellar basal body-associated FliI family protein	KUL73_14640	1.06	2.11E-02	1.13	1.12E-02	2.72	1.03E-09
<i>flgG</i>	flagellar basal-body rod protein FlgG	KUL73_14650	-5.77	2.36E-33	-7.48	2.46E-57	-1.41	5.68E-03
	flagellar basal body L-ring protein FlgH	KUL73_14660	-1.21	5.01E-03	-1.41	8.07E-04	1.12	1.26E-02
	hypothetical protein	KUL73_14700	-0.97	1.51E-03	-3.52	2.04E-33	-0.18	5.87E-01
	flagellin	KUL73_14725	-1.46	7.41E-39	-2.47	1.91E-110	-2.37	3.48E-101
MF GO term: GO:0016614 oxidoreductase activity, acting on CH-OH group of donors								
MF GO term: GO:0016616 oxidoreductase activity, acting on the CH-OH group of donors, NAD or NADP as acceptor								
<i>guaB</i>	IMP dehydrogenase	KUL73_01290	-1.21	1.53E-20	-0.64	6.71E-07	-0.99	5.12E-11
	NADP-dependent isocitrate dehydrogenase	KUL73_01875	-2.10	3.10E-50	-1.24	1.05E-18	-1.75	1.80E-28
<i>mdh</i>	malate dehydrogenase	KUL73_06315	-1.13	2.58E-06	-0.72	2.82E-03	-0.04	8.98E-01
	3-hydroxybutyryl-CoA dehydrogenase	KUL73_15880	-0.68	4.10E-04	-1.41	4.66E-14	-0.92	1.85E-06
MF GO term: GO:0051287 NAD binding								
	NADP-dependent isocitrate dehydrogenase	KUL73_01875	-2.10	3.10E-50	-1.24	1.05E-18	-1.75	1.80E-28
<i>nuoF</i>	NADH-quinone oxidoreductase subunit NuoF	KUL73_08085	0.41	3.86E-02	2.17	2.66E-31	0.45	2.96E-02

further statistics showing counts of differentially expressed sRNAs between these various combinations also distinguishes whether these elements are up/down regulated (see [Supplementary Fig. S16](#)).

4. Discussion

4.1. Genome and transcriptome

R. rubrum DSM 467^T hybrid assembly combining long Oxford

Table 5

Differential expression analysis results of selected genes related to significantly enriched GO terms for comparison between wild-type vs knock-out strain for cultivation on acetate (KO AY vs WT AY).

Gene abbr.	Putative physiological function	Locus tag	KO AY T1 vs WT AY T1		KO AY T2 vs WT AY T2	
			log2 fold change	p-adj	log2 fold change	p-adj
MF GO term: GO:0016887 ATP hydrolysis activity						
<i>bchl</i>	magnesium chelatase ATPase subunit I	KUL73_02545	-1.00	4.95E-04	-0.68	1.98E-02
	AFG1 family ATPase	KUL73_06305	-0.51	2.44E-03	-0.28	9.87E-02
<i>arsA</i>	arsenical pump-driving ATPase	KUL73_07505	1.89	4.00E-10	1.50	9.17E-07
<i>tsaE</i>	tRNA (adenosine(37)-N6)-threonylcarbamoyltransferase complex ATPase subunit type 1 TsaE	KUL73_17750	1.32	1.32E-05	1.58	2.37E-07
	AAA family ATPase	KUL73_19335	-0.63	3.15E-05	-0.68	7.10E-06
<i>mfd</i>	transcription-repair coupling factor	KUL73_08940	-0.96	6.99E-10	-1.07	7.44E-12
<i>uvrA</i>	excinuclease ABC subunit UvrA	KUL73_09080	1.09	1.59E-04	1.53	1.04E-07
<i>mutL</i>	DNA mismatch repair endonuclease MutL	KUL73_15195	-1.16	1.75E-09	-1.02	1.40E-07
<i>htpG</i>	molecular chaperone HtpG	KUL73_00370	1.06	1.54E-02	1.00	2.13E-02
<i>groL</i>	chaperonin GroEL	KUL73_00840	1.71	4.34E-04	1.88	1.06E-04
<i>clpB</i>	ATP-dependent chaperone ClpB	KUL73_03910	1.89	3.75E-11	2.55	2.95E-19
<i>dnaK</i>	molecular chaperone DnaK	KUL73_18350	0.77	1.04E-02	1.42	1.44E-06
<i>lon</i>	endopeptidase La	KUL73_08040	1.04	3.86E-12	1.14	3.20E-14
<i>hslU</i>	ATP-dependent protease ATPase subunit HslU	KUL73_18580	0.76	1.54E-05	1.06	9.12E-10

Table 6

Number of predicted small RNAs in *Rhodospirillum rubrum* DSM 467^T.

Feature	Predicted # in chromosome	Differentially expressed	Predicted # in plasmid	Differentially expressed
Total sRNAs	2,329	2,321	10	10
<i>trans</i>	55	55	0	0
<i>cis</i>	2,274	2,266	10	10

Nanopore and short Illumina reads reveal one circular chromosome and one plasmid contig as was expected based on the previously published type strain genome [10]. The chromosome exhibits GC content of 65.4 %, and the plasmid has 59.8 %, which is more than the average for Gram-negative bacteria [48]; nevertheless, it again corresponds to the previously published type strain as well as the chromosome length 4.35 Mbp and plasmid length 53.84 kbp. In comparison with type strain, the overall number of genes is similar; there is only a slight difference between the number of predicted protein-coding genes (DSM 467^T:3807 vs S1^T:3850), number of RNAs (DSM 467^T:71 vs S1^T:83) and predicted pseudogenes (DSM 467^T:49 vs S1^T:9). The differences may be due to the use of different sequencing platforms for strain sequencing and various tools for their assemblies and annotations.

RNA-Seq samples for WT cultivation on fructose (samples no. 1 – 9) were prepared slightly differently from the remaining ones (samples no. 10 – 33). Particularly, different rRNA depletion kit was used for the latter group, which resulted in very low contamination of reads belonging to 16S and 23S rRNA genes (see Supplementary Fig. S2) and UMIs were added to allow their deduplication which caused the difference in read lengths after pre-processing. While the reads without UMIs collected from the WT cultivation on fructose were 70 bp long, the remaining ones were only 66 bp long. Yet, these dissimilarities did not cause any significant differences in the overall high number of reads per sample or in the quality of the reads. Furthermore, reads were mapped to the reference wild-type strain and the results also proved a high quality

of our data as we were able to uniquely map the majority of the reads. In the worst case, WT_FY_sfb_T3 sample had 81 % of uniquely mapped reads. This particular sample also had the highest contamination of rRNA in reads with almost 47 %. On the other hand, many samples had over 98 % of the reads with unique hit within the genome, and the multimapping reads were identified only in 1–2 % of the reads. Multimapping reads could be a consequence of 523 existing overlaps between neighboring genes within the genome of *R. rubrum* (see Supplementary Fig. S3). Nevertheless, our RNA-Seq data proved to be of a very high quality, thus enabling further genome-wide study of changes in the expression levels between different cultivation conditions. As main differences were observed for particular substrates, we firstly used RNA-Seq data to detect differential expression on a genome-wide scale in particular time-points for acetate and fructose cultivation for WT (Supplementary Figs. S4 – S6) and KO (Supplementary Figs. S7 and S8) strains. Additionally, we compared differences between KO and WT strains in particular time-points on fructose (Supplementary Figs. S9 – S11) and acetate (Supplementary Figs. S12 and S13). All volcano plots showed expected relation among fold changes and statistically significant differences in expression of protein coding genes where number of down-regulated and up-regulated genes is roughly the same.

4.2. Comparative analysis of acetate and fructose growth

The GO enrichment analysis comparing acetate and fructose cultures

for the *R. rubrum* wild-type strain revealed insights into the different pathways of substrate assimilation. Fructose assimilation undergoes catabolism mainly via the Embden-Meyerhof-Parnas (EMP) [49]. This was highlighted by enriched GO terms involved in carbohydrate metabolism, such as carbohydrate biosynthetic process (GO:0016051), disaccharide metabolic process (GO:0005984) and trehalose metabolic process (GO:0005991). In contrast, acetate is assimilated via the ethylmalonyl-CoA (EMC) and methylbutanoyl-CoA (MBC) [50–52] pathways, both of which require an electron-transferring flavoprotein (ETF) to channel electrons to the membrane. The EMP pathway, in the step of glyceraldehyde-3-phosphate oxidation, produces NADH [53] which is then oxidized in the membrane through the electron transport coupled phosphorylation. Therefore, the type of electron carrier used is a key difference between the assimilation of acetate and fructose. The term NADH dehydrogenase (ubiquinone) activity (GO:0008137) could refer to the initial step in the electron transport chain [54,55] and reflect this difference in the assimilation pathway. This change in electron delivery to the membrane seems to also be found in the term electron transfer activity (GO:0009055). Investigation of the specific genes involved in these terms revealed that most of the genes for the subunits of the NADH-quinone oxidoreductase are more transcribed and regulated under fructose conditions, except for the subunit F (*nuoF*) (KUL73_08085) which is upregulated for acetate (see Table 3 and Supplementary File 1). The NADH-quinone oxidoreductase is an important element of the respiratory chain, therefore a different efficiency in the electron transport chain (ETC) between both carbon sources could be expected, potentially leading to different energy levels. This hypothesis is also supported by the different expression and regulation of genes involved in the electron transport chain, c-type cytochrome (KUL73_05315), ubiquinol-cytochrome c reductase complex (KUL73_06225), cytochrome b/b6 domain (KUL73_06485) found under GO:0009055 (see Table 3). It is also noteworthy to mention that the genes for the *cbb3*-type cytochrome c oxidase complex [56], usually involved in presence of low oxygen levels, seemed to be transcribed for both carbon sources. This observation may appear counterintuitive considering the aerobic culture conditions, but the poor gas exchange expected in shake flask without baffles could have led to lower dissolved oxygen levels. Furthermore, the gene coding for pseudoazurin *ppaZ* (KUL73_05930), which is encoded by *ppaZ* [57], is positively regulated by the Reg/Prr system, related to sensing low oxygen levels, was activated exclusively under acetate culture.

The GO enrichment analysis also identified genes that could be relevant in the interpretation of the lower maximum specific growth rate (μ_{\max}) measured for fructose (see Table 2). Genes involved in reactive oxygen species (ROS) scavenging and management of the oxidative response (peroxiredoxin (KUL73_07360), peptide-methionine (S)-S-oxide reductase *MsrA* (KUL73_12215) and endonuclease III (KUL73_00795)) were more expressed during the growth on fructose (see GO:0006950 response to stress in Table 3 and Supplementary File 1). While this upregulation of a limited set of three genes suggests a possible increase in ROS production, this alone does not establish a direct causal relationship with lower μ_{\max} . Further investigations are required to understand the nature of this upregulation. However, it has been observed that ROS production could be coming from electron leakage from the ETC [58,59]. The increased presence of ROS may indicate that electrons are diverted from the ETC, diminishing the efficiency of the proton motive force (PMF), subsequently leading to reduced energy production efficiency (ATP). Moreover, ROS is known to damage cells and negatively impact their metabolism. A study on *Rhodobacter sphaeroides*, a closely related organism, has shown that, in the absence of oxygen and aerobic respiratory chain, a higher growth rate was associated with lower ROS generation in autotrophically growing cells [60]. Consequently, higher ROS levels could negatively affect cell metabolism and indicate a less efficient respiratory chain, resulting in diminished PMF, both contributing to the lower μ_{\max} .

The enrichment of the high-level BP terms cell cycle (GO:0007049)

and DNA recombination (GO:0006310) could be the result of the different assimilation pathways between acetate and fructose explaining different measured specific growth rates ($\mu_{\max_{AY}} = 0.094 \text{ h}^{-1}$ and $\mu_{\max_{FY}} = 0.057 \text{ h}^{-1}$). This observation suggests a different metabolic reorganization around DNA transcription and RNA translation for acetate and fructose cultures. Such profound difference seems to also be highlighted under the term GO:0006950 response to stress. Various stress responses mechanisms activations were identified under both conditions regarding DNA repair functions, transcription repair, recombination mediator, elements related to DNA methylation and chaperone synthesis. Nevertheless, it is worth noting that the carbon source was most probably not the only parameter that changed between both cultures. For example, cultivation using acetate has been shown to result in the basification of the medium throughout the culture (data not shown). As mentioned previously different levels of dissolved oxygen and carbon dioxide should also be expected. These additional changes could nuance the role of fructose and acetate metabolization in these growth rate and transcriptomics observations.

Another important difference between acetate and fructose is that acetate is an excellent substrate for growth associated PHB biosynthesis, which is not the case for fructose. Although the GO analysis did not directly reflect any changes in PHB cycle activity, PHB granules were observed only under acetate conditions (see Fig. 3). Therefore, these observations led us to orient our interpretation of the GO terms also towards changes induced by the presence of PHB. For instance, the changes in DNA recombination could have been the result of the presence of PHB granules associated with growth on acetate. A multifunctional protein (*PhaM*) has been discovered in *C. necator* [61,62] whose role is to link PHB granules to the nucleoid (via *PhaC*), among other functions. It was further proposed that PHB granules are associated with the DNA and are segregated with the nucleoid during cell division [63]. However, no *phaM* homologs were so far found in *R. rubrum*.

4.3. Consequences of the deletion of the PHB biosynthesis on fructose metabolism

The comparison of the cultivation and transcriptomics data analysis between the polymerases mutant and wild-type strains grown on fructose led to unexpected results. It started with a significant decrease in the maximum specific growth rate for the mutant strain (see Table 2). This finding is intriguing given the observation that fructose-grown cells did not present PHB granules. Thus, cells' growth was not expected to be affected by knocking out the PHA synthases. This suggests an unanticipated activity of the PHB metabolism during aerobic growth with fructose. Indeed, *R. rubrum* has been described to produce PHB from fructose mainly in case of nutrient (nitrogen) or oxygen limitation [64, 65], suggesting that activation of the PHB biosynthesis from fructose necessitates these specific conditions.

The GO enrichment analysis (see Supplementary File 3) indicates that vitamin biosynthesis (GO:0009110) was influenced by the presence of both *phaC1* and *phaC2*. Vitamins are important cofactors in numerous enzymatic reactions and mutant cells could react to the disrupted PHB biosynthesis by adjusting its pool of vitamins to favor alternative metabolic pathways. Transcription of genes associated with vitamins biosynthesis (see Table 4) suggested that the pool of thiamine, pyridoxine, and cobalamin could be modified. Additionally, the terms GO:0009236 cobalamin biosynthetic process and GO:0033014 tetrapyrrole biosynthetic process, which are closely related, have also been enriched upon deletion of the PHA synthases. Overall, it seems that a majority of cobalamin biosynthesis genes presented higher expression for the KO strain, suggesting that the cell could potentially try to favor cobalamin production (Supplementary File 3). Cobalamin (vitamin B12) participates in various important reactions (e.g., cofactor of reductase, acetyltransferase, and isomerase) and can influence transcriptional regulation [66]. Interestingly, the cobalamin metabolism was also modified by the CO adaptation of *R. rubrum* [67]. Regarding

tetrapyrrole-related genes, it also seems that a majority of them were more expressed for the mutant strain compared to the WT strain (Supplementary File 3). Tetrapyrroles [68] are essential cofactors notably for light absorption, oxidative stress and electron transport. Therefore, the mutant cells may also redirect their metabolic effort towards tetrapyrrole production as a consequence of the altered PHB biosynthesis. Tetrapyrrole is an intermediate of the bacteriochlorophyll biosynthetic pathway and has been observed to be accumulated and excreted during high cell density culture of *R. rubrum* [69], growing on succinate and fructose. It can be speculated that the accumulation of these pigments may contribute to photosynthetic membrane repression [70]. Thus, these observations suggest that shifts in cobalamin and tetrapyrrole metabolism could hint at changes in the regulation of the photosynthetic apparatus [71]. Supporting this idea further, the literature also suggests that pigment synthesis and PHB production may be interconnected via the cell's energy and redox status. A recent study [72] indeed showed that the pigment synthesis regulatory protein HP1 is able to sense the intracellular redox state and adjust the pigment synthesis. However, the interpretation of these results remains challenging as *R. rubrum* was cultivated in the presence of oxygen and in the dark. It is known that in this organism, pigment synthesis is repressed by oxygen to avoid oxidative stress and therefore occurs at low or zero oxygen levels, which is not the case in our experiment. We thus hypothesize that an intracellular redox imbalance may have been generated by disrupting the PHB cycle, which leads to a redox state relatively similar to the one observed in the absence of oxygen, which consequently impacted pigment-related regulations.

Changes in oxidoreductase activity were also detected in the oxidation of CH-OH group (GO:0016614) and reduction of NAD(P)⁺ (GO:0016616) suggesting again that a perturbation of the PHB biosynthesis could influence redox-related cellular mechanisms. When looking closely into the transcription of genes representing these GO terms (Table 4 and Supplementary File 3), some of them were downregulated for the KO strain (encoding for IMP dehydrogenase (KUL73_01290), NADP-dependent isocitrate dehydrogenase (KUL73_01875), malate dehydrogenase (KUL73_06315), and 3-hydroxybutyryl-CoA dehydrogenase (KUL73_15880)). This interpretation is also confirmed by the enriched term NAD binding (GO:0051287). In particular, the wild-type strain presented an upregulation of the NADP-dependent isocitrate dehydrogenase gene (KUL73_01875) during the exponential growth phase, compared to the mutant strain (Table 4). This enzyme is responsible for the production of NADPH in the tricarboxylic acid (TCA) cycle and is responsible for the production of NADPH and is also important with respect to oxidative stress response [73]. Consequently, changes in the regulation of KUL73_01875 upon deletion of the PHA synthase could indicate a different intracellular redox state between the WT and KO strains. This difference could also be one of the reasons why different maximum specific growth rates were observed, as the lower growth rate of the mutant strain could be the result of an altered activation of the TCA cycle. In addition, the KO strain presented a higher transcription of *nuoF* (KUL73_08085) (Table 4), responsible for NADH-quinone oxidoreductase subunit NuoF, which was previously upregulated for acetate in the comparison between both carbon sources for the WT strain (Table 3 and Supplementary File 1). The relationship between the PHB cycle and redox metabolism could originate from a modification of the activity of acetoacetyl-CoA reductase (PhaB), which uses a reduced cofactor to produce PHA precursors. The control mechanism of the PhaB activity would need to be elucidated to establish the link between the absence of PHA synthases and the overall redox state of the cell.

Moreover, the enrichment analysis also revealed an unexpected impact of both PHA synthase deletions on the flagellum-mediated cell motility (GO:0071973 and GO:0097588). Further analysis of the genes involved (see Table 4 and Supplementary File 3) suggested that the assembly or composition of flagellar components is modified, impacting the cell motility. Interestingly, the relationship between the presence of flagellum and PHB accumulation has been studied in the past. For

instance, in the nutrient-limited accumulation of PHB in *C. necator*, cells appeared to be flagellated during the exponential growth phase and the flagellation became stagnant during the PHB accumulation phase [74]. This is followed by the complete loss of flagella during the subsequent PHB mobilization (after addition of a nitrogen source). The PHB⁻ mutant strain [75] unable to produce PHB also had a complete absence of flagellation under all conditions [74]. Additionally, other studies showed that the deletion of genes involved in flagellum formation resulted in enhanced PHA production in natural and unnatural PHA producers [76,77]. Our observation in *R. rubrum* is therefore coherent with the literature on other strains and demonstrated that the link between flagellum-mediated cell motility and PHB formation deserves further exploration.

4.4. Consequences of the deletion of the PHB biosynthesis on acetate metabolism

The strong growth inhibition observed in KO strain (see Table 2) is expected to be related to the enrichment of high-level GO terms regulation of biological process (GO:0050789) and regulation of cellular metabolic process (GO:0031323). Our findings also suggested that the growth defect correlates with changes in ATP hydrolysis activity (GO:0016887). Investigating the genes representing this term (see Table 5 and Supplementary File 4) suggested variations in expression of ATPase elements between strains (magnesium chelatase ATPase subunit I, AFG1 family ATPase, arsenical pump-driving ATPase, tRNA (adenosine(37)-N6)-threonylcarbamoyltransferase complex ATPase subunit type 1 TsaE, and AAA family ATPase). Genes involved in the dysfunction of DNA repair (transcription-repair coupling factor, excinuclease ABC subunit UvrA, and DNA mismatch repair endonuclease MutL) were also differently transcribed for the KO strain. In addition, genes involved in chaperones (molecular chaperone HtpG, chaperonin GroEL, ATP-dependent chaperone ClpB, and molecular chaperone DnaK) and proteases (endopeptidase La and ATP-dependent protease ATPase subunit HslU) were upregulated in the KO strain, indicative of stress response elements. These genes were also present in the GO:0016817 hydrolase activity, acting on acid anhydride. These observations are expected to be more related to an impaired substrate assimilation and carbon metabolism in the mutant strain, rather than to the PHB metabolism per se. Indeed, the similar *R. rubrum* Δ phaC1 Δ phaC2 strain has been reported to poorly assimilate acetate under aerobic conditions [52], compared with the wild-type strain. It was also observed elsewhere that the presence of intracellular PHB was associated with higher growth rates [12,16]. Thus, a perturbation of the EMC pathway and possibly also of the methylbutanoyl-CoA pathway (MBC) [50] acting under acetate assimilation as an anaplerotic pathway, could impact negatively the TCA cycle when its biosynthetic precursors are depleted. This diminished TCA cycle activity would lead to a weaker production of reduced electron carriers and a lower energy generation despite harbouring a functional aerobic respiratory chain. This could subsequently lead to the observed defect in ATP hydrolysis. It is also worth noting that the EMC pathway shares the 2 first steps with the PHB biosynthesis.

Based on the absence of expression differences in the redox regulation related terms, the redox metabolism may not be the reason for the growth defect, as it was previously suggested in *Rhizobium etli* [78] and *Azotobacter beijerinckii* [79] *phaC* mutants presenting NADH build-up. This supports the hypothesis that an impaired carbon flow is central to explain the differences in acetate assimilation between PHB⁻ mutants and wild-type *R. rubrum*. Indeed, several studies on PHB⁻ mutant organisms [80–83] observed a disrupted carbon flow under conditions conducive to PHB accumulation. This illustrates the role of the PHB cycle in the central metabolism and provides a framework for understanding how the *phaC* deletion affects acetate assimilation.

4.5. Small RNAs

Post-transcriptional bacterial regulation by small RNAs has been described as the fastest response to external stimuli under specific conditions related to the availability of regulatory elements at a given time [84]. Besides that, Reyer et al. [85] showed by modelling that sRNAs may be able to also act co-transcriptionally on nascent mRNA molecules and, thus, making the regulation by sRNAs even more efficient. Their regulatory role is primarily associated with the bacterial response to stress conditions [86] and thus making sRNAs the subject of studies in order to understand the principles of stress responses in organisms. Direct inference of sRNA from batch RNA-Seq is very sensitive to the choice of main thresholds used for the peak detection from the coverage signal. We left *low_cut_off* threshold on the default value as it seems reasonable value to both avoid low sequencing noise while not excluding true positives with lower transcription. *High_cut_off* threshold for the baerhunter's prediction was inferred from samples depth distributions as this threshold affects the number of predicted features and thus should be experiment specific. Therefore, it seems to be ideal using normalized sample depth as in our study. The length of sRNA can vary but it typically spans within the interval 40–500 bp [87]. Our predictions exceed both boundaries, which is expected. While the shortest length is given by the threshold *min_sRNA_length* of predicted elements, the longest length does not explicitly refer to the length of predicted element but rather it is situated in that region as the exact prediction from standard RNA-Seq can be misleading. The high number (>99 %) of regulated sRNAs (see Table 6) correlates with the regulatory role of these elements, however from the principle of baerhunter detection and the fact that differentially expressed elements do not explicitly imply that there cannot be the noise in the data, further analysis and experiments need to be done to get more precise information about the sRNA content. On the other hand, a larger number of non-coding regulatory elements in *R. rubrum* genome is expected as versatile bacteria were shown to possess more regulatory elements [88]. As our results demonstrated, intergenic regions in *R. rubrum* genome hide probably regulatory potential that remains to be further studied and understood.

5. Conclusions

In summary, our study focused on the genomic, transcriptomic, and metabolic aspects of *R. rubrum* DSM 467^T and its Δ *phaC1* Δ *phaC2* knock-out strain unable of PHA accumulation, shedding light on their responses to different carbon sources, namely acetate and fructose. Comparative analysis between cultures grown on these substrates unveiled distinct pathways for substrate assimilation. The growth on fructose resulted in an upregulation of carbohydrate metabolism due to the most likely activity of the Embden-Meyerhof-Parnas pathway. Indeed, acetate assimilation is believed to utilize the ethylmalonyl-CoA and methylbutanoyl-CoA pathways. These differences extended to electron transport chain components, impacting energy production efficiency and growth rates. Generally, the absence of PHA biosynthetic capability in KO strain affected fructose metabolism unexpectedly, indicating a broader impact on vitamin biosynthesis and tetrapyrrole metabolism. Changes in redox-related mechanisms and flagellum-mediated cell motility emphasized the intricate connections within cellular processes. Similarly, the PHB synthesis deletion's consequences on acetate metabolism elucidated growth inhibition mechanisms. It was suggested that impaired substrate assimilation resulted in reduced energy generation and, as a consequence, activated systems for cellular repair and stress response. Finally, the study delved into metabolism and stress response regulation by small RNA. A comprehensive analysis of differentially expressed sRNAs underscored their regulatory role emphasizing the need for further experiments to refine our understanding.

CRedit authorship contribution statement

Véronique Amstutz: Writing – original draft, Investigation. **Kristyna Hermankova:** Writing – original draft, Visualization, Formal analysis, Data curation. **Matej Bezdicek:** Investigation. **Katerina Mrázová:** Investigation. **Karel Sedlar:** Writing – original draft, Supervision, Resources, Data curation, Conceptualization. **Katerina Jureckova:** Writing – original draft, Visualization, Formal analysis, Data curation. **Marketa Nykrynova:** Writing – original draft, Visualization, Formal analysis, Data curation. **Eva Slaninova:** Writing – original draft, Investigation. **Hugo Fleuriot-Blitman:** Writing – original draft, Investigation, Formal analysis. **Stanislav Obruca:** Writing – original draft, Supervision, Resources, Funding acquisition, Conceptualization. **Kamila Hrubanova:** Investigation. **Manfred Zinn:** Writing – original draft, Supervision, Resources, Funding acquisition, Conceptualization.

Declaration of Competing Interest

None.

Data Availability

The genome assembly referred to in this paper is CP077803.1 for chromosome and CP077804.1 for plasmid. All sequencing data have been deposited in the NCBI Sequence Read Archive (SRA) under the project accession number PRJNA742260. Particular SRA accession numbers for genome sequencing are SRR28268536 (WT strain ONT), SRR28268535 (WT strain Illumina), and SRR28268534 (KO strain Illumina). Particular SRA accession numbers for RNA-Seq samples are listed in Supplementary Table S1.

Acknowledgments

This study was supported by the bilateral grant project GACR 21–15958L / SNSF 205321L_197275/1. The authors would like to thank Professor Kevin E. O'Connor and Professor Tanja Narancic, University College Dublin, Ireland for kindly providing the Δ *phaC1* Δ *phaC2* *R. rubrum* mutant strain and Jana Hanslikova for help with sequencing this knock-out strain. The authors thank Delong Instruments a.s., Czech Republic, namely Jaromir Bacovsky for providing the low voltage transmission electron microscopy analysis.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.csbj.2024.06.023.

References

- [1] Simon C, Daniel R. Metagenomic analyses: past and future trends. *Appl Environ Microbiol* 2011;77:1153–61. <https://doi.org/10.1128/AEM.02345-10>.
- [2] Gest H. A serendipic legacy: Erwin Esmarch's isolation of the first photosynthetic bacterium in pure culture. *Photosynth Res* 1995;46:473–8. <https://doi.org/10.1007/BF00032302>.
- [3] Revelles O, Tarazona N, García JL, Prieto MA. Carbon roadmap from syngas to polyhydroxyalkanoates in *Rhodospirillum rubrum*. *Environ Microbiol* 2016;18:708–20. <https://doi.org/10.1111/1462-2920.13087>.
- [4] Karmann S, Panke S, Zinn M. Fed-batch cultivations of *Rhodospirillum rubrum* under multiple nutrient-limited growth conditions on syngas as a novel option to produce poly(3-Hydroxybutyrate) (PHB). *Front Bioeng Biotechnol* 2019;7. <https://doi.org/10.3389/fbioe.2019.00059>.
- [5] Najafpour G, Younesi H. Hydrogen production from synthesis gas using the photosynthetic bacterium *Rhodospirillum rubrum*. *ASEAN J Chem Eng* 2005;5:35. <https://doi.org/10.22146/ajche.50162>.
- [6] Brandl H, Knee EJ, Fuller RC, Gross RA, Lenz RW. Ability of the phototrophic bacterium *Rhodospirillum rubrum* to produce various poly(β -hydroxyalkanoates): Potential sources for biodegradable polyesters. *Int J Biol Macromol* 1989;11:49–55. [https://doi.org/10.1016/0141-8130\(89\)90040-8](https://doi.org/10.1016/0141-8130(89)90040-8).
- [7] Kaftan D, Bina D, Koblížek M. Temperature dependence of photosynthetic reaction centre activity in *Rhodospirillum rubrum*. *Photosynth Res* 2019;142:181–93. <https://doi.org/10.1007/s11220-019-00652-7>.

- [8] Najafpour GD, Younesi H. Bioconversion of synthesis gas to hydrogen using a light-dependent photosynthetic bacterium, *Rhodospirillum rubrum*. *World J Microbiol Biotechnol* 2007;23:275–84. <https://doi.org/10.1007/s11274-006-9225-2>.
- [9] Ghosh R, Hardmeyer A, Thoenen I, Bachofen R. Optimization of the Sistrofem culture medium for large-scale batch cultivation of *Rhodospirillum rubrum* under semiaerobic conditions with maximal yield of photosynthetic membranes. *Appl Environ Microbiol* 1994;60:1698–700. <https://doi.org/10.1128/aem.60.5.1698-1700.1994>.
- [10] Munk AC, Copeland A, Lucas S, Lapidus A, Del Rio TG, Barry K, et al. Complete genome sequence of *Rhodospirillum rubrum* type strain (S1¹). *Stand Genom Sci* 2011;4:293–302. <https://doi.org/10.4056/sigs.1804360>.
- [11] Lehman LJ, Roberts GP. Identification of an alternative nitrogenase system in *Rhodospirillum rubrum*. *J Bacteriol* 1991;173:5705–11. <https://doi.org/10.1128/jb.173.18.5705-5711.1991>.
- [12] Jin H, Nikolau BJ. Role of genetic redundancy in polyhydroxyalkanoate (PHA) polymerases in PHA biosynthesis in *Rhodospirillum rubrum*. *J Bacteriol* 2012;194:5522–9. <https://doi.org/10.1128/JB.01111-12>.
- [13] Jin H, Nikolau BJ. Evaluating PHA productivity of bioengineered *Rhodospirillum rubrum*. *PLoS One* 2014;9:e96621. <https://doi.org/10.1371/journal.pone.0096621>.
- [14] Mastroleo F, Van Houdt R, Leroy B, Benotmane MA, Janssen A, Mergeay M, et al. Experimental design and environmental parameters affect *Rhodospirillum rubrum* S1H response to space flight. *ISME J* 2009;3:1402–19. <https://doi.org/10.1038/ismej.2009.74>.
- [15] Mastroleo F, Van Houdt R, Atkinson S, Mergeay M, Hendrickx L, Wattiez R, et al. Modelled microgravity cultivation modulates N-acetylhomoserine lactone production in *Rhodospirillum rubrum* S1H independently of cell density. *Microbiol* 2013;159:2456–66. <https://doi.org/10.1099/mic.0.066415-0>.
- [16] Klask C. Heterologous expression of various PHA synthase genes in *Rhodospirillum rubrum*. *Chem Biochem Eng Q* 2015;29:75–85. <https://doi.org/10.15255/CABEQ.2014.2249>.
- [17] Kouřilová X, Schwarzerová J, Pernicová I, Sedlář K, Mrázová K, Krzyžánek V, et al. The first insight into polyhydroxyalkanoates accumulation in multi-extremophilic *Rubrobacter xylanophilus* and *Rubrobacter spartanus*. *Microorganisms* 2021;9:909. <https://doi.org/10.3390/microorganisms9050909>.
- [18] Mrazova K, Bacovsky J, Sedrlova Z, Slaninova E, Obruca S, Fritz I, et al. Uranyl-less low voltage transmission electron microscopy: a powerful tool for ultrastructural studying of cyanobacterial cells. *Microorganisms* 2023;11:888. <https://doi.org/10.3390/microorganisms11040888>.
- [19] Leger A, Leonardi T. pycoQC, interactive quality control for Oxford Nanopore Sequencing. *J Open Source Softw* 2019;4:1236. <https://doi.org/10.21105/joss.01236>.
- [20] Kolmogorov M, Yuan J, Lin Y, Pevzner PA. Assembly of long, error-prone reads using repeat graphs. *Nat Biotechnol* 2019;37:540–6. <https://doi.org/10.1038/s41587-019-0072-8>.
- [21] Li H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 2018;34:3094–100. <https://doi.org/10.1093/bioinformatics/bty191>.
- [22] Vaser R, Sović I, Nagarajan N, Sikić M. Fast and accurate *de novo* genome assembly from long uncorrected reads. *Genome Res* 2017;27:737–46. <https://doi.org/10.1101/gr.214270.116>.
- [23] Ewels P, Magnusson M, Lundin S, Källner M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* 2016;32:3047–8. <https://doi.org/10.1093/bioinformatics/btw354>.
- [24] Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014;30:2114–20. <https://doi.org/10.1093/bioinformatics/btu170>.
- [25] Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 2009;25:1754–60. <https://doi.org/10.1093/bioinformatics/btp324>.
- [26] Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence Alignment/Map format and SAMtools. *Bioinformatics* 2009;25:2078–9. <https://doi.org/10.1093/bioinformatics/btp352>.
- [27] Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, et al. Pilon: An integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 2014;9. <https://doi.org/10.1371/journal.pone.0112963>.
- [28] Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, et al. NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res* 2016;44:6614–24. <https://doi.org/10.1093/nar/gkw569>.
- [29] Huerta-Cepas J, Szklarczyk D, Heller D, Hernández-Plaza A, Forslund SK, Cook H, et al. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res* 2019;47:D309–14. <https://doi.org/10.1093/nar/gky1085>.
- [30] Carver T, Thomson N, Bleasby A, Berriman M, Parkhill J. DNAPlotter: circular and linear interactive genome visualization. *Bioinformatics* 2009;25:119–20. <https://doi.org/10.1093/bioinformatics/btn578>.
- [31] Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream M-A, et al. Artemis: sequence visualization and annotation. *Bioinformatics* 2000;16:944–5. <https://doi.org/10.1093/bioinformatics/16.10.944>.
- [32] Biswas A, Staals RHJ, Morales SE, Fineran PC, Brown CM. CRISPRDetect: a flexible algorithm to define CRISPR arrays. *BMC Genom* 2016;17:356. <https://doi.org/10.1186/s12864-016-2627-0>.
- [33] Akhter S, Aziz RK, Edwards RA. PhiSpy: a novel algorithm for finding prophages in bacterial genomes that combines similarity- and composition-based strategies. *Nucleic Acids Res* 2012;40. <https://doi.org/10.1093/nar/gks406>.
- [34] Roberts RJ, Vincze T, Posfai J, Macelis D. REBASE: a database for DNA restriction and modification: enzymes, genes and genomes. *Nucleic Acids Res* 2023;51:D629–30. <https://doi.org/10.1093/nar/gkac975>.
- [35] Sayers EW,avanaugh M, Clark K, Ostell J, Pruitt KD, Karsch-Mizrachi I. GenBank. *Nucleic Acids Res* 2019. <https://doi.org/10.1093/nar/gkz956>.
- [36] Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, et al. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* 2015;31:3691–3. <https://doi.org/10.1093/bioinformatics/btv421>.
- [37] Kopylova E, Noé L, Touzet H. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* 2012;28:3211–7. <https://doi.org/10.1093/bioinformatics/bts611>.
- [38] Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 2012;41:D590–6. <https://doi.org/10.1093/nar/gks1219>.
- [39] Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 2013;29:15–21. <https://doi.org/10.1093/bioinformatics/bts635>.
- [40] Smith T, Heger A, Sudbery I. UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. *Genome Res* 2017;27:491–9. <https://doi.org/10.1101/gr.209601.116>.
- [41] Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 2014;30:923–30. <https://doi.org/10.1093/bioinformatics/btt656>.
- [42] Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014;15:550. <https://doi.org/10.1186/s13059-014-0550-8>.
- [43] van der Maaten L. Accelerating t-SNE using Tree-Based Algorithms. *J Mach Learn Res* 2014;15:3221–45.
- [44] Wickham H. ggplot2. New York, NY: Springer New York; 2009. <https://doi.org/10.1007/978-0-387-98141-3>.
- [45] Alexa A., Rahnenfuehrer J. topGO: Enrichment Analysis for Gene Ontology 2023. doi:10.18129/B9.bioc.topGO.
- [46] Ozuna A, Libertò D, Joyce RM, Arnvig KB, Nobeli I. baerhunter: an R package for the discovery and analysis of expressed non-coding regions in bacterial RNA-seq data. *Bioinformatics* 2020;36:966–9. <https://doi.org/10.1093/bioinformatics/btz643>.
- [47] Lawrence M, Huber W, Pagès H, Aboyoun P, Carlson M, Gentleman R, et al. Software for computing and annotating genomic ranges. *PLoS Comput Biol* 2013;9:e1003118. <https://doi.org/10.1371/journal.pcbi.1003118>.
- [48] Li X-Q, Du D. Variation, evolution, and correlation analysis of C+G content and genome or chromosome size in different kingdoms and phyla. *PLoS One* 2014;9:e88339. <https://doi.org/10.1371/journal.pone.0088339>.
- [49] Grammel H, Gilles E-D, Ghosh R. Microaerophilic cooperation of reductive and oxidative pathways allows maximal photosynthetic membrane biosynthesis in *Rhodospirillum rubrum*. *Appl Environ Microbiol* 2003;69:6577–86. <https://doi.org/10.1128/AEM.69.11.6577-6586.2003>.
- [50] De Meur Q, Deutschbauer A, Koch M, Bayon-Vicente G, Cabezas Segura P, Wattiez R, et al. New perspectives on butyrate assimilation in *Rhodospirillum rubrum* S1H under photoheterotrophic conditions. *BMC Microbiol* 2020;20:126. <https://doi.org/10.1186/s12866-020-01814-7>.
- [51] Leroy B, De Meur Q, Moulin C, Wegria G, Wattiez R. New insight into the photoheterotrophic growth of the isocitrate lyase-lacking purple bacterium *Rhodospirillum rubrum* on acetate. *Microbiol* 2015;161:1061–72. <https://doi.org/10.1099/mic.0.000067>.
- [52] Narancic T, Scollica E, Kenny ST, Gibbons H, Carr E, Brennan L, et al. Understanding the physiological roles of polyhydroxybutyrate (PHB) in *Rhodospirillum rubrum* S1 under aerobic chemoheterotrophic conditions. *Appl Microbiol Biotechnol* 2016;100:8901–12. <https://doi.org/10.1007/s00253-016-7711-5>.
- [53] Rudolf C, Grammel H. Fructose metabolism of the purple non-sulfur bacterium *Rhodospirillum rubrum*: effect of carbon dioxide on growth, and production of bacteriochlorophyll and organic acids. *Enzym Micro Technol* 2012;50:238–46. <https://doi.org/10.1016/j.enzmictec.2012.01.007>.
- [54] Erb TJ, Fuchs G, Alber BE. (2S)-Methylsuccinyl-CoA dehydrogenase closes the ethylmalonyl-CoA pathway for acetyl-CoA assimilation. *Mol Microbiol* 2009;73:992–1008. <https://doi.org/10.1111/j.1365-2958.2009.06837.x>.
- [55] Erb TJ, Berg IA, Brecht V, Müller M, Fuchs G, Alber BE. Synthesis of C₅-dicarboxylic acids from C₂-units involving crotonyl-CoA carboxylase/reductase: The ethylmalonyl-CoA pathway. *Proc Natl Acad Sci* 2007;104:10631–6. <https://doi.org/10.1073/pnas.0702791104>.
- [56] Keci S, Pawlik G, Lohmeyer E, Koch H-G, Daldal F. Biogenesis of *cbb3*-type cytochrome c oxidase in *Rhodobacter capsulatus*. *Biochim Et Biophys Acta (BBA) - Bioenerg* 2012;1817:898–910. <https://doi.org/10.1016/j.bbabi.2011.10.011>.
- [57] Jain R, Shapleigh JP. Characterization of *nirV* and a gene encoding a novel pseudoazurin in *Rhodobacter sphaeroides* 2.4.3. *Microbiol* 2001;147:2505–15. <https://doi.org/10.1099/00221287-147-9-2505>.
- [58] Larosa V, Remacle C. Insights into the respiratory chain and oxidative stress. *Biosci Rep* 2018;38. <https://doi.org/10.1042/BSR20171492>.
- [59] Korshunov S, Imlay JA. Two sources of endogenous hydrogen peroxide in *Escherichia coli*. *Mol Microbiol* 2010;75:1389–401. <https://doi.org/10.1111/j.1365-2958.2010.07059.x>.
- [60] Lee YR, Lee W-H, Lee SY, Lee J, Kim M-S, Moon M, et al. Regulation of reactive oxygen species promotes growth and carotenoid production under autotrophic conditions in *Rhodobacter sphaeroides*. *Front Microbiol* 2022;13. <https://doi.org/10.3389/fmicb.2022.847757>.

- [61] Wahl A, Schuth N, Pfeiffer D, Nussberger S, Jendrossek D. PHB granules are attached to the nucleoid via PhaM in *Ralstonia eutropha*. *BMC Microbiol* 2012;12:262. <https://doi.org/10.1186/1471-2180-12-262>.
- [62] Pfeiffer D, Wahl A, Jendrossek D. Identification of a multifunctional protein, PhaM, that determines number, surface to volume ratio, subcellular localization and distribution to daughter cells of poly(3-hydroxybutyrate), PHB, granules in *Ralstonia eutropha* H16. *Mol Microbiol* 2011;82:936–51. <https://doi.org/10.1111/j.1365-2958.2011.07869.x>.
- [63] Bresan S, Jendrossek D. New insights into PhaM-PhaC-mediated localization of polyhydroxybutyrate granules in *Ralstonia eutropha* H16. *Appl Environ Microbiol* 2017;83. <https://doi.org/10.1128/AEM.00505-17>.
- [64] Liu J, Zhao Y, Diao M, Wang W, Hua W, Wu S, et al. Poly(3-hydroxybutyrate-co-3-hydroxyvalerate) production by *Rhodospirillum rubrum* using a two-step culture strategy. *J Chem* 2019;2019:1–8. <https://doi.org/10.1155/2019/8369179>.
- [65] Godoy MS, de Miguel SR, Prieto MA. Aerobic-anaerobic transition boosts poly(3-hydroxybutyrate-co-3-hydroxyvalerate) synthesis in *Rhodospirillum rubrum*: the key role of carbon dioxide. *Micro Cell Fact* 2023;22:47. <https://doi.org/10.1186/s12934-023-02045-x>.
- [66] Cheng Z, Li K, Hammad LA, Karty JA, Bauer CE. Vitamin B₁₂ regulates photosystem gene expression via the CrtJ antirepressor AerR in *Rhodobacter capsulatus*. *Mol Microbiol* 2014;91:649–64. <https://doi.org/10.1111/mmi.12491>.
- [67] Cavazza C, Collin-Faure V, Pérard J, Diemer H, Cianféran S, Rabilloud T, et al. Proteomic analysis of *Rhodospirillum rubrum* after carbon monoxide exposure reveals an important effect on metallic cofactor biosynthesis. *J Proteom* 2022;250:104389. <https://doi.org/10.1016/j.jprot.2021.104389>.
- [68] Busch AWU, Montgomery BL. Interdependence of tetrapyrrole metabolism, the generation of oxidative stress and the mitigative oxidative stress response. *Redox Biol* 2015;4:260–71. <https://doi.org/10.1016/j.redox.2015.01.010>.
- [69] Zeiger L, Grammel H. Model-based high cell density cultivation of *Rhodospirillum rubrum* under respiratory dark conditions. *Biotechnol Bioeng* 2010;105:729–39. <https://doi.org/10.1002/bit.22589>.
- [70] Carius L, Carius AB, McIntosh M, Grammel H. Quorum sensing influences growth and photosynthetic membrane production in high-cell-density cultivations of *Rhodospirillum rubrum*. *BMC Microbiol* 2013;13:189. <https://doi.org/10.1186/1471-2180-13-189>.
- [71] Ghosh R, Roth E, Abou-Aisha K, Saegesser R, Autenrieth C. The monofunctional cobalamin biosynthesis enzyme precorrin-3B synthase (CobZ_{RR}) is essential for anaerobic photosynthesis in *Rhodospirillum rubrum* but not for aerobic dark metabolism. *Microbiol* 2018;164:1416–31. <https://doi.org/10.1099/mic.0.000718>.
- [72] Godoy MS, de Miguel SR, Prieto MA. A singular PpaA/AerR-like protein in *Rhodospirillum rubrum* rules beyond the boundaries of photosynthesis in response to the intracellular redox state. *MSystems* 2023;8. <https://doi.org/10.1128/msystems.00702-23>.
- [73] Obruca S, Snajdar O, Svoboda Z, Marova I. Application of random mutagenesis to enhance the production of polyhydroxyalkanoates by *Cupriavidus necator* H16 on waste frying oil. *World J Microbiol Biotechnol* 2013;29:2417–28. <https://doi.org/10.1007/s11274-013-1410-5>.
- [74] Raberg M, Reinecke F, Reichelt R, Malkus U, König S, Pötter M, et al. *Ralstonia eutropha* H16 flagellation changes according to nutrient supply and state of poly(3-hydroxybutyrate) accumulation. *Appl Environ Microbiol* 2008;74:4477–90. <https://doi.org/10.1128/AEM.00440-08>.
- [75] Schlegel HG, Lafferty R, Krauss I. The isolation of mutants not accumulating poly-β-hydroxybutyric acid. *Arch Für Mikrobiol* 1970;71:283–94. <https://doi.org/10.1007/BF00410161>.
- [76] Wang J, Ma W, Wang Y, Lin L, Wang T, Wang Y, et al. Deletion of 76 genes relevant to flagella and pili formation to facilitate polyhydroxyalkanoate production in *Pseudomonas putida*. *Appl Microbiol Biotechnol* 2018;102:10523–39. <https://doi.org/10.1007/s00253-018-9439-x>.
- [77] Wang J, Ma W, Fang Y, Zhang H, Liang H, Liu H, et al. Engineering the outer membrane could facilitate better bacterial performance and effectively enhance poly-3-hydroxybutyrate accumulation. *Appl Environ Microbiol* 2021;87. <https://doi.org/10.1128/AEM.01389-21>.
- [78] Cevallos MA, Encarnación S, Leija A, Mora Y, Mora J. Genetic and physiological characterization of a *Rhizobium etli* mutant strain unable to synthesize poly-β-hydroxybutyrate. *J Bacteriol* 1996;178:1646–54. <https://doi.org/10.1128/jb.178.6.1646-1654.1996>.
- [79] Senior PJ, Dawes EA. The regulation of poly-β-hydroxybutyrate metabolism in *Azotobacter beijerinckii*. *Biochem J* 1973;134:225–38. <https://doi.org/10.1042/bj1340225>.
- [80] Raberg M, Voigt B, Hecker M, Steinbüchel A. A closer look on the polyhydroxybutyrate- (PHB-) negative phenotype of *Ralstonia eutropha* PHB⁴. *PLoS One* 2014;9:e95907. <https://doi.org/10.1371/journal.pone.0095907>.
- [81] Cook AM, Schlegel HG. Metabolite concentrations in *Alcaligenes eutrophus* H 16 and a mutant defective in poly-β-hydroxybutyrate synthesis. *Arch Microbiol* 1978;119:231–5. <https://doi.org/10.1007/BF00405400>.
- [82] Jung Y-M, Yong-Hyun L. Investigation of regulatory mechanism of flux of Acetyl-CoA in *Alcaligenes eutrophus* using PHB- negative Mutant and transformants harboring cloned *phbCAB* Genes. *J Microbiol Biotechnol* 1997;7:215–22.
- [83] Steinbüchel A, Schlegel HG. Excretion of pyruvate by mutants of *Alcaligenes eutrophus*, which are impaired in the accumulation of poly(β-hydroxybutyric acid) (PHB), under conditions permitting synthesis of PHB. *Appl Microbiol Biotechnol* 1989;31:168–75. <https://doi.org/10.1007/BF00262457>.
- [84] Shimoni Y, Friedlander G, Hetzroni G, Niv G, Altuvia S, Biham O, et al. Regulation of gene expression by small non-coding RNAs: a quantitative view. *Mol Syst Biol* 2007;3. <https://doi.org/10.1038/msb4100181>.
- [85] Reyer MA, Chennakesavalu S, Heideman EM, Ma X, Bujnowska M, Hong L, et al. Kinetic modeling reveals additional regulation at co-transcriptional level by post-transcriptional sRNA regulators. *Cell Rep* 2021;36:109764. <https://doi.org/10.1016/j.celrep.2021.109764>.
- [86] Gottesman S, McCullen CA, Guillier M, Vanderpool CK, Majdalani N, Benhammou J, et al. Small RNA regulators and the bacterial response to stress. *Cold Spring Harb Symp Quant Biol* 2006;71:1–11. <https://doi.org/10.1101/sqb.2006.71.016>.
- [87] Huang H-Y, Chang H-Y, Chou C-H, Tseng C-P, Ho S-Y, Yang C-D, et al. sRNAMap: genomic maps for small non-coding RNAs, their regulators and their targets in microbial genomes. *Nucleic Acids Res* 2009;37:D150–4. <https://doi.org/10.1093/nar/gkn852>.
- [88] Bervoets I, Charlier D. Diversity, versatility and complexity of bacterial gene regulation mechanisms: opportunities and drawbacks for applications in synthetic biology. *FEMS Microbiol Rev* 2019;43:304–39. <https://doi.org/10.1093/femsre/fuz001>.