Re-Sequencing *Rhodobacter sphaeroides* strain 2.4.1 to facilitate understanding varying hydrogen evolution rates

Nadine Wappler, nwappler@hs-mittweida.de | Eric Zuchantke | Röbbe Wünschiers

University of Applied Science Mittweida, Faculty MNI

Abstract

The aim of sequencing a genome is to identify so far unknown genes and their function. So called single nucleotide polymorphisms (SNP) are responsible for the diversity of organisms. By re-sequencing *Rhodobacter sphaeroides* and a subsequent SNP analysis we try to understand experimental variations of hydrogen production rates among strains from different laboratories. We focused our analysis on SNPs of chromosome 1 and 2 as well as the five plasmids.

1 Introduction

About 80 % of the worldwide energy demand is covered by non-renewable fossil fuels. This results in emission of pollutants and global warming. Hydrogen (H_2), produced by renewable energy sources that often show daily or seasonal fluctuations, is an ideal future energy carrier.[i] Hydrogen is clean and releases huge amounts of energy per weight when combusted or used in fuel cells. By biological fermentation, hydrogen can be produced with renewable substrates as starting material.[ii] Some photosynthetic bacteria like *Rhodobacter sphaeroides* are capable of producing hydrogen from organic waste by photofermentation. *R. sphaeroides* belongs to the class of proteobacteria[iii] and it is ubiquitous in soil[vi]. It has various pathways for energy production, e. g. aerobic and anaerobic respiration or anoxygenic photosynthesis.[v] The strain 2.4.1 has two chromosomes and five plasmids (A-E).[iii] The organism was sequenced for the first time in 2001 by the University of Texas.[iv]

2 Material and Methods

2.1 Bacterial growth, harvest

Rhodobacter sphaeroides (DSM-No 158) was grown as batch culture in standard medium as a pre-culture at 30 °C and later as main-culture at 27 °C, pH 7 in a 1 litre bioreactor. Vitamins and carbon sources were added to the media. Both cultures were illuminated by halogen lamps. The cells were harvested by centrifugation, resuspended in RNAlater (life technologies, Darmstadt) and frozen at –80 °C until the next processing step.

2.2 Isolation of total DNA

DNA was extracted and isolated using the MasterPure™ Complete DNA and RNA Purification Kit (epicentre, Hessisch Oldendorf). Samples were thawed and placed on ice. 1 µl Proteinase K and after that 300 µl Lysis Solution Tissue & Cell were added. The tubes were mixed until the pellet was dissolved and homogenised at 65 °C for 2 minutes and subsequently incubated for 1 minute at – 20 °C. 250 µl MPC Protein Precipitation Reagent were added and the mix centrifuged for 10 minutes at 14 100 rcf. The upper clear phase was added to a new tube with 500 µl ice-cold Isopropanol, inverted 40 times and again centrifuged for 15 minutes at 14 100 rcf. Subsequently, the supernatant was removed and pellet dried for 1 minute at 65 °C. 50 µl of TE buffer and 2 µl RNAsecure (life technologies, Darmstadt) were added and the samples were resuspended for 15 minutes at 65 °C.

2.3 Digestion of RNA and sequencing

The isolated DNA had to be exempted from RNA contamination. We used RNase CocktailTM Enzyme Mix (life technologies, Darmstadt) and incubated for 30 minutes at 37 °C. The sequencing was performed by Illumina MiSeq (1 x 150 nt) at Eurofins Genomics (Ebersberg).

2.4 Data analysis

For genome data analysis (starting with fastq files), we solely used open source software. After a first assessment of quality by FastQC[vii] we trimmed and clipped the sequence files by FastX-Toolkit[viii]. The indexing and mapping to the reference sequence from NCBI (NZ_AKVW01000001.1) was performed with Bowtie2[ix]. SAMTOOLS[x] converted the incurred SAM- to a BAM-file which was sorted and indexed afterwards. From the resulting Fasta-file a consensus sequence was cre-

ated by BLASTing, which delivered a textfile with SNPs. We compared the resulting sequences with the reference sequences from the NCBI Nucleotide database [http://www.ncbi.nlm.nih.gov/nuccore/484336764].

3 Results and discussion

For a first insight into changes of genome sequence we focused our analysis on SNPs of chromosome 1 and 2 as well as all five plasmids (A-E). We found a total of 222 SNPs. This SNPs caused 71 changes in genes (see table 1).

	SNP	gene	unchanged genes	SNP in genes
chromosome 1	128	6022	5974	48
chromosome 2	0	1708	1708	0
plasmid A	5	217	217	0
pasmid B	0	209	209	0
plasmid C	2	177	177	0
plasmid D	83	200	188	12
plasmid E	4	68	67	1
total	222	8601	8540	71

Table 1: Overview about identified SNPs

Due to their importance in metabolism, we searched for changes in proteins, specifically by base substitutions. Some enzymes showed an amino acid substitution caused by a SNP. We found intergenic SNPs, which can be neutral or have an influence on expression if they are located for example in a promotor region. These can influence transcriptomic regulation. Genic SNPs can be neutral too, or lead to the substitution of amino acids that may affect enzyme activity (see figure 1).

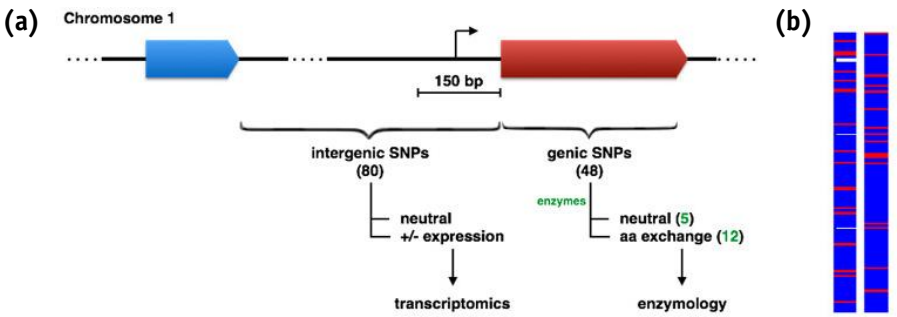

Figure 1: Schematic representation of effects caused by SNPs on chromosome 1 (a), which can have affects on transcriptomic regulation and enzymology. Such changes may lead to a new protein function. Found changes were mapped on chromosome 1 (b). White lines mark intergenic regions, blue lines regions with genes and red lines SNPs.

4 Conclusion

By re-sequencing the genome of *Rhodobacter sphaeroides* we found a total of 222 SNPs. Thereof, 71 SNPs caused changes in protein coding genes. We focused further analysis on enzymes which are only located on chromosome 1. Within some enzymes, the SNP caused an amino acid substitution.

5 Outlook

Our longterm aim is the optimisation of Rhodobacter fermentation for hydrogen production. A theoretical approach is shown in figure 2. By co-cultivating green algae and *R. sphaeroides*, hydrogen can be produced efficiently from carbon dioxide and light.

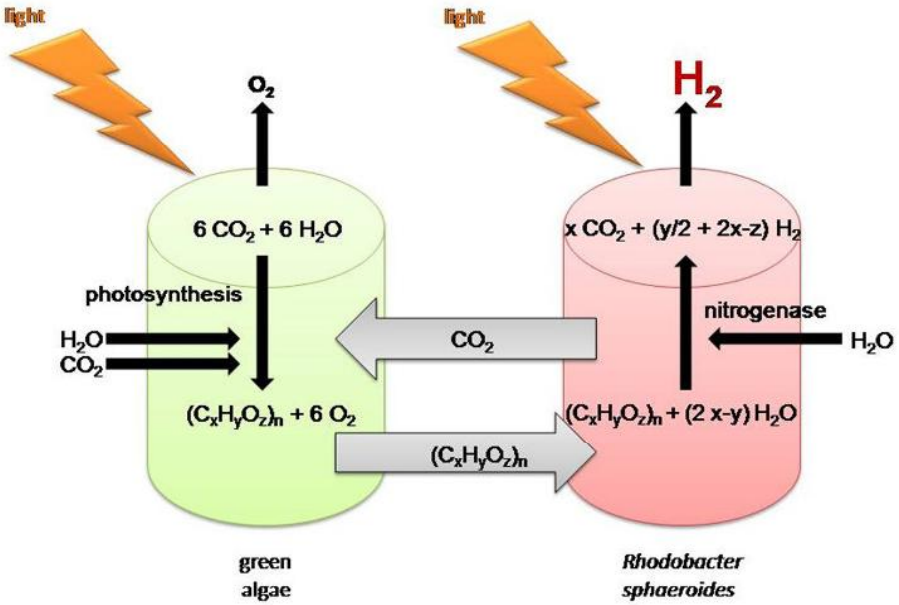

Figure 2: Schematic display of hydrogen production by *Rhodobacter sphaeroides* with the help of green algae. The bacteria use carbohydrate equivalents from the metabolism of the algae. Hydrogen is derived from bacterial fermentation and carbon dioxide is fed back to the algae.

Acknowledgement

This work is performed in cooperation with the group of Jost Weber from the Technical University Dresden/Germany. For financial support we thank the State Ministry of Higher Education, Research and the Arts as well as the European Social Fund.

[i] Franchi, Elisabetta; Tosi, Claudio; Scolia, Giuseppe; Penna, Gino Della.; Rodriguez, Francesco; Pedroni, Paola Maria (2004): Metabolically Engineered *Rhodobacter sphaeroides* RV strains for Improved Biohydrogen Photoproduction Combined with Disposal of Food Wastes. Mar. Biotechnol. 6, 552–565, DOI: 10.1007/s10126-004-1007-y.

[ii] Kim, Dong-Hoon Kim; Lee, Ji-Hye; Kang, Seoktae Kang; Hallenbeck, Patrick C.; Kim, Eui-Jin; Lee, Jeong K.; Kim, Mi-Sun Kim (2014): Enhanced photo-fermen-

tative H2 production using *Rhodobacter sphaeroides* by ethanol addition and analysis of soluble microbial products. Biotechnology for Biofuels, 7: 79.

[iii] Kontur, Wayne S.; Schackwitz, Wendy S.; Ivanova, Natalia; Martin, Joel; LaButti, Kurt; Deshpande, Shweta; Tice, Hope N.; Pennacchio, Christa; Sodergren, Erica; Weinstock, George M.; Noguera, Daniel R.; Donohue, Timothy J. (2012): Revised Sequence and Annotation of the *Rhodobacter sphaeroides* 2.4.1 Genome. J Bacteriol., 194(24): 7016–7017.

[iv] Mackenzie, Chris; Choudhary, Madhusudan ; Larimer, Frank W.; Predki, Paul F.; Stilwagen, Stephanie; Armitage, Judith P.; Barber, Robert D.; Donohue, Timothy J.; Hosler, Jonathan P.; Newman, Jack E.; Shapleigh, James P.; Sockett, R. Elizabeth; Kaplan, Jill Zeilstra-Ryalls & Samuel (2001): The home stretch, a first analysis of the nearly completed genome of *Rhodobacter sphaeroides* 2.4.1. Photosyn. Res. 70: 19–41.

[v] Metz, Sebastion (2010): Blaulichtabhängige Genregulation in *Rhodobacter sphaeroides*. Untersuchungen zur physiologischen Funktion der Blaulichtphotorezeptoren AppA, CryB und LOV. Inaugural – Dissertation zur Erlangung des Doktorgrades der Naturwissenschaften.

[vi] Porter, Steven L.; Wilkinson, David A.; Byles, Elaine D.; Wadhams, George H.; Taylor, Stephen; Saunders, Nigel J.; Armitage, Judith P. (2011): Genome Sequence of *Rhodobacter sphaeroides* Strain WS8N. J Bacteriol., 193(15): 4027–4028.

Used software and version

[vii] FastQC, Version 0.10.1

[viii] FastX-Toolkit, Version 0.0.13

[ix] Bowtie2, Version 2.1.0

[x] Samtools, Version 0.1.19-96b5f2294a