

PRELIMINARY GENOME ANALYSIS OF *PORPHYROBACTER* SP. NOV. ISOLATED FROM BRACKISH
SEDIMENT COLLECTED FROM THE TRUNK RIVER BANK, MA.

Kirsten Grond

Division of Biology, 116 Ackert Hall, Kansas State University, Manhattan, KS
Contact: kgrond@ksu.edu

INTRODUCTION

The Proteobacteria are one of the most diverse bacterial phyla, and incorporate a number of pathogens such as *Escherichia coli*, *Vibrio cholerae*, *Campylobacter jejuni* and *Helicobacter pylori* (Brenner et al. 2005). Proteobacteria are predominantly gram-negative and found in all environments, including fresh water, marine, terrestrial and even gastro-intestinal ecosystems (Brenner et al. 2005). The Proteobacteria include seven classes: the acidobacteria and the alpha-, beta-, gamma-, delta-, epsilon- and zeta-proteobacteria. The Alphaproteobacteria comprise a wide variety of organisms, including phototrophs, plant and animal (endo)symbionts and one of the most common and successful bacteria on earth, SAR11 (Giovannoni et al. 1990; Sayers et al. 2016).

Within the Alphaproteobacteria, the *Porphyrobacter* genus is part of the *Sphingomonadaceae* family, and includes gram-negative, rod-shaped bacteria. *Porphyrobacter* sp. have been isolated from a wide array of environments, including marine and freshwater sources, and even swimming pool and on one occasion a stadium seat (Hiraishi et al. 2002; Furuhashi et al. 2013; Li et al. 2013; Coil et al. 2015).

Despite the wide variety in habitats this organism occurs in, only 17 genomes are available in the National Center for Biotechnology Information (NCBI) database of *Porphyrobacter* species and strains obtained in six different research projects. *Porphyrobacter* sp. contain bacteriochlorophyll a, and use aerobic anoxygenic photosynthesis for energy generation. They are strict aerobic and contain a wide variety of carotenoids, resulting in bright pigmentation of different species ranging from yellow to deep red.

The goal of our study was to sequence and describe the genome of a bacterial isolate obtained from a sediment sample collected at the fresh-salt water interface in Trunk River, MA.

METHODS

Sample collection. – The novel species of *Porphyrobacter* was isolated from a sediment sample collected from under the bicycle path bridge at the mouth of Trunk River, MA (41°32'04.0"N, 70°38'30.3"W). Saline content was approximately 14ppm, indicating a brackish habitat. Trunk river is a small stream that connects the Atlantic Ocean to an eel-grass rich pond. We used the following protocol designed to isolate Actinomycetes.

One gram of air dried sample was incubated at 37°C for 24 hours, to remove all moisture. Subsequently, 1g of dried sample was mixed with 1g CaCO₃ in a sterile petridish with a moistened Whatman filter suspended above the sample, and incubated at 20C for 24 hrs. The soil-CaCO₃ mixture was mixed with 50mL 1x Fresh Water base (17.1mM NaCl, 6.71mM KCl, 1.97mM MgCl₂·6H₂O, 0.68mM CaCl₂·2H₂O) and 1:10 dilution series (1-10⁻⁵) was plated on Starch Histidine Phenylalanine (STAT) medium. STAT medium consisted of 1X Fresh Water base, 10mM MOPS (pH 7.2), 5.7mM arginine, 10 gL⁻¹ starch, 0.5mM tryptophan, 0.2 mL 1M Na₂S₂O₃·5H₂O, 1mL 100nM KH₂PO₄/K₂HPO₄ at 1:3.175 (pH 7.2) and 1.5% agar. EDTA trace elements, multivitamin mix and cyclohexamide were added after autoclaving. Plates were incubated under aerobic conditions at room temperature (20°C).

Colonies were isolated and restreaked for isolation five times on STAT and Sea Water Complete (SWC) medium until pure cultures were obtained. SWC consists of 1X sea water base, 5 gL⁻¹ bacto tryptone, 1 gL⁻¹ yeast extract, 0.3% glycerol and 1.5% agar.

PCR and 16S rRNA Gene Sequencing.— We picked individual colonies of one of our isolates from the SWC plates for sequencing. Cells were lysed by suspending colonies in 25 µl alkaline phosphatase, and heating to 95°C for 5 min. PCR reactions consisted of 12.5 µl GoTaq® Green Mastermix (Promega Corporation, Madison, WI), 2 µl of both the general bacterial 16S rRNA gene forward 515F (GTGCCAGCMGCCGCGGTAA-3') and reverse 1391R(5'-GACGGGCGGTGTGTRCA-3') primers, 7.5 µl H₂O and 1 µl template. PCR conditions consisted of an initial denaturation of 2 min at 95°C, followed by 20 cycles of: 30s at 95°C, 30s at 55°C and 90s at 72°C, and followed by an extension of 10 min at 72°C. PCR product was visualized with gel electrophoresis on a 1% agarose gel stained with the SYBR® Gold Nucleic Acid Stain (Thermo Fisher Scientific, Waltham, MA). PCR products were submitted for external sequencing.

Genome Sequencing and Analysis. – DNA from colonies was extracted using the QIAGEN Genomic-tip 100/G extraction column (QIAGEN, Germantown, MD), to obtain high molecular weight DNA for sequencing. We sequenced the genome of the isolate using the MinION sequencer (Oxford Nanopore Technologies, Oxford, UK). Sequencing libraries were prepped according to manufacturers' instructions, and the bacterial isolate was sequenced using two MinION flow cells.

The isolate's genome was assembled using the Nanopolish program (Loman et al. 2015), and putative protein clusters were identified using Prokka (Seemann 2014). We visualized protein cluster locations in Anvi'o (Eren et al. 2015), and aligned our isolate's genome with 17 genomes available in the NCBI database based on protein cluster locations, following the pan genome analysis pipeline.

RESULTS

Colony and Bacterial Characteristics. – The colony selected for our study had dark orange pigment (Fig. 1) and had a light absorbance peak at 810 nm, indicating the presence of bacteriochlorophyll a. Colony growth was slow, with an average colony size of 5mm in diameter after 10 days of incubation. Colonies grew on both fresh and salt water based media, indicating high salt tolerance.

16S rRNA Gene Sequencing. – We successfully amplified our isolate as was shown by clear bands on our agarose gel, but were unable to obtain 16S rRNA gene sequences from our isolate. After three failed sequencing attempts, we decided to use the MinION sequencer to obtain a full genome sequence. We blasted the 16S rRNA gene obtained from our genomic sequencing in NCBI nucleotide blast, and our isolate blasted to *P. cryptus* and *P. mercurialis* with 97-98% sequence similarity.

Genome description. – We obtained 95,932 reads from the two MinION flow cells. The isolate had a genome size of 3,733,775 bp, divided over two contigs. Prokka identified 7674 putative genes in the genome, while the estimate derived from Anvi'o was slightly lower at 7061 genes. The average length of protein clusters in our genome was 309bp. We were unable to close the genome with the 197X (?; *to be recalculated*) coverage we achieved. Our gene content was high at 1.9 genes/kb, compared to the other available genomes of *Porphyrobacter* species (range = 0.90-1.16 genes/kb). Our isolate had a high GC-content of 65%, which was comparable to other *Porphyrobacter* species in the genus. The majority of our putative genes aligned with the core genome identified from our pan genome analysis (Fig. 2&3).

DISCUSSION

We isolated a novel species of *Porphyrobacter* using a protocol designed to isolate Actinobacteria. Actinobacteria have a high G-content (~72%), which makes them heat resistant and able to survive the drying steps described in our methods. The relatively high GC-content of our isolate (65%) likely increased its heat tolerance, resulting in its survival through our harsh protocol. Known *Porphyrobacter* species occur in a wide range of environments, and our novel species' characteristics fit into this wide niche breadth.

The potential presence of chlorophyll a indicates photosynthetic capabilities of our isolate, which is concurrent with other species in the genus. We will identify the location and size of the photosynthetic gene cluster, and compare and contrast these data with the 17 species and strains available. In addition, we will examine the metabolic machinery present in our isolate.

We observed a large protein cluster consisting of 13,000 proteins, which was unique to our isolate. Up to present we were unable to identify the putative function or validity of this cluster, but we suspect it is an artifact of combining data from two different flow cells. Identification of the proteins and genes in this cluster will indicate whether this cluster is indeed an artifact, or if it is a large fragment incorporated into the genome from external sources.

In conclusion, we isolated a novel, drought resistant species of *Porphyrobacter* from a brackish sediment sample, which expands our knowledge of the environmental range this genus occupies. We will continue to explore our isolate's genome, and attempt to identify genes that are involved in its halotolerance, energy metabolism and drought resistance, which will be described in a genome report to be submitted to the American Society for Microbiology.

ACKNOWLEDGMENTS

Many thanks go to Lisa Johnson Cohen, Harriet Alexander and C. Titus Brown for providing the MinION, performing the library preparation, genome sequencing and genome assembly. You saved my project. I thank Diane Newman and Jared Leadbetter for letting me participate in the Microbial Diversity course at the Marine Biological Laboratory, which has been an amazing and valuable learning experience. Lastly, I thank the S.O. Mast Memorial Fund and the Kansas State University Graduate Student Council for their financial support, without none of this work would have been possible.

REFERENCES

- Brenner DJ, Krieg NR, Staley JT, Garrity GW, Eds. (2005) *Bergey's Manual of Systematic Bacteriology*. Volume 2 : The Proteobacteria. Springer US. ISBN978-0-387-95040-2.
- Coil DA, Flanagan JC, Stump A, Alexiev A, Lang JM, Eisen JA. (2015) *Porphyrobacter mercurialis* sp. nov., isolated from a stadium seat and emended description of the genus *Porphyrobacter*. *PeerJ* 3:e1400 <https://doi.org/10.7717/peerj.1400>.
- Eren AM, Esen ÖC, Quince C, Vineis JH, Morrison HG, Sogin ML, Delmont TO. (2015) Anvi'o: an advanced analysis and visualization platform for 'omics data. *PeerJ* 3:e1319 <https://doi.org/10.7717/peerj.1319>.
- Furuhata K, Edagawa A, Miyamoto H, Kawakami Y, Fukuyama M. (2013) *Porphyrobacter colymbi* sp. nov. isolated from swimming pool water in Tokyo, Japan. *J Gen Appl Microbiol.* 59:245-50.
- Giovannoni SJ, Britschgi TB, Moyer CL, Field KG. (1990) Genetic diversity in Sargasso Sea bacterioplankton. *Nature* 345:60–63.
- Hiraishi A, Yonemitsu Y, Matsushita M, Shin Y, Kuraishi H, Kawahara K. (2002) Characterization of *Porphyrobacter sanguineus* sp. nov., and aerobic bacteriochlorophyll-containing bacterium capable of degrading biphenyl and dibenzofuran. *Arch. Microbiol.* 178:45.
- Li X, Koblížek M, Feng F, Li Y, Jian J, Zeng Y. (2013). Whole-genome sequence of a freshwater aerobic anoxygenic phototroph, *Porphyrobacter* sp. strain AAP82, isolated from the Huguangyan Maar Lake in southern China. *Genome Announc.* 1(2):e00072-13. doi:10.1128/genomeA.00072-13.

Loman NJ, Quick J, Simpson JT. (2015) A complete bacterial genome assembled *de novo* using only nanopore sequencing data. *Nature Methods* 12:733–735.

Sayers et al. "Alphaproteobacteria". National Center for Biotechnology Information (NCBI) taxonomy database. Retrieved: 2016-08-16.

Seemann T. Prokka: rapid prokaryotic genome annotation. (2014) *Bioinformatics* 30:2068-2069.

FIGURES

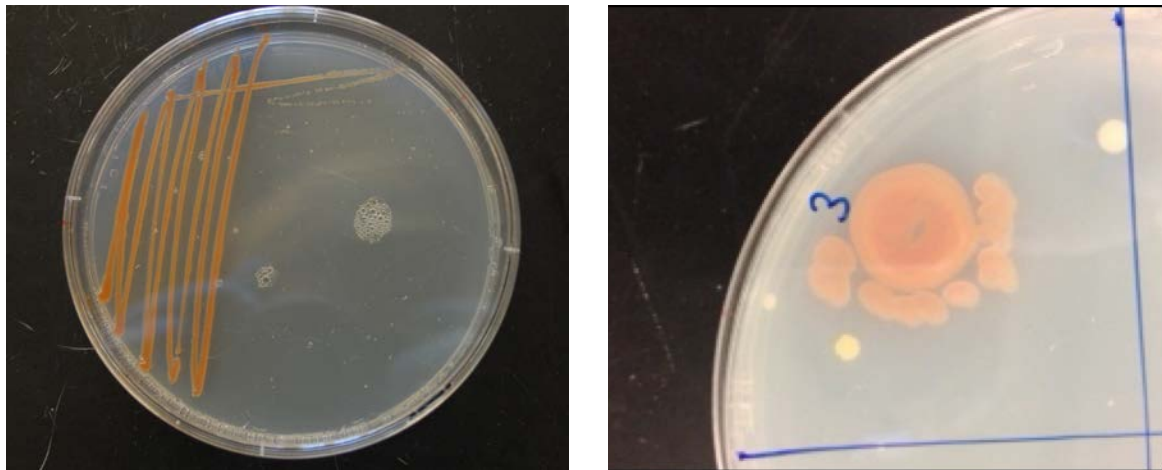


Figure 1. Colony morphology and pigmentation of the bacterial isolate of this study, grown on SWC (left) and STAT (right) media.

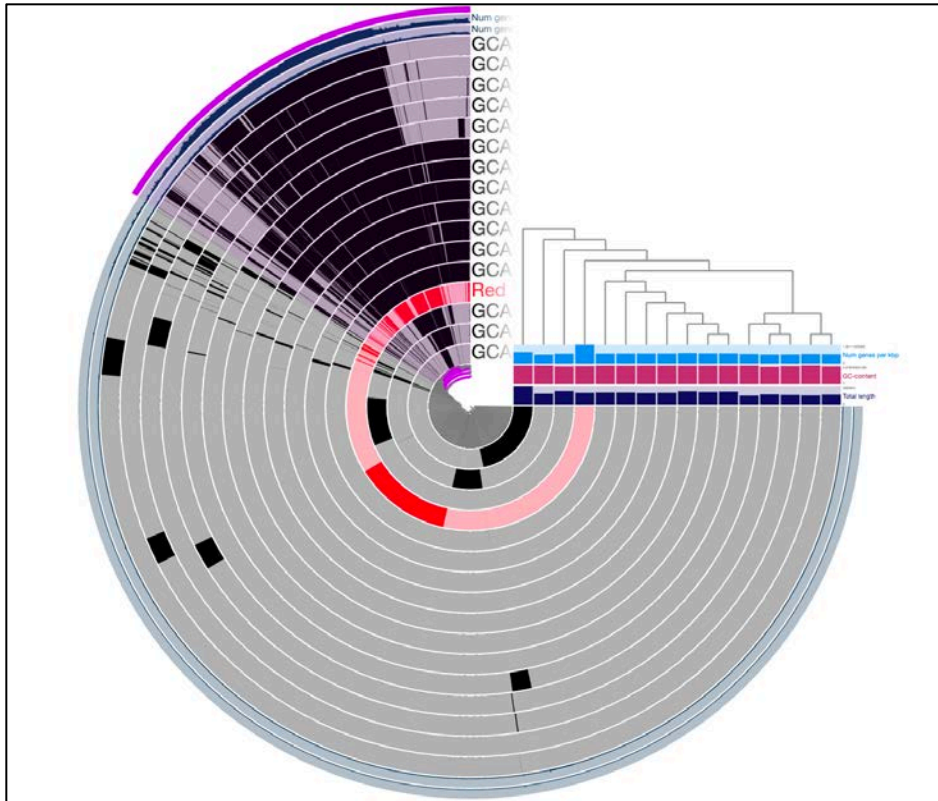


Figure 2. Pan genome of 17 *Porphyrobacter* species and strains (grey scale), including the novel species isolated in this study (red). Presence of putative genes or protein clusters are shown as black and darker red bars. The proposed core genome bin of the *Porphyrobacter* genus is highlighted in purple.

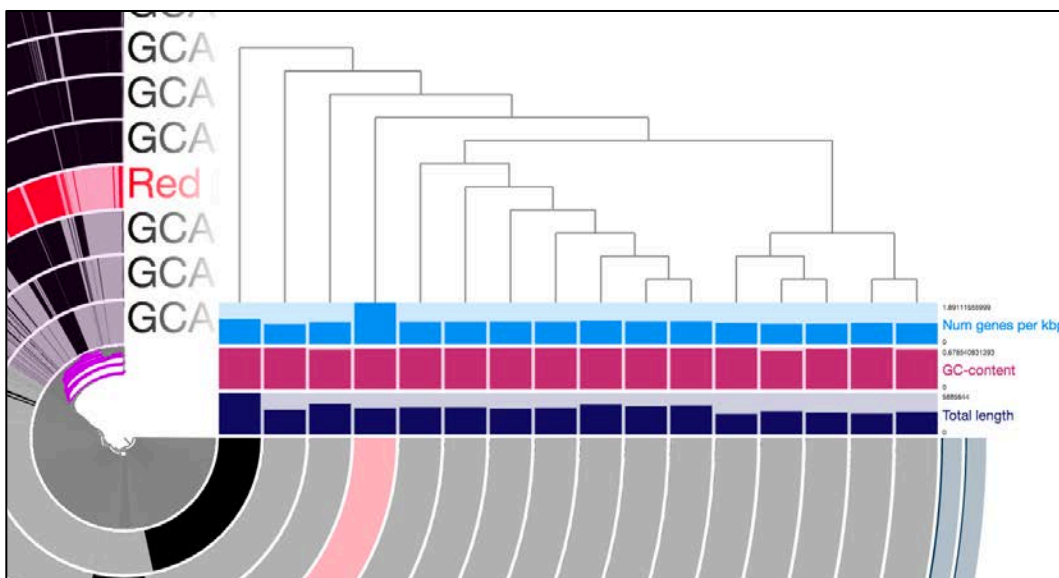


Figure 3. Genome length, GC-content and numbers of genes per kbp for 17 *Porphyrobacter* species/strains (grey scale) and the novel species isolated in this study (red).