The Full Genome Sequence of an Antarctic Microbe Constructed Using a Rapid, Portable Sequencer and a Hybrid Assembly

B. Boles[1], R. Murdoch[2], J. Mikucki[3]
[1] Department of Civil and Environmental Engineering, University of Tennessee
[2] Center for Environmental Biotechnology, University of Tennessee
[3] Department of Microbiology, University of Tennessee

Introduction

Blood Falls in the McMurdo Dry Valleys, Antarctica is home to a diverse makeup of microbial life. These microbes and their activities hold clues to ecosystem function and adaptation to extreme conditions. The genomes of these microbes can be sequenced to unlock these clues. A *Shewanella sp.* was isolated from Blood Falls and, because of is cultivability, serves as a model organism to study. In this investigation, the genome of *Shewanella* strain BF02_Shew was sequenced and analyzed to explore the lifestyle of Antarctic extremophiles. Here a genome sequencing pipeline for field use was developed. Ultimately, we aim to modify our pipeline for remote genome sequencing to extract information from microbial field samples in remote locations allowing us to understand microbial life in environmental conditions.

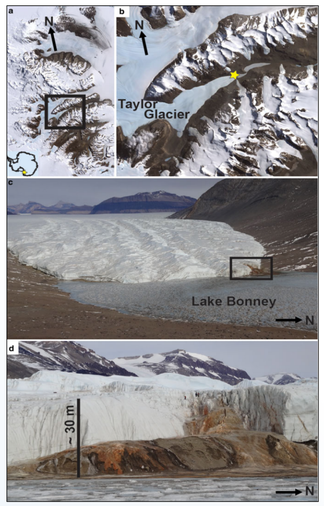


Figure 1 (right): Blood Falls; Taylor Glacier, Antarctica; Location from which *Shewanella* strain BF02_Shew was cultivated.[1]

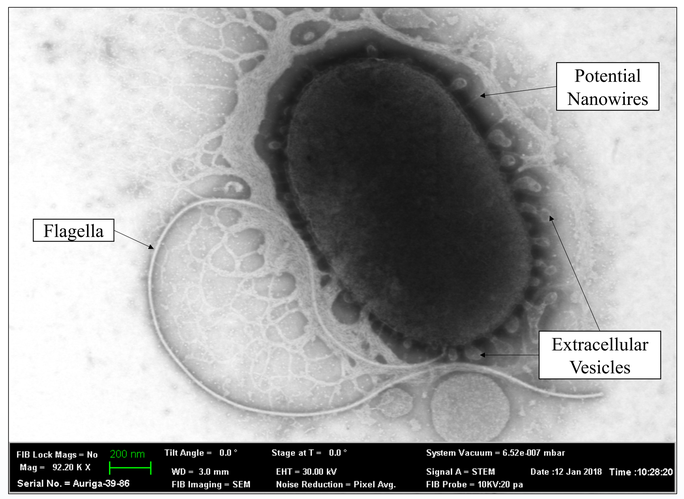
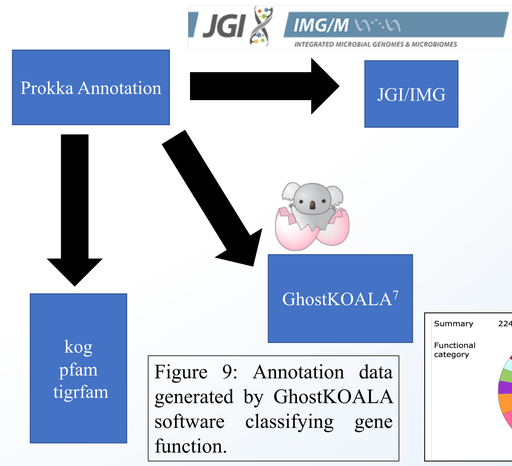


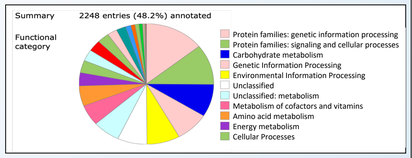
Figure 7: STEM image of the *Shewanella* strain BF02_Shew isolate showing a flagella, extracellular vesicles, and potentially nanowires.

Searching for Genes and Functions



After the genomic data of the *Shewanella* strain BF02_Shew was processed, the data was analyzed with prokka v1.13. The genes identified with prokka can then be searched for using different gene databases. These genes give clues to potential functions of the *Shewanella* strain BF02_Shew in the Blood Fall ecosystem.

Figure 9: Annotation data generated by GhostKOALA software classifying gene function.



Oxford Nanopore Sequencing

Genomic DNA was extracted from a pure culture of *Shewanella* strain BF02_Shew using the MoBio PowerSoil DNA isolation kit. An average DNA concentration was found to be 28.9 ng/µL with an optical density at 260/230 of 2.15 and an OD 260/280 of 2.01. Gel electrophoresis showed the average base length of the genomic DNA use to be approximately 9-10 kb. The recommended DNA quality for the Oxford Nanopore Technologies (ONT) pipeline is OD 260/280 of 1.8 and OD 260/230 of 2.0-2.2, an average fragment size to be >30kb, and 400 ng of genomic material to be loaded into the flowcell.

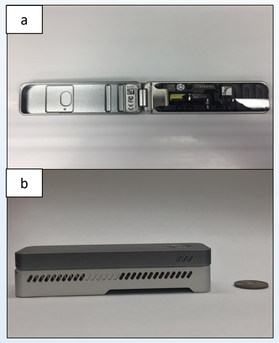


Figure 2 (above): a) The Oxford Nanopore open to show the flowcell, b) the Oxford Nanopore closed next to a penny to show size.

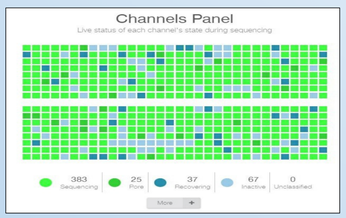


Figure 3: MinKNOW interface showing individual pores while sequencing genomic data.[2]

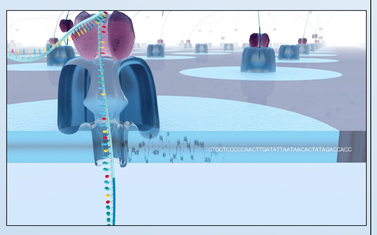


Figure 4: Nanopore sequencing graphic.[2]

Despite the genomic material failing to meet the ONT guidelines for sequencing, the material was prepared according to the ONT SKQ-RAD002 Rapid Sequencing protocol and loaded into the FLO-MIN106 (R9.4 SpotON) flowcell. Using the ONT MinKNOW software and the Metrichor live basecalling function, the sample was sequenced for 26 hours.

Illumina HiSeq

The genomic material was extracted using the CTAB method. The extracted genomic material was then sent to Mr. DNA for Illumina HiSeq processing.

Figure 5: Illumina HiSeq next to person for size comparison.[3]

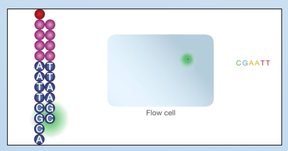
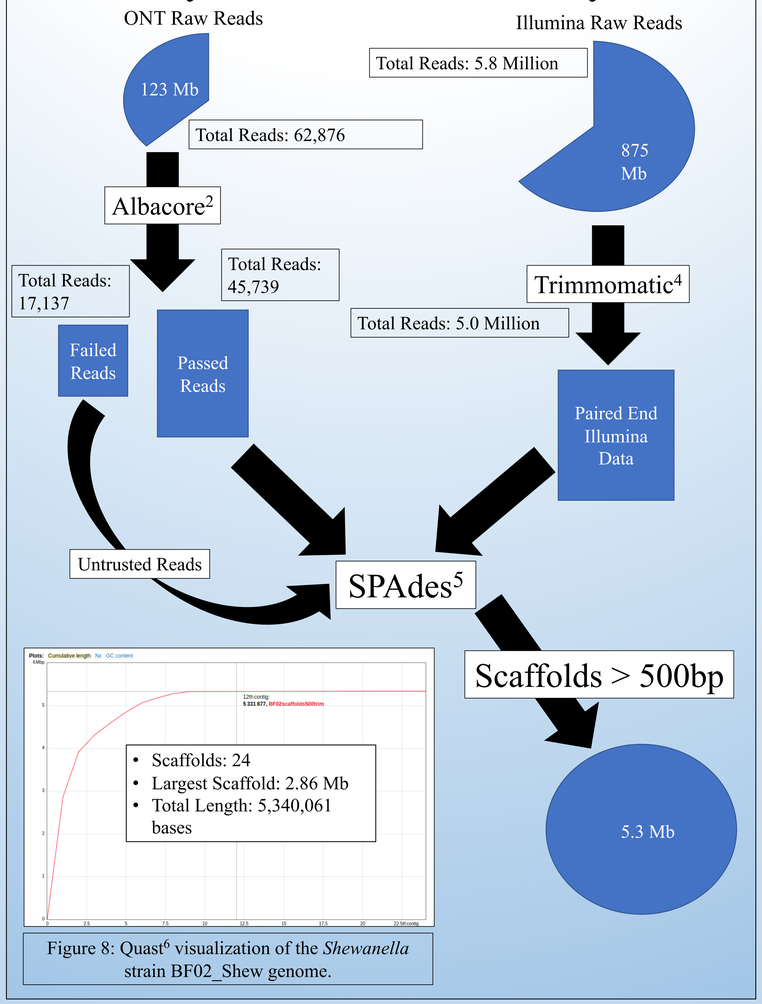


Figure 6: A graphic depicting Illumina sequencing.[3]

Hybrid Genome Assembly



- Scaffolds: 24
- Largest Scaffold: 2.86 Mb
- Total Length: 5,340,061 bases

Figure 8: Quast[6] visualization of the *Shewanella* strain BF02_Shew genome.

Links Between Phenotype and Genotype

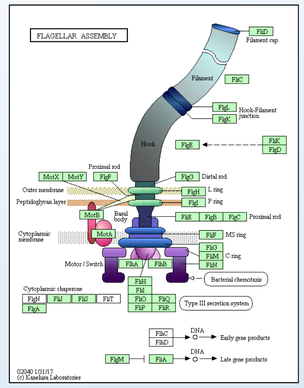


Figure 10: Genes associated with polar flagella assembly and function. Green boxes indicate genes identified by GhostKOALA.[7]

Confirmed by the STEM image of the *Shewanella* strain BF02_Shew. the microbe has the potential for motility. This could allow for *Shewanella* strain BF02_Shew to travel to areas of desired nutrients.

Flagellar Assembly

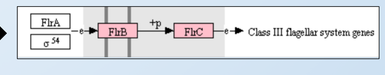


Figure 11: Two Component regulatory system leading to flagellar assembly. Red boxes indicate identified genes.[7]

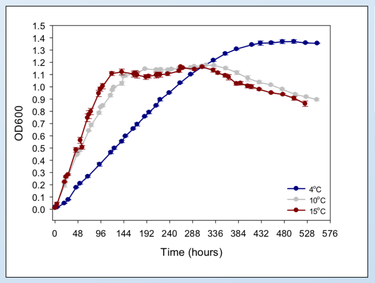


Figure 12: Growth curve of *Shewanella* strain BF02_Shew in marine broth.[8]

Cold Adaptations

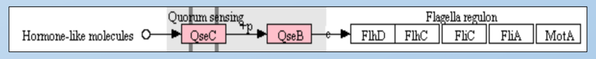
As seen in Figure 12, *Shewanella* strain BF02_Shew is a cold tolerant organism, and grows more quickly at 10 °C and 15 °C. The *fabF* gene, found in this genome, is related to cold and pressure tolerance in microorganisms. The gene has the potential to be horizontally transferred from *Marinobacter lipolyticus* to *Shewanella* strain BF02_Shew.

Figure 13 (left): Test tube showing the crystal violet stained biofilm of *Shewanella* strain BF02_Shew.[8]

Biofilm Formation

Figure 14: The two component regulatory system coding for quorum sensing.[7]

The gene coding for quorum sensing may allow *Shewanella* strain BF02_Shew to form biofilms and excrete extracellular vesicles. This would allow for sharing of substrates in an environment with an extreme lack of nutrients and energy turnover.



References:
[1]BADGELEY, J., et al. (2017). An englacial hydrologic system of brine within a cold glacier: Blood Falls, McMurdo Dry Valleys, Antarctica. *Journal of Glaciology, 63*(239), 387-400.
[2]Oxford Nanopore Technologies (nanoporetech.com)
[3]Illumina (youtube.com/user/IlluminaInc)
[4] Bolger AM, et al. (2014) Trimmomatic: A flexible trimmer for Illumina Sequence Data, *Bioinformatics, 30 (15): 2114-2120.*
[5] Bankevich, A., et al. (2012). SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *Journal of Computational Biology.*
[6]Gurevich, A, et al., (2013) QUAST: quality assessment tool for genome assemblies, *Bioinformatics* 29 (8): 1072-1075.
[7] Kanehisa, M., et al. (2016) BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. J. Mol. Biol. 428, 726-731.
[8] A. Jarratt (unpublished data).

Acknowledgements: Special thanks to Abigail Jarratt for the data on *Shewanella* strain BF02_Shew; Dr. John Dunlap for teaching me how to take images of cells on the STEM; Dr. Robert Murdoch for pointing me in the correct direction concerning all things bioinformatics; Dr. Jill Mikucki for support both in and out of the lab.