Kellen Tyrrell

Dr Brandon Briggs

BioA455

Identifying an unknown sequence using 16S rRNA

Methods

To identify the genus of the unknown organism, its 16S rRNA was identified and compared to known sequences. Quast (Alexey, 2013) was used to identify the 16S rRNA sequences of the unknown genome, and the percent completeness of the unknown genome. Bedtools (Aaron, 2010) was then used to extract the 16S rRNA gene from unknown sequence, and BLAST (Altschul, 1990) was used to find the closest known sequence to the extracted 16S rRNA sequence. In order to compare the unknown sequence with the reference sequence, BLAST was used to create a comparison file, and ACT (Carver, 2005) was used to graphically represent this comparison.

Results

The unknown genome was found to have a genome size of 3.51 Mbp, a gc content of 56.19%, and a BUSCO score of 85.14%. The unknown genome was predicted to contain 3030 unique genes, and 9 rRNA genes. The organism with the genome most closely paired with the unknown using NCBI BLAST was Marinobacter excellens HL-55. The Marinobacter excellens genome obtained had a size of 3.99 Mbp, a gc content of 56.26%, and was predicted to contain 3423 unique genes, with 9 of them being rRNA genes. Therefore, there were 393 genes missing

in the unknown genome as compared to the reference genome, which occurred from 1.8 Mbp to 2.3 Mbp in the genome (Figure 2).
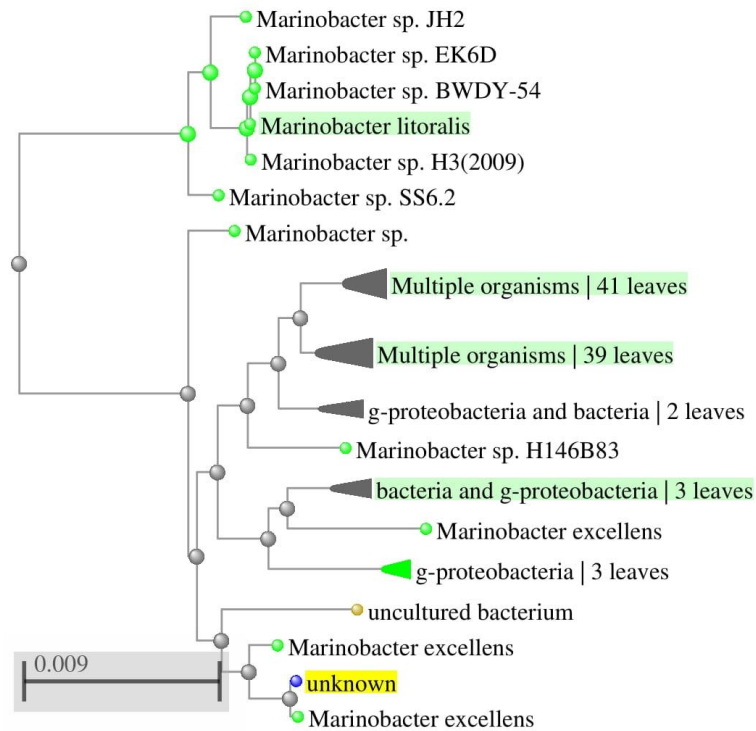

Figure 1. Blast Tree View visualization of blast search using unknown 16S rRNA sequence. The unknown 16S rRNA is highlighted in yellow
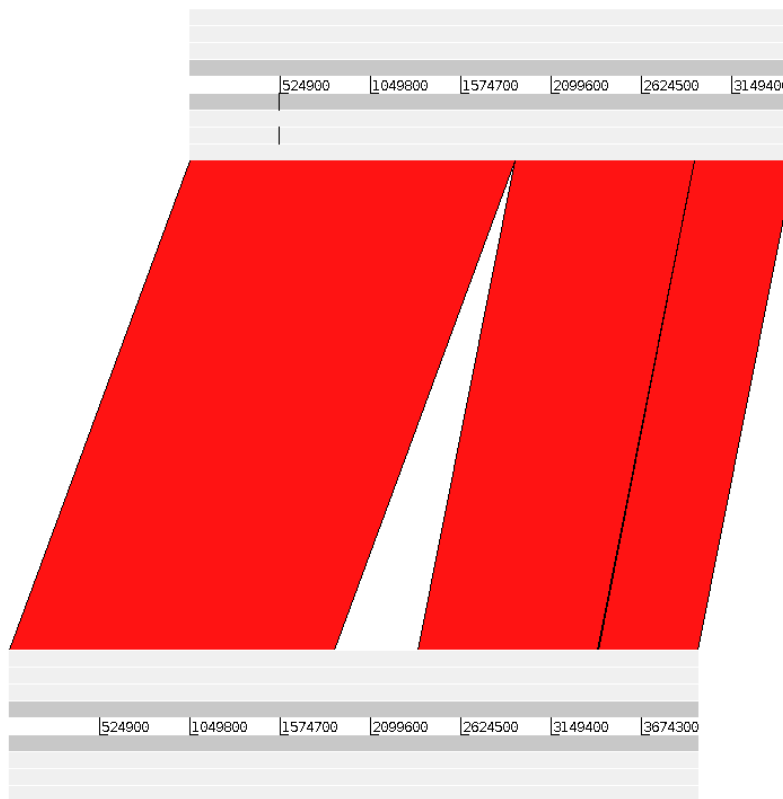
Figure 2. ACT visualization of BLAST alignment between the unknown sequence and Marinobacter excellens genome. The unknown genome is pictured above, and Marinobacter excellens genome pictured below. The gap in the genome occurs from 1.8 Mbp to 2.3 Mbp

References

Aaron R. Quinlan, Ira M. Hall, BEDTools: a flexible suite of utilities for comparing genomic features, Bioinformatics, Volume 26, Issue 6, 15 March 2010, Pages 841–842, https://doi.org/10.1093/bioinformatics/btq033

Alexey Gurevich, Vladislav Saveliev, Nikolay Vyahhi and Glenn Tesler, QUAST: quality assessment tool for genome assemblies, Bioinformatics (2013) 29 (8): 1072-1075. doi: 10.1093/bioinformatics/btt086 First published online: February 19, 2013

Altschul, S., Gish W., Miller W., Myers E., Lipman., D. (1990). Basic local alignment search tool. *Journal of Molecular Biology.* 215(3), 403-410

Carver TJ, Rutherford KM, Berriman M, Rajandream MA, Barrell BG, Parkhill J. ACT: the Artemis Comparison Tool. Bioinformatics. 2005 Aug 15;21(16):3422-3. doi: 10.1093/bioinformatics/bti553. Epub 2005 Jun 23. PMID: 15976072.