

POSTER PRESENTATION

Open Access

Sequence and annotation of the Wizard007 mycobacterium phage genome

Ejike Anyanwu, Kaitlyn Cole, Karlee Driver, Anthony Falcone, Elizabeth Farnsworth, Benjamin Howard, Brittney Howard, Courtney Howard, Rodney King, Jordan Olberding, Mackenzie Perkins, Claire Rinehart*, Heidi Sayre, Tyler Scaff, Sarah Schrader, Prasanna Tamarapu Parthasarathy, Cynthia Tope

From UT-ORNL-KBRIN Bioinformatics Summit 2010
Cadiz, KY, USA. 19-21 March 2010

Material and methods

The Wizard007 bacteriophage was isolated from a soil sample taken at N36° 49'36.8" W87° 29'42.5" in Hopkinsville, KY. It was plaque purified on *Mycobacterium smegmatis* and DNA was isolated using the Promega Wizard DNA purification kit. The DNA was sequenced at the Virginia Commonwealth University by Roche 454 DNA sequencing. The genome quality was confirmed in Consed [1] and found to be 51,034 bp in length. Genes were called from the assembled sequence with a workflow consisting of Glimmer [2], GeneMark [3], tRNAscan [4], and SDFinder to identify potential gene features. Gbrowse [5], Apollo [6], and BLAST [7] were used to call and annotate genes. A 10 base 3' overhang was found at the ends by ligating the genomic DNA and sequencing a PCR product produced across the ligated ends.

Results and conclusion

Wizard007 is 95% similar to the *Mycobacterium phage Peaches* (NC_013694) when comparing the number of gene BLAST hits. The first 41,000 bp and the genes contained therein match Peaches' organization and sequence very closely. The 3' end of Wizard007 lacks 4 coding sequences that are found in Peaches (CDS 69, 74, 75, & 78) and it has one coding sequence that is not called in Peaches (gp 76). Additionally there are two coding sequences in Peaches (CDS 83 & 84) that do not match Wizard007 (gp 77 & 79).

Acknowledgements

Funding for the DNA sequencing of Wizard007 was provided by the WKU Bioinformatics and Information Science Center. Supplies and technical

support for the isolation and genome annotation were provided by the Howard Hughes Medical Institute Science Education Alliance National Genome Research Initiative.

Published: 23 July 2010

References

1. Gordon D, Abajian C, Green P: **Consed: A Graphical Tool for Sequence Finishing.** *Genome Research* 1998, **8**:195-202.
2. Delcher AL, Harmon D, Kasif S, White O, Salzberg SL: **Improved microbial gene identification with GLIMMER.** *Nucleic Acids Research* 1999, **27**:4636-4641.
3. Besemer J, Lomsadze A, Borodovsky M: **GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions.** *Nucleic Acids Research* 2001, **29**:2607-2618.
4. Lowe TM, Eddy SR: **tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence.** *Nucleic Acids Research* 1997, **25**:955-964.
5. Stein LD, Mungall C, Shu S, Caudy M, Mangone M, Day A, Nickerson E, Stajich JE, Harris TW, Arva A, Lewis S: **The generic genome browser: a building block for a model organism system database.** *Genome Res.* 2002, **12**:1599-610.
6. Lewis SE, Searle SMJ, Harris N, Gibson M, Iyer V, Richter J, Wiel C, Bayraktaroglu L, Birney E, Crosby MA, Kaminker JS, Matthews B, Prochnik SE, Smith CD, Tupy JL, Rubin GM, Misra S, Mungall CJ, Clamp ME: **Apollo: a sequence annotation editor.** *Genome Biol* 2002, **3**:RESEARCH0082, research.
7. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: **Basic local alignment search tool.** *J. Mol. Biol* 1990, **215**:403-410.

doi:10.1186/1471-2105-11-S4-P15

Cite this article as: Anyanwu et al.: Sequence and annotation of the Wizard007 mycobacterium phage genome. *BMC Bioinformatics* 2010 **11** (Suppl 4):P15.

* Correspondence: claire.rinehart@wku.edu

Department of Biology, Western Kentucky University, Bowling Green, KY 42101, USA