Kapadia and Rinehart *BMC Bioinformatics* 2013, **14**(Suppl 17):A10
http://www.biomedcentral.com/1471-2105/14/S17/A10

BMC Bioinformatics

MEETING ABSTRACT Open Access

Ab Initio prediction of mycobacteriophages protein structure and function

Chiraag D Kapadia, Claire A Rinehart*

From 12th Annual UT-ORNL-KBRIN Bioinformatics Summit 2013
Buchanan, TN, USA. 22-24 March 2013

Background

Mycobacterium smegmatis is a soil bacterium. Over 448 mycobacteriophages specific for *M. smegmatis* have been sequenced and grouped into clusters of related genomes based on the similarity of their products and genome organization. Only 20% of mycobacteriophage genes have known function, as predicted by protein sequence level alignments [1].

Figure 1 Seven predicted pham 6714 structures superimposed into a single model matching methyltransferase structures. PDB files were rendered using UCSF Chimera (http://www.cgl.ucsf.edu/chimera/). Proteins include MeeZee gp 78, JN243856.1; Backyardigan gp 74, JF704093.1; and ShiLan gp 70, JN020143.1

* Correspondence: claire.rinehart@wku.edu
Department of Biology, Western Kentucky University, Bowling Green, KY, 42101, USA

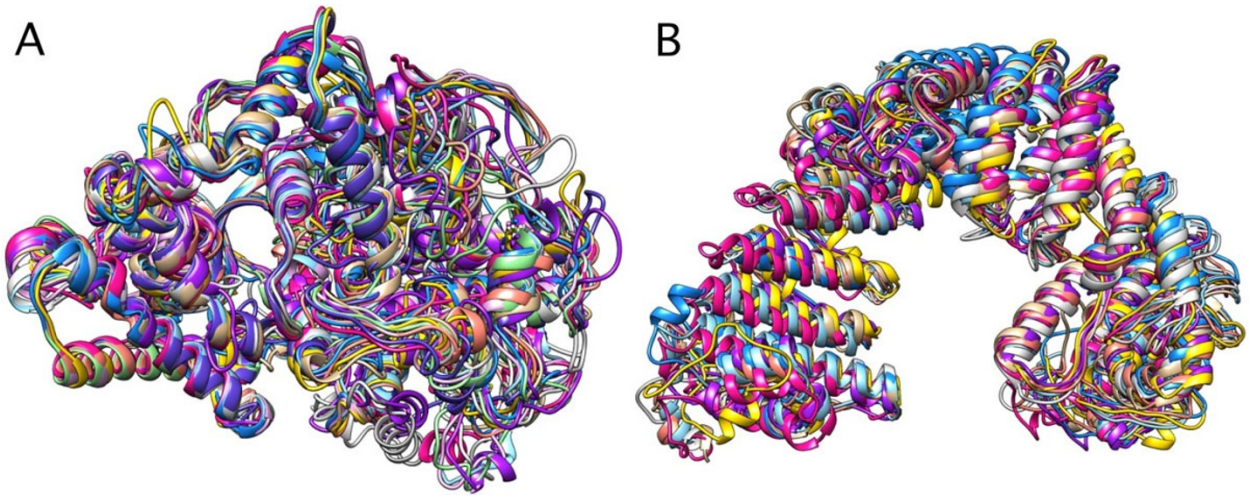

Figure 2 Ten predicted pham 2789 structures superimposed in two possible models matching either a neutral endopeptidase (A) or an importin protein (B). PDB files were rendered using UCSF Chimera (http://www.cgl.ucsf.edu/chimera/). Proteins include Ava3 gp 250, JQ911768.1; Rizal gp 242, EU826467.1; and Pio gp 260, JN699013.1

Materials and methods

Genes that are grouped together using BLAST at the protein sequence level have been assembled into loose groupings called phams [2]. The phagesdb.org/phams database contains the protein sequences organized by phams. From these data we used *ab initio* folding, using I-TASSER [3], to predict the structure of multiple phams across numerous mycobacteriophage clusters. Predicted models were grouped into structural families based upon RMSD scores from pairwise comparisons. Models from two structural families per pham were submitted to COFACTOR [4], which finds the best structural homologies to proteins in the PDB library and returns the matching structures along with GO terms, EC numbers and active site information.

Results

Based on COFACTOR output, we were able to suggest functions for the genes in each respective pham examined. Two notable results: 1) pham 6714 is predicted to be a methyltransferase (Figure 1), and 2) pham 2789 is predicted to be either a neutral endopeptidase (Figure 2a) or an importin protein (Figure 2b). These predicted functions will need to be confirmed experimentally.

Published: 22 October 2013

References

1. Hatfull GF: **Mycobacteriophages: genes and genomes.** *Annu Rev Microbiol* 2010, **64**:331-356.
2. Cresawn SG, Bogel M, Day N, Jacobs-Sera D, Hendrix RW, Hatfull GF: **Phamerator: a bioinformatics tool for comparative bacteriophage genomics.** *BMC Bioinformatics* 2011, **12**:395.
3. Roy A, Kucukural A, Zhang Y: **I-TASSER: a unified platform for automated protein structure and function prediction.** *Nature Protocols* 2010, **5(4)**:725-738.
4. Roy A, Yang J, Zhang Y: **COFACTOR: An accurate comparative algorithm for structure-based protein function annotation.** *Nucleic Acids Res* 2012, **40(Web Server issue)**:W471-W477.

doi:10.1186/1471-2105-14-S17-A10
Cite this article as: Kapadia and Rinehart: ***Ab Initio*** **prediction of mycobacteriophages protein structure and function.** *BMC Bioinformatics* 2013 **14**(Suppl 17):A10.