Cover Letter for Arthrobacterphage Amigo

Version 2

Baylor University

Dear Debbie,

I used the file that had the latest documentation, converted it to one that has the 32 extra bases, changed the start site for gene 1, and reblasted this file.

This file is named AmigoconvertedAdairfinalBLAST.dnam5

Notes for Amigo

* Amigo is a singleton Arthrobacterphage.
* The draft Phamerator file had 93 genes and all but a few were orphams in the Actino\_draft Phamerator database.
* We were not able to look at a heuristic CP graph because GenMark was down.
* When we annotated, Amigo was not in the Arthrophage Phamerator database.
* The final annotation has 109 genes including 7 tRNA genes.
* The inserted genes are between 90 and 240 bps.

**Gene Summary**

19 hypothetical proteins/unknown function

10 genes blasted to known functions

genes 42-48 tapemeasure;structural genes;lysin

genes 62-85 DNA polymerase, kinase, ligase, helicase, methylase, hydrolase

**Genome organization**

* The original assembly files indicated that it had a repeat of 1552 bp.
	+ Resequencing at Pitt resulted in adding 32 bases to the beginning of the genome, resulting in a repeat of 1584. Dan commented that this has now been seen in a few other Amigo like genomes and he chose a “bump” on the coverage to serve as base 1.
* The repeat region contains 7 called ORFS. With the addition of the 32 bases, ORF 1 matches ORF 103.
* Gene 7 was called in the reverse direction, but if you compare to the genes 103-109, gene 7 should match 109, which is a forward gene of 120bp.

**Other areas to debate**

Gene 14: to lengthen or not to lengthen…

Gene 21: forward or reverse?

Gap between 27/28: 93bp; we did not call

Inserted genes: 1,6,7,29,32,61,65,91,93,94,95,99,103,108,109

Since most of our called genes did not BLAST to any known genes and the coding potential chart used from *A. aurescens* has not been used with a large number of arthrobacterphage, we opted for filling gaps with an SD score over 200 and ORFs over ~ 90 bp even if there was no coding potential or BLAST result. During the summer workshop we saw some small peptides from other phage proteomic studies, so we opted to hypothesize that this may actually occur.

* Potential Frameshift: gene 77-78. We are fairly confident this is a -1 frameshift, but we did not call it due to lack of any protein evidence or comparable genomes.