Genomic Analysis of Sahara and SoilAssassin

Background:

Bacteriophages are viruses that target and kill bacteria. Under a national research project called SEAPhages, two students from the University of Pittsburgh discovered bacteriophages SoilAssassin and Sahara: Daniel Biery and Sahara Grinkewitz, respectively. Both Sahara and SoilAssassin are in the cluster CZ and subcluster CZ2 and attack the bacteria Gordonia. We began our analysis of the genomes of SoilAssassin and Sahara in hopes of finding promoter regions and conserved repeats which would help us understand the genomes of Sahara and SoilAssassin. Because Gordonia bacteria is closely related to the pathogen that causes Tuberculosis, conducting searches involving these phages may have the potential to treat TB.

Methods:

Research Question: What similarities do these two bacteriophages share and what do these similarities reveal about the importance of certain genes in Sahara and SoilAssassin?

Hypothesis: By comparing our bacteriophages' promoter regions and conserved repeats we would identify which genes are most important to their survival and function

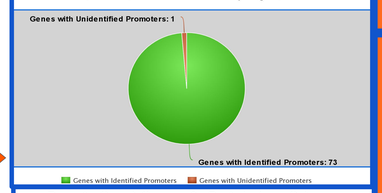
Procedure: First, we obtained background information for both bacteriophages through PhagesDB. PhagesDB provided us with information about the lysin production for both phages. Next, we used BPROM, which is a program that searches for gene promoters. BPROM yielded few results, so we moved to DNAMaster, the central bioinformatics software for our project. With DNAMaster, we obtained most of our data for promoters and for reverse strand similarity. Lastly, we utilized the BLAST tool in PhagesDB, which provided data on the genomic similarity between our phages and our matching nucleotide regions, called conserved repeats.

Results:

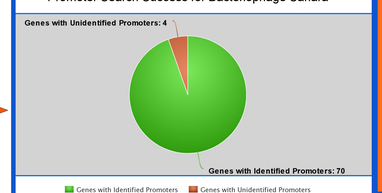
Sahara	SoilAssassin
95% Genomic Similarity (from BLAST)	
Found promoters for 94% of Genes	Found promoters for 98% of Genes
Matching Promoters found in Sahara's Genes 9, 10, 38, and 68	Matching Promoters found in SoilAssassin's Genes 9, 10, 39, and 66.
35.1% of Sahara's Genes had conserved repeats	33.7% of SoilAssassin's Genes had conserved repeats
High density of conserved repeats in genes 3, 15, and 28	
Conserved Repeats are more concentrated at the beginning and end of the Genome	Conserved Repeats are more concentrated at the center of the Genome
Lysin Production in Genes 18 and 23	Lysin Production in Genes 18 and 24
54% alignment of reverse running Genes	

Promoter Search Results:

Promoter Search Success for Bacteriophage SoilAssassin



Promoter Search Success for Bacteriophage Sahara



After analysis of the promoter regions, conserved repeats, and lysin-production of Sahara and SoilAssassin, it is clear that genes 3, 9, 15, 18, and 28 are of particular importance to our phages' genomes. Gene 18's production of Lysin A is especially noteworthy, as lysin producing bacteriophages have immense potential in the medical field.

By Logan Wilson, Fhasal Alam and Professor Sharon Gusky

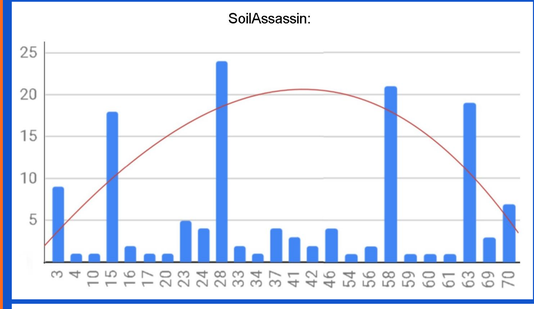
With contributions from Amanda Gregg, Renee Dunbar, and Amit Banik

Comparison: Sahara vs. SoilAssassin (Yellow represents genomic dissimilarities)

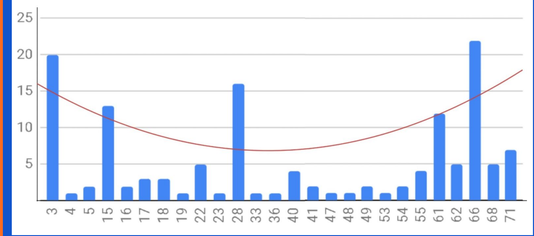
Supported by NSF ATE grant #1801062

Conserved Repeats Search Results:

SoilAssassin:



Sahara:



Discussion:

Analysis: We can conclude that genes 3, 9, 15, 18, and 28 are integral to the function of our phages. To summarize, in both Sahra and SoilAssassin, Genes 3, 15, and 28 had a high concentration of conserved repeats. This data led us to conclude that these genes are important to the function of our phages. Gene 9 had matching promoters between Sahara and SoilAssassin and also codes for the head-to-tail stopper. Head-to-tail stopper genes code for the proper insertion of DNA from virus to bacteria, emphasizing this gene's importance. Lastly, gene 18 is very important to both of our phages, as not only did it include an incredible amount of conserved repeats, but it also directly codes for Lysin A in both bacteriophages. Therefore, genes 3, 9, 15, 18, and 28 are all very important to both Sahara and SoilAssassin.

Implications: Our research highlights the importance of lysin production within bacteriophages. Lysin directly attack the cell walls of bacteria, so Lysin-producing bacteriophages could serve as a viable alternative to antibiotics. Bacteriophages are only able to target one type of cell, which is the bacteria that they are used to attacking. This means that lysin-producing phages can be used as a specialized treatment against certain pathogens, and antibiotics can be avoided. Further, our research leaves some hanging questions for future researchers to address, namely:

- Why do Sahara's genes 10, 38, and 68 and SoilAssassin's genes 10 and 66 have identical promoters but no identical function?
- Why does SoilAssassin have a defined function for gene 28 but Sahara does not?

Fortunately for future researchers, our identification of promoter positions gives them a head-start to their research.

Limitations: One limitation of our research is that it focuses on phages that attack Gordonia. Gordonia is not a common disease-causing agent, meaning that our research does not actively contribute to the fight against any mainstream diseases. However, because Gordonia is closely related to the pathogen causing Tuberculosis, our research may be applicable in the future in the fight against Tuberculosis.

Sahara: SoilAssassin: