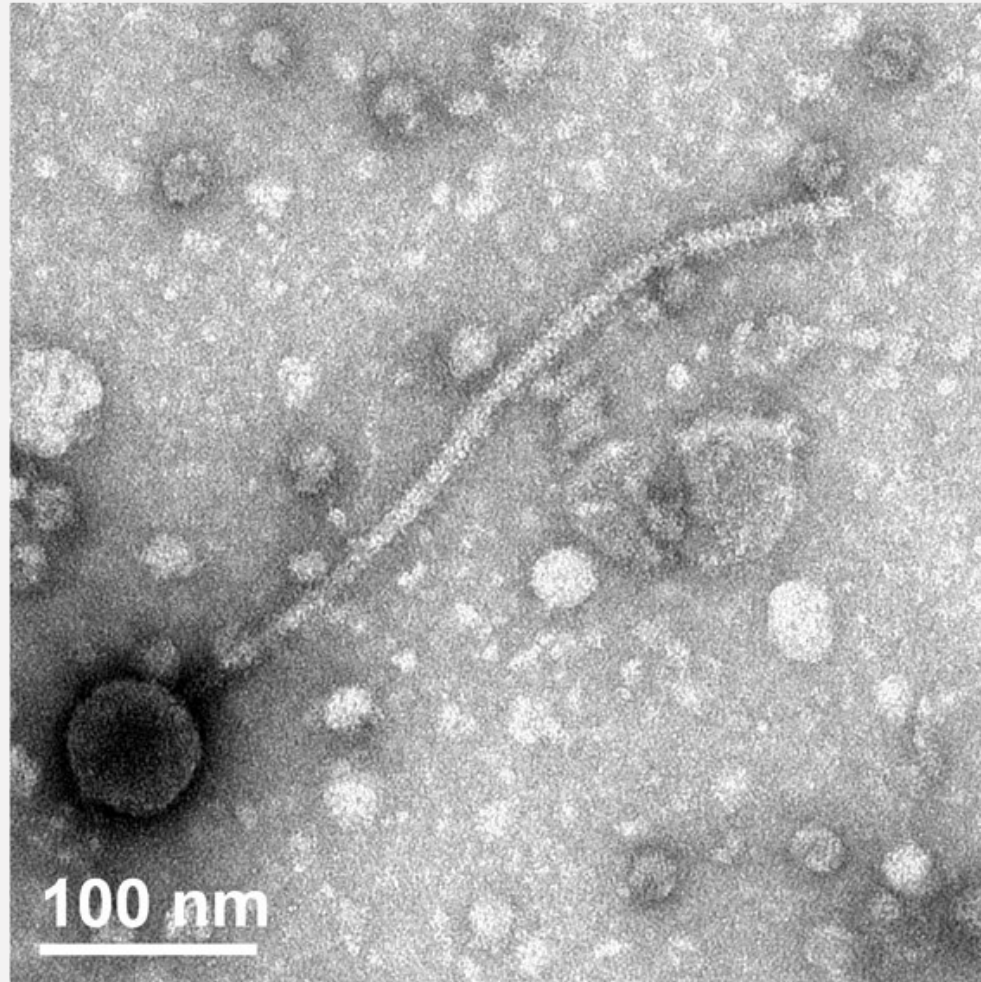


Genomic Annotation Research of Bacteriophage DoobyDoo

Ibrahim Abdulrehman and Preston Penny

April 19th, 2024



Outline

- Background on Bacteriophages and Bacteria
- Background on Genomic Annotation Process
- Interesting Finds: Possible Related Functions?
- Reflections and Summary
- Future Applications of Our Research

What's the DoobyDoo?

- DoobyDoo is a Bacteriophage (Phage)
- Bacteriophages are viruses that infect and kill a **specific** bacterial host
- DoobyDoo was purified and isolated in Fall 2023
- The genome was sequenced from the Pittsburgh Bacteriophage Institute

Background: The Threat of Resistance

- Widespread exposure to antibiotics has created mutations within bacterial populations resulting in the emergence of drug-resistant bacteria.
- The emergence of drug-resistant bacteria is now recognized as a global public health threat (Munita & Arias, 2016).
- The CDC estimates that roughly 23,000 individuals die from drug-resistant infections annually (Munita & Arias, 2016).

Background: The Bacteriophage

- Bacteriophages possess a unique ability to infect and kill a singular bacterial host.
- They are among the most abundant biological entities on the planet and can withstand extreme environments (Principi, et al., 2019).
- The bacteriophage's potential ability to eradicate drug-resistant bacteria has warranted a growing importance surrounding bacteriophage research (Pedulla et al., 2003).

Background: Genomic Annotation

- Central Dogma: DNA Genes → RNA → Protein
- Genomic annotation is the process of identifying unknown genes and predicting the function of the genes
- Steps of genomic annotation:
 - 1. Is it a gene?** (Evidence: Coding Potential & E-values)
 - 2. Where does it start?** (Evidence: RBS Values & Annotation Comparisons)
 - 3. What is its function?** (Evidence: Protein Database & Genomic Comparison)

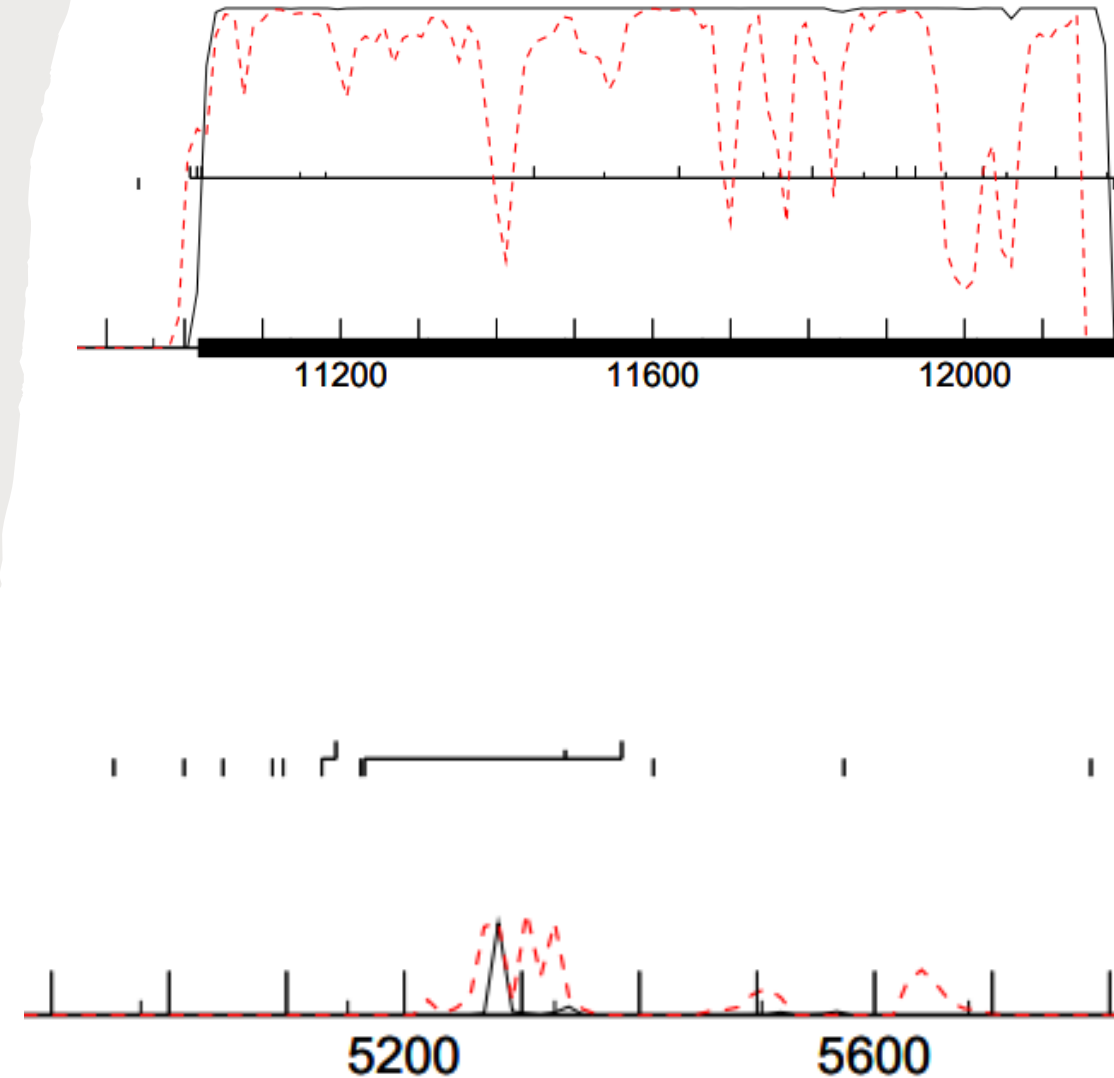
Step 1: Determining the Presence of a Gene

Bioinformatic tools: Genemark and BLAST

Step 1: Is it a Gene?

Evidence #1: Coding Potential

- Genemark software allowed us to visualize the coding potential of a predicted gene
- Utilizes probability scores to calculate coding potential
- Presence of strong coding potential suggests the presence of a gene in predicted region



Images retrieved from:

<https://genemark.bme.gatech.edu/GeneMark/tmp/genemarks.20240418.120926.1745575.gms.out.pdf>

Step 1: Is It a Gene?

Evidence #2: E-values

- BLAST program calculates an expect value (E-value).
- Runs gene sequence through a database to check alignment with other phage genes.
- Lower E-values indicates that a gene is found in DoobyDoo

QBlast High-Scoring Pairs (HSP)			
HSP Data		Alignment	
Bit Score	221.5	Identities	133
Score	563	%Identity	100.00
E-Value	0.0E0	Positives	133
Length	133	%Similarity	100.00
% Aligned	100.0 %	Gaps	0

QBlast High-Scoring Pairs (HSP)			
HSP Data		Alignment	
Bit Score	35.4	Identities	21
Score	80	%Identity	40.38
E-Value	8.41	Positives	29
Length	52	%Similarity	64.44
% Aligned	20.1 %	Gaps	7

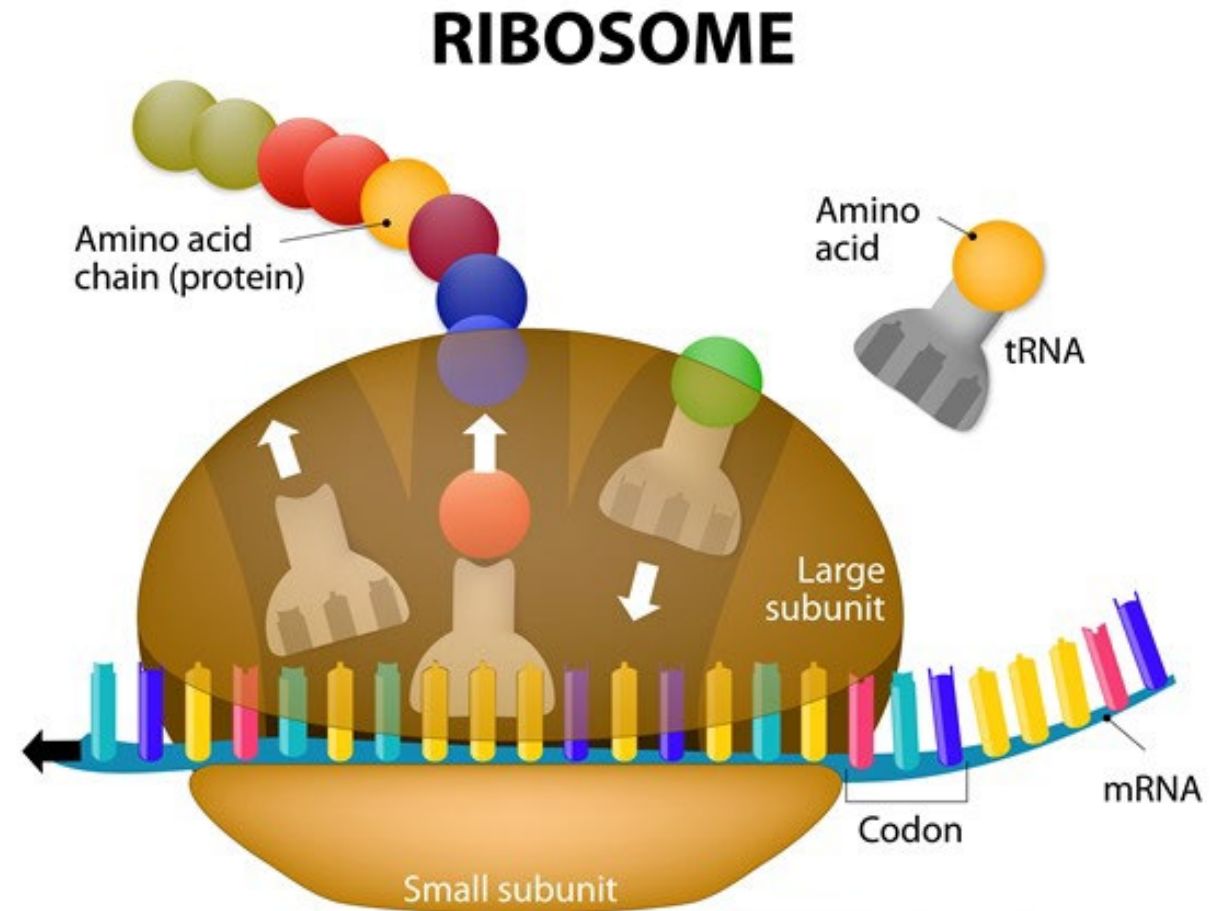
Step 2: Determining the Start Site of a Gene

Bioinformatic tools: RBS Values and Annotation Comparisons

Step 2: Where Does the Gene Start?

Evidence #1: RBS Values

- RBS: Ribosome Binding Site
- Shows the probability that a Ribosome will bind to a gene
- Values within range are suggestive of a start site



©Designua / Shutterstock.com

Image retrieved from: <https://www.news-medical.net/life-sciences/Ribosome-Function-in-Cells.aspx>.

Step 2: Where Does The Gene Start?

Evidence #2: Annotation Comparisons (Starterator)

- Comparison of completed phage genomes is a powerful tool allowing us to see:
 - A. How many genes from other phages start at the same position
 - B. How common start positions are in a sequence



Step 2: Combining Data

- Created comparison tables
- Needed to assess all the pieces holistically
- More evidence for a site = More likely

Option 1	Option 2
RBS values within range	RBS values not within range
10 phages with similar starts	4 phages with similar starts
Large gap from last gene	Small gap from last gene

Step 3: Predicting the Gene Function

Bioinformatic tools: Protein Databases and Genomic Comparisons

Step 3: What Is the Function?

Evidence #1: Protein Databases

- Amino acids of the genes were run through protein databases
- The databases aligned the sequence with similar amino acids
- Alignments tell us likely function

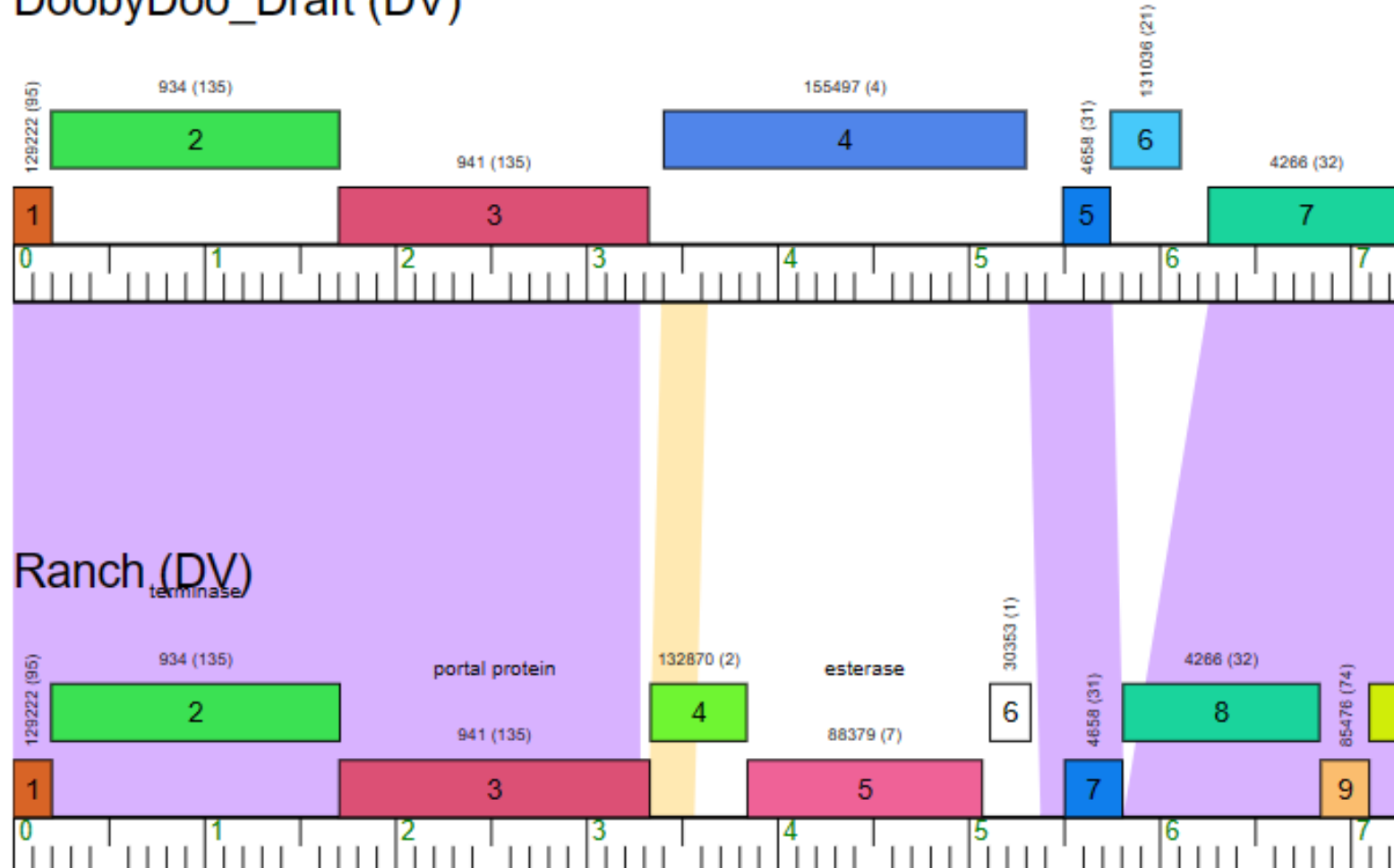
<input type="checkbox"/>	8	6L8O_A	DNA repair protein RAD5; DNA binding protein, DNA damage tolerance, Helicase , Snf2 family; 3.3A {Kluyveromyces lactis NR	100	2.3e-45	403.31	38.8	457	952
<input type="checkbox"/>	9	7XYF_K	ATP-dependent helicase fft3; DNA binding, remodeler, nucleosome, Fft3-nucleosome complex, DNA BINDING PROTEIN; HET: MSE;	100	5.5e-42	371.74	41.7	439	672
<input type="checkbox"/>	10	7XYG_K	ATP-dependent helicase fft3; DNA binding, remodeler, nucleosome, Fft3-nucleosome complex, DNA BINDING PROTEIN; 5.4A {Dro	100	1.8e-41	374.01	39.7	451	922

Step 3: What Is the Function?

Evidence #2: Genomic Comparisons

- The entire DoobyDoo genome was compared with phages in the same cluster
- Function could be inferred from previous completed genomes
- Can check for any “hidden genes”

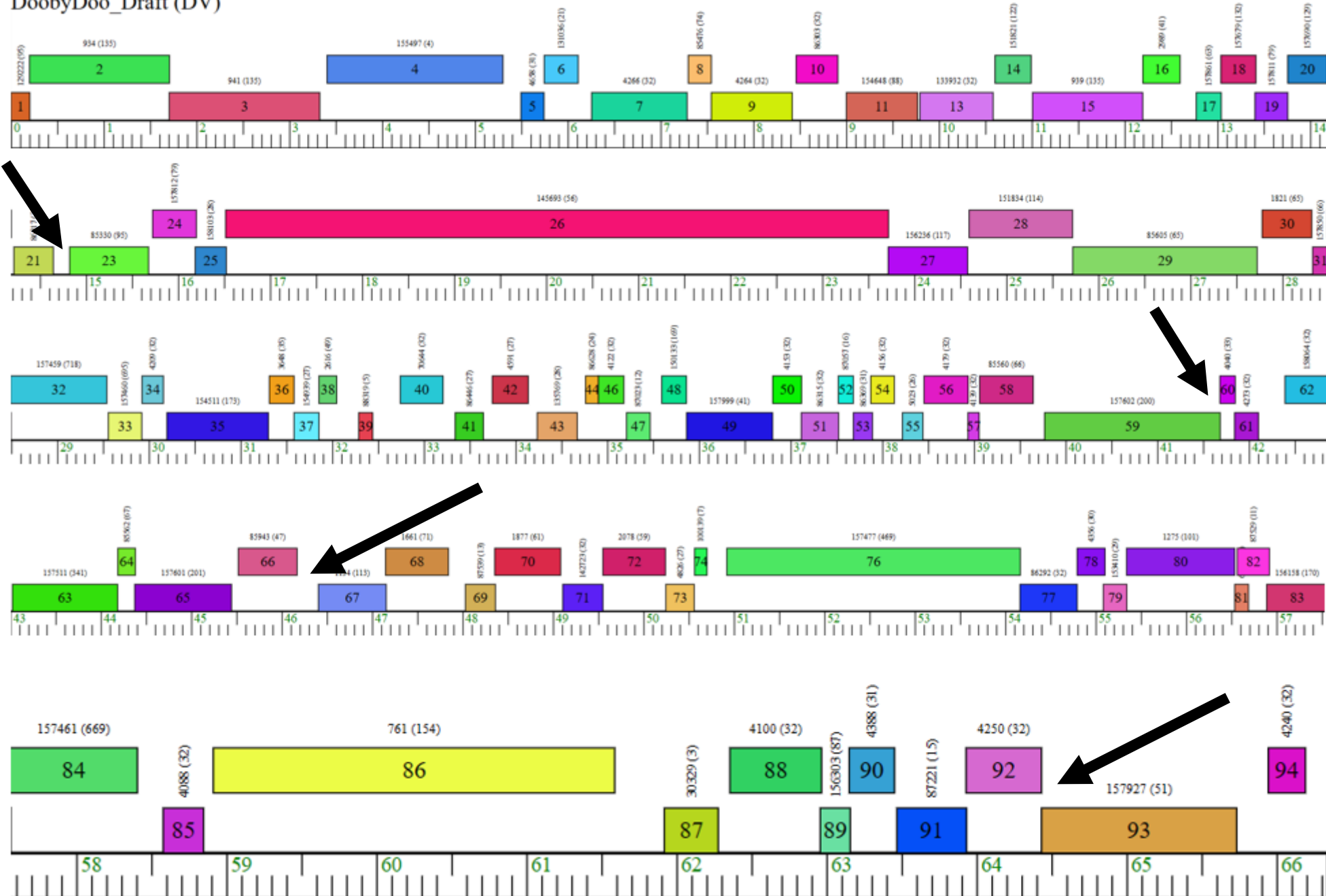
DoobyDoo_Draft (DV)

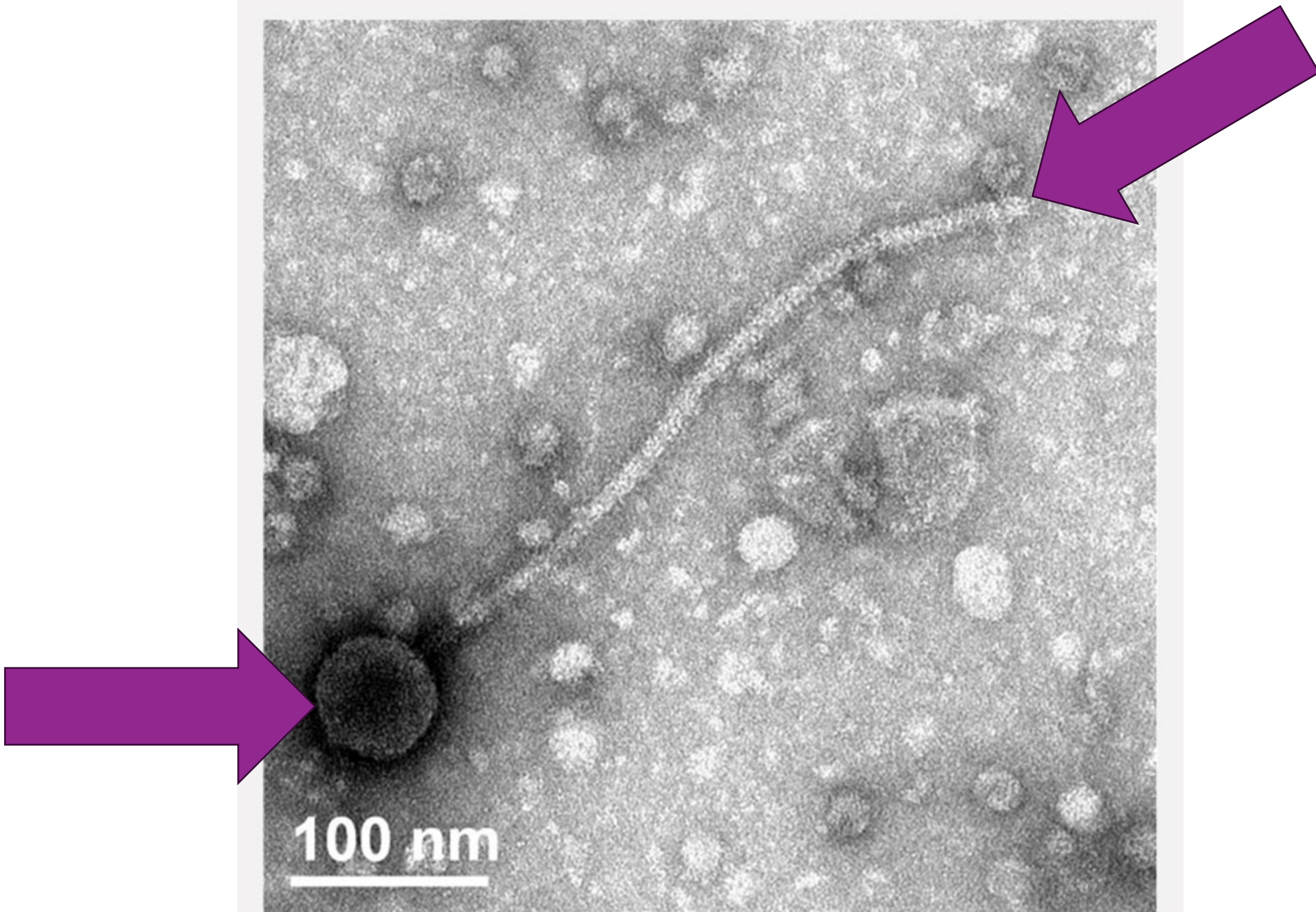


Results

- All 98 DoobyDoo genes have been manually annotated for:
 - The start site
 - The predicted function
- Located 4 hidden genes
- Most genes have hypothetical function
- Perfect overlap of genes 54/55/56; Related function?

DoobyDoo_Draft (DV)





TEM image captured from Dr. Barry Stein at the Indiana University Bloomington Electron Microscopy Center

Summary

Bacteriophage research is used to combat antibiotic resistant bacteria

In Fall 2023 DoobyDoo was found, isolated, and characterized

In Spring 2024 using bioinformatics, the genome was manually annotated

Goal is for the DoobyDoo genome to contribute to annotations of future phages/research

Future Directions

Peer-review	Finalize peer-review of each other's gene and protein annotations
Submit	Send completed annotation to SEA PHAGES for official review
Write	Write manuscript for publication

and...find more phages!

References

-
- Davey, Reginald. (2019, October 22). Ribosome Function in Cells. News-Medical. Retrieved on April 11, 2024 from <https://www.news-medical.net/life-sciences/Ribosome-Function-in-Cells.aspx>.
 - Munita, J. M., & Arias, C. A. (2016). Mechanisms of Antibiotic Resistance. *Microbiology spectrum*, 4(2), 10.1128/microbiolspec.VMBF-0016-2015. <https://doi.org/10.1128/microbiolspec.VMBF-0016-2015>
 - Pedulla, Marisa L., et al. (2003) Origins of Highly Mosaic Mycobacteriophage Genomes. *Cell*, 113(2), 171–82. *ScienceDirect*, [https://doi.org/10.1016/S0092-8674\(03\)00233-2](https://doi.org/10.1016/S0092-8674(03)00233-2)
 - Principi, N., Silvestri, E., & Esposito, S. (2019). Advantages and Limitations of Bacteriophages for the Treatment of Bacterial Infections. *Frontiers in pharmacology*, 10, 513. <https://doi.org/10.3389/fphar.2019.00513>

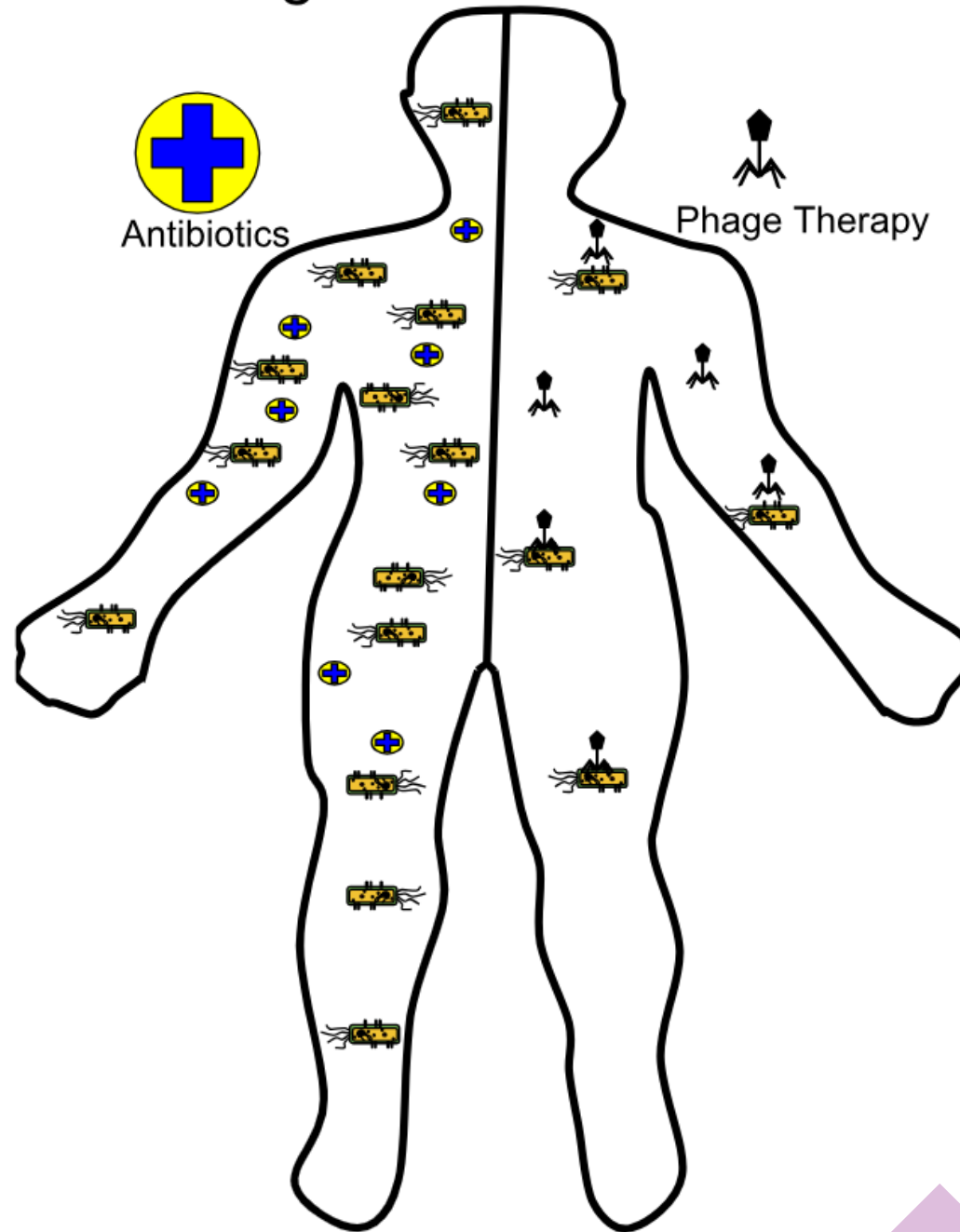


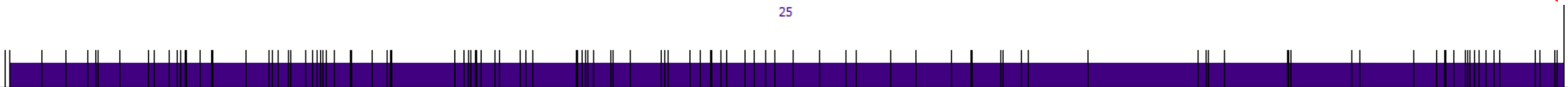
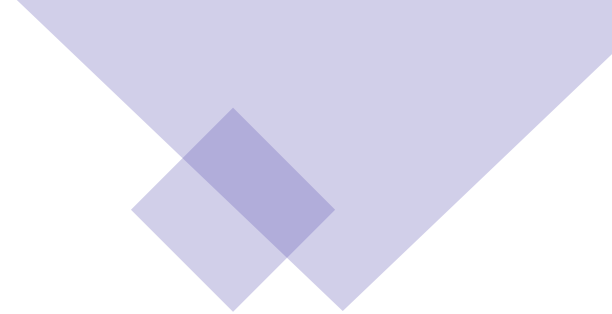
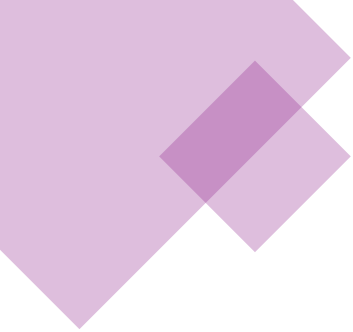
Thank You

Presenters: Ibrahim Abdulrehman and Preston Penny

Questions?

Drug Resistant Infection





25

